The Pennsylvania State University
The Graduate School
Department of Veterinary and Biomedical Sciences

PATHOGENOMICS OF NEONATAL MENINGITIS CAUSING

*ESCHERICHIA COLI*

A Dissertation in
Pathobiology
by

Dona Saumya S. Wijetunge

© 2014 Dona Saumya S. Wijetunge

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2014
The dissertation of Dona Saumya S. Wijetunge was reviewed and approved* by the following:

Subhashinie Kariyawasam
Associate Professor
Dissertation Advisor
Chair of Committee

Bhushan Jayarao
Professor

Chitrita DebRoy
Associate Professor

Edward Dudley
Associate Professor, The Department of Food Science

Anthony Schmitt
Associate Professor
Director, Pathobiology Graduate Program

*Signatures are on file in the Graduate School
ABSTRACT

Bacterial neonatal meningitis is one of the most devastating infections in early period of human life that accounts for high mortality and morbidity among infants. *Escherichia coli* is the most predominant Gram-negative bacterial pathogen associated with neonatal meningitis. Evidence from previous studies has shown a significant heterogeneity of virulence properties of neonatal meningitis causing *E. coli* (NMEC) which makes it difficult to define the NMEC pathotype and understand its pathogenesis. In this study, we have attempted to define the NMEC pathotype and identify novel virulence traits using a pathogenomic approach. First, we have evaluated genotypic and phenotypic characteristics of NMEC (n=53) in comparison to those characteristics of fecal commensal *E. coli* (HFEC, n=48) in order to identify a set of genotypic and phenotypic characteristics that can be used to distinguish NMEC from HFEC. We found that a typical NMEC can be defined as K1⁺, *sitA*⁺ and having at least two of the three genes, *vat*, *neuC* and *iucC*. Then, we sequenced the genome of prototypic NMEC strain *E. coli* RS218 and conducted a comparative genomic analysis with reference to the other sequenced extra intestinal pathogenic *E. coli* genomes. Analysis of the genome revealed that RS218 chromosome is 5.087 Mb in size with an average G+C content of 50.6%. Total of 51 genomic islands (GIs) have been identified in the RS218 genome which are absent from the *E. coli* K12 genome. Out of these GIs, 16 GIs were common to all NMEC strains whereas two GIs were common to all ExPEC. The GIs shared in all NMEC encode for several sugar uptake pathways, an acid tolerance operon, iron uptake systems and putative adhesins and invasins, indicating potential virulence associated
genes of *E. coli* RS218 genome that may contribute to NMEC pathogenesis. In addition to the chromosome, it contains one large plasmid, pRS218 which is 114,231 bp in size which belongs to the IncFIB/IIA incompatibility group, and contains a genetic load region that possesses several virulence and fitness traits such as enterotoxicity, iron acquisition and copper tolerance. Based on an *in vitro* invasion assay that used a human brain microvascular endothelial cell line and an *in vivo* assay that used a neonatal rat pup model of neonatal meningitis, the pRS218-cured strain was significantly attenuated as compared to RS218 wild-type strain. We also observed that the genes (*n*=45) located on pRS218 were overly represented in NMEC strains compared to HFEC indicating the importance of pRS218 to NMEC pathogenesis.
TABLE OF CONTENTS

List of Figures .............................................................................................................................. vii
List of Tables ............................................................................................................................... viii
Acknowledgements ..................................................................................................................... ix

Chapter 1 INTRODUCTION .................................................................................................... 1

References .................................................................................................................................. 3

Chapter 2 REVIEW OF LITREATURE ....................................................................................... 6

2.1 Neonatal Meningitis ............................................................................................................. 6
  2.1.1 Global incidence of neonatal meningitis ................................................................. 6
  2.1.2 Clinical picture, sequel and complications ........................................................... 8
  2.1.3 Bacterial pathogens associated with neonatal meningitis ................................... 11
  2.1.4 Diagnosis and management of neonatal bacterial meningitis ............................. 18
  2.1.5 Blood brain barrier and study models of bacterial meningitis .......................... 20

2.2 Neonatal Meningitis causing Escherichia coli (NMEC) ............................................... 24
  2.2.1 Clonality of NMEC ............................................................................................... 24
  2.2.2 NMEC and other extra-intestinal pathogenic E. coli .......................................... 25
  2.2.3 Current concepts of pathogenesis and virulence traits of NMEC ..................... 31

References .................................................................................................................................. 47

Chapter 3 DEFINING THE PATHOYPE OF NEONATAL MENINGITIS CAUSING E. COLI ...63

3.1 Abstract ............................................................................................................................... 63
3.2 Introduction ........................................................................................................................ 64
3.3 Materials and Methods ..................................................................................................... 65
  3.3.1 Bacterial strains and media .................................................................................. 65
  3.3.2 Phylogrouping of E. coli ...................................................................................... 66
  3.3.3 Serotyping ............................................................................................................ 66
  3.3.4 Virulence genotyping ........................................................................................... 66
  3.3.5 PFGE .................................................................................................................... 68
  3.3.6 Biofilm assay ........................................................................................................ 68
  3.3.7 Antibiotic susceptibility testing .......................................................................... 69
  3.3.8 In-vitro cell invasion assay ................................................................................... 69
  3.3.8 Statistical analysis ............................................................................................... 71

3.4 Results ............................................................................................................................... 71
  3.4.1 Phylogenetic typing of NMEC and HFEC ............................................................ 71
  3.4.2 Serotyping ............................................................................................................ 71
  3.4.3 Virulence genotyping ........................................................................................... 72
  3.4.4 PFGE .................................................................................................................... 76
  3.4.5 Biofilm assay ........................................................................................................ 77
  3.4.6 Antibiotic susceptibility testing .......................................................................... 77
  3.4.7 In-vitro cell invasion assay ................................................................................... 79
Chapter 4 GENOMICS OF *E. COLI* RS218................................................................. 92
  4.1 Abstract........................................................................................................... 92
  4.2 Introduction..................................................................................................... 93
  4.3 Materials and Methods .................................................................................. 95
    4.3.1 *Escherichia coli* RS218 strain ................................................................. 95
    4.3.2 Genomic DNA extraction........................................................................... 96
    4.3.3 Genome Sequencing .................................................................................. 96
    4.3.4 Construction of the whole genome optical map ....................................... 96
    4.3.5 Genome assembly and gap closure .......................................................... 97
    4.3.6 Genome Annotation .................................................................................. 98
    4.3.7 Identification of repetitive elements and prophage regions .................... 98
    4.3.8 Genome Annotation .................................................................................. 99
  4.4 Results ............................................................................................................ 99
    4.4.1 General characteristics of RS218 genome .............................................. 99
    4.4.2 Comparative genomics of RS218 ............................................................ 101
  4.5 Discussion ...................................................................................................... 106
  References........................................................................................................... 127

Chapter 5 THE VIRULENCE PLASMID OF *E. COLI* RS218................................................. 139
  5.1 Abstract........................................................................................................... 139
  5.2 Introduction..................................................................................................... 140
  5.3 Materials and Methods .................................................................................. 142
    5.3.1 Bacterial strains and media ...................................................................... 142
    5.3.2 Plasmid isolation, sequencing, assembly and annotation ......................... 143
    5.3.3 Comparative genomics and phylogenetic analysis .................................... 144
    5.3.4 Analysis of plasmid profiles of NMEC strains ........................................... 144
    5.3.5 Evaluation of prevalence of selected pRS218 genes in NMEC and HFEC ..... 144
    5.3.6 Plasmid curing and complementation ..................................................... 145
    5.3.7 Bacterial growth curve ............................................................................. 147
    5.3.8 *In vitro* invasion assay ............................................................................ 147
    5.3.9 Neonatal rat meningitis model ................................................................... 148
  5.4 Results ............................................................................................................ 149
    5.4.1 General properties of pRS218 ................................................................. 149
    5.4.2 pRS218 is remarkably similar to plasmids in UPEC ................................ 151
    5.4.3 Genes of pRS218 are overly represented in NMEC than in HFEC .......... 152
    5.4.4 Plasmid-cured strain demonstrated a marked attenuation ....................... 156
  5.5 Discussion ...................................................................................................... 157
  References........................................................................................................... 166

Chapter 6 SUMMARY AND FUTURE RESEARCH ................................................................ 171
  References........................................................................................................... 174
  Appendix............................................................................................................. 177
LIST OF FIGURES

Figure 2-1. Global causes of childhood deaths in 2010 .............................................................. 7
Figure 2-2. Structure of tight junction ......................................................................................... 22
Figure 2.3: Virulence factors and host factors involved in NMEC pathogenesis ...................... 33
Figure 3-1. Pulsotyping, phylogrouping, serotyping and biofilm assay ................................. 73
Figure 3-2. Distribution of virulence traits between NMEC and HFEC ................................. 75
Figure 3-3. Relative invasion frequencies observed between NMEC and HFEC ............ 79
Figure 3-4. Distribution of K1, sitA, vat, neuC, iucC among NMEC .................................. 87
Figure 4-1. Circular map of the chromosome of E. coli RS218 .............................................. 101
Figure 4-2. Pairwise alignment of E. coli RS218 genome to E. coli K12 .......................... 102
Figure 4-3. Characteristics of bacteriophages in E. coli RS218 genome .............................. 104
Figure 4-4. Distribution of RS218 genomic islands (GIs) ExPEC. ........................................ 105
Figure 4-5. Clustering of RS218 genomic islands (GIs) among ExPEC strains ................. 105
Figure 5-1. Graphical circular map of pRS218 ........................................................................ 150
Figure 5-2. Comparison of pRS218 sequence to some virulence plasmids of other E. coli ... 153
Figure 5-3. Plasmid profiles of some NMEC strains ............................................................... 153
Figure 5-4. Evolutionary relationship of IncFIB /IIA plasmids in pathogenic E. coli ........ 154
Figure 5-5. Confirmation of pRS218 curing ............................................................................ 159
Figure 5-6. Evaluation of virulence potential of pRS218 in vitro and in vivo ..................... 160
LIST OF TABLES

Table 2-1. Neonatal sepsis: an international perspective. .................................................. 12

Table 2-2. Capsular types and characteristics of E. coli.................................................. 34

Table 3-1. Distribution of different phylogroups between NMEC and HFEC. .................... 72

Table 3-2. Distribution of O and H antigen types among NMEC and HFEC ....................... 74

Table 3-3. Prevalence of virulence genes between NMEC and HFEC ............................... 76

Table 3-4. Antibiotic profiles of NMEC and HFEC ........................................................ 78

Table 3-5. Virulence traits distributed among NMEC and HFEC .................................... 82

Table 4-1. General characteristics of E. coli RS218 genome ........................................... 100

Table 4-2. Characteristics of genomic islands identified in E. coli RS218 genome .......... 103

Table 4-3. Comparison of E. coli RS218 genome with other ExPEC genomes ................. 105

Table 4-4. Genomic islands of E. coli RS218 ................................................................. 123

Table 5-1. Virulence and fitness associated traits in pRS218 .......................................... 151

Table 5-2. Prevalence of pRS218 genes among NMEC and HFEC ................................. 154
ACKNOWLEDGEMENTS

First and foremost, I wish to express my deepest appreciation to my supervisor, Dr. Subhashinie Kariyawasam, for accepting me to her lab at the first place, and for the constant support, guidance and encouragement. Thanks for the risk you have taken by having a student without much background knowledge in molecular biology and teaching every bit of it carefully and being patient! Your perpetual confidence in my abilities was at times the only remaining spark to fuel my efforts.

Also, I am extremely grateful to my thesis committee members; Prof. Bhushan Jayarao who helped me from the beginning with guiding me through selecting required courses and his continuous support on my research, Dr. Chitrita DebRoy for her altruistic guidance and support with results analysis and scientific writing and Dr. Edward Dudley for his technical insights, generous help and constructive criticism for successful completion of this thesis.

Also, I have had the fortune to become accustomed with very generous and talented group of professionals at the Department of Veterinary and Biomedical sciences. Sincerely, I’m thankful to Dr. Thomas Denagamage for all the support, guidance and care from the moment I reached to State College five years ago until now. He helped me tremendously not only with the lab work but also taught me how to survive and adapt to this new environment. Also, I’m grateful to Prof. Vivek Kapur and his lab members, Ling Lin, Ro, and Becky for their kind support in genome sequencing, Dr. Dunn, Dr. Hattel, and Dr. Pendleton at the Animal Diagnostic Laboratory for their assistance with the microscopy, Mrs. Valerie Linter, Ms. Traci Pierre, Mrs. Tammy Matthews and Mrs. Ellen King from the bacteriology section for tremendous amount of technical support, Mrs. Beth Roberts at E. coli reference center for helping me with PFGE and serotyping. Likewise, I would like to give my thanks to Mr. Derek Shaffer for helping me whenever I need help with computers and software. Also, my sincere thanks to Sarah Snyder and Kristin Cox at the front office for their kind support with paper works and encouragement throughout the time I have been here.

Many thanks to my current and former lab members, Dr. Atul Chaudhari, Eranda, Christine, Sudharsan, Megan and Matt for all your help in numerous ways whenever it was needed. Thanks for all the moments we shared together, it was really nice working with you all.
Also, I would like to give my sincere thanks to Dr. Eranda Rajapaksha and Dr. Saumya Wickramasinghe for introducing me to Dr. Kariyawasam at the first place and their support and encouragement.

I would also like to thank all my friends at State College, especially for dear friend Dr. Donald Henderson for his love and care during all these years. To all the friends from Sri Lankan community live in State College, for their love and care, and the remembrance of our motherland you bring about. Special thanks to Damitha and Padma, Sarath and Sasie, Sirapala and Ranjani and Dr. Yapa for their encouragement, love and care. Likewise, I would like to acknowledge the moral and spiritual support from the Monks at Pittsburgh Buddhist Center during this time.

I would also like to thank my family: my loving mother, Chandralatha Wijetunge and my father, Sarath Wijetunge. Your endless love, guidance and, motivations brought me this far. Your care and support were the ultimate energy to my journey in education. I am who I am now because of you! Also, my dearest mother-in-law, father-in-law, for all of the love and care that you have given me and your blessing was what sustained me this far.

Words cannot express how grateful I am to my husband, Eranda, for all of sacrifices that you have made on my behalf and who has, without hesitation, whole-heartedly supported me unconditionally during good and bad times. I truly thank you for standing by my side, even when I was irritable and depressed. I feel that what we both learned a lot about life and strengthened our commitment and determination to each other and to live the life to the fullest. Thank you and love you!

Finally, I must say that I was brought up by free education system of Sri Lanka. I honestly bow my head towards my nation in thanks.
Dedication……

To My Loving Parents and My Husband
Chapter 1

INTRODUCTION

Neonatal sepsis and meningitis is the third most common disease in neonates that accounts 0.393 million deaths worldwide (1). The incidence of bacterial meningitis in infants is about 0.3 to one case per 100,000 live births per year in developed countries and it is ten times higher in underdeveloped countries (2-4). Among many bacterial pathogens associated with neonatal meningitis (NM), Neonatal Meningitis causing *E. coli* (NMEC) has been identified as most predominant Gram-negative pathogen (5, 6). Despite advanced antimicrobial therapy and supportive care, mortality and morbidity rates of NM due to NMEC continue to be as high as 30-50 % (4). Besides the higher mortality, a large number (30 to 50%) of surviving infants manifest permanent neurological dysfunctions such as hearing loss, convulsive disorders, abnormal speech patterns, cortical blindness, and mental retardation (7). Furthermore, recent reports have shown the emergence of antibiotic-resistant strains of NMEC that challenge existing therapeutic strategies.

Transmission, port of entry and pathogenesis of NM is complex and is poorly understood (6). Neonates are thought to acquire NMEC strains that are present in mother’s gastrointestinal tract just before (vertically transmitted) or during (perinatally acquired) birth. Vaginal colonization of pathogenic *E. coli* has been observed in 3-20% of pregnant women (8). Molecular epidemiologic investigations of mother-infant pairs have demonstrated that the same strains of *E. coli* were isolated from the cerebrospinal
fluid (CSF) and stools of the infant with meningitis, and stools of the mother [8]. Bacteria are thought to enter the blood stream through the intestines or nasal mucosa to cause bacteremia, and then cross the blood-brain barrier (BBB) to cause meningitis and pleocytosis of CSF (9). It has been shown that E. coli can cross the BBB without affecting the integrity of the human brain microvascular endothelial cells (HBMEC) or the permeability of the BBB (9). Crossing of the BBB by circulating bacteria requires a high degree of bacteremia, and bacterial adherence to and invasion of the HBMECs.

Several factors have been previously implicated in bacterial attachment to and invasion of HBMECs, which include FimH of the type 1 pili, flagella, new lipoprotein I (NlpI), OmpA, cytotoxic necrotizing factor 1 (CNF1), invasion of brain endothelial cell (Ibe) proteins and arylsulfatase-like protein (AslA) (5, 10-14). However, most of these virulence factors are not universally present in NMEC, and other than the genes encoding FimH and K1 capsule which are present in between 30-75% of NMEC tested, the other virulence markers are present in a relatively small fraction of NMEC strains (15). Additionally, it is has been shown that NMEC strains harbor genes present on virulence plasmids of other extra-intestinal pathogenic E. coli (ExPEC) (15). These data indicate the difficulty of defining the NMEC pathotype as well as the necessity of characterizing novel virulence traits in NMEC pathotypes to better understanding of its pathogenesis.

The advent of whole-genome sequencing of bacterial pathogens has empowered the rapid progress in understanding the genetic basis of bacterial pathogens (16, 17). Specifically, it has been used to identify novel virulence factors encoded in chromosome or plasmids that transform nonpathogenic bacteria into pathogenic organisms with potential to invade and survive in host tissues. Currently, many E. coli genomes
representing different pathovars and commensals have been sequenced and available for
the public to use in such research strategies. Among them three NMEC strains, namely,
*E. coli* CE10, S88 and IHE3034 have been sequenced (18, 19). However, these strains
have not been studied in relation to NMEC pathogenesis. Nevertheless, genome sequence
of prototype strain *E. coli* RS218 which has been used to understand the NMEC
pathogenesis since 1980’s is not available to the public. Therefore, availability of the
complete genome sequence of RS218 strain may bolster the identification of novel
virulence determinants which are critically needed for a better understanding of the basic
mechanisms of pathogenesis and the rational development of methods for treatment and
prevention of neonatal meningitis.

The objectives of the present study are to identify phenotypic and genotypic
characteristics that distinguish NMEC pathotype from commensal *E. coli*, whole genome
sequencing of the NMEC prototype strain RS218 to reveal the potential virulence traits,
and to identify the genes that are conserved in NMEC and may contribute to NMEC
pathogenesis.

**References**


Chapter 2

REVIEW OF LITREATURE

2.1 Neonatal Meningitis

2.1.1 Global incidence of neonatal meningitis

Despite the advanced therapeutics, infant intensive care and prevention strategies, infections in the neonatal period account for over one million deaths annually worldwide (1). According to the recent data published in 2012 by the Child Health Epidemiology Reference Group of World Health Organization (WHO) and United Nations Children’s Emergency Fund (UNICEF), 3.7 million annual deaths of children under 5 years old in 2010 occurred in neonatal period (2). The most common reasons for neonatal deaths are preterm birth complications, intrapartum related complications, and neonatal sepsis and meningitis. As such, neonatal sepsis and meningitis is the third most common disease in neonates that leads to 0.393 million deaths worldwide (Figure1).

Incidence of neonatal meningitis and sepsis varies over time and different geographical regions. Until 1980’s, the incidence of neonatal bacterial meningitis was about 1 per 1000 live births (LBs) in industrialized countries (3, 4). However, Louis et al. in 1994 reported that the incident rate had been decreased to 0.22 -0.37 per 1000 LBs in England and Wales (5). Similarly, Babara et al. in 1995 reported a decline in deaths due to neonatal sepsis and meningitis in the United States from 1979 to 1994. This study,
which was undertaken over a period of 16 years revealed that the neonatal deaths due to sepsis has been decreased from 50.5% to 30% the during the study period. Further, a national cohort study done by Babara et al. 1999, which covered all the regions of the United States revealed that the overall incidence of bacterial sepsis and meningitis in newborns during the period between 2006 and 2009 was 0.98 cases per 1000 LBs indicating that the incident rate is fairly constant despite the death from neonatal sepsis and meningitis has been reduced over the past few decades (6).

**Figure 2-1.** Global causes of childhood deaths in 2010. Source: The Lancet Volume 379, Issue 9832 2012 2151 – 2161.

According to a report published by WHO in 1999, 98% of neonatal meningitis cases occurred in developing countries (7). In sub-Saharan Africa, the mortality rate of neonatal meningitis is 0.7-1.3 deaths per1000 LBs. A community based study done in Pakistan revealed the incident rate of neonatal meningitis was 0.8/1000 LBs. A similar study revealed the incidence rate of neonatal meningitis in India was 4.9/1000 LBs. In
Brazil, the incident rate reported was 4.2/1000 LBs whereas in Guatemala the rate is much higher as 6.1 / 1000 LBs (8). However, it is believed that the true incidence rate of neonatal meningitis is way much higher than the reported cases due to difficulties in hospital settings, differences among studies and unregistered deaths due to poor health care systems. (8).

Despite the reduction of mortality rate and incidence of neonatal meningitis during the past few decades, neonatal bacterial meningitis still remains as one of the major infections of neonates with significantly higher mortality and morbidity rates compared to other neonatal infections (2). Besides the high mortality and morbidity rates, neurological complications during the course of disease and long term neurological sequale among the survivors are other concerns associated with bacterial meningitis.

2.1.2 Clinical picture, sequel and complications

World Health Organization defines the neonate as a new-borne child who is less than 30 days of age. By definition, neonatal bacterial meningitis is a bacterial infection of leptomeninges (dura mater, pia mater and subarachnoid) covering the brain during the neonatal period. Typically, bacterial meningitis produces a suppurative inflammation in the meninges which commonly involves the cerebral hemispheres leading to meningoencephalitis. Depending on the time of onset of symptoms, neonatal sepsis and meningitis can be categorized into two basic clinical syndromes; early onset disease (EOD) where the symptoms occur less than 72 hours after birth and late onset disease
(LOD) where symptoms occur after 72 hours of birth (9). There are variations in clinical signs and causative agents associated with these two subtypes of the disease (10).

Early onset disease is usually associated with obstetrical complications and preterm births (10). The mode of transmission is from mother to the baby during the process of birth. In EOD, clinical presentation is frequently manifested as non-neurological symptoms. These non-pathognomic signs may be subtle or sometimes cannot be detected until later stage of the disease. Most common clinical signs ans symptoms in EOD are body temperature instability (fever or hypothermia), episodes of apnea or bradycardia, hypotension, jaundice, feeding difficulty (refuse foods or vomiting), and irritability (11, 12). Occasionally stupor or lethargy can be seen during EOD indicating a neurological involvement of the disease.

In comparison to EOD, LOD is not usually associated with obstetrical complications and prematurity (10). Most of the LOD patients are full-term infants that show clinical signs after the first week of life. Mode of transmission of LOD may be from mother to the infant. Other possible modes are human contacts, indwelling catheters or contaminated equipment (nosocomial). Late onset disease symptoms are predominantly neurological. More than 75% cases of LOD have manifested ‘stupor’ which is characterized by a marked diminution in the capacity to react to environmental stimuli (10). Other neurological symptoms such as seizures, bulging of anterior frontanelle or soft spot, extensor posture or opisthotonos, focal cerebral signs including gaze deviation and hemiparesis, cranial nerve palsies, and nuchal rigidity are also common with LOD.
Clinical complications of bacterial meningitis also play a pivotal role in severity and mortality rates of bacterial meningitis. Neurological complications are mainly due to uncontrolled inflammatory process against the bacterial infection rather than the damage caused by bacteria themselves. The predominant forms of neuronal injury are necrotic cortical injury, apoptotic hippocampal injury, occlusive vasculitis and thrombosis which lead to several neurological complications in patients (10, 13). The major complication observed is increased intra-cranial pressure which is manifested by bulging of anterior frontanelle or soft spot in the patients. Increased intra-cranial pressure is mostly due to pathological processes of ventriculitis which leads to ventricomegaly, subdural effusions, acute hydrocephalus, intra-cerebral hemorrhages, cerebral edema or effusions and abscess. Hydrocephalus is more common with type III Group B streptococci- (GBS) and *Escherichia coli* K1- associated NM than with other capsular groups of GBS and non-K1 *E. coli* infections, whereas brain abscesses are common in *Citrobacter* associated meningitis (12).

Despite the mortality, long-term adverse neurological outcomes among survivors are an important crisis associated with NM. Neurological sequel is present among 20-58% survivors (11). Severity and type of outcome of the sequels depend on the bacterial pathogen associated with NM and the area of neuronal damage (13). The common neurological sequels reported are mental retardation, hydrocephalus, seizures, cerebral palsy, developmental delay, hearing loss, learning problems, language disorders, vision impairment and behavioral impairments. Among these, bilateral hearing loss is the most common disability associated with NM and reports have shown that the presence of
permanent deafness among 5-30% of surviving patients with auditory nerve inflammation (11-13). Compared to GBS meningitis, survivors of Gram-negative bacterial meningitis, particularly *E. coli* had been shown to manifest more adverse sequelae (12, 14). For example, Pong *et al.* reported that 58% of survivors of *E. coli* meningitis showed more neurological sequelae than that of survivors of GBS infections (15%) (12). However, severe disabilities were more commonly associated with GBS infections than that of *E. coli* infections (14, 15). Instead of aforementioned sequelae, nonspecific conditions such as middle ear diseases, epilepsy, diabetes and asthma had also been reported among the survivors (14).

### 2.1.3 Bacterial pathogens associated with neonatal meningitis

Many Gram-positive and Gram-negative bacterial pathogens have been reported as causative agents of neonatal bacterial meningitis. Among these, most commonly encountered pathogens are Group B Streptococci, *Escherichia coli* and *Listeria monocytogenes* (14, 16, 17). Other than these organisms, Group D Streptococci, coagulase-negative staphylococci, *S. aureus*, *Citrobacter*, *Acinetobacter*, *Klebsiella*, *Enterobacter* spp. and *Salmonella* spp. have also been isolated from patients with NM (17). Atypical bacteria have also been reported to cause NM in some cases in particular developing countries (Table 2.1) (10).
Table 2-1. Neonatal sepsis: an international perspective. Source: Arch Dis Child Fetal Neonatal Ed 2005.

<table>
<thead>
<tr>
<th>Country</th>
<th>Type of study</th>
<th>Duration of study (months)</th>
<th>Total No of positive blood cultures</th>
<th>Early onset EOS %</th>
<th>Early onset Mortality %</th>
<th>Late onset LOS %</th>
<th>Late onset Mortality %</th>
<th>Most common isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malaysia</td>
<td>Prospective surveillance</td>
<td>0 (1001)</td>
<td>136</td>
<td>25 (35/136)</td>
<td>12</td>
<td>74 (20/69)</td>
<td>18</td>
<td>Acinetobacter, Klebsiella</td>
</tr>
<tr>
<td>Kenya</td>
<td>Prospective and retrospective survey</td>
<td>6 (1997 –8)</td>
<td>121</td>
<td>30 (21/69) (72 h)</td>
<td>4</td>
<td>30 (&gt;72 h)</td>
<td>10</td>
<td>Klebsiella, Citrobacter</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Prospective surveillance</td>
<td>11 (1994 –5)</td>
<td>62</td>
<td>47</td>
<td>8</td>
<td>53</td>
<td>5</td>
<td>Staph aureus, Pseudomonas</td>
</tr>
<tr>
<td>India</td>
<td>Prospective Surveillance</td>
<td>6 (1997)</td>
<td>96</td>
<td>50</td>
<td>9</td>
<td>50</td>
<td>4</td>
<td>Staph aureus, Klebsiella</td>
</tr>
<tr>
<td>Panama</td>
<td>Surveillance, retrospective</td>
<td>216 (1975 –92)</td>
<td>577</td>
<td>47 (&lt;5 days)</td>
<td>44</td>
<td>53 (&gt;5 days)</td>
<td>22</td>
<td>Klebsiella, Staph aureus</td>
</tr>
<tr>
<td>India</td>
<td>Surveillance</td>
<td>15 (1996 –7)</td>
<td>157</td>
<td>85 (6 days)</td>
<td>49</td>
<td>14 (&gt;6 days)</td>
<td>68</td>
<td>Klebsiella, Pseudomonas</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>Case control study</td>
<td>60 (1983 –8)</td>
<td>61</td>
<td>39</td>
<td>21</td>
<td>61</td>
<td>24</td>
<td>Staphylococci, Klebsiella, Enterobacter, Serratia</td>
</tr>
<tr>
<td>India</td>
<td>Prospective surveillance</td>
<td>24 (1995 –6)</td>
<td>131</td>
<td>23</td>
<td>4</td>
<td>77</td>
<td>10</td>
<td>Klebsiella, Enterobacter fecals</td>
</tr>
<tr>
<td>The Gambia, Papua New Guinea, Philippines, Ethiopia</td>
<td>Multicentre study (4 prospective surveillance studies)</td>
<td>Each study conducted over 24 months (1990 –1993)</td>
<td>167 (84 in the neonatal period)</td>
<td>30% (7 days)</td>
<td>n/a</td>
<td>70%</td>
<td>n/a</td>
<td>Staph aureus, Streptococcus pyogenes, E coli</td>
</tr>
</tbody>
</table>

Percentage of cases by age of onset and most common isolates. If not otherwise stated, early onset (EOS) is defined as first 48 hours, and late onset (LOS) as more than 48 hours.

*The data in the table refer to the blood culture results of the newborn (0–1 months old) of the four studies. In The Gambia the most common pathogen was Staph aureus; in Papua New Guinea S pyogenes; in the Philippines Salmonella, and in Ethiopia S pyogenes and E coli.

GBS, Group B streptococcus. All the data refer to blood cultures only. EOS %, EOS/all positive blood cultures × 100; LOS %, LOS/all positive blood cultures × 100. Mortality %, EOS/all positive blood cultures × 100 and LOS/all positive blood cultures × 100; n/a, not available.
a. Group B Streptococci (GBS)

*Streptococcus agalactiae* or Group B streptococci are Gram-positive bacteria that form β-hemolytic mucoid colonies on sheep blood agar plates. They were first isolated from vaginal microflora of a post-partum mother by Lancefield and Hare in 1935 (18). Group B streptococci inhabit as commensals in the gastrointestinal tract of humans and vaginal microflora of 10-40% of pregnant women (19-21). Also, it is a constituent of microflora of lower gastrointestinal and upper respiratory tracts of newborns (19, 21). Colonization of GBS in neonates occurs as *in-utero* ascending infection from the maternal genital tract microflora or direct contact during delivery (22, 23). The first case NM associated with GBS was reported in 1958 by Nyhan (24). Group B Streptococci are the most common pathogen associated with neonatal bacterial meningitis particularly in the United States and the Western Europe (22). The incidence of neonatal GBS infection in these countries is 1.8 per 1000 LBs (25). According to the recent report by Centers for Disease Control and Prevention (CDC), the incidence of GBS meningitis has been declined to 0.7 per 1000 LBs due to the universal screening of GBS colonization in the urogenital tract of pregnant mothers and prophylactic antibiotic treatment to prevent vertical transmission of GBS to newborns (26).

Early onset GBS meningitis occurs due to vertical transmission (prenatal or perinatal) of GBS from mother to baby whereas late onset disease is due to nosocomial transmission. Mortality of GBS associated EOD varies from 11 to 50% (23). Group B Streptococci associated LOD is relatively less fatal accounting only 3% mortality in affected neonates (23). Among the different capsular types of GBS, capsular type III is the most predominant type encountered in NM although types Ia and V have also been
reported (27). Early onset disease which represents 78-84% of GBS meningitis cases is known to cause by capsular types Ia, III and V whereas LOD is usually associated with capsular type III (27).

