Studies on colonization factors of enterotoxigenic *Escherichia coli*: Prevalence, Expression and Regulation during pathogenesis

Thesis Submitted for the degree of
Doctor of Philosophy (Science)
in
Life Science & Bio-technology

by

DEBJYOTI BHAKAT, M.Sc.
(JU Index No.: 48/18/Life Sc./25)

Department of Life Science and Bio-technology
Jadavpur University
Kolkata, India
2022

आई सी एम आर – राष्ट्रीय हैजा तथा आंत्ररोग संस्थान
स्वास्थ्य अनुसंधान विभाग, स्वास्थ्य और परिवार कल्याण मंत्रालय, भारत सरकार

ICMR - National Institute of Cholera and Enteric Diseases
Department of Health Research, Ministry of Health and Family Welfare, Government of India

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled **"Studies on colonization factors of enterotoxigenic *Escherichia coli*: Prevalence, Expression and Regulation during pathogenesis."** Submitted by Sri ~~/ Smt.~~ **DEBJYOTI BHAKAT** who got his / ~~her~~ name registered on 13.02.2018 For the award of Ph. D. (Science) degree of Jadavpur University, is absolutely based upon his own work under the supervision of **Dr. NABENDU SEKHAR CHATTERJEE** and that neither this thesis nor any part of it has been submitted for either any degree / diploma or any other academic award anywhere before.

NSChatterjee 22-02-2022

(Signature of the Supervisor date with official seal)

नबेन्दु शेखर चटर्जी / Nabendu Sekhar Chatterjee
(वैज्ञानिक - एफ / Scientist F)
आई.सी.एम.आर राष्ट्रीय कॉलरा और आंत्र रोग संस्थान
ICMR National Institute of Cholera & Enteric Diseases
पी-३३, सी.आई.टी. रोड, स्कीम-१०एम, वेलियाघाटा
P-33, CIT Road, Scheme-XM, Beliaghata
कोलकाता-७०००१० / Kolkata-700010

DECLARATION

I do, hereby, declare that the work embodied in this thesis entitled "**Studies on colonization factors of enterotoxigenic *Escherichia coli*: Prevalence, Expression and Regulation during pathogenesis**" submitted for the award of Doctorate of Philosophy (Science) in Life Science and Bio-technology, is the completion of work carried out under the supervision of Dr. Nabendu Sekhar Chatterjee, Scientist-F, at the Division of Biochemistry, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata. Neither this thesis nor any part of it has been submitted for either any equivalent degree/diploma or any other academic award elsewhere.

Date: 02/03/2022

Place: Kolkata

Debjyoti Bhakat

Signature of the candidate

Debjyoti Bhakat

to…

Maa, Baba,
Didi & Bhai

Contents

Acknowledgement

Arriving at this point I would like to take this opportunity to sincerely thank everyone who has supported me in different ways to reach this point. Without all of you, this thesis would never have happened. Perhaps no words are sufficient to express my gratitude, sincere appreciation and praise for all those who have contributed both professionally and emotionally in this journey and helped this dissertation take its present shape.

I would like to express my profound and sincere gratitude to my supervisor, Dr. Nabendu Sekhar Chatterjee, Division of Biochemistry, Indian Council of Medical Research – National Institute of Cholera and Enteric Diseases (ICMR-NICED), Kolkata, who has given me the opportunity to pursue my Ph.D. career and for his motivation, encouragement, untiring help and constant guidance throughout my thesis work tenure. I am fortunate enough to have a guide like him who has supported me with his patience and knowledge whilst allowing me space to work in my own way. He has taught me the intricacies of research with great patience while critically analyzing the results which eventually guided me to enter a fruitful and exciting field of research. He has always encouraged me to be involved with various scientific activities out of the sphere of my research which helped me to great extent to be matured as a student of science. His guidance has ultimately helped me to formulate my dissertation to its present status

I express my sincere gratitude to Dr. Shanta Dutta, Director, ICMR-NICED, Kolkata, India for giving me the opportunity to work at this esteemed institution. I am grateful to Dr. Asish Kumar Mukhopadhyay and Dr. Sushmita Bhattacharya for their valuable suggestion and critical insight regarding my work. I am very thankful to Dr. Hemanta Koley for his help, suggestion and guidance regarding the animal experiments. I am grateful to Dr. Sandipan Ganguly, Dr. Provash Chandra Sadhukhan, Dr. Amit Pal, Dr. Sulagna Basu, Dr. Alok Deb and Dr. Santasabuj Das for their help and allowing me to use their laboratory facilities.