Risk factors associated with early-onset GBS infection of neonates include pre-term delivery, prolonged rupture of membranes, intra-partum fever and peri-partum Group B streptococcal infections such as urinary tract infections (UTIs), chorioamnionitis, and peripartum bacteremia in pregnant women (26).

b. *Escherichia coli*

In 1885 the pediatrician Escherich isolated a bacterial species that he called "*Bacterium coli commune*" from the feces of a healthy baby. This was the first description of *E. coli*, which to the present day is recognized as a largely diverse group of bacteria that inhabits the lower intestinal tract of most warm-blooded animals as commensals as well as a well-known pathogen associated with variety of infections in humans and animals. Pathogenic nature of *E. coli* can be attributed to their acquisition of novel virulence traits by horizontal gene transfer. There are two types of *E. coli* infections among humans and animals; enteric diseases and extra-intestinal diseases.

*Escherichia coli* has been identified as the most predominant Gram-negative bacterial pathogen associated with NM, which is only second to Gram-positive GBS meningitis (10, 13, 28). Among different capsular serogroups of *E. coli*, K1 is the most common capsular serogroup associated with NM which represents approximately 80% of *E. coli* strains isolated from neonates with meningitis and 40% of strains causing bacteremia or sepsis (29). Association of *E. coli* K1 with NM had been reported since
The K1 capsular polysaccharide antigen is highly homologous to the capsular antigen of group B Neisseria meningitides indicating a common pathogenic mechanism between these two unrelated bacterial species (31, 32). It has been estimated that the prevalence of E. coli K1 associated NM is around 1 in 100 to 200 among infants that carry E. coli K1 in their gut. According to the surveillance data analyzed by National Institute of Child Health and Human Development Neonatal Research Network involving 16 U.S. academic neonatal centers, showed that the prevalence of E. coli K1 meningitis has been increased from 3.2 to 6.8 cases per 1000 LBs during the period between 1998 and 2000 (33). This increase was attributed to the implementation of extensive antibiotic prophylaxis particularly with β-lactam antibiotics, during prenatal period for the prevention of GBS infections which selectively enhances the survival of Gram-negative bacilli which are frequently resistant to penicillin. However, some evidence suggests that intra-partum antibiotic prophylaxis has not been associated with a concomitant increase of E. coli meningitis in neonates. Instead, full-term infants exposed to intra-partum antibiotic therapy for 4 hours or greater had shown decreased odds of early onset E. coli meningitis (5).

There are many risk factors associated with occurrence of E. coli meningitis. Among those, prematurity has been identified as a major risk factor associated with E. coli-induced meningitis. Orskov et al., 1985 observed that rectal colonization of E. coli K1 in pre-term neonates were relatively higher than that of full-term babies (34). Acquisition of E. coli K1 for early colonization of rectum in the neonates is mainly through the exposure to genitourinary and fecal microflora of the mother during birth (34, 35). Therefore, maternal microflora is a major source of infection in neonates. However,
some evidence suggest that equipment and nurses (nosocomial) also play a major role in *E. coli* colonization in newborns particularly at a later stage of neonatal period and subsequent infections (10). Other risk factors include maternal intra-partum infection such as urinary tract infections with *E. coli* K1, prolonged rupture of membranes, indwelling catheters and other complications during labor (10). Although the severity of the sequel is less with *E. coli* K1 than with GBS, many studies have demonstrated a higher incidence of neurological sequale among the survivors of *E. coli* meningitis (30-50%) than those with GBS infections (5-15-%) (14).

c. *Listeria monocytogenes*

*Listeria monocytogenes* is a Gram-positive bacterium that naturally inhabits environmental niches like soil and water as well as gastrointestinal tract of most animals. In humans, it is mainly associated with food-borne infections particularly in immunocompromised individuals, children and pregnant women. Mode of transmission of *Listeria* is through contaminated unpasteurized milk, soft cheese, deli meat, fruits and vegetables (36).

The prevalence of listeriosis among general population is 0.7/100,000 people whereas in pregnant women the prevalence is 12/100,000 which is 17-fold higher than that of the general population (37). Higher prevalence of listeriosis among pregnant women is mainly due to the suppressed Th1 cell activity during pregnancy which is required for defense against Listeria. According to a report published by CDC, 16% cases of listeriosis occurred during the period between 2004 and 2007 was pregnancy related, and it was more common during the third trimester (38). Even though the mother shows
mild flu-like symptoms, listeriosis has detrimental effects to the fetus as a result of abortion, still birth, premature births and neonatal infections (39).

*Listeria monocytogenes* is the third most common bacterial pathogen associated with NM which accounts for 5-20% of NM cases in western countries (17). As mentioned before, neonates acquire *Listeria* through maternal infections which represent two thirds of listerial meningitis in newborn. The mode of transmission can be either by inhalation of infected amniotic fluid or transplacentally from the maternal circulation to fetus (17). Other than the *in-utero* transmission, colonization during parturition or environmental origin also has been reported. Early onset disease is the most common form of listeriosis in newborns as a result of *in-utero* transmission; however, it has also been rarely associated with LOD in healthy babies without maternal infections and complications during delivery (17).

As for many other infections due to *Listeria* in human, the major serotype of *L. monocytogenes* associated with NM is 4b which represents 80% of the cases. Rest of the cases are caused by serotypes I and III (40). Clinical signs of listerial meningitis are more similar to GBS infections except very high mortality rates of 3-50% associated with *Listeria*. Additionally, granulomas (‘Granuloma infantispticemia’) in skin, liver, lungs, kidneys and brain are other findings exhibited by patients, particularly those with septicemia and meningitis (38). Prognosis is poor with EOD which has a mortality rate of 15-50%, compared to LOD which has only 10-20% mortality rate among the affected neonates (10).
2.1.4 Diagnosis and management of neonatal bacterial meningitis

Prompt diagnosis and treatment is required to reduce the adverse outcomes of neonatal bacterial meningitis (NBM). Since the clinical signs of NM are subtle and nonpathognomic, a diagnosis cannot be made with the clinical signs alone. Until laboratory confirmation is available, initial diagnosis is done with the clinical signs, risk factors associated with the patient such as low birth weight, peri-partum complications or prior history of mother having babies with meningitis, etc. Definitive diagnosis of NBM is done by cerebro-spinal fluid (CSF) examination via lumbar puncture (LP). Acute bacterial meningitis is indicated by the presence of bacteria along with abnormal CSF parameters such as reduced glucose level, increased white blood cells especially polymorphonuclear cells and increased protein level (13). Computed tomography (CT) and magnetic resonance imaging (MRI) scanning also can be used as indirect diagnostic methods to trace the cerebral inflammation, abscessation and ventricomegaly (10). Identification of bacteria using polymerase chain reaction has also been implemented particularly for GBS infections.

The first choice of treatments for the NBM is antimicrobial therapy to clear the bacteria from the central nervous system (CNS). The major constraint in antibiotic treatment in NBM is achieving appropriate concentration of antibiotics in CSF for complete clearance of bacteria due to the restricted permeability of blood brain barrier (BBB) for xenobiotics. Many clinical trials and studies have shown that the Minimal bactericidal concentration (MBC) of antibiotics is 10-30 times higher in CSF than that of in-vitro MBCs (17). Additionally, action of some antibiotics such as aminoglycosides can
be varied from bactericidal to bacteriostatic due the presence of low pH and higher protein levels in CSF during the infectious processes. Therefore, the goal of current antibiotic therapy is to achieve an adequate concentration of the antibiotic in CSF to kill bacteria while minimizing the toxic effects of the drug itself.

Given the possibility of higher mortality and neurological complications, ‘empirical antibiotic treatment’ before the identification of particular pathogen associated with the disease in the individual is the initial approach of the treatment strategy. Selection of antibiotics in the empirical treatment depends on the onset of the disease. For example, if the disease onsets early, ampicillin with either an aminoglycoside or cefotaxime is commonly recommended. Whereas in LOD, nafcillin or vancomycin, plus cefotaxime or ceftazidime with or without an aminoglycoside are recommended (41). Once the proper causative agent has been identified, treatment protocol and duration of the treatment is dependent on the pathogen involved and its antibiotic susceptibility pattern. Since most Gram-positives are sensitive to β-lactam antibiotics, GBS and Listeria meningitis are treated with a high dose of penicillin in combination with an aminoglycoside for 14 to 21 days (41). Gram-negative bacillary meningitis is treated with cephalosporins in combination with aminoglycosides for at least 21 days for the complete clearance of bacteria from the CSF. A recent study indicate the emergence of CTX-M-15-producing *E. coli* which renders the limitation of current treatment protocols for *E. coli* meningitis (42).

In spite of an appropriate antibiotic treatment, adjunctive therapy for management of NM is necessary to improve the survival rate in affected babies and to reduce neurological complications in survivors. The aims of adjunctive therapy are to reduce
both bacteriological and inflammation-induced damage to the brain tissues. These include parenteral fluid therapy with essential nutrients as a supportive care to alleviate the shock which may develop during early progression of the disease (41). Although the adjunctive therapy including dexamethasone has been shown to improve the survival of patients, routine use of dexamethasone is currently not recommended due to possible hippocampal apoptosis that leads to long-term neuropsychological and cognitive squeal in survivors (43). Osmotic diuresis with mannitol or glycerol is a common practice particularly with patients having elevated intra-cranial pressure. Other treatments include granulocyte transfusion, exchange transfusion and intravenous immunoglobulin administration or pathogen-specific polyclonal or monoclonal antibody.

2.1.5 Blood brain barrier and study models of bacterial meningitis

Blood brain barrier is a dynamic barrier that separates the CNS from the circulation which impedes and controls the influx of blood constituents to the brain. Existence of barrier between the CNS and periphery was first discovered by Paul Ehrlich in 1880s and later by Goldman who observed that the inability of an aqueous dye to diffuse from circulation to the brain and vice versa when it was injected via intravenous or intra-thecal route (44). The term BBB was first introduced by Lewandowsky in 1898 who observed that the toxic effects of neurotoxins were seen only when they were injected to the brain but not to the circulation (45). After several decades, the structural component involved in BBB was identified as brain microvascular endothelium. The brain microvascular endothelial cells (ECs) are distinguishable from
other vascular endothelia by the following structural differences; (i) absence of fenestrations, (ii) presence of intercellar tight junctions (TJ), (iii) low level of pinocytosis and transcytosis, (iv) polarized presence of active transport systems on the apical surface of cell membrane, and (v) harboring high levels of mitochondrial in the cells (46). These characteristics of BBB ensure the restriction of movements of hydrophilic compounds, toxic compounds, ions and blood cells to the ECs while allowing influx of nutrients and oxygen from blood and efflux of toxic chemicals from the cells. Tight junctions are the main structural element in the BBB. As demonstrated by electronmicrography, TJs are consisted of intra-membranous strands or fibrils on ECs that bind to the adjacent ECs forming a tight junction (Figure 2.2) (47). The major proteins involved in the TJ formation are occludins, claudins, junctional adhesion molecules (JAM) and cytoplasmic accessory proteins or zona occludins (ZO). Occludin and claudins are proteins having three transmembrane domains at the lateral surface of ECs with extracellular domain that bind with the same proteins on adjacent cells and intra cellular tail which binds to ZO and to the actin molecules in the cytoskeleton to ensure structural integrity of the ECs. Junctional adhesion molecules are filamentous proteins that stabilize the TJ and ZOs facilitate the connection of TJ fibrils to the cytoskeleton of the ECs to prevent the paracellular diffusion of compounds from the blood. Apart from TJs, basal lamina, higher number of pericytes that cover the vasculature and astrocytes underneath the pericytes also play a role in maintaining the integrity of BBB. All these components are essential to maintain trans-endothelial resistance (TER) which usually ranges from 300–600 Q/cm, a unique characteristic of the BBB.
Additional to the structural elements of the BBB, signal transduction also plays a pivotal role in the behavior of BBB. Recent studies have identified that the signal transduction of TJ is bidirectional; from cytoplasm to TJ to control the assembly and integrity of TJs, and from TJs to cytoplasm to modulate the gene expression in ECs. Although the exact mechanism of signaling are still to be elucidated, there are several pathways have been identified in relevant to TJ signaling. These include protein kinase A, protein kinase C, G proteins and Rho family GTPases which regulate jucntional assembly and permeability of TJs (48).

Despite the tight control over integrity and function of BBB, meningitic bacterial pathogens have evolved to evade the functional restriction of BBB and invade into the brain tissues. Recent studies have demonstrated two types of mechanisms used by the meningitic bacteria in the penetration of BBB; Trojan horse mechanism where the
intracellular bacteria invade the leukocyte and gain the access to the ECs via leukocyte transendothelial migration (eg. *N. meningitides*) and by harboring individual virulent traits to adhere and induce phagocytosis in ECs and to damage the TJ and the ECs for the invasion of underlying brain tissues (*E. coli*, GBS) (46).

Both *in vitro* and *in vivo* models of BBB are used extensively to understand the pathogenesis of bacterial meningitis. *In vitro* models include primary brain endothelial cells in bovine or human origin; Bovine Microvascular Endothelial Cells (BMECs) and Human Brain Microvascular Endothelial Cells (HBMECs) and Immortalized Human Cerebral Microvascular Endothelial Cells (hCMECs /D3 cells) (49, 50). These two cell culture models have been verified to have correct TER of 300–600 Q/cm which is a unique feature of BBB in addition to endothelial and BBB cell markers. Although the primary cell cultures are promising as models for BBB, the limited passage of cells due to the loss of properties of BBB and high cost are major constraints for their use in research. Recently, an immortalized cell line, hCMECs /D3 has been introduced as a model for BBB which doesn’t have aforementioned constrains. It is derived from the cerebral microvascular endothelium of human that is transduced with a lentiviral vector harboring human telomerase, SV40 T antigen. This D3 cell line has been used as a model for BBB in several studies of drug trials, ion transport, and fungal and bacterial meningitis.

Neonatal rat model of infection is the most extensively used *in-vivo* model to understand many bacterial pathogens associated with NM. It was first used by the Scannapieco *et al.* in 1982 to study the virulence and colonization capacity of *E. coli* K1 (51). Thereafter, many studies have used rat pup models successfully to study pathogenesis of bacterial meningitis (52). Here, neonatal rats were infected with bacterial
pathogens on day 5 to simulate human neonatal infection. Site of inoculum depends on the nature of the study which includes gastric gavage or subcutaneous, intraperitoneal and intra-thecal routes. In the study of neonatal *E. coli* meningitis, intraperitoneal route (usually 100 -200 CFU/100 µl) is the most common route of infection used to discover the virulence factors associated with *E. coli* survival in the blood and transmigration of BBB (52).

2.2 Neonatal Meningitis causing *Escherichia coli* (NMEC)

2.2.1 Clonality of NMEC

*Escherichia coli* encompass significant diversity in their genetic makeup that renders phenotypic and genotypic differences among different pathovars. Serotyping, phylo-grouping multi locus sequence typing (MLST) and pulsed-field gel electrophoresis techniques have been used widely to identify the clonal relationship of *E. coli*. Serotyping is the first typing method adopted to differentiate *E. coli* using serological evaluation depending on the diversity of O-antigen (O-LPS), K antigen (capsular type) and H antigen (flagella type) (53). Among the plethora of different serotypes, a few serotypes of *E. coli* have been predominantly associated with NM (54). There are 181 O antigens, 80 K antigens, and 56 H antigens identified in *E. coli*. Although there is no experimental evidence to suggest a relationship between the O-antigen type and pathogenicity of NMEC, *E. coli* belonging to O18, O83, O7, O12 and O1 have been commonly implicated in NM (55). Capsular types present in NMEC have been well
studied in relation to pathogenesis of *E. coli* meningitis. The most predominant capsular type associated with NMEC strains is K1 and the rest possess the K5 capsular type. The involvement of K1 capsular type in NMEC pathogenesis will be discussed in the next section. With regard to H antigens of NMEC, there are no adequate data reported but H7 has been identified in well studied NMEC strains such as S88, EC10 and RS218.

Phylogrouping is another well-established typing method which use three genetic loci, *chuA*, *yjaA* and TspE4 to group the *E. coli* strains into 4 phylogroups: A, B1, B2 and D (56). Most of the NMEC strains belong to the B2 phylogroup as is the case with ExPECs causing other diseases. However, a trivial limited number of NMEC strains also belong to A, B1 and D groups (57).

### 2.2.2 NMEC and other extra-intestinal pathogenic *E. coli*

Both human and animal infections caused by the pathogenic *E. coli* strains are broadly classified into two groups; intestinal infections and extra-intestinal infections. Pathogenic *E. coli* associated with each group have shown genotypic and functional dissimilarities between them to cause the distinctive clinical disease. For example, intestinal pathogenic *E. coli* strains usually do not present in the fecal flora in healthy individuals (59). They are obligate pathogens and their transmission to healthy individuals is through contaminated water or food, by which they gain access to and colonize the intestinal tract and induce gastroenteritis or colitis (59). Unlike the diarrhoegenic *E. coli*, ExPEC strains asymptotically colonize the intestinal tract showing commensalism in healthy individuals (58, 59). However, unlike commensal *E.
coli, they have the ability to establish the infection when they gain access to extra-intestinal sites such as urinary tract, blood, lungs, etc. (58).

Furthermore, under each category of pathogenic E. coli strains, different pathovars have been identified according to their distinct virulence profiles that attributes to discrete pathogenic mechanism to cause different infections. In intestinal pathogenic E. coli, there are six well-described pathovars that have been identified so far in relation to enteric diseases. These include, enteropathogenic E. coli (EPEC), enterohemorrhagic E. coli (EHEC), enterotoxigenic E. coli (ETEC), enteroaggregative E. coli (EAEC), enteroinvasive E. coli (EIEC) and diffusely adherent E. coli (DAEC).

Extra-intestinal pathogenic E. coli (ExPEC) are a group of E. coli that causes non-enteric infections in both animals and humans. They were first defined by Russo et al. in 2000 (59). These E. coli have ability to cause non-enteric diseases or infections outside of the gut such as respiratory tract infections, UTIs, NM, sepsis, peritonitis, osteomyelitis, soft tissue infections and pneumonia in human and animals. There are three pathovars that have been identified in ExPECs: avian pathogenic E. coli (APEC), uropathogenic E. coli (UPEC) and NM causing E. coli (NMEC).

Variety of diseases associated with ExPECs has been reported in livestock as well as in companion animals. Extra-intestinal E. coli of animal origin predominantly belongs to phylogroup B1 or A (60). Colisepticemia or septicemia associated E. coli is an important disease in calves and piglets particularly in colostrum deprived animals. Among the other bacterial pathogens associated with colisepticemia, E. coli represents 50 – 65% of the cases in calves which is cause of significant mortality and morbidity leading to a serious economic impact (61, 62). Although many O-serogroups have been
identified in *E. coli* associated colisepticemia, common O groups encountered were O78 and O119 which were also associated with higher mortality rates than other O serogroups (61). The most distinctive virulence traits associated with septicemic *E. coli* are F17, F165, CS31A and pyelonephritis-associated pilus (Pap or P fimbral) adhesins (63). In adult cattle, the most common disease associated with ExPEC is mastitis in dairy cows. More often, *E. coli* mastitis is a mild infection in dairy cows followed by rapid elimination of bacteria from mammary tissue. However, severe, per acute mastitis due to *E. coli* have been commonly associated with cows shortly after parturition which may result in toxemia and death. Unlike the septicemic *E. coli* strains, *E. coli* strains causing mastitis do not belong to a particular O-antigen type. Linton *et al.* in 1979 revealed 67 different O-serogroups among 279 *E. coli* isolates recovered from cases of dairy cow mastitis (64). Common virulent traits among mastitis *E. coli* are serum resistance rendered by group II capsule, cytotoxic necrotizing factor, and ‘curli’ adhesin which binds to fibronectin in mammary tissue. However, comparative studies between environmental *E. coli* and *E. coli* isolated from mastitis cases, suggest that mastitis causing *E. coli* strains are opportunistic pathogens present in the environment rather than a distinct pathotype (65).

In companion animals, particularly in dogs and cats, ExPECs are associated with a variety of diseases such as UTIs, prostatitis, otitis externa, vaginitis, pyometra and some perinatal infections (66). Among these diseases, *E. coli* associated with UTIs and pyometra have been extensively studied. As shown by many investigators, *E. coli* is the most predominant bacterial pathogen associated with uncomplicated UTIs in dogs and cats (67). The main virulence determinants of UPEC isolated from canine and feline
cases include type 1 fimbriae (F1 fimbriae), P fimbriae, S fimbriae (Sfa), afimbrial adhesin (Afa), α-hemolysin (Hly), aerobactin, cytotoxic necrotizing factor 1 (CNF1) and cytolethal distending toxin (Cdt) (68). Several studies have suggested that canine and feline UPEC strains share common genetic traits and pathogenic mechanisms with uropathogenic *E. coli* in humans.

Extra-intestinal pathogenic *E. coli* associated with diseases in poultry is collectively known as APEC. These strains have been associated with systemic infections in poultry which responsible for significant morbidity, mortality and carcass condemnation leading to severe negative economic consequences to poultry industry (69, 70). The systemic infections caused by APECs in avian species are manifested as different clinical diseases such as septicemia, omphalitis, salpingitis, peritonitis, airsacculitis, cellulitis, yolk-sac infection and swollen head syndrome, which are collectively termed ‘colibacillosis’ (70). For several years, APECs have been considered as opportunistic pathogens. However, several recent studies that characterized commensal *E. coli* in feces and cloaca of apparently health birds and *E. coli* from colibacillosis cases have demonstrated unique genotypic and phenotypic traits that make these pathogenic *E. coli* a distinct pathotype called APEC (70). The most common O-serogroups associated with colibacillosis are O78, O2 and O1 which represent 80% of the isolates recovered from cases (70). The common virulent traits in APEC include type I fimbriae, P fimbriae, iron acquisition systems (e.g. aerobactin, yersiniabactin and salmochelin) capsular polysaccharides, temperature sensitive hemaglutinins (*tsh*) and cytotoxins. Some of these virulence markers are encoded by genes harbored by large virulence plasmids (71).
Unike ExPEC of animal origin, ExPECs associated with human infections are predominantly belong to phylogroups B2 and D (60). In human, UTIs and NM are the major infections caused by ExPEC. Uropathogenic *E. coli* are a highly specialized subset of pathogenic *E. coli* that have the predilection to urinary tract and harbor specific virulence traits that are required to adhere to and invade the uro-epithelium). In women, *E. coli* is the most common bacterial pathogen associated with acute and recurrent UTIs. It has been documented that the total number of cases of uncomplicated cystitis, pyelonephritis and catheter-associated UTIs in the US is 6–8 million cases, 250,000 cases and 1–1.5 million cases per year, respectively (72). Interestingly, many studies have demonstrated that UPEC strains are predominantly associated with certain O-antigen types. These include O1, O2, O4, O6, O16, O18, O22, O25, and O75. Furthermore, some studies have shown a strong correlation between the serogroup and the presence of different virulence traits in UPEC. The virulence traits present and associated with UTI pathogenesis are fimbriae such as Type I, P and S, toxins such as CNF1, secretory autotransporter (Sat), vacuolating cytotoxin (Vat), haemolysins and iron acquisition systems (73).

Although several studies have demonstrated differences in ExPEC strains in relation to their host specificity, recent evidence suggests that certain similarities exist between the genotypes of ExPECs of animal and human origin, indicating the zoonotic potential of ExPECs of animal origin (74-76). Dezfulian *et al.* (2003) and Girardeu *et al.* (2003) demonstrated that ExPEC strains isolated from septicemic cases of piglets and calves share virulence traits with human UPEC and septicemic strains but not with the diarrhoeagenic *E. coli* (75). Similarly, Rodriguez-Siek *et al.* compared 524 isolates of
APEC and 200 isolates of human UPEC strains and found considerable overlapping of serogroups, virulence genes and phylogroups between these two groups of ExPEC (70). Furthermore, Kariyawasam et al. (2007) applied suppression subtractive hybridization to UPEC and APEC genomes and revealed that many genes including the uropathogenic-specific protein gene (Usp) known to possessed by UPEC is also present in APEC suggesting a genetic relatedness between UPECs and APECs (77).

Several studies have also demonstrated a close genetic relatedness between NMEC and APEC. Mora et al. (2009) revealed that certain ExPEC strains were not host specific and can establish infections in both animal and human hosts (78). The clonal group O1:K1:H7/ ST95 has been identified in E. coli isolated from both colibacillosis cases of poultry and NM cases in human, suggesting that some APEC isolates may have the potential to cause disease in human. Furthermore, IbeA which is a major invasin involved in invasion of BBB by NMEC is harbored by APEC as well. Germon et al. (2005) observed that ibeA was not present in fecal E. coli in poultry but commonly present in APECs belonging to O2, O18 and O88 serogroups (79). It has also been demonstrated that the APEC strain 53/213 which harbors ibeA has the ability to invade HBMECs and invasion has been attenuated by 30% in an isogenic mutant of 53/213 strain lacking ibeA. Similarly, Tivendale et al. (2010) has shown that some NMEC strains belonging to O18/ST95 clonal group was able to establish colisepticemia in a chick model of infection whereas some APEC strains were able to cause meningitis in 5-day-old rat pups in a rat pup model of NM (52). Additionally, the investigators observed that a large plasmid of 133 kb in size was present in all NMECs and APECs belonging to O78 serogroup suggesting a commonality between APEC and NMEC.
Similarly, Peigne et al. (2009) showed that the plasmid of NMEC strain S88 (O45:K1:H7) is closely related to APEC ColV plasmid and is associated with high-level bacteremia in rat pups in an infection model of neonatal meningitis (80).

### 2.2.3 Current concepts of pathogenesis and associated virulence traits of NMEC

As for many other bacterial pathogens associated with meningitis, pathogenesis of NMEC meningitis involves sequential events that depend on the virulence potential of bacteria. These include (i) initial colonization of neonates, (ii) invasion of primary site and spread to the circulation, (iii) intravascular survival and multiplication, (iii) invasion of BBB, and (iv) survival and multiplication in subarachnoid space and brain tissues (28, 81).

The neonatal intestine is sterile at birth as demonstrated by the analysis of meconium using culture-based techniques (82). During birth and shortly thereafter, maternal urogenital and fecal microflora acts as a major source for initial colonization of bacteria in the GI tract and other mucosae of the infant to establish the mucosal microbiome. A strong relationship had been identified between incidence of NBM and the maternal urogenitary microflora. Previous studies have shown that the *E. coli* K1 represents 5–20% of vaginal flora of pregnant mothers (83). Further, 50% of the neonates born to these carrier mothers were colonized with *E. coli* K1 during delivery. There was a stronger correlation between *E. coli* K1 colonization in the gut and the pre-term and low birth weight babies than that of full-term and normal weight babies (15). Some other studies have identified *E. coli* K1 colonization in maternal vagina and
amniotic fluid as a risk factor for delivery of pre-term and low birth weight babies. However, only a relatively low percentage of neonates colonized with *E. coli* K1 developed meningitis, suggesting that the presence of *E. coli* K1 bacteria themselves is not adequate for the progression of the disease. Although there is no strong experimental evidence to describe the mechanism involved, many risk factors have been identified by retrospective studies in relation to incidence of *E. coli* K1 meningitis. These include pre-term delivery, low birth weight and prolong rupture of amniotic membranes (15).

The exact mechanism and the virulence factors associated with bacterial transverse of mucosal barrier and access to the intravascular space are poorly understood. However, the routes of infection are identified as intranasal or gastro-intestinal. Golade et al. (1977) had observed a significantly higher incidence of bacteremia when infant rats were fed with *E. coli* K1 (56%) as compared to intranasal inoculation of rat pups (15%) indicating oral transmission route as the major route of infection (84). Bacterial survival in blood, bacteremia and subsequent events in the pathogenesis of MNEC meningitis have been studied extensively using *E. coli* RS218 strain which is considered as the prototype strain of NMEC. Virulence factors of *E. coli* K1 and the contributing host factors are summarized in Figure 2.3.

The hallmarks of progression of *E. coli* meningitis are high level of bacteremia which is a prerequisite for the penetration of BBB and transcytosis of ECs of BBB with a zipper-like mechanism (28). During these steps, bacteria must evade the host-defense mechanisms and functional barriers in the BBB. *Escherichia coli* K1 are evolved with a variety of survival strategies to counteract these deleterious events as discussed in detail below.
Figure 2.3: Sequential events, virulence factors and host factors involved in NMEC pathogenesis. Source: Nature Reviews. Neuroscience (May 2003): 376-85.

### a. K1 capsule

The capsule is the outermost extracellular slime layer made up of polysaccharides and peptides that act as a protective structure surrounding some bacteria and fungi. In *E. coli*, antigenic surface polysaccharides are present in two major structures in outer layers of the cell; lipopolysaccharides (LPS) in the cell wall (O antigen) and acidic polysaccharides in capsule which is called the K (Kapsel in German) antigen. Capsular polysaccharides are firmly attached to the underneath cell wall by covalent binding with phospholipid or lipid A moieties of the cell wall (85). Many functions of capsule have been identified. It can act as a permeability barrier for bacterial cell thereby resisting the osmotic shock. Also it protects the bacterial cell from harsh environmental conditions such as dessication, facilitates biofilm formation and acts as a reservoir for nutrition. In pathogenic *E. coli*, capsule promotes bacterial colonization of a variety of hostile niches
such as blood, kidney, meninges and lung (85). There are 100 different capsular types identified in *E. coli* and these are broadly categorized in to four groups according to its serological, genetic and biosynthetic make up (Table 1) (86). Relationship between K antigen and the virulence of *E. coli* was first shown by Smith *et al.* (30). This study demonstrated that capsule of bovine *E. coli* strains were more pathogenic to guinea pigs than its capsular mutant counterpart and the increased virulence was due to the evasion of agglutination and phagocytosis of bacteria put forth by the capsule. Similarly in mid-1970’s, several studies revealed the association of particular capsular types of *E. coli* with NM. According to these studies, *E. coli* K1 serogroup is the most predominant capsular type associated with NM. About 80-84% of *E. coli* that were isolated from CSF of infected neonates possessed the K1 capsule (30, 87).

**Table 2-2.** Capsular types and characteristics of *E. coli*. Source: 1999 Blackwell Science Ltd, Molecular Microbiology, 31, 1307-1319.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Former K-antigen group</td>
<td>IA</td>
<td>II</td>
<td>I/II or III</td>
<td>IB (O-antigen capsules)</td>
</tr>
<tr>
<td>Co-expressed with O</td>
<td>Limited range</td>
<td>Many</td>
<td>Many</td>
<td>Often O8, O9 but sometimes none</td>
</tr>
<tr>
<td>serogroups</td>
<td>(O8, O9, O2, O10, O1)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Co-expressed with colanic acid</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Terminal lipool moiety</td>
<td>Wzy-dependent</td>
<td>Processive</td>
<td>Processive</td>
<td>Processive</td>
</tr>
<tr>
<td>Direction of chain growth</td>
<td>Wzx (PST)</td>
<td>ABC-2 exporter</td>
<td>ABC-2 exporter</td>
<td>KpsE?</td>
</tr>
<tr>
<td>Polymerization system</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Translocation proteins</td>
<td>KpsD, KpsE (KpsF)?</td>
<td>KpsD, KpsE</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Elevated levels of CMP-Kdo synthetase</td>
<td>kps near: his, rbf</td>
<td>kps near: sevA</td>
<td>Kps near: sevA</td>
<td>rbf near: his</td>
</tr>
<tr>
<td>Genetic locus</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Thermoregulated (i.e. not expressed below 20°C)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Positively regulated by the Rcs system</td>
<td>Serotype K30</td>
<td>Serotypes K1, K5</td>
<td>Serotypes K10, K64</td>
<td>Serotypes K40, 01 11</td>
</tr>
<tr>
<td>Model system</td>
<td>Klebsiella, Erwinia</td>
<td>Neisseria, Haemophilus</td>
<td>Neisseria, Haemophilus</td>
<td>Many genera</td>
</tr>
<tr>
<td>Similar to</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

In NM, K1 antigen is particularly associated with survival of *E. coli* in blood thereby producing high level of bacteremia which is essential to invade the BBB (28). K1 capsule of NMEC is a homo-polymer of α-[2-8]-linked polysialic acid and N-
acetylneuraminic acid (also called polySia) which is interestingly identical to *Nisseria meningitides* capsular polysaccharide which also causes meningitis in humans (88). These polysaccharides are structurally similar to polysialic acids in embryonic neural membranes and extracellular matrix of many tissues in humans. Therefore, K1 antigen is poorly immunogenic and does not render formation of specific antibodies due to “antigenic mimicry”. This prevents the development of adaptive immunity against K1 *E. coli* and facilitates bacterial survival in the host (89). Additionally, the polyanionic capsule also acts as a barrier against phagocytosis by concealing surface structures such as LPS and membrane proteins to prevent activation of the alternate pathway of complement cascade. It also prevents the deposition of membrane attack complex on the bacterial cell wall (90).

Furthermore, *E. coli* K1 strains also have the ability to modify their capsule by phase variation. Phenotypic variation of the polysaccharide capsule of K1 was first observed by Orskov *et al.* in 1979 (91). These investigators revealed that the phase variation was due to O-acetylation of C7 or C9 carbon positions of sialic acid homopolymer. They also observed that O-acetyl-positive variants were more resistant to neuraminidase hydrolysis than the O-acetyl-negative variants, and O acetyl-negative variants were poorly immunogenic than the O-acetyl-positive strains. Interestingly, *E. coli* K1 isolated from blood possessed the O-acetyl negative phenotype conferring poor immunogenicity that facilitates their survival in blood.

The acetylation of capsular polysaccharides is catalyzed by O-acytyltransferase enzyme (O-AcTase) which uses acetyl CoA as a cofactor. Deszo *et al.* in 2005, and collaborators have unraveled the genetic basis and mechanism of phase variation of K1
capsular antigen encoded by *neuO* gene which is also present in prototypic NMEC strain RS218 (92). The K1-acetylation is catalyzed by an Ac-CoA dependent O-AcTase, encoded by a contingency locus called *neuO*. The location of *neuO* was identified as the K1-specific lysogenic bacteriophage-like element called CUS-3(93). Additionally, they demonstrated that the phase variation was due to a translational frame shift of the ‘hepatonucleotide region’ in the *neuO* due to a slipped strand DNA mispairing or unequal recombination. This genetic event results in three different allelic types that give rise to different capsular types.

b. **Outer membrane protein A (OmpA)**

Outer membrane protein A is the most abundant porin present in Gram-negative cell wall which remained highly conserved during the evolution of *Enterobacteriaceae* (94). Structurally, OmpA is a 35 KDa protein which consists of an eight-stranded beta-barrel structure connected by tight turns on the periplasmic side and four large mobile loops on the extracellular side (95). Power *et al.* in 2006 demonstrated that there are two alleles of *E. coli ompA* gene; *ompA1* and *ompA2* which encode different amino acid sequences especially at loops 2 and 3 rendering a polymorphism to *ompA* (58). Also *ompA2* allele is commonly associated with pathogenic *E. coli*. Both physiological and non-physiological functions were identified in relevant to OmpA of bacteria. It is an essential protein to maintain the integrity of the cell wall and it acts as an aqueous channel for influx of small molecules to the cell. Also, it is involved in conjugal transfer process of *E. coli* which is mediated by F-plasmid and acts as receptors for certain bacteriophages and bacteriocins.
The relative importance of OmpA in the pathogenesis of NM was first reported by Weiser and Gotschlich in 1991(96). Using an embryonic chicken model of bacterial virulence and neonatal rat model of *E. coli* K1 pathogenesis, they demonstrated a seven-fold decrease in the incidence of bacteremia associated with an OmpA-negative isoagenic mutant of NMEC as compared to its wild type counterpart. Further, they revealed that the wild type strain had a 10-fold increase in survival and growth in a chicken model of infection, and was more resistant to the bactericidal effects of pooled human serum than the OmpA-negative mutant.

The mechanism behind serum resistance driven by OmpA was elaborated by subsequent studies. Complement-mediated bacterial killing and phagocytosis are the most important first line defenses exerted by the host to prevent and control extracellular bacterial infections. In classical pathway, C3 convertase is one of the major components for subsequent formation of membrane attack complex. Prasadarao *et al.* (2002) revealed the association of OmpA of NMEC in relation to inhibition of C3 convertase (97). They found that extracellular loops or the N-terminus of OmpA binds with the CCP3 complement control molecule present in C4b protein. The C4b-bacteria complex then acts as a cofactor for Factor-I and cleaves the C4b into C4d which is an inhibitor of C3 convertase. Wooster *et al.* in 2006 demonstrated that OmpA of logarithmic phase *E.coli* was more effectively evaded the complement cascade activation than the OmpA of *E. coli* in post-exponential phase (98).

In addition to evasion of complement mediated killing, OmpA also plays a role in the interaction of *E. coli* with polymorphonuclear leukocytes (PMNs). Polymorphonuclear leukocytes-mediated killing is another important first line of defenses
against bacterial infections. Sukumaran et al. in 2004, observed that OmpA-positive *E. coli* K1 can survive and multiply inside phagosomes of human and murine monocytes and macrophages 24 h post-infection compared to OmpA-negative mutants which were degraded within the phagosomes 1 hr post-infection (99). Furthermore, they demonstrated that the survival of *E. coli* inside the macrophages was due to the interaction of OmpA with pg96 membrane protein on neutrophils. Interaction of gp96 and OmpA down regulates the Rac1, Rac2 and gp91phox of neutrophils which are major components of NADPH oxidase pathway by which the reactive oxygen species are formed to kill the bacteria engulfed by phagosomes. Nevertheless, OmpA-gp96 also down regulates the Toll-like receptor 4 (TLR4) and complement receptor 3 (CR3) (100).

Another study by the same group revealed that ompA K1 interaction with macrophages induces the formation of BclXL which is an inhibitor of apoptosis indicating that the bacteria hijack the macrophage function (101). Similarly, a separate study demonstrated that the OmpA of *E. coli* K1 can bind with the CD64a (Fc receptor) of macrophages and increase the expression of TLR4 and CR3. This interaction down regulates the pro-inflammatory cytokine production and prevents phosphorylation of IkB pathway which in turn inhibits LPS-induced pro-inflammatory cytokine secretion (102).

Moreover, the contribution of OmpA to adhere to and invade the BBB was analyzed by Prasadarao et al. in 2002 (103). This study which used BMECs cells to compare the invasiveness of OmpA-positive wild type *E. coli* K1 and its OmpA-negative isogeneic mutant revealed that the invasiveness of OmpA-positive *E. coli* K1 was 25-50 folds higher than that of the mutant, and the invasiveness was restored in the mutant strain when the *ompA* gene was provided in-trans. They further demonstrated that the
invasiveness of wild-type *E. coli* K1 was inhibited in the presence of anti-OmpA antibodies. Subsequent studies revealed that molecular interaction of OmpA is with GlcNAcb1-4GlcNAc epitopes of glycoproteins which are present only in HBMECs (104). Exact receptor on HBMECs for OmpA was identified later as Ecgp36 which interacts with the extracellular loops of OmpA (105).

Interestingly, OmpA has been implicated in the process of internalization of bacteria into the HBMECs by means of cytoskeleton rearrangements. Interaction of OmpA and Ecgp36 activates focal adhesin kinase (FAK) and phosphatidylinositol 3-kinase (PI3K) which are components of cellular signaling pathways to rearrange the actin cytoskeleton. This cytoskeleton rearrangement results in membrane ruffling and cytoplasmic protrusions which in turn facilitate bacterial entry into HBMECs (28, 106).

c. **Type I fimbrial adhesin**

Fimbriae are non-flagellar, filamentous organelle on the cell wall of bacteria. They are present in a vast majority of members belonging to *Enterobacteriaceae* family and are categorized into different types according to their ability to bind mannose on the erythrocytes (107, 108). In *E. coli*, 14 different fimbriae were identified based on their differences in serological properties, inhibition patterns, chemical properties, antigenicity, morphology, size and receptor binding abilities. Among those, Type I fimbriae are the most prevalent fimbriae present in both commensal and pathogenic *E. coli*. The major structural protein of *E. coli* Type I fimbriae is FimA which is embedded in the cell wall. Approximately 1000 subunits of FimA are combined together to form the base of the fimbria. The minor proteins include, FimF, FimG and FimH. FimF and FimG and are
involved in the formation of the shaft of the fimbria. The fimbrial tip is consisted of FimH which confers mannose-specific binding ability to the bacterium (109).

Type I fimbrial adhesins have been studied in relation to pathogenesis in different human infections caused by *E. coli* such as NM, UTIs and intestinal diseases (110). The contribution of fimbriae in *E. coli K1* in NM was first demonstrated by Guerina *et al.* in 1983 (111). This study showed that mannose-sensitive Type I fimbriae of *E. coli K1* specifically mediate the adherence of bacteria to the oro-pharyngeal mucosa of rat pups leading to bacteremia. This observation of colonization via fimbrial adhesins in NM was further proved by Cox *et al.* (1990) revealing that the topical vaginal application of methyl D-mannoside which binds to fimbriae, inhibited the colonization of *E. coli* resulting in complete prevention of the disease in neonates (112).

Apart from the involvement in initial colonization, *E. coli K1* fimbriae also have the ability to facilitate adherence to the BBB (113). It was shown that *fimH* expression of *E. coli K1* was upregulated in HBMECs. Furthermore, a *fimH*-negative mutant and a Type I fimbrial locked-off mutant showed relatively less adherence and invasion frequencies than that of *fimH* locked-on variants, confirming the role of Type I fimbriae in bacterial binding to and invasion of HBMEC (113). A subsequent study by the same group identified the receptor for *fimH* binding as CD48, a glycosylphosphatidyl-inositol-anchored receptor present on HBMECs (114). Also, they found an interaction between FimH and CD48 that leads to elevated cytosolic-free calcium which can alternate the cell signaling pathways and activate RhoGTPases. All these ultimately lead to actin cytoskeletal rearrangement which is essential for internalization of bacteria by non-immune somatic cells.
d. Invasins of brain endothelium (Ibe proteins)

Invasins of brain endothelium proteins are novel invasins that are associated with BBB adherence and translocation which were discovered by TnphoA mutagenesis (115, 116). These include *ibeA, ibeB* and *ibe T*. The *ibeA* and *ibeT* located on a 20.3 kb size pathogenicity island (PAI) called GimA (genomic island of newborn meningitis causing *E. coli* containing the invasion locus *ibeA*) in *E. coli* K1 prototype, RS218 strain (117). This PAI which is present in ExPECs including the NMEC is located adjacent to *fim* operon. The *GimA* consists four operons namely, GimA1: *ptnIPKC*, GimA2: *cglDTEC*, GimA3: *gcxKRCI* and GimA4: *ibeRAT*. The first three operons encode proteins that are related to substrate transportation and carbon source metabolism such as glucose and bacterial stress response. The fourth operon, *ibeRAT* encodes for 3 proteins, IbeR, IbeA and IbeT. Among these IbeR is the regulatory protein and the other two proteins are involved in invasion of BBB.

The *ibeA* (*ibe10*) encoded by a 1.37-kb DNA region of *ibeRAT* operon. Huang *et al.* (1995) demonstrated that an isogenic in-frame deletion mutant of *ibeA* was significantly less invasive to HBMECs than the parent strain (115). Furthermore, invasiveness of *E. coli* K1 had been decreased by 50 times when IbeA recombinant protein is present in the culture medium compared to a partial fragment of the IbeA protein, revealing a competitive inhibition occurring at the level of binding of *E. coli* K1 to HBMECs via IbeA protein. Subsequent studies identified two receptors on HBMECs that mediate binding of IbeA to HBMECs (118). The primary receptor for IbeA binding was identified as vimentin which is the major intermediate filament (IF) present in the
mesenchymal cells including endothelium and bone marrow originated cells such as PMNs. The secondary binding receptor was polypyrimidine tract-binding protein (PTB)-associated splicing factor (PSF) (119). Vimentin is constitutively expressed on the surface and is a structural protein in the cytoskeleton of HBMECs whereas PSF is an inducible membrane protein on HBMECs. Zou et al. (2006) demonstrated that surface vimentin acts as a ligand for IbeA protein of E. coli K1 and this interaction can be inhibited by vimentin inhibitors such as withaferin A and acrylamide (118). IbeA that binds to vimentin activates a signaling pathway named extracellular-signal-regulated kinase (ERK) in HBMECs that is actively involved in rearranging vimentin in the cytoskeleton to trigger membrane ruffling (120). Additionally, a new role of IbeA-vimentin interaction in stimulating the expression of adhesion molecules on HMBECs to enhance the transendothelial migration of PMNs has also been observed (121).

IbeT is a transmembrane protein with three putative extracellular domains which are encoded by ibeT located downstream to the ibeA in the GimA operon. Involvement of IbeA and IbeT in E. coli pathogenesis was also identified by TnphoA mutagenesis (122). Zou et al. (2008) demonstrated that an isogenic mutant of ibeT named BTD2-3 was less adherent and invasive to HBMECs than its parental strain. Further, they observed that the colonization of BTD2-3 in the gut of neonatal rat pups were significantly less and BTD2-3 was less able to cause bacteremia and meningitis in rat pups than that of the parental strain. A subsequent study done by Cortes et al. revealed the involvement of IbeT in augmenting the fimbrial expression by ExPEC which suggests an involvement of IbeT in E. coli K1 pathogenesis.
The protein IbeB is a 50 kDa which is a novel invasin identified in *E. coli* RS218 which does not belong to GimA operon. Noninvasive nature of RS218 mutant having an in-frame deletion of *ibeB* was first demonstrated by Huang *et al.* (1999) (116). They observed a 100 time reduction in invasiveness with the *ibeB* mutant of RS218 compared to the parental strain in a neonatal rat pup model of meningitis. There was no difference in the degree of bacteremia observed with both mutant and the parental strain. These evidence suggest that a unique role of *ibeB* in penetration of BBB. A parallel observation of contribution of *ibeB* to disease pathogenesis has been documented in APEC strain DE205B (123).

e. **Cytotoxic necrotizing factor 1 (CNF1)**

Cytotoxic necrotizing factors (CNFs) are A-B type toxins of pathogenic *E. coli* strains that consist of a C-terminal catalytic domain, a N-terminal cell binding domain and a transmembrane domain in the middle (124, 125). There are two types of CNFs namely, CNF1 and CNF2 (also called *Vir* toxin) identified in *E. coli*. Although the two toxins are 90% similar in structure, CNF1 is a chromosomally-encoded toxin commonly found in pathogenic *E. coli* of both human and animal origin whereas CNF2 is a plasmid-encoded toxin which is commonly present in *E. coli* strains associated with infections in cattle and sheep (126).

Cytotoxic necrotizing factor 1 was first described by Caprioli in 1983, who observed a necrotizing effect of a secretory toxin of *E. coli* on rabbit skin and CHO, Vero, and HeLa cells (127). It was later described as a toxin which has the ability to change the
cytoskeleton of the eukaryotic cells via Rho-GTPase signaling system (128). CNF1
involvement in NMEC pathogenesis is first identified by Khan et al. in 2002 (129). They
observed that an isogenic mutant of cnf1 was significantly less efficient in invading of
human BMECs and activating RhoA than of the parent E. coli K1 strain.

Cytotoxic necrotizing factor 1 does not contain a signal sequence of Sec-dependent
secretory pathways and is secreted by the E. coli strains that do not harbor a type III
secretory system such as E. coli JS96. For several decades, the secretory mechanism of
CNF1 was not identified. However, a recent study has shown that CNF1 in UPEC is
transported to the external environment via outer membrane vesicles (OMVs) (130,131).
Furthermore, Yu et al. (2012) has identified the YgfZ, a periplasmic protein which
facilitates the loading of CNF1 to OMVs but this phenomenon still warrants further
investigation (131). Once the CNF1 is secreted into the external environment, it is
internalized by eukaryotic cells via a receptor-mediated endocytosis. The receptor for
CNF1 was identified as a 37-kDa laminin receptor precursor (LRP) of BMECs (132).
This interaction facilitates endocytosis of CNF1 and the acidic environment of the
endosome activates the transmembrane domain to inject catalytic domain into the
cytoplasm. The catalytic domain activates Rho GTPases by deamination of glutamine 63
of RhoA. The glutamine residue is essential for GTP hydrolysis, and its modification
results in constitutively activated Rho GTPases by abolishing the conversion of GTP-
bound active form to GDP-bound inactive form (133). Another study demonstrated that
the continuous activation of Rho-GTPases stimulates the phagocytic nature of somatic
cells other than PMNs thereby enhances the internalization of bacteria and synthetic
beads by HBMECs (134). Therefore, CNF1 is considered as an essential virulence factor
that facilitates the invasion process of *E. coli* K1 into the HBMECs. Additionally, GTPase activates the cytokine secretion in leukocytes which is also a hallmark of neuronal injury in the NBM. It has also been shown that CNF1 involves in modifying the permeability of BBB thereby enhancing the trans-epithelial migration of leukocytes to the brain tissues and also contributes to the inhibition of apoptosis in infected cells (135, 136).

f. Arylsulfatases (AslA)

Arylsulfatases are enzymes containing conserved sulfatase motifs that catalyze the hydrolysis of sulfate esters. The *aslA* gene of *E. coli* closely resembles the *atsA* of *Klebsiella pneumoniae* which encodes a 52 kDa polypeptide with membrane spanning segments, two sulfatase motifs and a signal sequence at the N terminal (137). Although it is structurally similar to AtsA and other human sulfatases, AslA of *E. coli* K1 is unable to hydrolyze sulfate esters *in vitro* (138). Involvement of AslA in the penetration of the BBB by *E. coli* K1 was first discovered through Tn*phoA* mutagenesis of *E. coli* K1 during a study that involved screening of potential invasive traits (139). Hoffman *et al.* (2000) observed that the 27A-6 mutant, later identified as an *aslA*-deficient mutant of *E. coli* K1 was less invasive (approximately 50% reduction) in HBMECs than its parental strain. Interestingly, an *in vivo* study using a neonatal rat pup model of meningitis revealed a marked attenuation of the 27A-6 mutant compared to the parental strain as evidenced by less number of rat pups showing signs of meningitis in the group infected with the 27A-6 mutant strain (32%) than the group infected with the parental strain (82%) (138). Moreover, restoration of the wild type phenotype was observed, when the
mutant was complemented with the *aslA* gene *in trans*. However, there was no significant reduction in bacteremia caused by the mutant as compared to the wildtype strain suggesting a unique involvement of AslA in the penetration of BBB.

**g. Other virulence factors of NMEC**

Several other virulence factors have been identified using the signature-tagged mutagenesis (STM) of RS218 strain to identify novel genes involved in invasion of BBB (139). These include (i) *traJ*, a gene that encodes a protein required for the activation of a major promoter called *pY* of the conjugal transfer plasmid F' in *E. coli*, (ii) *yaiU*, that encodes a flagellin-like protein, (iii) *cigA*, that encodes a putative protein which possesses a tyrosine kinase-like activity, and (iv) some other genes which are homologous to unknown genes of *E. coli* K12. Isogenic mutants of each of these genes have shown a 2.1 to 4.6-fold reduction in invasiveness to HBMECs as compared to the wild-type strain. However, the exact mechanism and host factors involved in each of these genes have to be investigated further.

A recent study using NMEC strains isolated from different geographical locations indicate a number of pitfalls in studies that used only the prototypic strain of NMEC (i.e. *E. coli* RS218). Louge *et al.* (2012) observed that most of the virulence factors found in NMEC studies using RS218 strain were not universally present or predominant in the other NMEC isolates. Also, the common virulence factors identified in previous studies such as fimbriae and capsule were also present in nonpathogenic fecal strains of *E. coli* (57). Furthermore, the genes that differentiate NMEC from fecal *E. coli* and more than
50% of genes possessed by NMEC were unknown and located in plasmid. All these suggest that there are many virulent traits that are hitherto unknown and yet to be identified for comprehensive understanding of NMEC pathogenesis. Availability of the complete genome sequence of NMEC RS218 strain together with its plasmid sequence will greatly bolster the studies directed at unraveling virulence mechanisms possessed by NMEC. Such studies will undoubtedly shed light on to our current understanding of NMEC pathogenesis and will identify rational vaccine and therapeutic targets for the cure of NM associated with NMEC.

References


29. Mulder CJJ, van Alphen L, Zanen HC: Neonatal Meningitis Caused by 
*Escherichia coli* in The Netherlands. *Journal of Infectious Diseases* 1984, 
150(6):935-940.

30. McCracken Jr G, Glode M, Sarff L, Mize S, Schiffer M, Robbins J, Gotschlich E, 
Orskov I, Orskov F, Cooperative Neonatal Meningitis Study G: Relation 
between *Escherichia coli* K1 Capsular Polysaccharide Antigen and Clinical 

31. Frosch M, Edwards U, Bousset K, Krauße B, Weisgerber C: Evidence for a 
common molecular origin of the capsule gene loci in Gram-negative bacteria 
expressing group II capsular polysaccharides. *Molecular Microbiology* 1991, 

32. Grados O, Ewing WH: Antigenic Relationship between *Escherichia coli* and 

33. Stoll BJ, Hansen NI, Sánchez PJ, Faix RG, Poindexter BB, Van Meurs KP, 
Bizzarro MJ, Goldberg RN, Frantz ID, Hale EC et al: Early Onset Neonatal 
Sepsis: The Burden of Group B Streptococcal and *E. coli* Disease Continues. 
*Pediatrics* 2011.

34. Ørskov I, Ørskov F: *Escherichia coli* in extra-intestinal infections. 
*Epidemiology & Infection* 1985, 95(03):551-575.

35. Sarff L, McCracken G, Schiffer M, Glode M, Robbins J, Ørskov I, Ørskov F: 
Epidemiology of *Escherichia coli* in Healthy and Diseases Newborns. *The 

for Microbial Contamination in Fruits and Vegetables at the Preharvest 
Level: A Systematic Review. *Journal of Food Protection* &#174; 2012, 
75(11):2055-2081.

37. Jackson KA: Pregnancy-associated listeriosis. *Epidemiology and infection* 2010, 
138(10):1503-1509.

38. Girard KF, Gavin WF: Listeriosis in the newborn. *The Journal of Pathology and 


49. Weksler B: The hCMEC/D3 cell line as a model of the human blood brain barrier. Fluids and barriers of the CNS 2013, 10(1):16.

50. Vernon H, Clark K, Bressler J: In Vitro Models to Study the Blood Brain Barrier. In: In Vitro Neurotoxicology. Volume 758y Editor]. City]: Publisher|;


68. Johnson JR, Kaster N, Kuskowski MA, Ling GV: Identification of Urovirulence Traits in Escherichia coli by Comparison of Urinary and Rectal E. coli


meningitis-causing Escherichia coli: How closely related are they?


Escherichia coli meningitis: induction of bacteremia and meningitis in infant
rats fed E. coli K1. Infection and Immunity 1977, 16(1):75-80.
85. Silver RP, Aaronson W, Vann WF: The K1 capsular polysaccharide of
86. Whitfield C, Roberts IS: Structure, assembly and regulation of expression of
87. Ørskov I, Ørskov F: Escherichia coli in Extra-Intestinal Infections. The
89. Silver RP, Aaronson W, Vann WF: The K1 Capsular Polysaccharide of
Escherichia coli. Reviews of Infectious Diseases 1988, 10(ArticleType: research-
article / Issue Title: Supplement 2. Microbial Surfaces: Determinants of Virulence
and Host Responsiveness / Full publication date: Jul. - Aug., 1988 / Copyright ©
90. Horwitz MA, Silverstein SC: Influence of the Escherichia coli capsule on
complement fixation and on phagocytosis and killing by human phagocytes.
JB: Form variation in Escherichia coli K1: determined by O-acetylation of
the capsular polysaccharide. The Journal of Experimental Medicine 1979,
149(3):669-685.
92. Deszo EL, Steenbergen SM, Freedberg DI, Vimr ER: Escherichia coli K1
polysialic acid O-acetyltransferase gene, neuO, and the mechanism of capsule
form variation involving a mobile contingency locus. Proceedings of the
National Academy of Sciences of the United States of America 2005,
102(15):5564-5569.
ER: Escherichia coli K1-Specific Bacteriophage CUS-3 Distribution and


111. Guerina NG, Kessler TW, Guerina VJ, Neutra MR, Clegg HW, Langermann S, Scannapieco FA, Goldmann DA: **The Role of Pili and Capsule in the**


Chapter 3

DEFINING THE PATHOYPE OF NEONATAL MENINGITIS CAUSING E. COLI

3.1 Abstract

Neonatal meningitis causing E. coli (NMEC) is the most predominant Gram-negative bacterial pathogen associated with meningitis in newborn infants. High levels of heterogeneity and diversity have been observed in the repertoire of virulence traits and other characteristics among NMEC, and therefore, it has been difficult to define the NMEC pathotype. In this study, we characterized NMEC (n=53) by comparing the characteristics with fecal E. coli isolated from healthy individuals (HFEC) (n=48). Genotypic characterization reflected that the majority (67.92%) of NMEC belonged to B2 phylogenetic group which more frequently carries virulence-associated genes than do groups A and B1 strains. Of the HFEC surveyed, 59% belonged to groups A or D, which are considered to be nonpathogenic. Serotyping revealed that the most common O and H antigen types present in NMEC were O1 (15%), O8 (11.3%) and O18 (13.2%), and H7 (25.3%). These antigen types except the O8 were not observed among HFEC strains. Virulence genotyping clearly indicated that >70% of NMEC carried kpsII, K1, neuC, iucC, sitA and vat of which the prevalence was very low prevalence among HFEC (range 0 - 27%). All NMEC were invasive in their capability to invade human brain microvascular endothelial cells. However HFEC of which only 79% also showed the invasive phenotype suggesting that NMEC pathogenicity cannot be determined purely on
the basis of invasiveness. The NMEC strains demonstrated an increased ability to form biofilm in Luria Bertani broth medium relative to HFEC (79.2% vs 39.9%). Taken together, our study results clearly indicated that virulence gene profiling along with phylogrouping can be used to define the NMEC pathotype.

3.2 Introduction

*Escherichia coli* is a versatile bacterium that exists as a commensal in the lower gastrointestinal tract of human and animals as well as a pathogen that causes a variety of diseases (1, 2). Unlike commensals, pathogenic *E. coli* harbor various virulence traits which is the basis for categorizing them into different pathovars and each pathovar of *E. coli* has the ability to establish a distinct infection (2). The virulence traits include adhesins, iron acquisition systems, toxins, invasins and serum resistant components of the cell wall that are encoded by the chromosome or/and the plasmids (1-5).

The distinct pathotype of *E. coli* known as neonatal meningitis causing *E. coli* (NMEC) are extra-intestinal pathogenic *E. coli* (ExPEC) that have the ability to survive in blood and invade meninges of infants to cause meningitis (6, 7). *Escherichia coli*-associated neonatal meningitis is one of the most common infections that accounts for high mortality and morbidity rates (10-30%) during the neonatal period (8,9). Although the distinct sets of virulence traits have been identified in other ExPEC pathovars and diarrheagenic *E. coli*, no such definitive set of virulence traits has been identified to define the NMEC pathotype. Several studies have attempted to identify the characteristics and the meningo-virulent clones of NMEC using phenotypic and genotyping methods
such as serotyping, multi-locus sequence typing (MLST), phylogrouping, pulsed-field gel electrophoresis (PFGE), antibiotic resistance gene profiling and virulence genotyping to (6, 8-11). A recent study showed that NMEC strains are diverse in their virulence gene repertoire (11) and specific genes that are associated with NMEC could not be identified. None of these studies have evaluated the ability of various genotypes to invade the blood brain barrier (BBB) to determine if a particular genotype/s is/are more invasive than the others to human brain microvascular endothelial cells. We attempted to determine the pathotype of NMEC based on various genotypic and phenotypic characteristics and found a set of virulence genes that were commonly associated with NMEC strains and relatively low in fecal E.coli from healthy humans.

3.3 Materials and Methods

3.3.1 Bacterial strains and media

E. coli isolates comprising of NMEC (n =53) and HFEC (n= 48) were characterized. The NMEC strains included 51 strains isolated from the cerebrospinal fluid of neonates with meningitis that were presented to the Johns Hopkins Medical School, Maryland, USA, and two well-studied NMEC strains, RS218 (O18:K1:H7) and EC10 (O7: K1). The archetypical NMEC strains, E. coli RS218 and EC10 strains have been isolated from the cerebrospinal fluid of neonates with meningitis in 1970’s and were kindly provided by Dr. James Johnson, University of Minnesota, St. Paul, MN, USA and Dr. David Klumpp, Northwestern University, Evanston, IL, USA ,respectively. Fecal E.
coli (n= 48) isolated from healthy individuals were obtained from the E. coli Reference Center. Unless otherwise mentioned, all the bacterial strains were grown in Luria Bertani (LB) broth or LB agar.

3.3.2 Phylogrouping of E. coli

All strains were assigned to four main phylogenetic groups; A, B1, B2 and D according to the method described by Clemont et al. (8). This classification of phylogenetic groups was based on the amplification of two genes, chuA and yjaA and the DNA sequence, TspE4C2 by using a multiplex polymerase chain reaction (PCR) assay. The primers used are listed in Table A (Appendix A)

3.3.3 Serotyping

Serotyping was performed at the E. coli Reference Center (Pennsylvania State University, University Park, PA). Standard serum agglutination assay was conducted (Orskov et al, 1977) for the presence of all designated O groups (O1 to O187) except O13, O22, 31, 47 and 94 which cannot be typed by this method, and H types (H1 to H56).

3.3.4 Virulence genotyping

NMEC and HFEC strains were examined for the presence of 26 virulence factors (VFs) which have previously been recognized as the traits of NMEC that are important
for the establishment of meningitis in neonates (22). These VFs included adhesins, invasins, toxins, siderophores, and structural components of *E. coli*. Virulence genes were detected by PCR amplification carried out on a Master Cycler Pro (Eppendorf, Hamburg, Germany). All primers were obtained from Integrated DNA Technologies (Coralville, IA) and are listed in Table A (Appendix A).

Crude DNA was extracted by a rapid boiling method (12). All PCRs were performed in 25-μl reactions containing 1 U of *Taq* DNA polymerase (Denville Scientific Inc., Metuchen, NJ), 25 pmol of the forward and reverse primers, and 5 nmol of each deoxynucleoside triphosphate (Denville) in 1× buffer (15 mM MgCl₂, 100 mM KCl, 80 mM (NH₄)₂SO₄, 100 mM Tris-HCl, pH 9.0, 0.5% NP-40). The cycling conditions were as follows: initial denaturation at 94°C for 3 min, followed by 30 cycles of three steps consisting denaturation at 94°C for 1 min, primer annealing at the temperatures indicated in Table 1 for 1 min, and extension at 72°C for at least 30 s, according to the size of the amplified fragment (approximately 1 min/kbp), and followed by a final extension at 72°C for 10 min. Products were electrophoresed in a 1.5% agarose gel (Denville) for 1 hour at 120 V, stained with ethidium bromide (Bio-Rad Laboratories, Hercules, CA), and photographed under UV light using a gel documentation system (AlphaImager® HP, Alpha Innotech Corporation, San Leandro, CA). Each PCR included a negative control that contained all reagents except template DNA and a positive control that contained an *E. coli* that is known to possess the respective VF.
3.3.5 PFGE

Pulsed-field gel electrophoresis was conducted according to the method described by PulseNet (13). Chromosomal DNA was digested with the XbaI restriction enzyme. The electrophoresis was performed using CHEF DRII system (Bio-Rad, Marnes-la-Coquette, France) and the conditions consisted of an initial time of 2.2 s, a final time of 54.2 s at a gradient of 6 V cm\(^{-1}\) and an included angle of 120\(^\circ\). The gels were electrophoresed for 24 hours. *Salmonella enterica* serotype Braenderup strain H9812 (ATCC\textsuperscript{®} BAA664\textsuperscript{TM}, Manassas, VA) was used as a standard for normalization of gels. A dendrogram was constructed using the Dice similarity coefficient and the unweighted-pair group method by average linkages (UPGMA) or neighbor joining algorithm with 3% position tolerance using Bionumerics 4.0 software (Applied Maths, Austin, Texas).