Apart from the faculty, I wish to convey my gratitude to my seniors, friends and colleagues for their continuous help and co-operation that created a perfect laboratory environment. Amongst all, first and foremost I express my heartfelt thanks to my lab seniors Dr. Anusuya Debnath, Dr. Rhishita Chourashi, Dr. Sudipto Mandal, Dr. Amarshi Mukherjee, Dr. Moumita Mondal, Mr. Rakesh Naik and my all-time lab partner-cum-colleague Mr. Suman Das, all of whom taught me how to approach any scientific experiment and provided continuous advice, assistance and encouragement to complete my work, even at the darkest hours of scientific failures. A special thank to Mr. Indranil Mondal my friend-cum-colleague for being there with me in all the experiments with his constant help, support and co-operation. A special thanks to Dr. Priyadarshini Mukherjee for her help in animal experiments. I would also like to thank my lab friends and juniors Mrs. Priyanka Basak, Ms. Uzma Khan, Ms. Priyanka Maitra, Ms. Sushmita Kundu, for their unquestioned support and assistance throughout the course of this study. I desire to extend my appreciation to the lab interns Smritikana, Priyanka, Symphony, Anwesha, Swastik, Saurav, Rounack, Pritha, Ditam who have been exceptionally helpful throughout and I sincerely appreciate their friendship and assistance.

I shall treasure the camaraderie that has developed amongst us. I would also like to thank my other seniors Sagar da, Abhishek da, Amlan da, Asim da, Abhijeet da, Arindam da, Piyali di, Prosenjit da, Bipul da, Tanmoy da, Gautam da, Subhankar da, Subhoshree di, Tanushree di, Sraborni di, Tania di, Sayan Da, Nirmalya da, Shreya di, Rituparna di and Punam di and my collegues Rudra Narayan, Sangita, Debjani, Sreeja, Puja, Sanjeeb da, Tapas, Maruf, Ajanta, Amrita, Priyanka, Krishnendu da, Suparna, Mainak, *Sohini, Pralay, Suhrid, Vivek, Supradip and Sagnik for their selfless encouragement and support which helped to overcome the tough times during my work.*

I shall remain grateful to Bibhas da for his technical assistance in sample preparations of histological slide preparations. The staff members of this institute Ananda da, Alam da, Chinmoy da, Malay da, Manas da, Dipak da, Subho da, Prosenjit da and Gouri di have always come forward with helping hands, for which I sincerely thank them all. I am grateful to all the members and the staff of the Institute for their friendly behaviour, cooperation, good wishes throughout the course of my study.

This dissertation would not be possible without the financial assistance provided by the Indian Council of Medical Research, Govt. of India. This investigation was also supported in part by grants from the Department of Biotechnology, Govt. of India.

My heartfelt gratitude goes to Dr. Keinosuke Okamoto and Mrs. Miki Hidaka of Okayama University, Japan, for extending their hands of assistance and providing partial financial support to our lab.

I owe my deepest gratitude to all the authors of literatures reviewed, as their work provided me information and knowledge to carry out my thesis work. I earnestly appeal them to please excuse me for using some of their illustrations with proper citation and acknowledgement in my writing for academic purpose only.

I would like to thank all my teachers in school, college and university who have allowed me to grow and learn with them.

This dissertation would not have been possible without the generous support, encouragement and patience of my family members during the course of this journey. I pay my sincere respect to my Baba, Mr. Uttam Kumar Bhakat and Maa, Mrs. Lalita Bhakat who had to sacrifice a routine life in order to accommodate my long hour at the laboratory and who have been the unequivocal pillars of strength throughout my life. Words cannot express my heartiest gratitude for my brother, Mr. Shubhajyoti Bhakat, my sister, Mrs. Riya Bhattacharjee (Bhakat), my brother-in-law, Sangram Bhattacharjee for their unquestioned and sustained support throughout this endeavour. Ultimately, I thank the almighty and the departed soul of my Dida, Late Mrs. Neela Bhagat, for her eternal love and blessings which helped me sail through the tumultuous journey. I express my heartfelt and deepest gratitude to my family for their constant support, patience, love, encouragement and blessings throughout this journey.

And finally a special thanks to all my friends, their presence never allow me to feel low at any point in my life. Thank you everybody (including all the people I have mentioned above and if I have forgotten anyone else to talk about) for your help and cooperation, what made me reach this height in my career.

Dated:

Place: Kolkata Debjyoti Bhakat

Abbreviations

°C	Degree Celsius
β-ME	β-mercaptoethanol
μg	Microgram
μl	Microlitre
μM	Micromolar
μm	Micrometre
A_{260}	Absorbance at 260 nm
A_{280}	Absorbance at 280 nm
A_{420}	Absorbance at 420 nm
A_{550}	Absorbance at 550 nm
A_{600}	Absorbance at 600 nm
AcOH	Glacial acetic acid
APS	Ammonium Persulphate
BSA	Bovine serum albumin
cfu	Colony forming unit
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CF	Colonization Factor
CFA	Colonization factor antigen
cm	Centimetre
CTAB	Hexadecyltrimethylammonium bromide
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethylsulfoxide
DNA	Deoxy ribonucleic acid
dNTP	Deoxynucleoside triphosphate
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immune sorbent assay
ETEC	Enterotoxigenic *Escherichia coli*
FA	Fluid accumulation
FBS	Fetal bovine serum
Fe	Iron
GOI	Gene of interest

grm	Grams
h	hour
HRP	Horseradish peroxide
HT29	Human colorectal adenocarcinoma cell line
IPTG	Isopropyl β- d-1-thiogalactopyranoside
kDa	Kilo Dalton
kV	Kilo Volt
LB	Luria Bertani
LT	Heat- labile enterotoxin
M	Molar
M9	Minimal Media
MeOH	Methanol
MeOH	Methanol
$MgCl_2$	Magnesium chloride
mL	Millilitre
mM	Milli molar
mm	Millimeter
N	Normal
Na_2CO_3	Sodium carbonate
NaCDC	Sodium chenodeoxycholate
NaCH	Sodium cholate hydrate
NaCl	Sodium Chloride
NaDC	Sodium deoxycholate
NaGCH	Sodium glycocholate hydrate
NaHCO3	Sodium bicarbonate
NaOH	Sodium hydroxide
NBT	Nitro Blue Tetrazolium
NCVF	Non-classical Virulence Factor
ng	Nanogram
nm	Nanometer
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline

PCR	Polymerase Chain Reaction
pmole	Pico molar
PVDF	Polyvinylidene difluoride
qRT-PCR	Real-Time Quantitative Reverse Transcription PCR
RNA	Ribonucleic acid
RT	Room temperature
SDS	Sodium dodecyl sulfate
SEM	Standard error mean
SOB	Super Optimal Broth
SOC	Super Optimal broth with Catabolite repression
ST	Heat-stable enterotoxin
TAE	Tris acetate EDTA
TB	Transformation Buffer
TCA	Taurocholic acid sodium salt hydrate
TCA	Trichloroacetic acid
TE	Tris EDTA
TEMED	Tetramethylethylenediamine
Tm	Temperature of Melting
TMB	3,3',5,5'-Tetramethylbenzidine
UV	Ultraviolet
VF	Virulence factor
WT	Wild type strain
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter I

Introduction

In developing and underdeveloped countries, especially in children under the age of 5 years, Enterotoxigenic *Escherichia coli* (ETEC) infection is the primary cause of diarrhoea and is responsible for causing about 1.4 billion diarrhoeal bouts and 380,000 fatalities in children. ETEC is the bacterial pathogen accounting for approximately 20% of cases of acute watery diarrhoea and important etiological agent for traveller's diarrhoea.

The binding of colonization factors (CFs) with their associated receptors on the intestinal epithelium initiates ETEC pathogenesis, allowing the bacteria to tolerate intestinal peristalsis. The disease is triggered by the release of enterotoxins after colonization.

ETEC has been found to have over 27 different colonization factors, as well as heat-labile (LT) and heat-stable (ST) toxins. According to recent research, the pathogenicity of ETEC is also influenced by the presence of non-classical virulence genes such as *eatA, etpA, tia, tibA, leoA*, and others.

To date, there is no ETEC vaccine available in the market. This is mainly due to the fact vaccine development is centred on classical virulence factors alone and neglects the other important factors. Hence it is important to develop a robust surveillance system for the detection of classical as well as non-classical virulence factors for further improvisation in ETEC vaccine development efforts.

CS6 has been found to be a prevalent CF among clinical isolates of ETEC in our country and also worldwide. Although a large amount of information has been generated on CS6 at the gene and protein levels, little is known about the expression of the protein during pathogenesis. Not much information is available on the expression of CS6 under the influence of different environmental as well as host factors. It is possible that CS6 along with the non-classical virulence factors work synergistically during ETEC pathogenesis. However, studies related to the coordinated expression of non-classical virulence factors along with CS6 in response to the intestinal signals during pathogenesis have not been studied so far. Till date, regulation of CS6 expression remains an unanswered question. Expression of different virulence factors alters due to changes in the environmental conditions and optimum expression of these virulence factors are necessary for pathogenesis.

This project aims to better understand the distribution of most prevalent combination of classical and non-classical virulence factors among clinical isolates of

ETEC in our settings. We will also focus on their regulation of expression by different host and environmental factors during pathogenesis.