3.3.6 Biofilm assay

Qualitative biofilm assay was performed following the procedure described previously (14). Briefly, overnight bacterial cultures were diluted in 1:100 ratio with M9 minimal medium (BD Technologies, Research Triangle Park, NC) containing 10 \(\mu g/ ml\) niacin and LB medium followed by inoculation into U-bottom 96-well plates (Denville) in triplicates. The plates were incubated at room temperature for 24 hours at 37\(^\circ\)C. The plates were washed three times with distilled water and biofilms were stained with 0.1% solution of crystal violet for 15 min. After three washes with distilled water, the presences or absence of biofilms was evaluated.
3.3.7 Antibiotic susceptibility testing

Kirby-Bauer (KB) disk diffusion assay was performed for the antibiotic susceptibility testing using commercially available antibiotic paper disks (BD Technologies). The antibiotics used in KB test were amikacin (30 µg), amoxicillin with clavulanic acid (20/10 µg), chloramphenicol (30 µg), kanamycin (30 µg), nalidixic acid (30 µg), sulfamethoxazole with trimethoprim (23.75/1.25 µg), sulfisoxazole (0.25 µg) and tetracycline (30 µg). Susceptibility or resistance profiles were interpreted according to Clinical Laboratory Standard Institute (CLSI) standards. Extended spectrum β-lactamase (ESBL) sensitivity profiles were screened using confirmatory ESBL plates according to manufacturer’s instructions (Sensititre; Trek Diagnostics, Cleveland, OH). The dilution range of antibiotic concentrations used were as follows; Cefazolin (8 – 16 µg/ml), Cefepime (1 - 16 µg/ml), Cefoxitin (4 - 64 µg/ml), Meropenem (1 - 8 µg/ml), Cephalothin (8 - 16 µg/ml), Cefpodoxime (0.5 - 64 µg/ml), Ceftriaxone (1 - 128 µg/ml), Ciprofloxacin (1 - 2 µg/ml), Gentamicin (4 - 16 µg/ml), Ampicillin (8 - 16 µg/ml), Imipenem (0.5 - 16 µg/ml), Piperacillin/Tazobactam (4/4 - 64/4 µg/ml), Ceftazidime (0.25 - 128 µg/ml), Ceftazidime/Clavulanic Acid (0.25/4 - 128/4 1 µg/ml) and Cefotaxime (0.25 - 64 µg/ml). Plates were incubated at 37ºC for 24 hours and read automatically using ARIS® Sensititre System (Trek Diagnostics).

3.3.8 In-vitro cell invasion assay

The hCMEC/D3 cell line, which is known to possess the main characteristics of primary brain endothelial cells, was kindly provided by Dr. Babette Weksler at the Weill
Invasion assays were performed in triplicates using the method described (16). Briefly, hCMEC/D3 cells were seeded on 96-well tissue culture plates coated with rat collagen (R&D Systems, Inc., Minneapolis, MN) and maintained at 37°C in a humidified chamber containing 5% CO₂ to grow into a confluent monolayer for six days. The complete medium used for culturing of hCMECs was EBM-2 endothelial basal medium (Lonza, Allendale, NJ) containing 5% bovine serum (GE Healthcare Bio-Sciences, Pittsburgh, PA), 1.4 μM hydrocortisone (Sigma-Aldrich, St. Louis, MO), 5 μg/ml ascorbic acid (Sigma), 1% chemically defined lipid concentrate (Invitrogen, Carlsbad, CA), 10 mM HEPES (GE Healthcare Bio-Sciences), and 1 ng/ml human basic fibroblast growth factor (Sigma). Cells were infected with *E. coli* strains with a multiplicity of infection (MOI) of 100 (100 bacteria per hCMEC) and incubated for 2 hours at 37°C in 5% CO₂. Following three washes with phosphate buffered saline (PBS, pH7), infected hCMEC cultures were incubated with complete medium containing 100 μg/ml gentamicin (Sigma) for one hour to kill extracellular bacteria. Subsequently, to release the intracellular bacteria, 0.2 ml of 0.01% Triton X-100 (Sigma) was added to each well. Cells were disrupted by repeated pipetting and viable bacteria were enumerated by plating onto LB agar. Invasion frequencies were calculated comparing the ratio between the number of colony forming units (CFU) invaded and the number of CFU inoculated per well. To standardize the comparison of invasion frequencies, relative invasion was calculated as a percent of invasion compared with the well characterized NMEC strain RS218, which was arbitrarily set at 100%. 

Cornell Medical College, Cornell University (New York City, NY) (15).
3.3.8 Statistical analysis

Data were statistically analyzed by the Fisher’s exact test (virulence genotyping and phylogrouping), two-tailed student’s t-test (invasion frequency). A \( p \) value < 0.05 was considered statistically significant. The heat map of virulence gene profiles was constructed using the R code (http://www.R-project.org).

3.4 Results

3.4.1 Phylogenetic typing of NMEC and HFEC

All four major phylogenetic lineages (A, B1, B2, and D) were represented by both NMEC and HFEC with a difference in distribution of each lineage among two different isolate sources (i.e. NMEC and HFEC) (Table 3.1). When the distribution of NMECs and HFECs in each phylogroup is considered, phylogroup B2 was overly represented by NMEC (67.92%) compared to HFEC (29.17%) (\( p \) value < 0.0001) whereas the phylogroups A and B1 were more common in HFEC (31.25% and 12.5%, respectively) than in NMEC (9.43% and 3.77%, respectively) (\( p = 0.0001 \) and \( p = 0.0225 \), respectively). No statistical difference was observed with the distribution of NMEC and HFEC within group D (\( p \) value = 0.2395).

3.4.2 Serotyping

Serotyping of NMEC revealed a marked variation among both NMEC and HFEC isolates (Figure 3.1 and Table 3.2). Eighteen different O serogroups were observed.
among NMEC strains tested whereas 19 different O antigens were observed in HFEC. Three isolates of NMEC and six isolates of HFEC demonstrated the O antigen-negative phenotype and four isolates of NMEC and two isolates of HFEC demonstrated multiple O antigen types. Despite several O serogroups represented in both NMEC and HFEC, a clear difference in common O types associated with each group was observed. For example, common O antigen types seen among NMEC tested were O:1 (n=9), O:8 (n=6) and O:18 (n=7) whereas in HFEC, O:20 (n=8) and O:148 (n=5) were more common than the other O types. Similarly, among NMEC strains, ten different H antigen types were observed. However, most of the NMEC isolates (n=24) belonged to the H:7 flagella antigen type. Notably, HFEC revealed more variation in H antigen typing than NMEC, and of the 48 isolates of HFEC tested, 19 different H types were observed. Nine isolates of HFEC demonstrated the H antigen-negative phenotype.

Table 3-1. Distribution of different phylogroups between neonatal meningitis E. coli (NMEC) and fecal E. coli from healthy individuals (HFEC). A p value of <0.05 reflects a statistical significance.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Phylogroup (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>NMEC</td>
<td>9.43</td>
</tr>
<tr>
<td>HFEC</td>
<td>31.25</td>
</tr>
<tr>
<td>p value</td>
<td>0.001</td>
</tr>
</tbody>
</table>

3.4.3 Virulence genotyping

Of the 26 VFs detected in the study, prevalence of 21 genes was significantly higher (p <0.05) in NMEC than in HFEC (Figure 3.2 and Table 3.4). Similar to the
observations made by previous studies, some virulence genes thought to be involved in penetration of BBB such as \textit{fimH}, \textit{npl} and \textit{ibeB} were equally present in both NMEC and

**Figure 3-1.** Results of pulsotyping, phylogrouping, serotyping and ability to form in vitro biofilm assay for neonatal meningitis \textit{E. coli} (NMEC) and fecal \textit{E. coli} from healthy individuals (HFEC) used in the study. A substantial diversity in pulsotyping and O serogrouping was observed among the \textit{E. coli} studied.
Table 3-2. Distribution of O and H antigen types among NMEC and HFEC. M, multiple; POS, positive; NEG, negative.

<table>
<thead>
<tr>
<th>O Type</th>
<th># Isolates(n)</th>
<th>H Type</th>
<th># Isolates</th>
<th>O Type</th>
<th># Isolates(n)</th>
<th>H Type</th>
<th># Isolates(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>7</td>
<td>24</td>
<td>10</td>
<td>1</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>20</td>
<td>8</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>10</td>
<td>1</td>
<td>36</td>
<td>1</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>18</td>
<td>1</td>
<td>38</td>
<td>1</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>19</td>
<td>3</td>
<td>46</td>
<td>1</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>27</td>
<td>1</td>
<td>68</td>
<td>3</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>7</td>
<td>6/41</td>
<td>6</td>
<td>73</td>
<td>1</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>POS</td>
<td>2</td>
<td>104</td>
<td>2</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td></td>
<td></td>
<td>117</td>
<td>1</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td></td>
<td></td>
<td>137</td>
<td>1</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>75</td>
<td>3</td>
<td></td>
<td></td>
<td>146</td>
<td>1</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>78</td>
<td>1</td>
<td></td>
<td></td>
<td>148</td>
<td>5</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td>92</td>
<td>1</td>
<td></td>
<td></td>
<td>153</td>
<td>1</td>
<td>6/41</td>
<td>1</td>
</tr>
<tr>
<td>107</td>
<td>1</td>
<td></td>
<td></td>
<td>167</td>
<td>2</td>
<td>40/44</td>
<td>1</td>
</tr>
<tr>
<td>12/16</td>
<td>2</td>
<td></td>
<td></td>
<td>178</td>
<td>1</td>
<td>NEG</td>
<td>9</td>
</tr>
<tr>
<td>M</td>
<td>4</td>
<td></td>
<td></td>
<td>91/167</td>
<td>2</td>
<td>POS</td>
<td>13</td>
</tr>
<tr>
<td>NEG</td>
<td>3</td>
<td></td>
<td></td>
<td>M</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NEG</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HFEC strains. Likewise, certain virulence genes, such as $cnf1$ and $traJ$, which were considered to be essentially required for NMEC pathogenesis, were possessed by only a minority of NMEC ($cnf1$ and $traJ$ in 27.45% and 43.40% of NMEC, respectively). However, these genes were more frequently observed in NMEC than in HFEC (Table 3.4). Overall, as shown in the heat map (Figure 3.2), the number of virulence traits contained in NMEC was higher (13+3.84) than that of HFEC (5.50+2.49) revealing a huge variation between the virulence gene profiles of two different isolate sources.
Figure 3-2. Heat map demonstrating the distribution of virulence traits between neonatal meningitis causing *E. coli* (NMEC) and fecal *E. coli* from healthy individuals (HFEC). Red, presence of a virulence trait; Black, absence of a virulence trait; A, virulence traits prevalent in both NMEC and HFEC; B, virulence traits highly prevalent and overly represented in NMEC; C, virulence traits with low prevalence in both NMEC and HFEC.

Notably, K1 capsular type, *sfa/foc*, *sat*, *hlyA/D*, *iutA*, *papG III* and *afa* genes were only represented in NMEC although their prevalences, except the K1 capsular type, were low. Virulence traits *kpsII*, K1, *sitA*, *neuC*, *iucC* and *vat* were more commonly associated with NMEC than HFEC (prevalence >70% and representation was >2-fold higher in NMEC than in HFEC).
Table 3-3. Prevalence (%) of virulence genes between neonatal meningitis *E. coli* (NMEC) and fecal *E. coli* (HFEC); *p* value of <0.05 reflects a statistical significance. Note. The virulent traits that were highly prevalent and overly represented in NMEC compared to HFEC are marked in bold.

<table>
<thead>
<tr>
<th>Gene</th>
<th>NMEC</th>
<th>HFEC</th>
<th>Chi Square value</th>
<th><em>p</em> value</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>afa</em></td>
<td>9.43</td>
<td>0.00</td>
<td>7.446</td>
<td>0.0064</td>
</tr>
<tr>
<td><em>aslA</em></td>
<td>88.68</td>
<td>45.83</td>
<td>41.814</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>aufA</em></td>
<td>62.26</td>
<td>20.83</td>
<td>32.952</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>cnf</em></td>
<td>27.45</td>
<td>4.17</td>
<td>0.105</td>
<td>0.7456</td>
</tr>
<tr>
<td>* fimH*</td>
<td>92.45</td>
<td>100.00</td>
<td>6.380</td>
<td>0.0115</td>
</tr>
<tr>
<td><em>hyl</em></td>
<td>30.19</td>
<td>0.00</td>
<td>32.980</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>hylD</em></td>
<td>35.85</td>
<td>0.00</td>
<td>45.995</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>ibeA</em></td>
<td>35.85</td>
<td>20.83</td>
<td>6.881</td>
<td>0.0087</td>
</tr>
<tr>
<td><em>ibeB</em></td>
<td>100.00</td>
<td>91.67</td>
<td>6.380</td>
<td>0.0115</td>
</tr>
<tr>
<td><em>iss</em></td>
<td>25.49</td>
<td>4.17</td>
<td>16.132</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>iucC</em></td>
<td><strong>74.51</strong></td>
<td><strong>2.08</strong></td>
<td><strong>109.471</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
<tr>
<td><em>iutA</em></td>
<td>30.19</td>
<td>0.00</td>
<td>32.980</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>K1</em></td>
<td><strong>75.47</strong></td>
<td>0.00</td>
<td><strong>116.821</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
<tr>
<td><em>kpsII</em></td>
<td><strong>90.57</strong></td>
<td><strong>27.08</strong></td>
<td><strong>82.038</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
<tr>
<td><em>neuC</em></td>
<td><strong>71.70</strong></td>
<td><strong>18.75</strong></td>
<td><strong>54.522</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
<tr>
<td><em>nlpI</em></td>
<td>100.00</td>
<td>97.92</td>
<td>0.505</td>
<td>0.4773</td>
</tr>
<tr>
<td><em>ompA</em></td>
<td>66.04</td>
<td>60.41</td>
<td>0.536</td>
<td>0.464</td>
</tr>
<tr>
<td><em>papC</em></td>
<td>3.77</td>
<td>0.00</td>
<td>2.296</td>
<td>0.129</td>
</tr>
<tr>
<td><em>papGII</em></td>
<td>31.37</td>
<td>2.08</td>
<td>25.833</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>papGIII</em></td>
<td>21.57</td>
<td>0.00</td>
<td>22.523</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>sat</em></td>
<td>49.02</td>
<td>0.00</td>
<td>62.279</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>sfa/foc</em></td>
<td>49.06</td>
<td>0.00</td>
<td>62.279</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>sitA</em></td>
<td><strong>92.45</strong></td>
<td><strong>25.00</strong></td>
<td><strong>89.713</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
<tr>
<td><em>traJ</em></td>
<td>43.40</td>
<td>10.41</td>
<td>26.287</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>vat</em></td>
<td><strong>76.47</strong></td>
<td><strong>18.75</strong></td>
<td><strong>32.829</strong></td>
<td>&lt; <strong>0.0001</strong></td>
</tr>
</tbody>
</table>

3.4.4 PFGE

Out of 101 *E. coli* isolates studied, only 98 were typeable by PFGE. The HFEC isolate 17 and two NMEC isolates (14 and 20) were nontypeable with PFGE that used *XbaI* restriction digestion. Typeable isolates of both NMEC and HFEC revealed a high degree of variability based on PFGE fingerprint patterns (pulsotypes) where the similarity
of banding patterns varied from 60% to 100% (Figure 3.1). There were 86 distinct pulsotypes identified for the isolates tested. Of these 86 pulsotypes, 79 were unique and each was represented by only one isolate. The other 7 pulsotypes were shared by more than one isolate. Even though certain pulsotypes were shared by more than one isolate in a given source of *E. coli* (NMECs or HFEC), not a single pulsotype was shared between NMEC and HFEC.

3.4.5 Biofilm assay

The ability to form biofilm on a microtitre plate by NMEC and HFEC was assessed using both M9 minimal medium supplemented with niacin and LB broth. The 52.8% of NMEC strains cultured in M9 minimal medium were able to form biofilm whereas only 39.9% of HFEC possessed this phenotype (Figure 3.1), though the difference was not statistically significant (*p*=1.000). A higher percentage of NMEC strains (79.2%) formed biofilm when they were cultured in LB medium than in the M9 medium (52.8%). The HFEC isolates exhibited the same degree of biofilm formation regardless of which medium (M9 or LB broth) was used.

3.4.6 Antibiotic susceptibility testing

As depicted in Table 3.6, the majority of NMEC and HFEC were susceptible to most of the antimicrobials used in the study. All isolates tested were susceptible to gentamicin, imipenam, meropenem, piperacillin/tazobactam and amikacin. Different antibiotic sensitivity profiles (susceptible, resistance or intermediate) were exhibited for
the remaining antibiotics. Furthermore, five isolates of NMEC and three isolates of HFEC were resistant to all four third generation cephalosporins (cefotaxime, cefpodoxime, ceftazidine and ceftrixone). The minimal inhibitory concentration values of cefotaxime and ceftazidine alone and in combination with clavulanic acid for these eight isolates were increased by ≥3 folds revealing that these isolates possess the ESBL resistance phenotype.

Table 3-4. Antibiotic profiles of neonatal meningitis *E. coli* (NMEC) and fecal *E. coli* from healthy individuals (HFEC). Number of isolates (stated as percentage) in NMEC and HFEC were categorized as Sensitive (S), Intermediate (I) or Resistant (R) according to CLSI standards.

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>NMEC</th>
<th></th>
<th></th>
<th>HFEC</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>I</td>
<td>R</td>
<td>S</td>
<td>I</td>
<td>R</td>
</tr>
<tr>
<td>Amoxicillin</td>
<td>50.94</td>
<td>3.773585</td>
<td>45.28</td>
<td>64.58</td>
<td>4.167</td>
<td>31.25</td>
</tr>
<tr>
<td>Cefazolin</td>
<td>96.23</td>
<td>0</td>
<td>3.77</td>
<td>93.75</td>
<td>0</td>
<td>6.25</td>
</tr>
<tr>
<td>Cefepime</td>
<td>96.23</td>
<td>0</td>
<td>3.77</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>90.57</td>
<td>0</td>
<td>9.43</td>
<td>93.75</td>
<td>0</td>
<td>6.25</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>90.57</td>
<td>0</td>
<td>9.43</td>
<td>95.83</td>
<td>2.08</td>
<td>2.08</td>
</tr>
<tr>
<td>Cefpodoxime</td>
<td>90.57</td>
<td>0</td>
<td>9.43</td>
<td>9.16</td>
<td>0</td>
<td>8.33</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>90.57</td>
<td>0</td>
<td>9.43</td>
<td>93.75</td>
<td>0</td>
<td>6.25</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>88.68</td>
<td>1</td>
<td>9.43</td>
<td>91.66</td>
<td>2.08</td>
<td>6.25</td>
</tr>
<tr>
<td>Cephalothin</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gentamicin</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Imipenem</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Meropenem</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Piperacillin/Tazobactam</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ceftazidime/Clav. Acid</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Amikacin</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Amoxicillin</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>60.41</td>
<td>2.083</td>
<td>37.5</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>94.34</td>
<td>0</td>
<td>5.66</td>
<td>97.91</td>
<td>2.083</td>
<td>0</td>
</tr>
<tr>
<td>Sulf/trim</td>
<td>81.13</td>
<td>0</td>
<td>18.86</td>
<td>93.75</td>
<td>0</td>
<td>6.25</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>86.79</td>
<td>2</td>
<td>9.433</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nalidixic acid</td>
<td>96.23</td>
<td>0</td>
<td>3.773</td>
<td>81.25</td>
<td>0</td>
<td>18.75</td>
</tr>
<tr>
<td>Sulfisoxazole</td>
<td>67.92</td>
<td>3</td>
<td>14</td>
<td>68.75</td>
<td>16.66</td>
<td>14.58</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>81.13</td>
<td>0</td>
<td>10</td>
<td>91.66</td>
<td>0</td>
<td>8.33</td>
</tr>
</tbody>
</table>
3.4.7 *In-vitro* cell invasion assay

Analysis of invasion frequencies displayed a significant difference between NMEC and HFEC as expected (85.22 ± 18.75 Error vs 57.08 ± 20.53, respectively; *p* = 0.0367) (Figure 3.4). Out of 48 HFEC, ten isolates were noninvasive whereas all NMEC tested were able to invade hCMEC albeit at different frequencies. However, some HFEC strains showed invasion frequencies similar or higher to that of the prototypic NMEC strain RS218. Nevertheless, some NMEC strains exhibited lower invasion frequencies than the mean invasion frequency observed for HFEC.

Figure 3-3. Relative invasion frequencies observed between NMEC and HFEC populations. ** denotes a statistically significant difference.
3.5 Discussion

*Escherichia coli* consists of a diverse group of bacteria which are highly diverse in their pathogenic potential which ranges from commensals in the gut and pathogens causing a wide variety of diseases in both humans and animals (1, 2). Both diarrhegenic *E. coli* (a.k.a intestinal pathogenic *E. coli*) and ExPEC are grouped into many pathotypes or pathovars depending on the clinical outcome of the disease and distinct virulence properties (8, 19). Unlike intestinal pathogenic *E. coli*, which are frank pathogens acquired by the host through contaminated food or water, ExPEC inhabit the gastro-intestinal tract of healthy individuals (14, 19). These ExPEC possess the ability to infect extra-intestinal sites of the host due to acquired genetic determinants or virulence genes, thereby acting as opportunistic pathogens. Several studies have attempted to identify unique virulence traits of uropathogenic *E. coli* (UPEC), NMEC and avian pathogenic *E. coli*, which all are ExPEC, to understand the genetic makeup and pathogenic potential of each ExPEC pathovar. However, these studies indicate that a substantial genotypic and phenotypic heterogeneity exist within each pathotype, and a pathotype cannot be defined merely based on one single gene or a characteristic (25). Nevertheless, certain virulence properties are shared between the pathotypes making it difficult to delineate ExPEC pathotypes (6, 9-11, 17). Here, we comparatively analyzed genotypic and phenotypic characteristics of NMEC and HFEC to identify the properties that would aid in defining the NMEC pathotype.
The PCR-based phylogrouping has been used in several studies to identify the clonality within *E. coli* pathotypes (18). Previous studies have identified group B2 as the most common phylogroup among ExPEC proposing that phylogrouping can be used as a rapid typing method to identify potential ExPEC (9, 11, 19, 20). In the present study, we also observed that 67.9% of NMEC falling into B2 phylogroup. Phylogroups A and B1 were more commonly associated with HFEC than NMEC, and only 29.2% of HFEC belonged to B2. We noticed a significant disparity in terms of virulence gene profile between NMEC and HFEC belonging to same phylogroup. For instance, on their virulence gene profiles, the prevalence of virulence genes was statistically higher in NMEC belonging to phylogroups A, B2 and D than the HFEC belonging to the corresponding phylogenetic group (Table 3.5). Taken together, our observations emphasize that phylogrouping alone is not suitable to predict an *E. coli* as a potential NMEC although NMEC strains predominantly belong to B2 group.

Serotyping also revealed a considerable heterogeneity within NMEC and HFEC groups. Despite this diversity, we have observed certain O antigen types such as O18, O1 and O8 are more commonly present in NMEC than in commensal *E. coli*. Furthermore, we have observed a higher prevalence of H7 antigen among NMEC than in HFEC predicting an association between the H7 flagella antigen and NMEC. Previous studies have also noted that the H7 antigen is present in major virulent clones of ExPEC pathotypes, making it a potential target for future diagnostic and therapeutic interventions. Two subtypes of H7 antigen have been demonstrated for enterohemorrhagic *E. coli* (EHEC) according to the *flicC* (H7) allelic polymorphism, a phenomenon which possibly can be used to type ETEC (21, 22). For example, H7a and
H7c subtypes have been identified in *E. coli* O157: H7 irrespective of their geographical origin or genetic variation observed with other molecular typing methods (21). However, no study has been carried out to ascertain if the genetic polymorphism of H7 does exist in ExPEC pathotypes as well.

We also determined PFGE profiles and serotypes of NMEC and HFEC strains to ascertain their association with the NMEC pathotype. Pulsotyping has been used in various studies to identify pathogenic clones to distinguish between pathogenic and nonpathogenic strains of a given bacterial species (23, 24). Interestingly, in the present study, we did not come across a single pulsotype that shared between NMEC and HFEC. The closest similarity between NMEC and HFEC was observed for HFEC44 and NMEC4. However, even these two isolates were not genetically related according to Trenovor’s criteria because the similarity between the isolates was 93% with a difference in >9 bands.

Table 3-5. Number of virulence traits distributed among neonatal meningitis *E. coli* (NMEC) and fecal *E. coli* (HFEC) belonging to different phylogroups. * denotes a statistically significant difference between NMEC and HFEC. ** denotes no statistical significance.

<table>
<thead>
<tr>
<th>Phylogroup</th>
<th>No. Virulence associated traits (Mean + SD)</th>
<th><em>p</em> value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NMEC</td>
<td>HFEC</td>
</tr>
<tr>
<td>A</td>
<td>10+2.74</td>
<td>5.13+2.33</td>
</tr>
<tr>
<td>B1</td>
<td>8.0+2.83</td>
<td>6.5+-3.02</td>
</tr>
<tr>
<td>B2</td>
<td>15.19+3.21</td>
<td>4.64+1.95</td>
</tr>
<tr>
<td>D</td>
<td>12.4+3.92</td>
<td>6.38+2.79</td>
</tr>
</tbody>
</table>

Ability to form biofilms in different environmental conditions is one of the major phenotypic characteristics observed in pathogenic bacteria that help them to survive in
abiotic surfaces, and combat antimicrobials and antiseptics (25, 26). Given that biofilm formation is reported to assist in bacterial survival, colonization, and protection from host immune responses and antimicrobial agents, the ability of NMEC to produce biofilms may represent an important virulence mechanism in relation to disease pathogenesis. Therefore, we compared the ability of biofilm formation by NMEC and HFEC. The NMEC and HFEC isolates did not show a significant difference in the ability to form biofilms when they were grown in M9 medium. When LB broth was used as the medium, NMEC demonstrated a marked increase in the ability to form biofilm whereas no change was observed with HFEC. In consistent with our observations, several previous studies have demonstrated that growth medium has an effect on biofilm formation by pathogenic bacteria (13). Jackson et al., noted that central carbon flux has a regulatory effect on formation of biofilms (27). Readily available sugars have a negative regulatory effect on biofilm formation suggesting that the ability to form a biofilm is a survival strategy by bacteria when the carbon source is limited or laborious (13). The M9 medium used in current study contained 2% glucose which is a readily available carbon source whereas in LB broth the catabolizable amino acids provide the carbon source required for bacterial growth. Statistically significant increase in biofilm formation by NMEC strains in LB compared to M9 might be due to the availability of glucose in M9 implying that NMEC are more capable of responding to external stimuli to form biofilm than HFEC.

Beta-lactam antibiotics are one of the major treatment options available for controlling meningitis in infant patients (28). We evaluated the antibiotic sensitivity profiles of NMEC and HFEC along with ESBL resistance profiles. Although the majority of strains were sensitive to most antibiotics tested, 8 isolates showed an ESBL
resistance phenotype. Interestingly, three of these isolates that were ESBL resistant were fecal commensal isolates implying the fecal flora as a source for the emergence of ESBL resistance. There are several studies that indicate rapid distribution of ESBL phenotype in commensal microflora of human and animals since the most of the ESBL genes are encoded in mobilizable elements such as transposons and broad host range plasmids (29). Although a positive correlation between the presence of ESBL-positive HFEC and ESBL-positive UPEC has been observed by several investigators, there is no evidence to suggest an association between the presence of ESBL-producing bacteria in vaginal or fecal flora of pregnant women and corresponding ESBL-producing bacterial infections in neonates (30). However, it has been demonstrated that the risk factors such as premature delivery, premature membrane rupture and low birth weight which are known to predispose infants to E. coli meningitis are also involved in the colonization of ESBL-producing Enterobactericeae in the infant gut. Therefore, studies addressing these things will be necessary for adopting preventive strategies in the future.

The genetic heterogeneity of NMEC pathotype is one of the major constraints posing today to understand the pathogenesis of neonatal meningitis (6,9,11). Concordance with previous studies, we have also observed a marked heterogeneity in virulence gene profiles of NMEC. Obviously, NMEC possessed more VFs than HFEC. As aforementioned, the prevalence of some of the virulence genes (ibeA, sfa/foc, traJ and cnf1), which have previously been implicated in pathogenesis of neonatal meningitis was below 50% in NMEC. Some other virulence traits such as fimH, ompA and nlp1 that were recognized as essentially required for NMEC survival in blood and penetration of BBB were equally prevalent in both NMEC and HFEC questioning their suitability to define
the NMEC pathotype (31-36). We also observed that those NMEC and HFEC strains which did not harbor these “essential” virulence factors were still able to invade hCMEC. These data indicate that NMEC pathogenicity is poorly understood and highlight the need for future research directed at mining novel virulence traits that are truly involved in the penetration of BBB.

Interestingly, we have observed that iron siderophore system-associated traits such as *iutA, hylA, hylD, sitA* and *iucC* were highly associated with NMEC since those genes were present only in NMEC or overly represented in NMEC than HFEC. Although these siderophores were designated as virulent determinants essential for iron acquisition and enhancing the permeability of vascular endothelium due to reactors released by hemolysis, some studies have indicated that some of these siderophores are involved in other pathogenic mechanisms than iron acquisition (38). For example, sublytic amounts of *E. coli* hemolysin have been shown to increase the permeability of endothelial cell monolayers in a time- and dose-dependent manner in cultured pulmonary artery endothelial cells (38). Also, the same study demonstrated that a low dose of hemolysin was able to induce a toxin-mediated loss of endothelial barrier function. However, there is no scientific evidence explaining the involvement of siderophores to increase the permeability of BBB aiding NMEC pathogenesis. Such studies addressing the relevancy of these mechanisms in NMEC may invariably enrich the understanding the NMEC pathotype.

After observing a considerable heterogeneity in the prevalence of virulence factors in both NMEC and HFEC, we employed an *in vitro* invasion assay to correlate the virulence traits with the isolate’s ability to penetrate BBB. As expected, NMEC strains
which carried more virulence genes than HFEC were also more invasive to hCMEC/D3 cells than HFEC (Figure 3.4). Interestingly, there were 10 isolates of HFEC which showed an avirulent phenotype on in vitro cell invasion assay. All these isolates were negative for K1 capsular type, ibeA, traJ and sfal/foc indicating the importance of these genes in NMEC pathogenesis even though they are less prevalent in NMEC than some other virulence genes tested in the study. These “avirulent” isolates were positive for some other major virulence traits such as fimH, ibeB and nlpl suggesting that either these genes do not qualify to define the NMEC pathotype or although the genes are present they are not expressed in these isolates. Therefore, future studies analyzing the differences in gene expression will be a logical approach to differentiate NMEC from HFEC. Except these ten “avirulent” isolates, all other HFEC possessed the ability to invade the BBB barrier in vitro indicating fecal commensal E. coli as a potential source of NMEC. It is known that high level of bacteremia is a prerequisite for the attachment and invasion of BBB by NMEC (37, 38). Since the present study did not assess the ability of these E. coli to survive and multiply in blood, evaluation of these virulence attributes might be required to ascertain if these invasive HFEC strains indeed have the potential to cause neonatal meningitis.

We also analyzed NMEC depending on the most prevalent (>70% of the NMEC population) and overly represented virulence traits in the NMEC population as compared to HFEC (>3-fold) for further defining the NMEC pathotype. Based on these criteria, we have identified six genes namely, kpsII, K1, neuC, iucC, sitA and vat which are highly associated with NMEC pathotype. Interestingly, none of the HFEC isolates surveyed
possessed all 6 genes together whereas 40% of NMEC isolates harbored all six genes suggesting that these genes might be used to predict the NMEC pathotype (Figure 3.5).

**Figure 3-4.** Distribution of K1, sitA, vat, neuC, iucC among NMEC. Red designates presence of gene; Grey, absence of gene. Encircled area indicates the isolates and the genes used to define a typical NMEC.

In this context, we further analyzed the distribution of these six genes among the NMEC population. As depicted in Figure 3.5, depending on the capsular type, NMEC could be divided into two major groups, K1\(^+\) and K1. The majority of NMEC possessed the K\(^+\) capsular type (n=76\%) and all K1\(^+\) NMEC were also positive for sitA. Of the six genes selected to define the NMEC pathotype, vat, iucC and neuC were absent only in a minority of K1\(^+\) NMEC (0\% to 27\%). Interestingly, we noticed that among K1\(^+\) NMEC, vat and iucC were present either simultaneously or alternatively whereas no pattern was identified for neuC. In conclusion, our study results suggest that a typical NMEC pathogen might be described as K1\(^+\), sitA\(^+\) and having at least two of the three genes, vat, neuC and iucC. Using these criteria, we observed that 74\% of NMEC used in this study can be designated as a typical NMEC whereas none of the HFEC fallen within
this category. However, more NMEC isolates representing different geographical regions must be characterized for further confirmation of this phenomenon.

In conclusion, current study results indicate regardless of the genotypic and phenotypic heterogeneity observed among NMEC, some genotypic characteristics can still be used to distinguish NMEC from commensal *E. coli*.

References


Chapter 4

COMPLETE GENOME AND COMPARATIVE GENOMICS OF E. COLI RS218

4.1 Abstract

*Escherichia coli* RS218 is the prototypic neonatal meningitis causing *E. coli* (NMEC) strain that has been used in many studies relevant to NMEC pathogenesis. However, the sequencing of the RS218 genome has not been completed hampering the efforts to fully elucidate NMEC virulence and pathogenesis. In this study, the entire genome of *E. coli* RS218 including its plasmid was sequenced. Analysis of the genome revealed a circular chromosome of 5.087 Mb in size and a 114-Kbp plasmid both with an average G+C content of 50.6%. The RS218 chromosome contains 4,658 coding sequences, 88 transfer RNAs, 22 ribosomal RNAs (rRNA), one Clustered Regularly Interspaced Short Palindromic Repeats array and five noncoding rRNAs. The genome was compared with eight other completed genomes of extra intestinal pathogenic *E. coli* and the laboratory strain of *E. coli*, K12. Complete sequence analysis of *E. coli* RS218 revealed 98% similarity to cystitis causing *E. coli* strain UTI89. A total of 51 genomic islands (GIs) were present in *E. coli* RS218 which were absent from *E. coli* K12. Out of these 51 GIs, 16 were common to all NMEC studied whereas 2 GIs were common to all ExPEC. The GIs shared by all NMEC encode for several sugar uptake pathways, an acid tolerance operon, iron uptake systems and putative adhesins and invasins divulging potential virulence associated genes in *E. coli* RS218 genome that may contribute to NMEC pathogenesis.
4.2 Introduction

The whole genome sequencing of pathogens has opened up a new era of understanding the genetic basis of bacterial pathogens in terms of their clinical, pathological, epidemiological, diagnostic and evolutionary aspects (1-4). The first bacterial genome that was sequenced is the genome of Gram-negative coccobacillus, *Haemophilus influenzae* which is a human-restricted opportunistic pathogen (5). Since then, with the advances of rapid DNA sequencing technologies, DNA assembly software and annotation pipelines, many bacterial genomes have been sequenced (5). These include a variety of species representing both commensal and pathogenic strains of bacteria enabling scientists to develop testable hypotheses on genetic events that drive the bacterial adaptations to different habitats and evolution of bacterial virulence.

The first *E. coli* genome, a laboratory strain called *E. coli* K-12 MG1655, was published in 1997 (6). The circular chromosome of *E. coli* K-12 is of 4,639,221 bp in size which contains 4,288 open reading frames (ORFs) that are categorized into 2,584 operons. Additionally, this genome was observed to contain a significant number of transposable genetic elements, repeat elements, cryptic prophages and bacteriophage remnants which comprise about 18% of the total genome (6). Currently, there are more than 1000 *E. coli* genome sequencing projects reported to GenBank and of which 63 genomes have been completed and available to public (http://www.ncbi.nlm.nih.gov/genome/) (167). Moreover, comparative genomics of *E. coli* has revealed that the *E. coli* genome is open in the sense of acquiring novel genes from other bacterial species and consists of a core genome which is conserved in all *E. coli* as
well as strain- and isolate-specific genetic regions that explain genotypic and phenotypic diversity among different *E. coli* pathovars (7, 8). Altogether, it has been reported that *E. coli* pangenome (both core genome and the strain-specific regions/genes) consists approximately 13,000 genes (7). Interestingly, only 2,200 genes out of these 13,000 genes belong to the core genome, explaining the genetic basis for marked diversity observed in *E. coli* as a pathogen (7). Similarly, comparative genomics of *E. coli* with other closely related enteric bacterial genera revealed that the step-wise acquisition of foreign DNA by *E. coli* from *Salmonella*, *Yersinia* and *Vibrio* species (9). Moreover, it has been proven that genes that are horizontally transferred can be used to define a pathovar, particularly in intestinal pathogenic (diarrheagenic) *E. coli* (10). Most of these studies have been focused on intestinal pathogenic and uropathogenic *E. coli* (UPEC), and only few studies were conducted to study pathogenomics of other extra-intestinal pathogenic *E. coli* (ExPEC) strains (7, 11). The major reason for this limitation is the unavailability of adequate numbers of complete genomes of ExPEC pathotypes other than UPEC. For instance, there are only three genomes of neonatal meningitis-causing *E. coli* (NMEC) and two genomes of avian pathogenic *E. coli* (APEC) that have been completed and available for public (12-15).

As described in previous chapters, NMEC is the most ill-defined pathotype of all ExPEC even though it has been considered as one of the major pathogens associated with meningitis during early period of human life. *Escherichia coli* RS218 (hereafter referred as RS218) has been isolated from cerebrospinal fluid of a neonatal patient affected with meningitis in early 1980’s (16). This strain is considered as the prototypic strain for NMEC related research and most of the virulence traits identified with relevance to
NMEC pathogenesis have been studied using *E. coli* RS218 (17-20). Although sequencing of RS218 genome has been started in 2006 by a group of investigators at the University of Wisconsin (Madison, WI) under a National Institutes of Health (NIH)-funded genome project, the complete sequence of the genome is still either not completed and/or not available to the public (http://www.genome.wisc.edu/sequencing/rs218.htm).

Considering that the studies of NMEC aimed at identifying virulence traits that can fully explain NMEC pathogenesis are hampered by the unavailability of RS218 genome sequence, we determined to pursue that task. It is expected that the genome sequence of RS218 will aid in unrevealing potential virulence traits of RS218 that might fill the knowledge gaps pertaining to NMEC pathogenesis.

**4.3 Materials and Methods**

**4.3.1 *Escherichia coli* RS218 strain**

*Escherichia coli* RS218 strain was kindly provided by Dr. James Johnson (University of Minnesota, St Paul, MN). It has been isolated from the cerebrospinal fluid of an infant with meningitis in the 1980’s (16). This strain belongs to the B2 based on phylogrouping and sequence type 95 (ST95) according to multilocus sequence typing. The serotype of RS218 is O18:H7:K1 (16).
4.3.2 Genomic DNA extraction

Genomic DNA was isolated using Promega Genomic Wizard Kit (Promega Corporation, Madison, WI). Briefly, 1 mL of overnight culture was centrifuged for 5 min at 15,000 x g in a microcentrifuge. The pellet was resuspended in 480 μl of 50 mM ethylenediaminetetraacetic acid (EDTA) and processed for DNA extraction according to the manufacturer’s instructions. The genomic DNA pellet was dissolved in sterile distilled H2O, quantified using Nanodrop (Thermo Scientific, Wilmington, DE) and stored at -20°C.

4.3.3 Genome Sequencing

Genome sequencing was performed by Ion Torrent PGM Sequencing Technology (Life Technologies, Grand Island, NY) at the Genomics Core Facility of the Pennsylvania State University (University Park, PA) using a 318 chip to provide over 100-fold coverage of the genome. Approximately, 100 ng of high purity genomic DNA was submitted for sequencing.

4.3.4 Construction of the whole genome optical map (OpMap)

A whole genome restriction optical map was generated using NcoI digestion by OpGen, Inc. (Gaithersburg, MD) (21). In brief, high molecular weight DNA was extracted linearized, immobilized on to a Mapcard containing microchannels which hold single chromosomes and subsequently digested with NcoI restriction enzyme. Resultants
genomic DNA fragments were stained with a fluorescence dye, lengths were measured using fluorescent microscopy and assembled by overlapping fragments to generate a map of restriction cut sites. Restriction map was visualized using MapSolver software version 3.0 (OpGen).

### 4.3.5 Genome assembly and gap closure

Whole genome Ion Torrent sequencing generated approximately six million reads with an average length of 250 ± 31 bp. These short reads were assembled using both de novo and reference-guided assembly by SeqMan NGen 11.0 (DNASTAR, Madison, WI) and visualized using SeqMan Pro 11.0 software (DNASTAR). *Escherichia coli* UTI89 (CP000243) was selected as the reference genome for reference-guided assembly based on its closest similarity to *Nco*I restriction map of RS218 map. The genomic contigs obtained by both assemblies were ordered according to the OpMap to acquire correct orientation in the chromosome. Plasmid contigs were identified using Basic Local Alignment Search Tool for nucleotides (BLASTn) algorithm from National Center for Biotechnology Information (NCBI, [http://blast.ncbi.nlm.nih.gov/Blast.cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi)) and assembled using the reference plasmid sequence, pUTI89 (CP000243.1). Ambiguous or low coverage areas and gaps in both the chromosome and the plasmid were closed by primer walking method using Universal Primer Walking Kit (Clontech Laboratories, Mountain View, CA) followed by Sanger sequencing at the Penn State University’s Genomics Core Facility (University Park, PA).
4.3.6 Genome Annotation


4.3.7 Identification of repetitive elements and prophage regions

Initially predicted mobile elements through annotation pipeline were blasted using ISFinder web tool (https://www-is.biotoul.fr/) to identify the types of insertion sequences (IS) within the RS218 genome (23). Phage related sequences were searched using PHAge Search Tool (PHAST, http://phast.wishartlab.com) (24).
4.3.6 Genome Annotation

The complete sequence of RS218 chromosome was aligned to other ExPEC chromosomes, namely, IHE3034 (CP001969), S88 (CU928161), CE10 (NC_017646.1), CFT073 (AE014075), 536 (CP000247), UTI89 (CP000243), APECO1 (CP000468), APECO78 (CP0004009) and the laboratory strain E. coli K12 MG155 (U00096) using BLASTn algorithm and Mavue alignment tool (25). Alignments were visualized using Blast Ring Image Generator (BRIG) and Circular Genome Viewer (CG View) (26, 27). Each GI identified in RS218 genome in comparison to E. coli K12 was blasted against the selected ExPEC genomes with the cutoff values of 90% similarity and 90% identity.

4.4 Results

4.4.1. General characteristics of RS218 genome

The RS218 genome consists of a single chromosome (approximately 5.087 Mb in size) and a large 114,231-bp plasmid (pRS218). The RS218 chromosome is considerably larger than the E. coli K12 chromosome (4.6 Mb) and similar in size to most of the other ExPEC chromosomes such as UPEC strains CFT073 (5.23 Mb), UTI89 (5.21 Mb) and 536 (4.94 Mb), and APEC strain APECO1 (5.5 Mb). The RS218 chromosome contains 4,658 coding sequences (CDS), 88 transfer RNAs (tRNA), 22 ribosomal RNAs (rRNA), one Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) array and five noncoding rRNAs. The average GC content is 50.6% which is common to most E. coli
that have been sequenced (7, 28). However, a marked variation in GC content of some coding sequences which ranged from 46% to 54% was observed suggesting a horizontal gene transfer (Figure 4.1).

Apart from the chromosome, RS218 harbors a large conjugative plasmid with a similar GC content (51.02%). However, there are certain regions in the plasmid which are flanked by Insertion Sequence (IS) elements with different GC contents and showed homology to some genes identified in Salmonella, Yersinia and Vibrio species suggesting an IS-mediated acquisition of foreign DNA. Descriptive analysis of plasmid sequence and its virulence are discussed in Chapter 5. The complete annotated genome of RS218 was deposited in NCBI Genbank (http://www.ncbi.nlm.nih.gov/bioproject) under the project accession ID PRJNA78291. General features of the genome and plasmid are shown in Table 4.1, and schematic representations are shown in Figure 4.1.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>5,087,638 bp</td>
</tr>
<tr>
<td>GC content</td>
<td>50.06%</td>
</tr>
<tr>
<td>Genes</td>
<td>5043</td>
</tr>
<tr>
<td>Coding sequences</td>
<td>4658</td>
</tr>
<tr>
<td>rRNA (5S,16S,23S)</td>
<td>22</td>
</tr>
<tr>
<td>tRNA genes</td>
<td>88</td>
</tr>
</tbody>
</table>

**Table 4-1. General characteristics of E. coli RS218 genome.**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>114,231 bp</td>
</tr>
<tr>
<td>GC content</td>
<td>51.06%</td>
</tr>
<tr>
<td>Genes</td>
<td>160</td>
</tr>
<tr>
<td>Coding sequences</td>
<td>130</td>
</tr>
</tbody>
</table>
Figure 4-1. Circular map of the chromosome of *E. coli* RS218. The innermost ring represents the *E. coli* K12 chromosome used as a reference and its coordinates. The second ring (in purple/green) plots the GC skew of the reference, followed by the black ring which plots the G+C content. Red ring indicates the BLASTn comparison between the chromosomes of *E. coli* RS218 and *E. coli* K12. Blue arrows depict the coding sequences of RS218 in forward and reverse strands (inner and outer rings, respectively). The outermost red arrows indicate the genomic islands (GIs) identified in RS218 which were absent in *E. coli* K12.

4.4.2. Comparative genomics of RS218

The overall sequence architecture of RS218 genome was compared to the genome of *E. coli* laboratory strain K12 MG1655. Pairwise alignment of two genomes identified 51 genomic islands (GIs) which were unique to *E. coli* RS218 chromosome (Figure 4.2 and Table 4.2). Some of these GIs encode for phage regions (Figure 4.3). Furthermore, whole genome sequence comparison revealed that the closest sequence similarity of *E.
coli RS218 chromosome was with the UPEC strain UTI89 (Table 4.4). Distribution of these GIs among ExPEC strains used in this study revealed that the majority of GIs were commonly associated with ExPEC belonging to B2 phylogroup and in particular, with the ExPEC belonging to the MLST ST95 clone (Figure 4.4). Out of 51 GIs, two GIs were common to all ExPEC and 16 GIs were common to all NMEC strains analyzed (Figure 4.4).

**Figure 4-2.** Pairwise alignment of *E. coli* RS218 genome to the laboratory strain *E. coli* K12 using Mauve. Locally collinear blocks (LCBs) of DNA are depicted in the same color and connected via corresponding connection lines. Inverted regions of *E. coli* K12 genome compared to the orientation of *E. coli* RS218 genome are drawn below the axis. White and grey regions in each genome indicate strain-specific regions.
Table 4-2. Characteristics of genomic islands (GIs) identified in *E. coli* RS218 genome.

<table>
<thead>
<tr>
<th>GI #</th>
<th>Start</th>
<th>Stop</th>
<th>Size(Kbp)</th>
<th>Proteins encoded by the gene/region</th>
<th>Flanking tRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150613</td>
<td>160018</td>
<td>9.4</td>
<td>Fimbrial proteins Yad</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>228804</td>
<td>266534</td>
<td>37.7</td>
<td>Type IV secretory system</td>
<td>tRNA Asp-GGT</td>
</tr>
<tr>
<td>3</td>
<td>289950</td>
<td>321781</td>
<td>31.8</td>
<td>Vac, CFA/I fimbrial cluster</td>
<td>tRNA-Thr-CGT</td>
</tr>
<tr>
<td>4</td>
<td>328912</td>
<td>333812</td>
<td>4.9</td>
<td>Hypothetical</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>352727</td>
<td>357854</td>
<td>5.1</td>
<td>Ribose ABC transport</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>666141</td>
<td>672069</td>
<td>5.9</td>
<td>Putative transporter, Alcohol dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>833360</td>
<td>867256</td>
<td>33.9</td>
<td>Phage (intact)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>905310</td>
<td>913855</td>
<td>8.5</td>
<td>CRISPR-associated proteins</td>
<td>tRNA-Ser-GGA</td>
</tr>
<tr>
<td>9</td>
<td>1060208</td>
<td>1123164</td>
<td>63.0</td>
<td>S fimbrial cluster, Salmochelin, Antigen 43</td>
<td>tRNA-Ser-GGA</td>
</tr>
<tr>
<td>10</td>
<td>1219583</td>
<td>1268928</td>
<td>49.3</td>
<td>Phage, SitABCD</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1298752</td>
<td>1307372</td>
<td>8.6</td>
<td>Ton B dependent OMP, Hemin ABC transporter, Trehalase</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1381227</td>
<td>1416358</td>
<td>35.1</td>
<td>Phage</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1453629</td>
<td>1461121</td>
<td>7.5</td>
<td>RND efflux system</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1628585</td>
<td>1643663</td>
<td>15.1</td>
<td>Type I fimbriae, Phage related proteins</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1658146</td>
<td>1668394</td>
<td>10.2</td>
<td>PTS system-chitobiase-specific components</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1872257</td>
<td>1910488</td>
<td>38.2</td>
<td>Phage</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2090634</td>
<td>2128906</td>
<td>38.3</td>
<td>Yersiniabactin, Adhesin and Invasin</td>
<td>tRNA-Asn-GTT</td>
</tr>
<tr>
<td>18</td>
<td>2133298</td>
<td>2235245</td>
<td>101.9</td>
<td>Phage proteins, Yersiniabactin</td>
<td>tRNA-Asn-GTT</td>
</tr>
<tr>
<td>19</td>
<td>2259035</td>
<td>2266332</td>
<td>7.3</td>
<td>O antigen modification protein, Glycosyltransferase</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2322143</td>
<td>2331637</td>
<td>9.5</td>
<td>Putative inner membrane and cytoplasmic proteins</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2412032</td>
<td>2422963</td>
<td>10.9</td>
<td>Sugar fermentation and transport</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>2661646</td>
<td>2658013</td>
<td>41.6</td>
<td>Phage</td>
<td>tRNA-Arg-CCT</td>
</tr>
<tr>
<td>23</td>
<td>2919732</td>
<td>2924825</td>
<td>5.1</td>
<td>Putative proteins</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>3307311</td>
<td>3335791</td>
<td>28.5</td>
<td>Capsular polysaccharide island</td>
<td>tRNA-Phe-GAA</td>
</tr>
<tr>
<td>25</td>
<td>3121760</td>
<td>3157032</td>
<td>35.3</td>
<td>Putative type VI secretory proteins, Sugar fermentation</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>3336285</td>
<td>3341820</td>
<td>5.5</td>
<td>Prepilin, Acessory colonization factor (AcfD) precursor</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>3554984</td>
<td>3362306</td>
<td>7.3</td>
<td>Conserved TRP repeat protein</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>3408660</td>
<td>3418262</td>
<td>9.6</td>
<td>Putative iron ABC transporters, ferrochrome-iron receptor</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>3431868</td>
<td>3438964</td>
<td>7.1</td>
<td>Uncharacterized fimbrial like proteins</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>3635820</td>
<td>3644315</td>
<td>8.5</td>
<td>Ribose ABC transport system</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>3795890</td>
<td>3804697</td>
<td>8.8</td>
<td>Putative fimbrial operon</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>3838120</td>
<td>3845038</td>
<td>6.9</td>
<td>PTS system, Mannose-specific components</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>3864570</td>
<td>3870170</td>
<td>5.6</td>
<td>Putative carbohydrate PTS system</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>3891923</td>
<td>3900874</td>
<td>9.0</td>
<td>Hemin transport proteins, ChuA</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>4012262</td>
<td>4017740</td>
<td>5.5</td>
<td>Putative lipoprotein.</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>4036410</td>
<td>4045507</td>
<td>9.1</td>
<td>O antigen glycosylation proteins</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>4069493</td>
<td>4076436</td>
<td>6.9</td>
<td>PTS system, Fructose-specific components</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>4259240</td>
<td>4275085</td>
<td>15.8</td>
<td>PTS system, Glucose-specific IIIB components, Sialic acid utilization</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>4276986</td>
<td>4282425</td>
<td>5.4</td>
<td>Tricarboxylate transport</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>4325074</td>
<td>4333727</td>
<td>8.7</td>
<td>Sugar kinase, Putative transport protein</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>4385586</td>
<td>4395426</td>
<td>9.8</td>
<td>Nucleotidase</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>4492336</td>
<td>4500704</td>
<td>8.4</td>
<td>PTS system, Sorbose-specific components</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>4529460</td>
<td>4576064</td>
<td>46.6</td>
<td>Phage</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>4575338</td>
<td>4582368</td>
<td>7.0</td>
<td>Hydroxybutyryl-CoA dehydratase, Propionate CoA-transferase</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>4588020</td>
<td>4601367</td>
<td>13.3</td>
<td>Succinyl-CoA ligase, Dihydrolipoamide dehydrogenase, Malate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>4629636</td>
<td>4637226</td>
<td>7.6</td>
<td>Puatative hemin protease, Formate dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>4762663</td>
<td>4772351</td>
<td>9.7</td>
<td>Hydroxybutyryl-CoA dehydratase, CoA-transferase</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>4815175</td>
<td>4821558</td>
<td>6.4</td>
<td>Arginine/ornithine metabolism</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>4840815</td>
<td>4963833</td>
<td>123.0</td>
<td>Pap operon, CNF1, RTX toxin transporter, Putative hemin receptors</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>4982442</td>
<td>5003447</td>
<td>21.0</td>
<td>GimA, lbeA</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>5014081</td>
<td>5018158</td>
<td>4.1</td>
<td>Type I restriction-modification system, Specificity subunit S</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4-3.** Characteristics of bacteriophages in *E. coli* RS218 genome. A, schematic diagram of relative positions of phages inserted in *E. coli* RS218 genome. Red, intact phage; Green, phages lacking few essential genes; Grey, incomplete phage. B, open reading frames present in each phage region with relevance to phage-associated functions.
Table 4-3. Comparison of general characteristics of *E. coli* RS218 genome with other sequenced ExPEC genomes used in this study.

<table>
<thead>
<tr>
<th><em>E. coli</em> strain</th>
<th>Pathotype</th>
<th>Phylogroup</th>
<th>Serotype</th>
<th>ST type</th>
<th>Size (Mb)</th>
<th>GC%</th>
<th>Gene</th>
<th>Protein</th>
<th>Plasmids</th>
<th>Coverage %</th>
<th>Identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS218</td>
<td>NMEC</td>
<td>B2</td>
<td>O18:K1:H7</td>
<td>95</td>
<td>5.08</td>
<td>50.06</td>
<td>5,043</td>
<td>4658</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IHE3034</td>
<td>NMEC</td>
<td>B2</td>
<td>O18:K1:H7</td>
<td>95</td>
<td>5.11</td>
<td>50.7</td>
<td>4,966</td>
<td>4,753</td>
<td>1</td>
<td>96</td>
<td>99</td>
</tr>
<tr>
<td>S88</td>
<td>NMEC</td>
<td>B2</td>
<td>O45:K1:H7</td>
<td>95</td>
<td>5.17</td>
<td>50.7</td>
<td>5,187</td>
<td>4,823</td>
<td>1</td>
<td>93</td>
<td>99</td>
</tr>
<tr>
<td>CE10</td>
<td>NMEC</td>
<td>D</td>
<td>O7:K1:H7</td>
<td>62</td>
<td>5.38</td>
<td>50.6</td>
<td>5,269</td>
<td>5,080</td>
<td>4</td>
<td>86</td>
<td>98</td>
</tr>
<tr>
<td>CFT073</td>
<td>UPEC</td>
<td>B2</td>
<td>O6:K2:H1</td>
<td>73</td>
<td>5.23</td>
<td>50.5</td>
<td>5,574</td>
<td>5,364</td>
<td>0</td>
<td>92</td>
<td>99</td>
</tr>
<tr>
<td>UPEC536</td>
<td>UPEC</td>
<td>B2</td>
<td>O6:K15:H31</td>
<td>92</td>
<td>4.94</td>
<td>50.5</td>
<td>4,779</td>
<td>4,619</td>
<td>0</td>
<td>91</td>
<td>99</td>
</tr>
<tr>
<td>UTI89</td>
<td>UPEC</td>
<td>B2</td>
<td>O18:K1:H7</td>
<td>95</td>
<td>5.18</td>
<td>50.6</td>
<td>5,272</td>
<td>5,162</td>
<td>0</td>
<td>98</td>
<td>99</td>
</tr>
<tr>
<td>APEC O1</td>
<td>APEC</td>
<td>B2</td>
<td>O1:K1:H7</td>
<td>95</td>
<td>5.50</td>
<td>50.3</td>
<td>4,968</td>
<td>4,853</td>
<td>2</td>
<td>93</td>
<td>99</td>
</tr>
<tr>
<td>APEC O78</td>
<td>APEC</td>
<td>B1</td>
<td>O78:K80</td>
<td>23</td>
<td>4.80</td>
<td>50.7</td>
<td>4,695</td>
<td>4,588</td>
<td>0</td>
<td>80</td>
<td>98</td>
</tr>
<tr>
<td>K12</td>
<td>Lab</td>
<td>A</td>
<td>OR:H48:K</td>
<td>08</td>
<td>4.64</td>
<td>50.8</td>
<td>4,288</td>
<td>4,141</td>
<td>-</td>
<td>79</td>
<td>98</td>
</tr>
</tbody>
</table>

**Figure 4-4.** Distribution of RS218 genomic islands (GIs) among other ExPEC strains used in the study. Blue, presence of GI; Grey, absence of GI; GI designations correspond to Table 4.2.

**Figure 4-5.** Clustering of RS218 genomic islands (GIs) among ExPEC strains used in this study for comparison.
4.5 Discussion

*Escherichia coli* RS218 is the reference strain that has been used mainly in the studies for understanding the pathogenesis of NMEC (17-20). Here, we performed the whole genome sequencing of *E. coli* RS218 and analysed its genome in detail. Then, the RS218 genome was compared with the genomes of *E. coli* K12, NMEC and APEC strains of which the complete genome sequences are available to public and three reference strains of UPEC (UTI89, UPEC536 and CFT073).

*Escherichia coli* K12 is originally isolated from a convalescent diphtheria patient in Palo Alto, CA in 1922, and used as a reference genome sequence to compare many pathogenic *E. coli* genomes because of its noninvasive phenotype in the intestinal epithelium (4, 7, 29). Alignment of RS218 genome with K12 showed the existence of ten locally collinear blocks revealing the overall conservation of core genome and architecture of *E. coli* species in general (Figure 4.2). However, it was observed that some of these blocks were inverted, and/or rearranged in orientation compared to each other. Also, the strain-specific regions are located within or outside of these collinear blocks. Both these observations explain the genomic differences between two strains that might be responsible for their differences in virulent (RS218) and avirulent (K12) *E. coli* phenotypes. Therefore, we analyzed RS218-specific CDS depending on the orthology explained by cutoff values of >90% and a coverage of >90%. We identified 1249 unique CDS in RS218 compared to K12. For further analysis, these ortholog genes were clustered into genomic islands (GIs) according to their locations in the RS218 genome, and more than 4 kb regions were designated as GIs.
Accordingly, we found 51 GIs with sizes ranged from 4.07 Kbp to 123.08 Kbp (Table 4.2). Interestingly some of these islands encode virulence proteins such as fimbrial adhesins, invasins, cytotoxins and iron acquisitions systems, and are flanked by tRNA sites with some containing phage-related or mobile genetic elements (Table 2). Some of these traits (capsular polysaccharides, invasin of brain endothelium, S fimbriae, etc.) were previously described as essential virulence factors in *E. coli* K1 (*E. coli* with K1 capsular type) pathogenesis whereas some others such as Vat, Salmochelin, SitABCD and invasins of Peyer’s patches have been recognized as potential virulence markers of other pathogenic *E. coli* (17-19, 30). Significance of these factors in *E. coli* K1 pathogenesis will be an interesting area of future research.

Phage regions in bacterial genome are interesting because they provide information regarding pathogen evolution. Nevertheless, some phage-associated genes are regarded as essential elements of pathogenesis of certain bacterial diseases. For example, major virulence determinants of enterohemorrhagic *E. coli*, Shiga toxins and type III secretory apparatus have been acquired by *E. coli* from *Shigella* via phages by means of horizontal gene transfer (31). Therefore, the RS218 genome was analyzed for the presence of phages and phage-associated genes using PHAST to identify the genes/regions that might have acquired by RS218 from bacteriophages. A total of eight phage regions were identified. These include five intact phages which encode all essential proteins for a complete phage as well as three incomplete or partial phage areas that lack some of the essential elements (Figure 4.3). The relative location and the details of each phage region are depicted in Figure 3. Interestingly, three of these
phage regions indicate horizontally-acquired virulence genes; sitABCD from *Salmonella*, HPI from *Yersinia* and the capsular polysaccharide operon (Table 4.2). Some of the virulence proteins associated with these areas such as capsular polysaccharide has already been demonstrated in relation to pathogenesis of NMEC (30). However, virulence traits encoded by these regions other than the capsular polysaccharides have not been studied in NMEC. These traits might have a strong association with virulence enhancing the pathogenic potential of RS218 and other NMECs as demonstrated in other pathogenic *E. coli* including UPEC.

There were several virulence or fitness-associated traits such as acquisition systems, adhesins, invasins, toxins, secretory systems present in the RS218 chromosome. Among those traits, iron is a vital element for survival and multiplication of bacteria because of its role in respiration, DNA replication and protection against oxidative stress (32-34). Since iron is tightly bound to proteins such as hemoglobin, transferrin, lactoferrin and ferritin, scarcity of free iron is one of the major obstacles encountered by bacterial pathogens within the mammalian host (33). Therefore, the acquisition of iron that is already bound to iron-binding proteins is a prerequisite for pathogens particularly at extraintestinal locations (35). In ExPEC, multiple iron acquisition systems encoded in the chromosome and/or plasmids have been identified (13, 34). These uptake systems can be divided into three main categories: siderophore-based systems, heme acquisition systems, and transferrin/lactoferrin receptors (35). Recent studies have shown that harboring multiple siderophore systems are more common in ExPEC than in commensal strains of *E. coli* (36-38). The chromosome of *E. coli* RS218 encodes six iron acquisition systems (salmochelin, enterobactin, yersiniabactin, FhuBCD, SitABCD and EfeU) and
several hemin, ferritin and ferrochrome receptors (e.g., ChuA, PiuC, CorA, etc).
Interestingly, most of these proteins are encoded by genes located on GIs and are discussed in detail below (Table 2).

Salmochelins are glucosylated derivatives of enterobactin which were initially identified in Salmonella sp. (39). Salmochelin operon encodes five (iroBCDEN) proteins, IroB; which glycosylates a portion of enterobactin to produce salmochelin that enhances hydrophobicity to siderophore and inability to be bound by the host proteins, IroC which is an ABC transporter that exports salmochelin, IroN which is a receptor for iron-bound salmochelin that transports iron back to periplasm, IroE which is a periplasmic esterase which degrades salmochelin to the linear trimeric form and IroD which degrades salmochelin (39). Recent studies have shown that the salmochelin is also prevalent in all subpathotypes of ExPEC (40). It has been well studied as an urovirulence trait that is implicated in urinary tract infections (36, 40, 41) Besides its role in iron acquisition, IroN, the siderophore receptor in salmochelin operon has also shown to facilitate internalization of bacteria into uroepithelial cells in vitro (41). Moreover, high prevalence of salmochelin (64%) has also been perceived in NMEC strains (42). In E. coli RS218, the complete iroBCDEN operon is located on GI 9 together with S fimbrial /Antigen 43 cluster. Negre et al. 2004, demonstrated that the iroN itself plays a key role in the virulence of NMEC strain C5 (O18:K1:H7) compared to other two iron acquisition systems, yersinibactin and ChuA, and implicated in high level of bacteremia which is a prerequisite for causing meningitis (42). Additionally, involvement of salmochelin encoded by the virulence plasmid of NMEC strain E. coli S88 (O45:K1:H7) in bacteremia has been demonstrated by Peigne et al. 2009 (15). Taken together, high
prevalence of salmochelin in NMEC strains including prototypic NMEC strain RS218 and previous observations relating to its key role in the occurrence of high level of bacteremia, indicate that salmochelin is not only a urovirulent trait but might also be a meningitic trait.

Yersiniabactin is another type of siderophores which are encoded by the high pathogenicity island (HPI) in highly pathogenic Yersinia sp. (43). In E. coli, the same HPI was found frequently in phylogenetic group B2 which is considered as a virulence clone of ExPEC (44). Schubert et al 2002, has demonstrated that HPI is widely distributed among pathogenic E. coli (45). Particularly, it is overly present in ExPEC compared to other pathogenic E. coli (45). Moreover, they found a higher correlation of HPI virulence in ExPEC compared to other important virulence traits such as pap, sfa/foc, aer, hly, K1 antigen, afa, and ibeA using mouse infection models of ExPEC infection providing a strong evidence for the role of HPI in ExPEC pathogenesis (45). The complete operon of yersiniabactin is located on the GIs 17 and 18 of RS218 flanked by tRNA^Asn-GTT which is similar to HPI in pathogenic Yersinia sp. (Table 1). Interestingly, although there are no studies that explain specific involvement of HPI in NMEC virulence, previous observations with ExPEC suggest that HPI might be an important virulence trait of RS218 involved in its pathogenesis (45, 46).

Another interesting finding of RS218 genome is the presence of Salmonella Pathogenicity Island 1-associated iron acquisition operon, sitABCD in the genomic island 10 of RS218 (Table 2). The sitABCD operon on this island is located just downstream to a lambda family intact phage revealing phage-associated introduction of foreign DNA to the RS218 genome. The functional significance of sitABCD operon has been well studied
in APEC (47,48). Unlike RS218, sitABCD operon in APEC is located on the virulence plasmid pAPEC-1 (47). Functional analysis using isotope uptake experiments has shown that sitABCD in APEC strain χ7122 mediates transport of iron and manganese (47). Both of these ions are essential for bacterial growth as well as to resist detrimental effects of hydrogen peroxide (H$_2$O$_2$) which is one of the major bactericidal mechanisms of phagocytes (47). It has been proven that SitABCD is involved in the resistance to H$_2$O$_2$ in Salmonella, Shigella and APEC (49, 50). However, there are no such experiments done to identify the functional correlation of sitABCD in human ExPEC although high prevalence of sitA has been observed in NMEC and UPEC. Therefore, it will be an interesting candidate to study immune evasion mechanisms of NMEC.

All iron acquisition systems discussed above are present in RS218 but not in E. coli K12. However, there are some iron acquisition systems in RS218 genome which are also present in K12 which may have a role in providing fitness to the bacterium for pathoadaptation. Siderophore enterobactin is a triscatechol derivative which enables bacteria to recover iron from their environment (51). Enterobactin is synthesized under low iron conditions and excreted into the environment where it binds Fe (III) with high affinity and specificity (51). The ferric siderophore complexes are taken up into the cells by specific ATP-binding cassette transporter complex (52). This siderophore system is common to many Gram-negative bacteria including both pathogenic and commensal E. coli and also recently found in two Gram-positive Streptomyces species (51).

Additionally, RS218 genome codes for FhuB, C and D which are required for the uptake of iron from aerobactin and other siderophores. However, RS218 does not encode an aerobactin synthesis system such as IucA, B, C and D or IutA indicating that FhuBCD is
involved in iron acquisition mechanisms in place of aerobactin. In *E. coli*, it has been demonstrated that FhuCDB transporter also accepts at least four different siderophores except aerobactin and it also provides binding sites for colicin and phages (53). However, more studies will be required to implicate the relevance of FhuCDB to NMEC pathotype or ExPEC virulence. Moreover, RS218 genome also contains genes that encode EfeU iron acquisition proteins which are specifically induced under low pH and anaerobic or microaerophilic environmental conditions (54). It is believed that neonates acquire pathogenic *E. coli* during the delivery via vaginal microflora of the mother (55). Human vagina is acidic due to the presence of lactobacilli, and under low pH, the ferric irons (Fe$^{3+}$) availability is very low since at low pH iron exists as ferrous (Fe$^{2+}$). Therefore, harboring the EfeU iron acquisition system which is induced by acidic environmental conditions and binds to Fe$^{2+}$ might be an important strategy exploited by RS218 in vaginal colonization to enhance its fitness in this particular niche. Other than these well studied iron acquisition receptors, RS218 also harbors several putative hemin receptors, iron-binding proteins and ferritin-like proteins that might be involved in enhancing the fitness of RS218 in iron-limited host tissues.

Bacterial adherence to host cells via adhesins represents a crucial step during the establishment of an infection (56). Most of these structures found in pathogenic bacteria have a dual role as invasins and adhesins while some act only as either adhesins or invasins (56). Mainly there are two types of adhesins present in pathogenic bacteria, fimbrial and nonfimbrial adhesins (56). Fimbrial adhesins provide a receptor-mediated contact with host tissues which facilitate microbial colonization at mucosal surfaces, biofilm formation as well as initiation of internalization of some pathogenic bacteria by
host cells (56). Regulatory cross-talk between fimbrial operons has been described for different fimbrial determinants of pathogenic E. coli (57). In RS218 genome, we have identified 8 fimbrial clusters, namely, mannose-specific type I fimbriae, sialic acid-specific fimbriae (S fimbriae), pyelonephritis-associated fimbriae (P fimbriae), meningitis-associated and temperature-regulated fimbriae (Mat fimbriae), curli fimbriae, Yad fimbriae, type IV fimbriae and YgiL fimbriae. Except type I fimbriae, Curli and type IV fimbriae, all the other fimbrial gene clusters in RS218 genome are located within genomic islands (Table 4.2). Interestingly, all these fimbriae have been implicated in pathogenesis of various infections caused by E. coli (58-60).

In neonatal meningitis, type I fimbriae have been implicated in binding to and invasion of human brain microvascular endothelial cells (BMEC) (55). The tip adhesin of type I fimbriae, FimH, binds with CD48 which increases the Rho-dependent changes in the cytoskeleton that facilitates internalization of bacteria by the BMEC cell (55). Similarly, S fimbriae have also been identified in relation to NMEC virulence (60). Saukkonen et al. 1988 has observed that S fimbrial bacterial clones were more virulent in a rat pup model of neonatal meningitis compared to type I fimbriae (61). However, Yang et al. 2004 found that S fimbriae did not play a major role in E. coli K1 binding to human BMEC in vitro and crossing the blood-brain barrier in vivo suggesting an alternative role of S fimbriae in NMEC virulence rather than acting as an adhesin (62).

Another three fimbrial gene clusters present in RS218, Yad, Pap and YgiL have been identified as urovirulence markers (58). For example, Yad fimbriae adhere to bladder epithelial cells whereas YgiL and Pap adhere to renal tubular cells (58). Possession of these adhesins implies an ability of RS218 to adhere to uroepithelial cells
which might act as a fitness factor for the survival of bacteria in the urogenital tract (58). In contrast, Mat and Curli fimbriae are specifically involved in adherence of bacteria to the intestinal epithelial cells and formation of biofilms which is another strategy of environmental adaptation seen in most pathogenic bacteria revealing that these adhesins might be involved in RS218 colonization of maternal intestinal and uroepithelial cells as well as initial colonization of neonates by the bacteria.

In addition to fimbrial adhesins, RS218 also possesses several non-fimbrial adhesins. Unlike fimbrial adhesins, these are embedded in bacterial cell wall and often associated with self-recognition (autoaggregation) and biofilm formation (63). The RS218 chromosome encodes two well characterized non-fimbrial adhesins, adhesin involved in diffuse adherence-I (AIDA-I) and Antigen 43, and several putative adhesins. Both AIDA-I and antigen 43 are autotransporter proteins associated with Type V secretion system (64, 65).

The autotransporter protein, Antigen 43 has been identified in many pathogenic and nonpathogenic bacteria that are known to form biofilms, and adhere to and internalize into intestinal epithelial cells (65). Several variant phenotypes of Antigen 43 and their functional differences have also been observed in E. coli (66). For example, Antigen 43 encoded by E. coli K12 has shown to provide an autoaggregation phenotype whereas Antigen 43 in UPEC CFT073 and RS218 has shown to lack this functionality while maintaining its ability to form biofilms.

The autotransporter protein, AIDA-I was first identified in E. coli strain 2787 which was isolated from a patient having infantile diarrhea (67). Subsequently, this protein was identified in many swine enterotoxigenic E. coli strains causing edema and
post-weaning diarrhea. It has been implicated as a virulence factor that facilitates adherence of these pathogenic \textit{E. coli} to intestinal epithelium (68). Interestingly, AIDA-I in RS218 genome is located as a single insertion just adjacent to tRNA\textsuperscript{Pro} gene without associated mobile elements or phage sequences. This gene is not present in the K12 genome. However, the heptosylation gene (\textit{aah}) which has been demonstrated as an essential factor for the formation of a functional AIDA-I adhesin was not observed in the RS218 genome suggesting that AIDA-I encoded by RS218 genome is not functional and has no significant role in RS218 adherence to epithelia (69). However, AIDA-like adhesins have been linked to APEC and UPEC pathogenesis (70). For instance, AIDA-I like adhesin UpaH in uropathogenic \textit{E. coli} involved in biofilm formation and colonization of the bladder epithelial cells (70). The RS218 genome also encodes for two AIDA-like adhesins. Therefore, these autotransporter proteins encoded by RS218 might be an interesting area for future studies.

We found another adhesin/invasin-like protein associated with GI 1 which belongs to the autotransporter (AT) family of proteins. Blasting of its protein sequence identified a conserved domain at the C terminus similar to adhesins/invasins such as YadA in pathogenic \textit{Yersinia}, Saa in LEE-negative Shiga toxin-producing \textit{E. coli} (STEC), and \textit{E. coli} immunoglobulin-binding (Eib) adhesin of virotoxigenic \textit{E. coli} and LEE-negative STEC (71-73). Regardless of structural difference at the N terminus of these AT proteins, they all have been implicated in enterocyte attaching and effacing phenotype as demonstrated by studies that used HeLa cells (73). Since no virulence trait of NMEC has been linked to initial colonization of NMEC, these AT proteins seem to be promising targets to look for such mechanisms.
Invasins are virulence proteins of bacteria those promote internalization of bacteria into host epithelial cells. In neonatal meningitis, there are two protective epithelial barriers encountered by NMEC in the process of bacterial entry into the meninges; mucosal epithelial cells and BMEC (blood brain barrier or BBB). In NMEC possessing the K1 capsular type, there are quite a few invasins have been identified with regard to penetration of BBB (30). These include invasins of brain endothelium (Ibe proteins), outer membrane protein A (OmpA), arysulfatase (AslA), conjugal transfer protein J (TraJ) encoded in plasmid and new lipoprotein 1 (NlpI) (17, 19, 74, 75). The functions of all of these invasins have been studied in depth to understand their relevance to NMEC pathogenesis and are described in Chapter 2. However, there are no studies on virulence traits that are involved in the penetration of intestinal mucosal barrier. In the RS218 genome, several putative adhesins, particularly related to intestinal epithelial adhesion and invasion, were detected. Some of these invasins such as YadA adhesin/invasin have been shown to have both adhesive and invasive properties, and are discussed above with adhesins. Other putative invasins encoded by RS218 at locus numbers W817_11190 (coordinates 1329217 to 1330611) and W817_07205 (coordinates 1329217 to1330611), and invasin-like protein encoded by GI 17 with yersiniabactin cluster contained bacterial Ig-like domains which have been implicated in bacterial adherence to extracellular matrix and internalization of bacteria into enterocytes (76). These bacterial Ig-like domains are found in many invasins and adhesins such as intimin of attaching and effacing E. coli, invasin of Yersinia species that are involved in internalization of bacteria into M cells, and adhesins of Yersinia and Salmonella that are
involved in colonization of Peyer’s patches (76). However, exact mechanistic pathways of these putative invasins are yet to be elucidated.

Pathogenic *E. coli* produce several exotoxins that are imperative for their pathogenesis. In RS218, there are several toxins that have either already been recognized in NMEC and other pathogenic *E. coli* or putative toxins. Among the toxins that are well characterized, cytotoxic necrotizing factor 1 (CNF1) belongs to the RTX family of toxins that has been implicated in the penetration of BBB by *E. coli* K1 (77). In the RS218 genome, it is encoded by a gene operon located on the GI 49. This gene operon is flanked by IS110 family transposase and contains the full set of genes including those encoding the proteins required for RTX toxin transport and activation. Other than CNF1, RS218 also carries an alpha hemolysin A (HlyA), its activator HlyC and the transporter HlyD.

In addition to HlyA, RS218 also encodes another hemolysin, HlyE which is a novel pore-forming toxin found in *E. coli*, *Salmonella* Typhi, and *Shigella flexneri* (78, 79). It has been shown that these HlyE family hemolysins are structurally different from HlyA and do not require activation or RTX transporters (79). Instead, they are secreted in the active form and the expression is induced under anaerobic conditions such as in the intestinal environment (79). The HlyE has also been demonstrated in APEC, where it does not contain the classical RTX operon, and in some enterohemorrhagic *E. coli* (78).

Another cytotoxin present in RS218 is vacuolating cytotoxin (Vat) which is located within the GI3. The Vat belongs to serine protease autotransporters of *Enterobacteriaceae* (SPATE) family which includes several virulence-associated proteins identified in *E. coli* (e. g. Sat and PicU), *Shigella* (PicU), and *Neisseria* and *Haemophilus* (e. g. IgA1 proteases and Hap) (80). It has been found that Vat is more prevalent in
ExPEC than in commensal *E. coli* indicating its pathogenic involvement in extra-intestinal infections (80). Although there are many studies that describe the prevalence of *vat* in ExPEC, the mechanistic role of Vat in relation to pathogenesis has only been demonstrated in APEC (80, 81). It has been demonstrated that Vat in APEC strain Ec222 is responsible for vacuolating cytotoxicity observed in chicken embryo fibroblast cells *in vitro*, and severe septicemia and increased mortality in a chicken respiratory infection model of colibacillosis (80). Interestingly, we also found (Chapter 3) that *vat* was more prevalent in NMEC than the commensal fecal *E. coli*, suggesting Vat can be an interesting target to unravel the pathogenic mechanisms of NMEC. The RS218 genome also encodes two bacteriocins, namely colicin E (ColE) and ColV, and their immunity proteins (82, 83). These toxic proteins act as antibacterial peptides that kill or inhibit the growth of colicin-negative bacteria while maintaining immunity for colicinogenic bacteria (83).

Apart from these cytotoxins, RS218 produces several toxin-antitoxin operons (TA) which are not present in *E. coli* K12. These TA systems are proteins that are encoded by the chromosome and/or plasmids that are involved in enhancing the fitness of bacteria (84). For instance, plasmid-encoded TA systems endorse the existence of plasmid and plasmid-encoded fitness in progeny (84). The exact mechanism of chromosome-encoded TA systems was not fully understood until recently. Norton et al, 2012 described the importance of TA systems in pathoadaptation of *E. coli* (85). Since one bacterium encodes several TA systems (RS218 encodes for 7 TA systems), they observed that in UPEC strain CFT073, each of these TA systems is niche-specific and behaves in such a way that some TA systems are required for intestinal colonization
while others are required for colonization of bladder epithelial cells and renal tubular cells (85). They also found that PasTI TA system is involved in the formation of persister cells which resist antibiotics, nutrient deprivation and reactive oxygen species by ExPEC (85). The chromosome of RS218 also carries all these TA systems except PasTI. Therefore, the role of each of these TA systems of NMEC in different host tissues will be an interesting area for future studies to address current pitfalls of NMEC pathogenesis.

Protein secretion plays a pivotal role in bacterial communication, environmental adaptation and survival (86). There are six secretory systems identified in Gram-negative bacteria (Type I-VI or T1SS-T6SS) (86). The RS218 genome codes for all these secretory systems, except a T3SS. It also codes for SecYEG and twin arginine pathway that are functional components of some of the secretory systems (86). Interestingly, a T3SS was absent in NMEC belonging to B2 phylogroup (IHE3034 and S88) whereas the group D NMEC strain CE10 possessed it (14).

A type I secretory system (T1SS) or ABC transporters are important secretory components in pathogenic E. coli which are involved in secretion of toxin-like hemolysins and colicins (87). As discussed earlier, RS218 genome encodes hemolysins as well as ColV and E which might be some of the effector molecules of RS218 that use T1SS. A type II secretory system (T2SS) which secretes A-B toxins such as the heat labile enterotoxin of ETEC also has been identified in RS218 (88). However, identifying the effector molecules secreted through T2SS in RS218 may warrant further studies. The type IV secretory system (T4SS) which shares structural similarity with T2SS is involved in transportation of large proteins and DNA molecules across the cell membranes in Gram-negative bacteria (86, 88). It has been observed that both T2SS and
T4SS are involved in UPEC virulence, particularly for the bacterial persistence in the urinary tract and renal tissues (89). Therefore, identifying involvement of these secretory systems in the penetration of BBB by NMEC and their persistence in brain tissues might open up new opportunities for understanding NMEC virulence.

The type V secretory system (T5SS) or autotransporters are Sec-dependent apparatus involved in toxin secretion in pathogenic E. coli (90). As discussed in the adhesin and invasin section above, RS218 encodes many autotransporters which potentially could be secreted by T5SS. Additionally, a virulence-associated Type VI secretory system is also present in the RS218 genome. Zhou et al, 2012 demonstrated that RS218 genome contains T6SS in a genomic island which is flanked by tRNA<sup>Asp</sup> and its effector molecules, Hcp (hemolysin-coregulated protein) and VgrG (valine glycine repeat) are located on the same island (18). Also they observed that Hcp proteins facilitate invasion of BBB, therefore, it acts as a virulence trait in NMEC (18).

Interestingly, we have found that there are four copies of Hcp protein in RS218 genome, of which two are located on the T6SS pathogenicity island (PAI) and the remainder is located on another PAI. Similarly, RS218 harbors two copies of VgrG clusters with and without T6SS. Some studies have shown that Hcp and VgrG proteins encoded by the same gene cluster with T6SS are structural components rather than the effectors, and real effectors of T6SS are clustered elsewhere on the genome. Further studies are essential to clarify if this observation is true for NMEC as well (90). The RS218 chromosome encodes for the extracellular nucleation-precipitation (ENP) pathway. This secretory system is responsible for the secretion and assembly of Curli fimbriae which are involved in biofilm formation, host cell adhesion and invasion and activation of the host immune
system (91). Other than these protein secretory systems, RS218 also contains genes that encode ion transporters such as cation transporters, tripartite ATP-independent periplasmic transporters (TRAP transporters) and antiporters.

The other sequenced ExPEC strains that were used in the study are depicted in Table 4. Blasting of the chromosome of *E. coli* RS218 against these *E. coli* chromosomes revealed that 79% of RS218 genome was similar to the *E. coli* laboratory strain K12 with 98% identity at the nucleotide level. Of the ExPEC genomes included in the study, UPEC strain UTI89, which was isolated from urine of a patient with acute cystitis, showed the closest similarity to the RS218 genome. Among NMEC strains, IHE3034 strain showed the greatest similarity with NMEC strain RS218. This *E. coli* IHE3034 strain has been isolated from a neonate with meningitis in Finland in 1976. Despite this geographical difference between the strains RS218 and IHE3034 with regard to their origin, both stains belong to the same serogroup (O18), phylogroup (B2) and the sequence type (ST95) suggesting a close genetic relatedness (92).

We also analyzed the distribution of RS218 GIs among these ExPEC strains to identify pathotype-specific GIs. Of the 51 GIs found in RS218, 48 islands were present in at least in one other ExPEC. All three GIs (7, 16, and 43) which were unique to RS218 were bacteriophage regions (Table 4). Two GIs (29 and 38) were shared by all ExPEC revealing that these islands might play an important role in ExPEC virulence regardless of their pathotype. Interestingly, one of these islands, GI 38, encoded for core oligosaccharide (core OS) of bacterial lipopolysaccharide (LPS) which is crucial for barrier function of the outer membrane and serves as a virulence factor (93). Five distinct core structures (K-12 and R1–R4), which produce various outer core OS of LPS, were
identified in *E. coli* (93). The genes present in the GI 38 resembles the R1 core OS type suggesting that all ExPEC might belong to the R1 core OS type (94). In consistant with this observation, the R1 core OS has been frequently found in pathogenic *E. coli*. The GI 29 encodes for putative polysaccharide biosynthesis proteins with tetratricopeptide repeats (TPRs). These types of proteins which include new lipoprotein I (Nlp I) also have been found in pathogenic bacteria. Such genes might be crucial for ExPEC to cause extraintestinal diseases or to survive in extraintestinal environments since these genes are particularly involved in functions such as membrane barrier, virulence and protein transport (95). These characteristics might play a role in ExPEC virulence and/or fitness.

When considering the NMEC pathotype, we identified 16 GIs including the previously discussed two islands which were common to all NMEC pathogens used for comparison purposes in the study (Table 4.3). The GI 5 encodes for a sugar transport system for multiple sugars belonging to lignocelluloses which are not fermentable by *Enterobactericeae* but by many anaerobic bacteria that inhabit the intestinal tract and produce sugar end products such as ribose, xylose, arabinose and galactoside (96). Therefore, harboring a transport system for these sugars may enhance the survival of NMEC in nutrient-limited and competitive environments like those in the host tissues. The GI 11 encodes DNA methylase enzyme (Ade-MTase) along with putative ABC membrane transport system and hemin receptor (Table 3). Interestingly, DNA methylation has been observed as a virulence phenotype of pathogenic bacteria such as *Salmonella* and *Yersinia* which required for expression of several virulence traits (97, 98). Since all NMEC possess this GI, it might be interesting to see whether this locus is involved in virulence gene expression in NMEC as is the case with other pathogenic
bacteria. The GI 13 encodes for putative integral membrane transport system which is similar to the multidrug resistance efflux system. However, we have observed that RS218 strain is susceptible to most antimicrobials screened (Chapter 2). Therefore, functional aspects of these efflux systems, in addition to their already known antimicrobial properties, in NMEC might open up a new path to identify novel properties of these systems in pathogenic bacteria.

Another common GI in NMEC is GI 15 which encodes for chitobiose-specific phosphotransferase system (PTS-Chitobiose). Chitobiose is a glycoprotein which is abundant in BBB (99). It has previously been demonstrated that OmpA of E. coli K1 interaction with chitobiose present in BBB is a crucial step for the internalization of bacteria into brain endothelial cells (100).

**Table 4-4. Genomic islands (GIs) of E. coli RS218 that are shared among all other NMEC with >90% coverage and >90% similarity.**

<table>
<thead>
<tr>
<th>GI #</th>
<th>Size (bp)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5127</td>
<td>Ribose/xylose/arabinose/galactoside periplasmic ABC transport</td>
</tr>
<tr>
<td>11</td>
<td>8620</td>
<td>DNA methylation (Ade-MTase), Ton B dependent OMP, hemin ABC transporter, Trehalase</td>
</tr>
<tr>
<td>13</td>
<td>7492</td>
<td>RND efflux system</td>
</tr>
<tr>
<td>15</td>
<td>10248</td>
<td>PTS system-chitobiose-specific components, molroprotein</td>
</tr>
<tr>
<td>17</td>
<td>38272</td>
<td>Yersiniabactin, adhesin and invasin (Peyer’s patches colonization)</td>
</tr>
<tr>
<td>24</td>
<td>28480</td>
<td>Capsular polysachcharide biosynthesis and secretion</td>
</tr>
<tr>
<td>26</td>
<td>5535</td>
<td>Preppillin, accessory colonization factor (AcfD) precursor</td>
</tr>
<tr>
<td>30</td>
<td>8495</td>
<td>Ribose ABC transport system</td>
</tr>
<tr>
<td>33</td>
<td>5600</td>
<td>Putative carbohydrate PTS system</td>
</tr>
<tr>
<td>34</td>
<td>8951</td>
<td>Hemin transport proteins, ChuA</td>
</tr>
<tr>
<td>37</td>
<td>6943</td>
<td>PTS system, fructose-specific components</td>
</tr>
<tr>
<td>39</td>
<td>5439</td>
<td>Tricarboxylate transport</td>
</tr>
<tr>
<td>42</td>
<td>8368</td>
<td>PTS system, sorbose-specific components</td>
</tr>
<tr>
<td>48</td>
<td>6383</td>
<td>Arginine/ornithine metabolism</td>
</tr>
<tr>
<td>27</td>
<td>7322</td>
<td>Conserved tetr tricopeptide repeats protein</td>
</tr>
<tr>
<td>36</td>
<td>9097</td>
<td>O antigen glycosylation proteins</td>
</tr>
</tbody>
</table>
Shifting of sugar fermentation pathways have also been characterized as fitness traits in pathogenic bacteria (101, 102). Therefore, harboring PTS-Chitobiose may be involved in the survival of NMEC in cytoplasmic vacuoles since bound chitobiose acts as an energy source within the vacuole. Also, GIs 30, 33, 37, and 42 encode for ribose transport systems, putative carbohydrate PTS system, fructose specific PTS, and sorbose-specific PTS which may enhance the fitness of NMEC allowing them to survive in nutrient-limited environments. The GI 39 encodes for tricarboxylate (TCA) transport operon (TctCBA) similar to Salmonella Typhimurium enabling bacteria to use metabolites like citrate (103). Harboring of numerous carbohydrate transport systems in NEMC might enhance metabolic fitness of NMEC. Therefore, functional and mechanistic properties of these operons, availability of these sugars in host environment, and involvement of metabolic shifting and gene regulation of NMEC inside the host, might be an interesting area to understand the niche-specific pathoadaptation of NMEC.

The GIs 14 and 24 encode for yersiniabactin, adhesin/invasin and the K1 capsular polysaccharide biosynthesis pathway proteins which are already discussed in previous sections with relevance to NMEC pathogenesis. The GI 26 which is located just downstream to the capsular polysaccharide island encodes for prepellin of type IV pili, accessory colonization factor D, and putative YghG proteins similar to ETEC and Vibrio cholerae T2SS structural components which are required for colonization of the mouse intestines and toxin production (104). The GI48 encodes for arginine/ornithine antiporters and deaminases which are considered as one of the acid resistance mechanisms in bacteria (105). It has been observed that pathogens which enter through oral route have evolved several acid resistance mechanisms to survive the stomach acidity. Major route
of transmission of NMEC is believed to be the oral route (106). Nevertheless, human vaginal environment is also acidic. Hence, bearing an acid resistance operon in NMEC may enhance their survival in mother’s vaginal environment, and in both the mother and the infant gastric environments enhancing the fitness of NMEC as a pathogen.

We also analyzed the distribution of RS218 GI islands among other ExPEC. More RS218 GIs were present in ExPEC belonging to B2 phylogroup (range 28-48), which is known to contain pathogenic *E. coli*, than in ExPEC belonging to groups A and D (range 3-16) (Table 4.4 and Figure 4.4). Accordingly, clustering of ExPEC strains depending on the presence or absence of RS218 GIs revealed two major clusters which separate phylogroup B2 from A and D revealing that there is a higher incidence of horizontal gene transfer among B2 phylogroup strains compared to the strains belonging to other phylogroups. Furthermore, we observed an overlapping of virulence traits in ExPEC belonging to B2 group. None of the GIs were specific to NMEC pathotype although RS218 harbors three unique islands which contain mostly phage regions. Fascinatingly, we have observed that the ST95 strains clustered together shared the highest number of GIs. This sub cluster includes all B2 NMEC, cystitis-causing UTI89 strain and APECO1 which has been isolated from the lung of a turkey with colisepticemia. The *E. coli* belonging to this sub cluster share eight RS218 GIs (5,8,12,23,24,25,32, and 35) which are not shared among ExPEC belonging to other sequence types. These islands encode for T6SS, putative lipoproteins, capsular-associated polysaccharides and putative fimbrial adhesins which might be involved in ST95 *E. coli*-specific virulence. The presence of RS218 GIs in APECO1 in high numbers indicates that APEC strains might be linked to human disease, revealing a zoonotic potential.
In conclusion, the whole genome sequence of RS218 revealed many potential virulence traits that might be useful for the complete understanding of NMEC pathogenesis, particularly, in relation to initial colonization of mucosal epithelia in both the mother and the infant as well as penetration of mucosal barrier and BBB, the mechanisms which are not yet understood. These traits include metabolic pathways, putative adhesins, invasins, iron acquisition systems, and cytotoxins which are novel or already demonstrated as virulence traits in other pathogenic *E. coli*. Additionally, distribution of RS218 GIs in other ExPEC strains implicates that acquisition of GIs might be associated with phylogeny, although horizontal gene transfer is thought to be a random event.

**Figure 4-6.** Pangenome analysis of human ExPEC genomes with *E. coli* RS218 genome.
References


88. Nivaskumar M, Francetic O: Type II secretion system: A magic beanstalk or a protein escalator. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research (0).


Chapter 5

THE VIRULENCE PLASMID OF *E. coli* RS218

5.1 Abstract

*Escherichia coli* is the most predominant Gram-negative bacterial pathogen associated with neonatal meningitis. Here, we present the complete nucleotide sequence and virulence potential of a large plasmid from the prototypic neonatal meningitis-associated *E. coli* (NMEC) strain RS218 (O18:K1:H7). This plasmid is 114,231 bp in size, belongs to the IncFIB/IIA incompatibility group, and contains a genetic load region that possesses several virulence and fitness traits such as enterotoxicity, iron acquisition and copper tolerance. The nucleotide sequence of pRS218 showed a 41-46% similarity to other NMEC plasmids. Interestingly, pRS218 demonstrated a remarkable nucleotide sequence similarity (up to 100%) to large virulence plasmids of *E. coli* associated with acute cystitis. We also observed that the genes located on pRS218 were overly represented by NMEC strains compared to fecal *E. coli* isolated from healthy individuals. An in vitro invasion assay that used a human brain microvascular endothelial cell line and an in vivo assay that used a neonatal rat pup model of neonatal meningitis were employed to examine the contribution of pRS218 in NMEC pathogenesis. Significant attenuation of plasmid-cured strain as compared to RS218 wild-type was observed in vitro as determined by invasion potential and in vivo as resolved by mortalities and bacterial recovery from the cerebrospinal fluid of infected rat pups. These data indicate that pRS218 plays an important role in NMEC pathogenesis.
5.2 Introduction

Neonatal meningitis (NM) and sepsis is the third most common disease in neonates that accounts for approximately 393,000 deaths worldwide (1). *Escherichia coli* has been identified as the most predominant Gram-negative pathogen associated with neonatal meningitis (2-5). Despite advanced antimicrobial therapy and supportive care, mortality and morbidity rates of NM due to neonatal meningitis-associated *E. coli* (NMEC) continue to be as high as 30-50% (6). Other than high mortality, adverse consequences of NM in surviving neonates is also a major medical concern (7, 8).

Plasticity of *E. coli* genomes has led to the identification of different pathovars of *E. coli* each of which is associated with a particular form of animal and/or human disease (9,10). Genomic plasticity of *E. coli* is mainly due to the acquisition of ‘genomic islands’ through horizontal gene transfer by means of plasmids, phages and insertion sequences (IS) (9). Of these elements, bacterial plasmids are self-replicating extra-chromosomal genetic material which have the potential to transmit a variety of phenotypic characteristics among the same or different species of bacteria (9,10,11). These phenotypic characteristics include novel metabolic capabilities, antibiotic resistance, heavy metal tolerance, virulence traits important for bacterial adherence, invasion and survival in host tissues (10, 11). Plasmid that encodes such phenotypic characters may provide competitive advantages to the bacterium for their survival and adaptation to novel niches.

Many virulence associated plasmids have been identified in pathogenic *E. coli* (10). A vast majority of these plasmids belong to IncF compatibility group. Structurally,
IncF plasmids consist of a conserved region that is common to all IncF plasmids which encodes conjugal transfer proteins, replication proteins and plasmid stability proteins and a ‘genetic load region’ or a variable region that encodes different virulence and fitness traits (10). A recent study that analyzed over 40 completed genomic sequences of IncF plasmids of *E. coli* revealed that these plasmids have evolved as virulence plasmids by acquiring novel virulence traits to their ‘genetic load regions’ through IS-mediated site specific recombination (10). Also, comparative genomic analysis of virulence plasmids in each pathovar of *E. coli* has shown that these genetic load regions encode virulence traits that are essential for and specific to their respective pathotype (10). These data suggest that acquisition of plasmid-encoded genes may play a significant role in the emergence of pathogens and different pathotypes of *E. coli*.

Although many virulence-associated plasmids in different intestinal pathogenic *E. coli* have been sequenced and studied, only a few virulence plasmids associated with each pathotypes of extra-intestinal pathogenic *E. coli* (ExPEC) causing human infection have been sequenced (10), (12, 13). For example, at the time of preparing this manuscript, only two plasmid sequences from NMEC strains are available in the public domain (12, 13). These two strains represent two of three major serogroups (O:18, O:45 and O:7) of *E. coli* that have been implicated in NM; pECOS88 from *E. coli* S88 (O45:K1) and pEC10A-D from *E. coli* CE10 (O7:K1). Despite the fact that NMEC prototypic strain RS218 which belongs to O:18 serogroup is the most commonly used *E. coli* strain used to study NMEC pathogenesis since 1980’s, its genomic sequence including the plasmid, has not been reported (14). It has been documented that the NMEC RS218 strain harbors a
large plasmid and similar sized plasmids have been observed in other NMEC and avian pathogenic *Escherichia coli* (APEC) belonging to the O:18 serogroup (15, 16).

Therefore, the objectives of the present study were to: (i) evaluate the contribution of pRS218 to NMEC pathogenesis by comparing invasive ability of plasmid- cured and wild-type strains *in vitro* and *in vivo*, and (ii) to analyze the plasmid sequence in order to understand the kinetics of the plasmid itself, its genetic and evolutionary relationship with virulence-associated plasmids in other pathogenic *E. coli* and distribution of pRS218 genes among NMEC.

### 5.3 Materials and methods

#### 5.3.1 Bacterial strains and media

Prototype NMEC strain *E. coli* RS218 (O18: H7: K1) and *E. coli* EC10 (O:7 K:1) were kindly provided by Dr. James Johnson (Department of Medicine, University of Minnesota, Minneapolis, MN). Both *E. coli* RS218 and EC10 strains have been isolated from cerebrospinal fluid of neonates affected with bacterial meningitis (15). A total of 51 NMEC strains which were isolated from neonatal meningitis cases were obtained from Dr. K.S. Kim (School of Medicine, John Hopkins University, Baltimore, MD) and 49 fecal *E. coli* strains isolated from feces of healthy individuals were obtained from the *E. coli* Reference Center (Pennsylvania State University, University Park, PA). All *E. coli* were stored in Luria Bertani broth (LB) at -80°C until further use. Bacteria were grown in
McConkey agar, LB broth or LB agar (LBA) as described below. All bacteriologic media were purchased from Becton, Dickinson and Company (BD), Sparks, MD.

5.3.2 Plasmid isolation, sequencing, assembly and annotation

Sequencing of pRS218 was performed as part of a project that sequenced the whole genome of *E. coli* RS218. The genomic DNA including the plasmid DNA was extracted using phenol:chloroform method as described previously. The DNA prep was further cleaned using Genomic Tips (Qiagen, Valencia, CA) (17)). Whole genome sequencing was performed using Ion Torrent PGM Technology (Life Technologies, Carlsbad, CA) at the Genomic Core Facility (Pennsylvania State University, University Park, PA). After initial *de novo* assembly of short reads using SeqManNGen 10 (DNASTAR Inc, Madison, WI), plasmid contigs were identified using BLAST algorithm (Blastn; www.ncbi.nlm.nih.gov) and subsequently aligned to the sequence of the reference plasmid, pUTI89 (CP000244). Gap closure was performed using primer walking into the gaps with the LongRange PCR Kit (Qiagen). The complete sequence of the plasmid was annotated using RAST (Rapid Annotation using Subsystem Technology) (18).
5.3.3 Comparative genomics and phylogenetic analysis

Comparative genomics of pRS218 with closely related IncFIB/FIIA plasmids of other *E. coli* was performed using Mauve 3.2.1 genome alignment web tool (http://gel.ahabs.wisc.edu/mauve/) (19). An evolutionary relationship of 26 plasmids belonging to the IncFIB/FIIA group based on repA1 gene sequence was performed using the neighbor-joining method. A neighbor joining tree was constructed by using the MEGA4 web tool (http://www.megasoftware.net/mega4/mega.html) (20, 21).

5.3.4 Analysis of plasmid profiles of NMEC strains

Extraction of large plasmids from NMEC strains were performed using an alkaline lysis method described previously (17). In brief, 1 ml of overnight culture of each *E. coli* strain was subjected to alkaline lysis using 10% sodium hydroxide followed by phenol-chloroform extraction of plasmid DNA. Plasmid profiles of NMEC strains were evaluated by electrophoresis on a 0.7% agarose gel containing 0.5 μg/ml ethidium bromide.

5.3.5 Evaluation of prevalence of selected pRS218 genes in other NMEC and fecal *E. coli*

Specific polymerase chain reactions (PCRs) were performed to determine the presence of selected gene coding regions (n=59) of pRS218 in other NMEC and fecal *E. coli* strains. Primers were designed using the Primer 3.0 web tool.
(http://bioinfo.ut.ee/primer3-0.4.0/) (Table B, Appendix A). PCR amplifications were performed using crude DNA extracted by the rapid boiling method (22). The PCR mixture contained 1 U of Taq polymerase (Qiagen), 1× Taq polymerase buffer, 3.5 mM MgCl₂, 125 μM each deoxynucleotide triphosphate (dNTP) and 150 nM each primer pair. PCR conditions were as follows: 1 cycle of 95°C for 1 min, followed by 30 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1.5 min, and a final extension at 72°C for 10 min. Amplicons were visualized on a 1.5% agarose gel containing 0.5 μg/ml ethidium bromide.

5.3.6 Plasmid curing and complementation

The plasmid stability gene, stbA of pRS218 was mutated by using a phage lambda Red recombinase system to facilitate plasmid curing. (23). Briefly, the chloramphenicol resistance cassette (cat) was amplified with PCR by using pKD3 plasmid as the template and primers consisted of 36 nucleotides extensions at 5’ and 3’ ends of stbA (forward primer 5’- ATG AAC GTA TAC TGC GAT GAT GGT TCA ACA ACA ATC GTG TAG GCT GGA GCT GCT TC-3’ and reverse primer 5’-TAC TCC TCT TTG AAA GCC GCG ATA GCT TCA ACC AGT CAT ATG AAT ATC CTC CTT AG-3’).

Amplified product was gel purified (MiniElute Kit, Qiagen) and electroporated to E. coli RS218 carrying the red helper plasmid, pKD119. Mutants (RS218:pRS218ΔstbA::cat) were selected for chloramphenicol resistance and confirmed by PCR using primers specific to stbA. Plasmid curing was done as described previously by 10% sodium dodecyl sulphate treatment (24). Plasmid curing was confirmed by comparing the
plasmid profile of the cured strain with the parent strain. The isolates which did not possess the plasmid was further verified for curing by PCR amplification of 5 genes, senB (forward primer 5’- GCA GAT TCG CGT TTT GAG CA-3’ and reverse primer 5’- CGG ATC TTT CAA CGG GAT GG-3’), scsD (forward primer 5’- CAT ACG CTG GAC GGG GAA AC-3’ and reverse primer 5’- GAC GCT CTC CCC TTC CGA CT-3’), traU (forward primer 5’- TTC CTT CTC GCC GGT CAT GT-3’ and reverse primer 5’- CCA GCG AGA GCG GGA AAA TA-3’), transposase (forward primer 5’- GCT TCG GGA ACG CTG TAA CG-3’ and reverse primer 5’- AGA AGG CTG CGG TGC TGA AG-3’), pRS218_113 (forward primer 5’- TGG GGG CTG AAA ACC AGA GA-3’ and reverse primer 5’- ACC GAA GGC ACG AAC TGC AT-3’), and ycfA (forward primer 5’- CGC CTG GTG GTG AAG GAA AG-3’ and reverse primer 5’- GAC CAC CTC CCG CAG AAC AC-3’) of pRS218. Isolates with no positive reaction in the PCR assays were considered to be cured of pRS218.

The plasmid complementation was performed using conjugation as described previously (25). The main obstacle for complementation was that the absence of an antibiotic resistance marker in pRS218 which can be used for subsequent selection. Therefore, pRS218 was first tagged with cat using the one step inactivation method (23). Briefly, the cat was amplified using pKD3 plasmid and primers consisted of 36 nucleotides extensions at 5’ and 3’ ends of a putative noncoding region of pRS218 located between base pairs 591 and 831 in the plasmid sequence (Forward primer 5’- CGC GTC GTT CAG TTG TCC AAC CCC GGA AAC GTG TAG GCT GGA GCT GCT TC-3’ and reverse primer 5’- CTC CTC AAT ACT CAA ACA GGG ATC GTT
TCG CAG AGG ACA TAT GAA TAT CCT CCT TAG-3'). Purified PCR product was electroporated to *E. coli* RS218 carrying the Red helper plasmid pKD119 to construct the pRS218::*cat*. The *E. coli* RS218 carrying pRS218::*cat* was then used as the donor to perform mating experiments. *Escherichia coli* DH5α used as an intermediate recipient to transfer pRS218::*cat* from the donor strain to the recipient plasmid-cured strain.

5.3.7 Bacterial growth curve

Bacteria were grown in LB broth at 37°C shaking overnight. Cultures were diluted 1:100 with LB broth or M9 medium with 10 µg/ml niacin and incubated at 37°C with shaking. Optical density at 600 nm (OD600) was taken in triplicate for every 20 min for 6 hours. The OD values from each time points were averaged and graphed to obtain the growth curve.

5.3.8 *In vitro* invasion assay

Invasion assays were performed using human brain microvascular endothelial cells (hBCMEC/D3) provided by Dr. Weksler B, Cornell University, NY. The hBCMEC/D3 cells were grown in endothelial basal medium (Lonza, Walkersville, MD) containing 5% fetal bovine serum (PAA The Cell Culture Company, Piscataway, NJ), 1.4µM hydrocortisone (Sigma-Aldrich, St. Louis, MO.), 5µg.ml⁻¹ acid ascorbic (Sigma), 1% chemically defined lipid concentrate (Gibco, Carlsbad, CA), 10mM HEPES (PAA The cell culture company), and 1ng.ml⁻¹ human basic fibroblast growth factor (Sigma),
The invasion assay was performed as described previously (26)). Briefly, endothelial cells were seeded at about $1 \times 10^5$ cells per well in 12-well tissue culture plates (Corning Life Sciences, Manassas, VA.) coated with rat collagen (R&D Systems, Trevigen, Gaithersburg, MD) and incubated at 37°C with 5% CO$_2$ in a humid chamber. Once the monolayer was confluent, it was washed with phosphate buffer saline (PBS, pH 7) and incubated with cell culture medium containing bacteria with a multiplicity of infection (MOI) of 100 for 2 hours at 37°C with 5% CO$_2$ to allow cellular invasion (26)). The extracellular bacteria were eliminated by incubation of the monolayers with a culture medium containing gentamicin (100 μg/ml) for 1 h. The monolayers were washed three times with PBS and lysed with 0.1% Triton X-100. The released intracellular bacteria were enumerated by plating on LB agar plates. Invasion frequencies were calculated by dividing the number of invaded bacteria with initial inoculum and expressed as a percentage to invasion frequency of wild type RS218 (wtRS218). The assays were performed three times in triplicate.

5.3.9 Neonatal rat meningitis model

All animal experiments were performed according to the protocol reviewed by Institutional Animal Care and Use Committee (IACUC), Pennsylvania State University, PA. Five-day-old Sprague-Dawley rat pups (n=10) were used in each experimental group. Rat pups were injected with approximately 200 CFU (range 160 to 210 CFU) of E. coli (wtRS218, plasmid-cured and plasmid-complemented strains) by the intraperitoneal route. For negative control group, PBS was injected intraperitoneally. The pups were
euthanized 24 hours post-infection to collect blood, cerebrospinal fluid (CSF) and brain tissues. For bacterial enumeration, blood was collected by intra-cardiac puncture and plated on MacConkey agar to detect septicemia. Cerebrospinal fluid was collected by cisternal puncture, and plated on MacConkey agar to demonstrate meningitis. Brain tissues collected from each group were fixed in 10% neutral buffered formalin, routinely processed for histopathology, stained with hematoxylin-eosin, and examined for lesions consistent with bacterial meningitis. Experiments were done in triplicates and paired $t$ test was used to compare the experimental groups.

5.4 Results

5.4.1 General properties of pRS218

Initial *de novo* assembly of short reads generated with Ion Torrent PGM technology identified 26 plasmid contigs ranging from 253 to 7,521 bp in length. These contigs were aligned to the reference plasmid sequence pUTI89 of uropathogenic *E. coli* UTI89 which was selected as the reference according to the sequence similarity of contigs (>90%). Complete sequence of pRS218 revealed that it is a circular plasmid of 114,231 bp in size with a G+C content of 51.02 % (Figure 5.1). A total of one hundred and sixty open reading frames (ORFs) were annotated including IncFIB and FIIA replicons. Based on the blast analysis, nearly one third of the ORFs (n=51) represents the genes involved in plasmid replication and conjugal transfer, along with 20 and 7 genes encoding mobile genetic elements (MGEs) and products involved in DNA repair,
respectively. Of the remaining ORFs, 59 encode unknown or hypothetical proteins, and 23 represent genes previously characterized in other bacteria. The plasmid does not harbor any antibiotic resistance genes that may provide a selective advantage in the face of antibiotic therapy. Genetic load region of the pRS218 encodes several virulence and fitness associated genes which have been characterized in other bacteria (Table 5.2). The annotated sequence of pR218 was deposited in NCBI GenBank (Accession number not yet assigned).

**Figure 5-1.** Graphical circular map of pRS218. From outside to the center: ORFs in forward strand, ORFs in reverse strand and GC skew. Plasmid genes were color coded as follows; Blue, conjugal transfer genes; Green, virulence or fitness-associated genes; Orange, plasmid replication genes; Red, IS elements; Black, plasmid stability genes; Light blue, hypothetical and putative genes. In the GC skew Lime indicates the areas where the GC skew above average (51%) and purple indicates the areas below the average.
Table 5-1. Virulence and fitness associated traits in pRS218.

<table>
<thead>
<tr>
<th>Virulence/ fitness associated genes</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putative Na-translocating NADH dehydrogenase</td>
<td>Na$^+$efflux (NrqC subunit)</td>
<td>(37)</td>
</tr>
<tr>
<td>High affinity Fe$^{2+}$ permease</td>
<td>Iron acquisition</td>
<td>(31)</td>
</tr>
<tr>
<td>High affinity Fe$^{2+}$ periplasmic component</td>
<td>Iron acquisition</td>
<td>(31)</td>
</tr>
<tr>
<td>High affinity Fe$^{2+}$ protein, membrane component</td>
<td>Iron acquisition</td>
<td>(31)</td>
</tr>
<tr>
<td>High affinity Fe$^{2+}$ binding protein, permease</td>
<td>Iron acquisition</td>
<td>(31)</td>
</tr>
<tr>
<td>Putative ABC transport sys, permease</td>
<td>Type I secretion</td>
<td>(31)</td>
</tr>
<tr>
<td>Putative ABC transport system, ATP-binding</td>
<td>Type I secretion</td>
<td>(31)</td>
</tr>
<tr>
<td>TonB-dependent heme/hemoglobin receptor</td>
<td>Iron acquisition</td>
<td>(31)</td>
</tr>
<tr>
<td>SenB</td>
<td>Enterotoxin in EIEC/Shigella</td>
<td>(38)</td>
</tr>
<tr>
<td>Putative GTP binding protein, YihA</td>
<td>Cell signaling and membrane ruffling</td>
<td>(41)</td>
</tr>
<tr>
<td>Haemolysin expression modulating protein</td>
<td>Thermo-osmotic regulation of hemolysin expression</td>
<td>(31)</td>
</tr>
<tr>
<td>Suppressor for copper sensitivity ScsC and ScsD</td>
<td>Copper tolerance</td>
<td>(35)</td>
</tr>
<tr>
<td>Glucose-1-phosphatase</td>
<td>Virulence regulator</td>
<td>(37)</td>
</tr>
<tr>
<td>Lytic transglycosylases</td>
<td>Cell division</td>
<td>(43)</td>
</tr>
</tbody>
</table>

5.4.2 pRS218 is remarkably similar to plasmids in acute cystitis causing *E. coli* strains.

The BLAST nucleotide algorithm (BLASTn) showed that pRS218 is 99% identical to *E. coli* plasmids pUTI89 (CP000244), p1ESCUM (CU928148) and pEC14_114 (GQ398086) of strains of *E. coli* causing acute cystitis, pUM146 (CP002168) of a strain associated with Crohn’s disease, and pECSF1(AP009379) of an *E. coli* strain belonging to the phylogenetic group B2 which was isolated from feces of a healthy adult (Figure 2) (27). Analysis of the repA1 sequence of FIIA replicon of 24
IncFIB/IIA plasmids in pathogenic *E. coli* revealed three main lineages of virulence plasmids (Figure 3). All NMEC virulence plasmids were clustered into one lineage based on the *repA1* sequence suggesting a common origin. Interestingly, pRS218 showed an identical origin with many virulence plasmids of *E. coli* strains causing cystitis (pUTI89 and pEC14_114), pECSF1 of the commensal phylogenetic group B *E. coli* strain SE15 and pCE10A of NMEC strain CE10.

5.4.3 Genes of pRS218 are overly represented in NMEC strains compared to fecal *E. coli*.

*E. coli* plasmid profiles of NMEC strains revealed 27 of 53 (51%) of NMEC strains examined harbored a plasmid similar to pRS218 in size (Figure 5.4). Furthermore, PCR analysis revealed that a vast majority of pRS218-associated genes tested (n=60) were overly represented among NMEC strains (n=52) as compared to commensal *E. coli* (Table 5.3).
Figure 5-2. Comparison of pRS218 sequence to some virulence plasmids of other *E. coli*. Each code indicates a plasmid sequence. From top to bottom; pRS218, pUTI89 (a plasmid of acute cystitis causing *E. coli* strain UTI89), pEC14_114 (a plasmid of uropathogenic *E. coli* strain EC14), pUM146 (a plasmid of adherent invasive *E. coli* strain UM146), p1ESCUM (a plasmid of acute cystitis causing *E. coli* strain UMN026) and pECSF1 (plasmid of commensal *E. coli* strain SE15). Each color box indicates clusters of ortholog genes present in plasmid sequences. White spaces in the blocks indicate the sequences that are not present in other plasmid sequences.

Figure 5-3. Evolutionary relationship of IncFIB/IIA plasmids in pathogenic *E. coli* based on the *repA1* sequence. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches.
Figure 5-4. Plasmid profiles of some NMEC strains

Table 5-2. Prevalence of pRS218 genes among NMEC strains and fecal commensal E. coli.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Predicted function</th>
<th>NMEC %</th>
<th>FEC %</th>
<th>Chi square value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRS218_007</td>
<td>Copper sensitivity</td>
<td>98.11</td>
<td>46.94</td>
<td>65.229</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_008</td>
<td>Copper sensitivity</td>
<td>96.23</td>
<td>22.45</td>
<td>113.187</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_010</td>
<td>Na+ traslocation</td>
<td>100.00</td>
<td>18.37</td>
<td>133.182</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_013</td>
<td>Iron permease</td>
<td>98.11</td>
<td>28.57</td>
<td>105.105</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_014</td>
<td>Iron transport</td>
<td>100.00</td>
<td>57.14</td>
<td>51.864</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_015</td>
<td>Membrane protein</td>
<td>96.23</td>
<td>18.37</td>
<td>124.113</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_016</td>
<td>ABC transporter</td>
<td>100.00</td>
<td>24.49</td>
<td>117.051</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_017</td>
<td>Membrane protein</td>
<td>94.34</td>
<td>77.55</td>
<td>12.706</td>
<td>0.0004</td>
</tr>
<tr>
<td>pRS218_018</td>
<td>ABC transporter</td>
<td>98.11</td>
<td>55.10</td>
<td>51.425</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_019</td>
<td>Putative thioredoxin precursor</td>
<td>83.02</td>
<td>18.37</td>
<td>20.529</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_020</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>18.37</td>
<td>133.182</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_022</td>
<td>Glucose-1-phosphatase</td>
<td>100.00</td>
<td>75.51</td>
<td>24.428</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_023</td>
<td>Glucose-1-phosphatase</td>
<td>98.11</td>
<td>16.33</td>
<td>137.169</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_031</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>26.53</td>
<td>107.541</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_034</td>
<td>Colicin immunity</td>
<td>84.91</td>
<td>91.84</td>
<td>2.407</td>
<td>0.1208</td>
</tr>
<tr>
<td>pRS218_035</td>
<td>ColicinJ production</td>
<td>66.04</td>
<td>100.00</td>
<td>49.668</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_036</td>
<td>ColicinJ production</td>
<td>77.36</td>
<td>97.96</td>
<td>20.16</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_038</td>
<td>ColicinJ production</td>
<td>100.00</td>
<td>26.53</td>
<td>112.012</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_039</td>
<td>Enterotoxin</td>
<td>100.00</td>
<td>71.43</td>
<td>33.918</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_042</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>44.90</td>
<td>68.924</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_056</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>6.12</td>
<td>177.358</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_057</td>
<td>ColicinJ production</td>
<td>100.00</td>
<td>100.00</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>pRS218_060</td>
<td>Hypothetical protein</td>
<td>96.23</td>
<td>10.20</td>
<td>148.454</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_063</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>24.49</td>
<td>120</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_064</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>0.00</td>
<td>197.04</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>pRS218_073</td>
<td>Hypothetical protein</td>
<td>94.34</td>
<td>53.06</td>
<td>43.152</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_074</td>
<td>Stability protein stbA</td>
<td>90.57</td>
<td>20.41</td>
<td>102.055</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_079</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>22.45</td>
<td>120.333</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_080</td>
<td>Unknown</td>
<td>100.00</td>
<td>100.00</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>pRS218_082</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>34.69</td>
<td>96.296</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_083</td>
<td>Transposase</td>
<td>98.11</td>
<td>22.45</td>
<td>120.333</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_086</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>22.45</td>
<td>120.333</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_088</td>
<td>Adenine-specific methyltransferase</td>
<td>100.00</td>
<td>13.33</td>
<td>151.027</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_089</td>
<td>Cytoplasmic protein</td>
<td>83.02</td>
<td>73.47</td>
<td>2.914</td>
<td>0.0878</td>
</tr>
<tr>
<td>pRS218_090</td>
<td>Hypothetical protein</td>
<td>30.19</td>
<td>48.98</td>
<td>7.553</td>
<td>0.006</td>
</tr>
<tr>
<td>pRS218_091</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>55.10</td>
<td>43.152</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_092</td>
<td>Putative antirestriction protein</td>
<td>73.58</td>
<td>83.67</td>
<td>3.014</td>
<td>0.0826</td>
</tr>
<tr>
<td>pRS218_093</td>
<td>Phage mubC</td>
<td>100.00</td>
<td>81.63</td>
<td>16.986</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_094</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>57.14</td>
<td>48.201</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_095</td>
<td>Hypothetical protein</td>
<td>75.47</td>
<td>6.12</td>
<td>98.786</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_099</td>
<td>Hypothetical protein</td>
<td>90.57</td>
<td>34.69</td>
<td>67.267</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_100</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>34.69</td>
<td>96.296</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_105</td>
<td>Cytoplasmic protein</td>
<td>75.47</td>
<td>93.88</td>
<td>13.781</td>
<td>0.0002</td>
</tr>
<tr>
<td>pRS218_106</td>
<td>Hypothetical protein</td>
<td>96.23</td>
<td>32.65</td>
<td>86.669</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_107</td>
<td>Adenine-specific methyltransferase</td>
<td>100.00</td>
<td>32.65</td>
<td>100.086</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_109</td>
<td>Hok/gef cell toxic protein</td>
<td>100.00</td>
<td>93.88</td>
<td>0</td>
<td>0.9944</td>
</tr>
<tr>
<td>pRS218_110</td>
<td>Hypothetical protein</td>
<td>98.11</td>
<td>26.53</td>
<td>107.541</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_113</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>83.67</td>
<td>17.391</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_113</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>73.47</td>
<td>31.214</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_114</td>
<td>Unknown</td>
<td>100.00</td>
<td>44.90</td>
<td>72.93</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_116</td>
<td>X polypeptide</td>
<td>97.96</td>
<td>46.94</td>
<td>65.229</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_118</td>
<td>TraJ/ conjugal transfer</td>
<td>43.40</td>
<td>10.20</td>
<td>27.955</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_131</td>
<td>Hypothetical protein</td>
<td>100.00</td>
<td>93.88</td>
<td>6.186</td>
<td>0.0129</td>
</tr>
<tr>
<td>pRS218_136</td>
<td>TraU/ conjugal transfer</td>
<td>100.00</td>
<td>42.86</td>
<td>79.72</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_154</td>
<td>TraI/ conjugal transfer</td>
<td>81.13</td>
<td>53.06</td>
<td>17.73</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_156</td>
<td>Dienelactone hydrolase</td>
<td>90.57</td>
<td>73.47</td>
<td>20.195</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>pRS218_159</td>
<td>Hypothetical protein</td>
<td>90.57</td>
<td>93.88</td>
<td>1.087</td>
<td>0.2971</td>
</tr>
<tr>
<td>pRS218_190</td>
<td>Hemolysin expression modulating protein</td>
<td>90.57</td>
<td>12.24</td>
<td>124.932</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
5.4.4 Plasmid-cured strain demonstrated a marked attenuation \textit{in vitro} and \textit{in vivo}.

To analyze the virulence potential of pRS218, the plasmid was cured from the wild type strain by mutating \textit{stbA} followed by 10\% SDS treatment. Curing of plasmid was confirmed by the absence of the plasmid in the plasmid prep and absence of 5 selected genes of pRS218 by PCR in a crude DNA extract made from the plasmid-cured strain (RS218\textsubscript{cured}). Figures 5A and 5B show the plasmid profiles and PCR amplification results of wtRS218 and RS218\textsubscript{cured}. No difference was observed in the growth rates between wtRS218 and RS218\textsubscript{cured} (Figure 5.5C). Virulence potential of pRS218 was determined by comparing RS218\textsubscript{cured} with wtRS218 based on their ability to invade hBCMEC/D3 cells \textit{in vitro} and to cause septicemia, meningitis and mortality \textit{in vivo} in a neonatal rat pup model of neonatal meningitis. \textit{In vitro} invasion assays using hBCMEC/D3 cells revealed a significant attenuation (p<0.05) of RS218\textsubscript{cured} (relative invasion 38+9.6\%) as compared to the wild type strain (100\%) (Figure 5.6 A).

Furthermore, invasiveness was restored after complementation (of RS218\textsubscript{cured} strain with pRS218 indicating its contribution to NMEC pathogenesis. Similar results were observed with \textit{in vivo} experiments. In detail, bacteria were re-isolated from 84\% CSF collected from rat pups infected with wtRS218 whereas only 29\% CSF samples collected from rat pups infected with RS218\textsubscript{cured} strain contained bacteria suggesting a role of pRS218 in translocation of bacteria through the blood brain barrier (BBB) to cause meningitis. The bacterial loads in CSF were 4.57+3.02 in rat pups infected with wtRS218 strain and 3.77+2.24 in rat pups infected with RS218\textsubscript{cured} strain. Although fewer pups were died in the groups infected with RS218\textsubscript{cured} as compared to the groups
infected with wtRS218 there was no statistically significance in mortality rates between two groups (Figure 6B). No mortalities were detected in the negative control group treated with PBS. In groups infected with wtRS218 approximately 84% of rat pups that survived 24 hours post infection showed bacteremia, whereas with RS218 cured groups it was only 33%. In both groups the number of bacteria in the blood was too numerous to count.

5.5 Discussion

Virulence plasmids in bacterial pathogens have been shown to play a major role in pathogenesis of many bacterial diseases (10, 28-30). In pathogenic E. coli, virulence associated large plasmids that are required to establish distinct disease phenotypes have been characterized using in vitro and in vivo studies (10, 13, 15, 29-31). Recently, it has been suggested that the plasmids may play a role in NMEC pathogenesis since most of the NMEC strains harbor plasmid-associated genes as compared to commensal E. coli (32). Escherichia coli RS218 which has been isolated from CSF of a neonate with meningitis in 1974 is considered as the prototype strain of NMEC. This strain has been used in the studies since then to identify the virulence traits that are particularly involved in NMEC pathogenesis (14). Here, we sequenced and analyzed the complete nucleotide sequence of pRS218, a large plasmid of E. coli RS218, and studied its contribution to NMEC pathogenesis.

The pRS218 sequence revealed a backbone typical to IncFIB/IIA-like plasmids in other pathogenic E. coli which possess both repA and repA1 replicons (10). In addition
to the replication proteins, the constant region of the plasmid encodes proteins involving conjugal transfer (Tra locus) and plasmid stability/inheritance. The tra locus comprises a 34.9 kb region containing 34 tra genes from traM to finO similar to F-like plasmids of E.coli and R100 plasmid of Shigella (33). The plasmid SOS inhibition protein (PsiAB), plasmid stabilizing proteins StbAB and CcdAB, toxin-antitoxin proteins involved in post segregation killing are also present in the constant region that confers stability and inheritance of the plasmid in progeny cells. Parallel to these findings, we have observed that the curing of pRS218 is very difficult with chemical methods alone such as ethidium bromide and SDS treatments. Therefore we mutated the stbA gene which has been identified as an essential gene for stable inheritance of IncF plasmids to achieve successful curing of pRS218 from E. coli RS218.

Genetic load region or the variable region of the pRS218 contains IS elements, virulence-associated genes, and several putative and hypothetical genes. The pRS218 contains 20 IS elements belonging to twelve different types. Previous studies have shown that IS mediated recombination might play a major role in acquiring novel genes into plasmids thereby allowing the plasmid to act as a “pathogenicity island precursor” (10, 13, 30). Interestingly, IS elements of pRS218 are located upstream or downstream of virulence/fitness-associated genes in genetic load regions providing more evidence for such speculation (Figure 5.1).
Figure 5-5. Confirmation of pRS218 curing. A, plasmid profiles of wtRS218 and RS218\textsubscript{cured}. B, Presence or absence of selected pRS218 genes in wtRS218 and RS218\textsubscript{cured}. Lanes 1, 100 bp ladder; 2, \textit{senB}; 3, \textit{scsD}; 4, \textit{transposase}; 5, \textit{traU}; 6, pRS218\textsubscript{113}; 7, \textit{ycfA}; 8, \textit{ompA}. C, Growth of wtRS218 and RS218\textsubscript{cured} \textit{E. coli} in LB broth (dashed line) and M9 containing 10-\mu g/ml niacin broth (solid line).
Figure 5-6. Evaluation of virulence potential of pRS218 in vitro and in vivo; A, Involvement of pRS218 in invasion of hBCMEC cells; B, Comparison of mortality, septicemia and meningitis among the groups of rat pups infected with wtRS218 and RS218cured; C, Bacterial counts in blood and CSF of rat pups which survive 24 hrs post-infection; ** denotes statistical significance and * denotes no statistical significance.

Types of virulence or fitness genes in the genetic load region of pRS218 are depicted in Table 5.2 and are mainly located upstream and downstream of IncFIB replicon. Upstream to the IncFIB replicon, are the ‘secreted copper-sensitivity suppressor
proteins C and D (scs\text{C} and scs\text{D}). Copper is an essential trace element for the growth of bacteria and as well as it acts as a toxic compound when available in excess. Antibacterial properties of ionic copper has been studied and used for many hospital settings to prevent nosocomial origin of infections (34). The pRS218 encoded scs\text{C} and scs\text{D} are 100% similar to copper suppressor proteins in the genomic island GI-DT12 of \textit{Salmonella enterica} subsp. \textit{enterica} serovar Typhimurium str. T000240 which have been studied in relation to conferring copper resistance in recombinant \textit{E. coli} carrying GI-DT12 providing a fitness advantage to the pathogen (35). Additionally, this region codes for several iron acquisition proteins, hemoglobin receptors and a putative ABC transporter which may be involved in the survival of bacteria in an iron limited milieu inside the host. Furthermore, pRS218 also encodes an enterotoxin called SenB, which has been found in enteroinvasive \textit{E. coli} and \textit{Shigella} spp that accounts for 50% of their enterotoxic activities (38). Interestingly, \textit{senB} is also present in the genomes of \textit{E. coli} CE10 and the \textit{Citrobacter koseri} which are associated with meningitis in newborns. Moreover, \textit{senB} is located just downstream to the \textit{cjr} operon which is an iron- and temperature-regulated operon that is expressed only during the pathogenic process of \textit{E. coli} suggesting that \textit{senB} may be involved in NMEC pathogenesis. A recent study reported that mutation of \textit{cjr} area of pUTI89 (which is >99% similar to pRS218) significantly decreased in bacterial invasion and intra-cellular bacterial community (IBC) formation in infected bladders. However, the association and specific mechanisms by which these genes and other genes harbored by pRS218 confer the NMEC phenotype is yet to be identified. Other than these putative virulence-associated genes, many
hypothetical proteins of unknown functions are present both upstream and downstream of IncFIB replicon.

Furthermore, we have screened 59 pRS218 genes among 53 NMEC strains and fecal *E. coli* isolated from healthy individuals. Interestingly, a vast majority of pRS218-associated genes tested were overly represented among NMEC strains as compared to commensal *E. coli* (Table 5.2) further suggesting a positive selection of pRS218 genes in NMEC pathotype. These overly represented genes included many hypothetical proteins and some virulence associated genes present in pRS218 such as copper sensitivity, iron acquisition, ABC transporter components, *traJ* and *senB*.

Additionally, we analyzed the sequence similarity and the evolutionary relationship of pRS218 with other NMEC plasmids, namely pECOS88 and pCE10A, and some other IncFIB/IIA plasmids of pathogenic *E. coli* (Figures 5.2 and 5.3). The pRS218 showed a remarkable sequence similarity to four plasmids found in *E. coli* associated with acute cystitis (pUTI89, pEC14_114, p1ESCUM and pUMN146 -) and a plasmid present in an enteroinvasive *E. coli* (pECSF1) (Figure 5.2). The differences detected among pRS218, pUT89, pEC14_114 and pUMN146 revealed only SNPs whereas an insertion of *tetABCD* antibiotic cassette was seen in p1ESCUM and pECSF1 (Figure 5.2). However, the nucleotide sequence of pRS218 showed a marked difference to other two NMEC plasmid sequences currently available in public domain. For example, pECOS88 shares similarity only with *tra* locus, *repA* and *repA1* regions of pRS218 revealing that the genetic load regions of these plasmids harbor different putative virulence and hypothetical genes to that of pRS218. Compared to pECOS88, pCE10A plasmid showed a relatively higher nucleotide sequence similarity to pRS218 genetic
load region containing the copper resistance-associated genes (\textit{scsDC}), \textit{cjrABC} and \textit{senB}. However, pCE10A lacks the \textit{tra} locus that makes the plasmid incapable of conjugal transfer.

Among many capsular serogroups of \textit{E. coli}, K1 serogroup is the most commonly associated with NM which represents approximately 80\% of \textit{E. coli} strains isolated from meningitis (4, 5). Neonates acquire \textit{E. coli} K1 mainly from the urogenital microflora of the mother. Although there are no studies regarding factors that facilitate the vaginal epithelial colonization and survival of the NMEC strains in urogenitary tract of women, it has been well documented that cystitis causing \textit{E. coli} can survive and persist inside bladder epithelial cells as IBCs which is a dormant stage that becomes activated and shed when the immunity of the carriers is suppressed as is the case during pregnancy (32). As aforementioned, a recent study has also indicated that the pUTI89 plasmid is essential for filamentation of IBCs which is the first event of reactivation from the dormant stage. The higher sequence similarity of pRS218 to other cystitis-associated plasmids and their close evolutionary relationships suggest that \textit{E. coli} RS218 might use the same strategy to survive in the urogenitary tract. However, the ability of \textit{E. coli} RS218 to invade bladder epithelial cells and to survive within the urogenitary tract remains to be investigated.

Pathogenesis of NMEC meningitis involves three main sequential events that are governed by the virulence potential of bacteria. These include initial colonization and invasion of gastro-intestinal tract, survival and multiplication in the blood, and invasion of the BBB (5). Pathogenic potential of pRS218 to penetrate intestinal epithelial cells and BBB was analyzed \textit{in vitro} and \textit{in vivo} using hCMEC/D3 cells and a neonatal rat pup model of neonatal meningitis, respectively. Curing of pRS218 from \textit{E. coli} RS218 has
not shown any effect on the growth rate revealing that differences observed between wild
type and plasmid cured strains in *in vitro* and *in vivo* studies were not due to the
differences in their growth rates (Figure 5.5C).

It is believed that the high level of bacteremia is a prerequisite for the penetration
of blood brain barrier by NMECs to establish neonatal meningitis (39). We have
observed a higher incidence of bacteremia among the rat pups infected with wild type
RS218 strain (84%) than the RS218<sub>cured</sub> strain indicating that plasmid encoded genes
might be involved in developing bacteremia. Iron is a major limiting factor that restricts
the survival and multiplication of bacteria inside the host. The genetic load region of
pRS218 encodes several high affinity iron acquisition proteins, hemolysin modulation
factor and hemoglobin receptor which may be involved in iron acquisition. Interestingly,
these genes were highly prevalent in NMEC strains as compared to fecal *E. coli* (Table
5.3). Furthermore, *in vitro* and *in vivo* study results clearly demonstrated that RS218<sub>cured</sub>
strain is far less capable of invading epithelial and endothelial cells as well as establishing
meningitis in neonatal rat pups as compared to its wild type strain, suggesting that
pRS218 might play a role in NMEC pathogenesis. The *traJ* which is present in pRS218
has been previously identified as a potential virulence trait in NMEC by signature-tagged
mutagenesis and *in vitro* endothelial invasion assays (36). The mutation of *traJ* was
shown to be attenuated in terms of invasive ability to penetrate BBB. However, more
than 50% of the NMEC strains used in this study did not possess *traJ* even though the
gene was more prevalent in NMEC than in fecal *E. coli* (Table 5.3). The present study
demonstrated that the curing of pRS218 offered a greater attenuation to RS218 strain than
did the mutation of *traJ* only suggesting that plasmid genes other than *traJ* might be
involved in NMEC pathogenesis. Interestingly, as shown in Table 5.3, pRS218 carries several genes that encode hypothetical proteins which are also more prevalent in NMEC than in fecal commensal *E. coli*. Most gene prevalence studies carried out to identify potential virulence markers of NMEC have based their studies on already known virulence genes of other ExPEC and only a limited number of studies have attempted to explore novel traits that might be helpful in defining the NMEC pathotype (4, 26, 32,40). Therefore, future studies aimed at delineating the mechanistic aspects of such hypothetical proteins which are more commonly occurring in NMEC than in fecal commensal *E. coli* may invariably help to close the knowledge gaps pertaining to our understanding of NMEC pathogenesis.

Although the RS218<sub>cured</sub> strain was significantly attenuated in terms of *in vitro* (bacterial invasion) and *in vivo* (reduction in bacterial counts in CSF and blood from infected rat pups) assays as compared to the wild type strain, it is not completely avirulent. This finding suggests that the full virulence of *E. coli* RS218 requires both chromosomal and plasmid-located genes. Further studies including in depth analysis of RS218 chromosome will advance our understanding of NMEC pathogenesis.

**Concluding Remarks**

Incomplete understanding of NMEC pathogenesis is a major hindrance that has been identified and pointed out by many scientists particularly in relation to formulation of novel therapeutic and prevention strategies for neonatal meningitis. The plasmid pRS218 in *E. coli* K1 RS218 strain belongs to IncFIB/IIA subset of virulence plasmids in pathogenic *E. coli*. These plasmids harbor many virulence traits
that are required for survival of bacteria inside the host. The nucleotide sequence of pRS218 showed a greater similarity to plasmids of *E. coli* associated with acute cystitis than plasmids from NMEC. However, the prevalence of pRS218 located genes was significantly higher in NMEC strains tested than fecal commensal *E. coli*. We have also demonstrated that the pRS218 is involved in NMEC pathogenesis using both *in vivo* and *in vitro* experiments. Future studies on pRS218 transcriptome analysis, identification of plasmid-located genes that are responsible for current observations and in-depth analysis of *E. coli* RS218 whole genome will invariably broaden our knowledge of NMEC pathogenesis.

References


Meningitis E. coli Strains and Are Able To Cause Meningitis in the Rat Model of Human Disease. *Infection and Immunity* 2010, 78(8):3412-3419.


Chapter 6
SUMMARY AND FUTURE RESEARCH

Increasing incidence, emergence of antibiotic resistance clones and high level of neurological complications in survived babies of neonatal meningitis caused by *E. coli* indicates the necessity of proper screening procedure to screen pregnant women carriers of NMEC and an effective antibiotic regime and/or vaccine to prevent the transmission of NMECs from the mother to the baby (1-3). The major constraints that hamper these approaches are the inability to define NMEC pathotype due to its significant heterogeneity in virulence traits and incomplete understanding of NMEC pathogenesis (4-6). Therefore, current study focused on identifying genotypic and phenotypic characteristics that might be useful in defining the NMEC pathotype and identifying potential virulence traits in the genome of NMEC that might be valuable in the future for strengthening the current understanding of NMEC pathogenesis.

First, we analyzed genotypic and phenotypic characteristics of NMEC in comparison to HFEC in order to identify a set of genotypic and phenotypic characteristics that can be used to distinguish NMEC from HFEC (Chapter 3). There, we used four phenotypic methods (biofilm assay, invasion assay, antimicrobial profiles, and serotyping) and three genotypic methods (phylogrouping, PFGE and virulence genotyping). In accordance with previous studies, we observed a substantial heterogeneity in terms of phenotypes and genotypes of NMEC. However, we were able to develop a criterion (K1+, *sitA*+ and having at least two of the three genes, *vat*, *neuC* and *iucC*) with virulence genotyping that might be useful in defining a typical NMEC. It
should be noted that some of the genes sitA, iucC and vatA described in this criterion have not yet been studied in relation to NMEC pathogenesis. Therefore, definition of NMEC using these genes might be considered as tentative that needs to be confirmed by mechanistic studies in future to understand the role of these traits in NMEC pathogenesis. Furthermore, screening of large collection of NMEC strains including isolates from different sources and geographical areas might be required for further validation and utility of these gene profiles in defining the NMEC pathotype.

There are several molecular biological methods that are available for identifying virulence genes in pathogenic bacteria (7-9). Out of these, whole genome sequencing of pathogens is one of the most popular methods, since it reveals the complete genetic makeup of a pathogen (10). There are only 3 complete genomes of NMEC are available through public databases (11, 12). Therefore, to increase the pangenome of NMEC for better understanding of genetic makeup of the NMEC pathotype, we performed whole genome sequencing of NMEC prototype strain RS218 (Chapter 4). Whole genome sequence of E. coli RS218 revealed several potential virulence factors that might be involved in basic steps of NMEC pathogenesis. Specially, we found several adhesins and invasins that might be associated with intestinal attachment and penetration by bacteria. Since there are no studies reported on the analysis of virulence traits required for initial adherence to and invasion of enterocytes in neonatal gut, it will be very interesting to study these virulence factors in the future for filling the knowledge gaps in NMEC pathogenesis. Additionally, we observed that several iron acquisition systems, cytotoxins and protein secretory systems in E. coli RS218 which might play a role in NMEC pathogenesis. Transcriptome analysis of these genes under in vitro and in vivo conditions
that mimic the host environmental conditions and site directed mutagenesis of the genes to ascertain the functional aspects of these traits will be necessary to broaden our understanding of NMEC virulence and pathogenesis.

Moreover, we were able to identify common genomic regions among NMEC used in this study by a comparative genomics approach (Chapter 4, Table 4-4). Some of these regions encoded for different sugar utilization and transport systems that are not present in the laboratory strain of *E. coli* K12. Ability to grow on different energy sources has been used to identify some bacterial species or strains apart from other closely resembled bacteria (13-15). Therefore, it will be useful to study further these sugar fermentation patterns common to NMEC which may lead to developing a screening procedure based on selective media to identify NMEC carrier mothers. Moreover, it has been observed that certain sugar transport systems and phosphorylation pathways have an effect on virulence gene expression of some *Enterobacteriaceae* species such as *Salmonella* and *Yersinia* species revealing that similar studies in NMEC may be beneficial to understand the regulation of virulence in NMEC pathotype.

Likewise, screening of GIs identified in the study in a large collection of NMEC from different sources representing different geographical regions, site directed mutation of these GIs in NMEC and comparing the mutants with the wild-type strains might be an elegant study to understand the role of these GIs in NMEC pathogenesis. Moreover, we found two genomic regions common in all ExPEC strains used in this study which encode for the oligosaccharide core of bacterial cell wall and a large uncharacterized protein containing tetratricopeptide repeat motif that might be suitable targets to add into the existing genetic pool of ExPEC.
In addition to the chromosome, we also analyzed pRS218, a large plasmid of *E. coli* RS218. This plasmid encodes several potential virulence and fitness traits such as enterotoxicy, iron acquisition and copper tolerance. Based on an *in vitro* invasion assay that used a human brain microvascular endothelial cell line and an *in vivo* assay that used a neonatal rat pup model of neonatal meningitis, the pRS218-cured strain was significantly attenuated as compared to the RS218 wild-type strain revealing that plasmid is involved in virulence. We also observed that the genes (n=45) located on pRS218 were overly represented in NMEC strains compared to HFEC indicating the importance of pRS218 to NMEC pathogenesis. Another study published recently has also observed similar findings revealing that some plasmid genes are overly represented in NMEC as compared to fecal *E. coli* suggesting that future studies aimed at plasmid transcriptome analysis and mutagenesis of these overly represented plasmid genes of NMEC might be the key steps required to ascertain the contribution of these plasmid genes to NMEC pathogenesis.

References


# Appendix

Table A: Details of the primers used in Chapter 3. References in the table correspondance to the references cited in Chapter 3.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene/Function</th>
<th>Sequence 5’-3’</th>
<th>Product size</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>chuA.1</td>
<td>Phylogrouping</td>
<td>gacgaaccaacgggtcaggat</td>
<td>279</td>
<td>(39)</td>
</tr>
<tr>
<td>chuA.2</td>
<td></td>
<td>tcgcgcctaccaagaaca</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yjaA.1</td>
<td>Phylo-rouping</td>
<td>tgaagtctcaggagaacgctg</td>
<td>211</td>
<td>(39)</td>
</tr>
<tr>
<td>yjaA.2</td>
<td></td>
<td>atgggaaatgctcccttcacac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TspE4C2.1</td>
<td>Phylogrouping</td>
<td>gatgaatctcgaggcaacctca</td>
<td>152</td>
<td>(39)</td>
</tr>
<tr>
<td>TspE4C2.2</td>
<td></td>
<td>cgcgcctaccaaaatgttaacg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>afaF</td>
<td>afimbrial adhesin</td>
<td>ctcagacctcagaaaacagatac</td>
<td>750</td>
<td>(40)</td>
</tr>
<tr>
<td>afaR</td>
<td>gctgggcagcaaactgatctc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aslAF</td>
<td>Arylsulfatase</td>
<td>gctgttgctcctggcaacc</td>
<td>463</td>
<td>(41)</td>
</tr>
<tr>
<td>aslAR</td>
<td></td>
<td>atccgcctacttcatataagc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aufAF</td>
<td>Auf fimbriae</td>
<td>tccagactcaggaacagatac</td>
<td>350</td>
<td>This study</td>
</tr>
<tr>
<td>aufR</td>
<td>ggcctacagatgaggtacaa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cnfF</td>
<td>Cytotoxic necrotizing factor1</td>
<td>ctcagactcctcgtccttcatttt</td>
<td>498</td>
<td>(42)</td>
</tr>
<tr>
<td>cnfR</td>
<td></td>
<td>aagatggagttctctatgcaggag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fimHF</td>
<td>Type I fimbriae</td>
<td>gcagctcacctgcctggctga</td>
<td>508</td>
<td>(42)</td>
</tr>
<tr>
<td>fimHR</td>
<td></td>
<td>tgcagacaaggtagatatgcgcag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hlyAF</td>
<td>Alpha hemolysin</td>
<td>accataataaggtcctctaggctca</td>
<td>1177</td>
<td>(42)</td>
</tr>
<tr>
<td>hlyAR</td>
<td></td>
<td>aacaaagagaacaggttcttcgct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hlyDF</td>
<td>secretion protein for HlyA</td>
<td>ctctggactctgaaaaggaactc</td>
<td>904</td>
<td>(43)</td>
</tr>
<tr>
<td>hlyDR</td>
<td></td>
<td>gccctgattactgagtccttgcttactc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ibeAF</td>
<td>Invasin of brain endothelium A</td>
<td>ctttcgctctcgcgaacgt</td>
<td>171</td>
<td>(44)</td>
</tr>
<tr>
<td>ibeAR</td>
<td></td>
<td>taccggtcctgtatgttactcata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ibeBF</td>
<td>Invasin of brain endothelium B</td>
<td>gcagattttgcctgtcttgttagtctagcttgatgtcagatgtcggtgctcttcggtttactcattttactc</td>
<td>660</td>
<td>This study</td>
</tr>
<tr>
<td>ibeBR</td>
<td></td>
<td>gttctgcgacttcgcgtctcggtcgtctgactgcgcg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>issF</td>
<td>Increased serum survival</td>
<td>cagacaacccgacacactgttgatgg</td>
<td>323</td>
<td>(43)</td>
</tr>
<tr>
<td>issR</td>
<td></td>
<td>acctgtcgcacactgttgatg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iutAF</td>
<td>Aerobactin synthesis</td>
<td>atgagctatctcggagcgtctg</td>
<td>587</td>
<td>(40)</td>
</tr>
<tr>
<td>iutAR</td>
<td></td>
<td>caggtgagaaaaactgtggcttgctgcttgctctgcgtctctgcgtctctcgcgt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iucCF</td>
<td>Aerobactin synthesis</td>
<td>acccgtctcgcacactgttgatg</td>
<td>269</td>
<td>(44)</td>
</tr>
<tr>
<td>iucCR</td>
<td></td>
<td>aaccctgctctcgcacactgttgcttgctctgcgtctctcgcgt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sitAF</td>
<td>Periplasmic iron-binding protein</td>
<td>agggggcacaacgttcgacttgctttgtctgtcacttgcttactc</td>
<td>608</td>
<td>(45)</td>
</tr>
<tr>
<td>sitAR</td>
<td></td>
<td>taccggtcctgtatgttactcata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Sequence</td>
<td>Length</td>
<td>References</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------</td>
<td>---------------------------------------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>kpsMTII F</td>
<td>Capsular type II</td>
<td>gcgcattgtcgtgatactgtgg</td>
<td>272</td>
<td>(46)</td>
</tr>
<tr>
<td>kpsMT II R</td>
<td></td>
<td>catccagacgataagcatgagca</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kpsMT K1 F</td>
<td>K 1 capsule</td>
<td>tagcaaaagttcattgtggtc</td>
<td>153</td>
<td>(46)</td>
</tr>
<tr>
<td>kpsMT IIR</td>
<td></td>
<td>catccagacgataagcatgagca</td>
<td></td>
<td>(46)</td>
</tr>
<tr>
<td>neuCF F</td>
<td>N-acetylneuraminic acid synthesis</td>
<td>gtggtagacattcgggatgtc</td>
<td>676</td>
<td>(40)</td>
</tr>
<tr>
<td>neuCR F</td>
<td></td>
<td>agtgtaaagccctgggtggtt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nlp I F</td>
<td>New lipoprotein I</td>
<td>aatatacttctctgcgtgaaagttgga</td>
<td>662</td>
<td>This study</td>
</tr>
<tr>
<td>Nlp I R F</td>
<td></td>
<td>aaataagttggtttcactgagttg</td>
<td></td>
<td>This study</td>
</tr>
<tr>
<td>ompAF F</td>
<td>Outer membrane protein A</td>
<td>cactaaatccaaagttatgtgtaaaa</td>
<td>550</td>
<td></td>
</tr>
<tr>
<td>ompAR F</td>
<td></td>
<td>gatacettttaacctgctctctacgc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>papCF F</td>
<td>Usher protein of P pilli</td>
<td>atatctcttcgcagggatgcaata</td>
<td>205</td>
<td>(46)</td>
</tr>
<tr>
<td>papCR F</td>
<td></td>
<td>gacggtcgtacgtgcaggggtgggtgc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>papG IF</td>
<td>Fimbrial adhesin PapG allele I</td>
<td>ctatatagctagctagtcaggg</td>
<td>474</td>
<td>(43)</td>
</tr>
<tr>
<td>papG IR</td>
<td></td>
<td>cctgatctctctcactattcagta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>papG IIF</td>
<td>Fimbrial adhesin PapG allele II</td>
<td>cccagettgttttcatcttgg</td>
<td>190</td>
<td>(43)</td>
</tr>
<tr>
<td>papGHIR F</td>
<td></td>
<td>ttctcactatgtgtgattgcg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>papGIIIF</td>
<td>Fimbrial adhesin PapG allele III</td>
<td>ggccgtcactggattttcactgg</td>
<td>258</td>
<td>(43)</td>
</tr>
<tr>
<td>papGIIIR</td>
<td></td>
<td>ccacaaatgacactgacgacac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>satF F</td>
<td>Secreted autotransporter toxin</td>
<td>gcaagaatatgtatataca</td>
<td>630</td>
<td>(47)</td>
</tr>
<tr>
<td>satR F</td>
<td></td>
<td>gtggtagacactgaaagggga</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfa/focF F</td>
<td>S fimbrial adhesin</td>
<td>cggagggaatattaaacacgttgca</td>
<td>410</td>
<td>(46)</td>
</tr>
<tr>
<td>sfa/focR F</td>
<td></td>
<td>ctcggagaactgggtgcattttac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trajF F</td>
<td>Conjugal transfer</td>
<td>caatggggcttttattgaactc</td>
<td>369</td>
<td>This study</td>
</tr>
<tr>
<td>trajR F</td>
<td></td>
<td>tggacaaacccacccagcattataaa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vat1 F</td>
<td>Vacuolating cytotoxin</td>
<td>tcctgagacataatggctcagg</td>
<td>900</td>
<td>This study</td>
</tr>
<tr>
<td>vat2 F</td>
<td></td>
<td>gtgctcagaaagggatgtt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordinates</td>
<td>Predicted function</td>
<td>Primer F (5'-3')</td>
<td>Primer R (5'-3')</td>
<td>Product size (bp)</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>4107- 4265</td>
<td>copper sensitivity</td>
<td>gagacggttgagcaccaatctg</td>
<td>acgeccagtitaltttcttcac</td>
<td>140</td>
</tr>
<tr>
<td>4255- 4761</td>
<td>copper sensitivity</td>
<td>catacgcctggacgaggggaac</td>
<td>gacgccttcctccccgcaet</td>
<td>143</td>
</tr>
<tr>
<td>4998- 5759</td>
<td>Na+ traslocation</td>
<td>atcaatgtgatgtgctttgtgct</td>
<td>ccegttaactgtgatgataaacct</td>
<td>378</td>
</tr>
<tr>
<td>6052-7992</td>
<td>iron permease</td>
<td>ggtctcagaactgtggaag</td>
<td>cgggtgttctgtgagggacat</td>
<td>401</td>
</tr>
<tr>
<td>8033- 8560</td>
<td>iron transport</td>
<td>cgtgcaccatagtgaataggaag</td>
<td>ctcaacacaagcgtttcac</td>
<td>400</td>
</tr>
<tr>
<td>8664-10043</td>
<td>membrane protein</td>
<td>tcgtgaeagtaaatctgactc</td>
<td>gceggcataagctgtatggt</td>
<td>400</td>
</tr>
<tr>
<td>10046-11329</td>
<td>ABC transporter</td>
<td>aagggttggtgtgactgataaat</td>
<td>ctcacagacacccgcaagata</td>
<td>399</td>
</tr>
<tr>
<td>11319-12449</td>
<td>membrane protein</td>
<td>aggtcaecgtaatgtggatt</td>
<td>aatgacagacttcccctcag</td>
<td>400</td>
</tr>
<tr>
<td>12454-13149</td>
<td>ABC transporter</td>
<td>cttcccattcttcctcctta</td>
<td>acctgatggagcttggttgct</td>
<td>379</td>
</tr>
<tr>
<td>13136-13621</td>
<td>putative thioredoxin precursor</td>
<td>gcggtgtaaagaaggtact</td>
<td>agacgcctacagctaaaacct</td>
<td>401</td>
</tr>
<tr>
<td>13646-14131</td>
<td>hypothetical protein</td>
<td>atagegacagtctgctcacta</td>
<td>acgctctgatgatggttttg</td>
<td>303</td>
</tr>
<tr>
<td>14253-14702</td>
<td>glucose-1-phosphatase</td>
<td>agacacgcgtcaggggaagtattt</td>
<td>ttctcgatagctgctggaa</td>
<td>354</td>
</tr>
<tr>
<td>14677-14997</td>
<td>glucose-1-phosphatase</td>
<td>acgatgacgacccggttaaaat</td>
<td>ataggctgattcgatgtgcttgg</td>
<td>311</td>
</tr>
<tr>
<td>18173-17826</td>
<td>hypothetical protein</td>
<td>attgctctgtgagcagc</td>
<td>gtcgccagccgtaacctt</td>
<td>301</td>
</tr>
<tr>
<td>20251-20775</td>
<td>colicin immunity</td>
<td>ttaataatgtggtggtggtact</td>
<td>atgaaacagtaaccctgtaataaccg</td>
<td>250</td>
</tr>
<tr>
<td>21065-21982</td>
<td>colicinJ production</td>
<td>tggcattcctgtagcagc</td>
<td>tggtgcagctgctggatcttc</td>
<td>350</td>
</tr>
<tr>
<td>21990-22766</td>
<td>colicinJ production</td>
<td>ctgattttcctcgtttatctg</td>
<td>agcctttatctagaggtggac</td>
<td>294</td>
</tr>
<tr>
<td>22935-25196</td>
<td>colicinJ production</td>
<td>ttagagcaggttgttgttgg</td>
<td>tgtgctctgctgctgggttgc</td>
<td>393</td>
</tr>
<tr>
<td>25265-26440</td>
<td>enterotoxin</td>
<td>gcagactcggtttgagca</td>
<td>cggatctttcaaggggatgg</td>
<td>302</td>
</tr>
<tr>
<td>28517-27762</td>
<td>hypothetical protein</td>
<td>tgaecatgtcagctaatgaaagat</td>
<td>tgcatacgacggagatcttcacc</td>
<td>399</td>
</tr>
<tr>
<td>38291-37500</td>
<td>hypothetical protein</td>
<td>cgtcacaaggatatattctaaac</td>
<td>gtagacaggagtagtattcagaaac</td>
<td>373</td>
</tr>
<tr>
<td>40184-38298</td>
<td>colicinJ production</td>
<td>ctgtggataacagcctcatcactaa</td>
<td>atggatcagctggctttc</td>
<td>301</td>
</tr>
<tr>
<td>43799-42630</td>
<td>hypothetical protein</td>
<td>ctctccccatcagcctttat</td>
<td>accecaactaatcaggggaa</td>
<td>600</td>
</tr>
<tr>
<td>46748-46975</td>
<td>hypothetical protein</td>
<td>tggacgtcttggtaggactcat</td>
<td>cctgtgacgacagctcagaaaaa</td>
<td>224</td>
</tr>
<tr>
<td>48251-47610</td>
<td>hypothetical protein</td>
<td>tgcacataacctctgtagcgtt</td>
<td>tatagcagatcaggtcagcgt</td>
<td>385</td>
</tr>
<tr>
<td>52321-52046</td>
<td>hypothetical protein</td>
<td>cagccagcaagctattaaaaca</td>
<td>gctaaggctcacttcgagc</td>
<td>276</td>
</tr>
<tr>
<td>53188-54159</td>
<td>stability protein stbA</td>
<td>ttgctgaccaactctctcctg</td>
<td>cgacagacagcagaaaaac</td>
<td>400</td>
</tr>
<tr>
<td>56513-56265</td>
<td>hypothetical protein</td>
<td>cgcattgaaactttctgcac</td>
<td>tctgcctgcagctattcctt</td>
<td>249</td>
</tr>
<tr>
<td>56648-57166</td>
<td>unknown</td>
<td>gtgtcctgatcttgcttgagta</td>
<td>tgcactcctttcatctacttcc</td>
<td>351</td>
</tr>
<tr>
<td>58824-59654</td>
<td>hypothetical protein</td>
<td>acaaatgaagattattatgcgttc</td>
<td>cgacagttcagctgctgacagc</td>
<td>372</td>
</tr>
<tr>
<td>60445-59648</td>
<td>transposase</td>
<td>gcttcgggagctgtaacgc</td>
<td>agaagacctgggtgctgtaag</td>
<td>414</td>
</tr>
<tr>
<td>61858-62169</td>
<td>hypothetical protein</td>
<td>tttcctgcgtaaagaggtgctg</td>
<td>gtcctgcagcgcagcttgtc</td>
<td>223</td>
</tr>
<tr>
<td>62245-62928</td>
<td>Ade-specific methyltransferase</td>
<td>cgtgtgatattagatgacttg</td>
<td>gttgagactcctaatcagttg</td>
<td>356</td>
</tr>
</tbody>
</table>
62929-63150 cytoplasmic protein
tatgcgcggacacgaaaaac
gaagcaggaatccagttcca
208
63230-63598 hypothetical protein
gttatcggtgccccggaaga
cattcacgttcacacatgc
254
63643-64614 hypothetical protein
atgaatgaaatgctgaatgcac
catcttctgccacctggtaact
406
64828-65253 putative antirestriction protein
gttaagagttgacgacgtct
agtcgaagctgcctaaatc
400
65300-65722 phage mubC
catcgccgatgtactggatac
ctgtaacacaagtcctgctc
373
65719-65910 hypothetical protein
cacagaaaccgcgaaat
cgtttctgctgccctgtaag
177
66381-66887 hypothetical protein
cggctgcgtcctatcattc
tcctgatgttatgttttgttact
256
67155-68516 hypothetical protein
tctccaaactacagcagt
gtaatttggcgttgtgactgaa
385
68528-69126 hypothetical protein
gttgaagagtgcgaccgtct
agtcaagtgccgcgtaaatc
400
70556-70789 hypothetical protein
gcgaatatttcagaatact
tcaggaaattccggatgacatggttc
213
70848-72806 hypothetical protein
gtacatcgggctgtgctcgt
taatgtttacattccaggctgattt
400
72861-73271 ade-specific methyltransferase
tagttcaagccacaggtaa
gatctgggctgtgtaacatga
371
74286-74444 hok/gef cell toxic protein
tagaaactaccagcagctctc
taccggattcgtaagccatga
154
75360-75647 hypothetical protein
gagctgctgaaacccgaga
taatctgacgcaggaactgttt
251
75691-76587 hypothetical protein
tcggtattttccggtgataaac
tataacctgcccgacaatatcag
359
77473-77883 X polypeptide
tagcccgggattacaaaatagat
ccggtataacccggtaaaaacct
354
78394-79080 TraU/ conjugal transfer
caatgggccgtttattgactc
tgaccaacccagcatatataaa
369
85396-85614 hypothetical protein
tgtcaaccttttatctgattgc
tcaagtgttacatcaggttgc
210
89620-90612 TraU/ conjugal transfer
ttccttcgcgcttcagttc
tccagcggagccgaaata
111
111369-112229 dienelactone hydrolase
tcgttgcgggagagatgata
tgattcagaagcagccatca
343
113415-113939 hypothetical protein
tggtcattcatctgatatgggaaat
ctgtgtttgtgtctgcatatacccc
387
113985-114194 Hly expression modulator
caaaaacaggttggctgattca
tattccatatcttttgggatctg
190
VITA
DONA SAUMYA SEWWANDI WIJETUNGE

EDUCATION

2009-2014 The Pennsylvania State University, USA
PhD in Pathobiology
Advised by Dr. Subhashinie Kariyawasam, GPA: 3.8

2001-2006 University of Peradeniya, Sri Lanka
BVSc in Veterinary Medicine and Animal Science
First Class Honors

PUBLICATIONS


ORAL AND POSTER PRESENTAIONS AT SCIENTIFIC CONFERENCES

1. The large virulence plasmid from meningitis-associated Escherichia coli strain RS218 is required for efficient invasion of human brain microvascular endothelial cells. 114th General Meeting of the American Society for Microbiology (2014) poster session.
3. Acute septicemia associated with Streptococcus pasteurianus in commercial turkey poult: Oral presentation at 84th Northeastern Conference in Avian Diseases, September 2012.
4. Acute septicemia caused by Streptococcus pasteurianus in turkey poult. Poster presentation at Association for American Avian Pathologist annual convention, August 2012