1.1. Keywords:

ETEC; Enterotoxigenic *Escherichia coli*; enterotoxins; LT; ST; classical colonization factors; CS5; CS6; non-classical virulence factors; EatA; virulence; gene expression; *in vivo* and *in vitro*; promoter.

Chapter 2

Review of Literature

2. *Escherichia coli*

Escherichia coli, the omnipresent microorganism in laboratories around the world, belongs to the genus *Escherichia.* It is a Gram-negative, rod-shaped, facultative anaerobic, coliform bacterium. It is a versatile organism existing both as a commensal in the lower intestine of warm-blooded organisms as well as a pathogenic form (Makvana *et al.,* 2015).

2.1. Discovery

The bacterium *Escherichia coli* was originally characterized by Dr. Theodor Escherich, the German-Austrian pediatrician in 1885 and he named it *"bacterium coli commune"* (Escherich, 1988) but later Castellani and Chalmers named the bacterium as *Escherichia coli* in 1919, which was officially accepted in 1958 (Minodier, 2011).

2.2. *Escherichia coli:* The bacterium

Escherichia coli (*E. coli*) (***Figure 2.1***) belongs to the Enterobacteriaceae family and is a Gram-negative, rod-shaped, facultative anaerobic bacterium. *Escherichia coli* cells usually occur as single straight rod and typical dimensions are 1.1–1.5 µm broad by 2–6 µm in length. They are either motile or non-motile, and mobility is due to lateral, rather than polar flagella. In addition to flagella, many strains have proteinaceous appendages or fibres extending outwards from the cell surface known as fimbriae or pili, which mainly plays a role for the initial binding to host tissues (Croxen *et al.,* 2013). The majority of *E. coli* strains are commensal colonizing humans and animals' gastrointestinal tracts as natural flora.

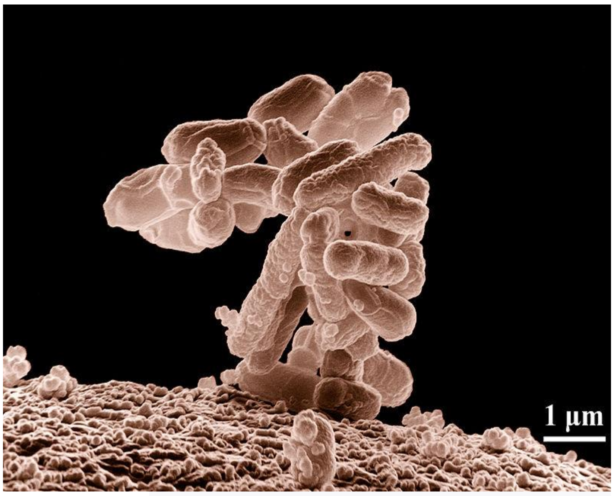

Figure 2.1: A cluster of *E. coli* bacteria enlarged 10,000 times in a low-temperature electron micrograph. Each individual bacterium is oblong shaped. **Source:** The Agricultural Research Service, the research agency of the United States Department of Agriculture.

However, certain strains have evolved into pathogenic *E. coli* by acquired virulence factors via plasmids, transposons, bacteriophages, and/or pathogenicity islands. Serogroups, pathogenicity mechanisms, clinical signs, and virulence factors are used to classify *E. coli* pathovars. These adapted and opportunistic pathogens are able to generate a variety of ailments, from gastrointestinal to the urinary system, circulation, and central nervous infections by acquiring a combination of mobile genetic elements (Kaper *et al.,* 2004).

2.3. Evolution of *Escherichia coli*

The pathogenic variants of *E. coli* differ from commensals by having extra genetic material which can be up to million base pairs. These extra base pairs harbour genes for virulence and fitness. *E. coli* genomes can be divided into two sets: core genome, having the conserved genes and flexible genome, having the pathogenic genes. This pathogenic gene pool had been acquired through the uptake and depletion of genetic materials by horizontal gene transfer and homologous recombination via conjugation and transduction of the species at integration hotspots throughout the genome (Touchon *et al.,* 2009; Croxen *et al.,* 2013). This large cluster of virulence genes is called pathogenicity islands (PAIs), which can either integrate with the chromosome or in the plasmid. By obtaining virulence elements via transposable elements, plasmids, bacteriophages, and/or pathogenicity islands, some strains have transformed into pathogenic *E. coli* (Lim *et al.,* 2010). The core genome of a bacterial species develops primarily through mutation and recombination, while the remainder of the genome undergoes horizontal gene transfer. Epistatic associations between virulence genes and the genetic origin could also be critical for *E. coli* evolution (Touchon *et al.,* 2009). All of these findings suggest that extraintestinal virulence is a multigenic phenomenon involving several gene combinations and redundancy. Furthermore, the lack of an extraintestinal infection-specific gene supports the notion that extraintestinal virulence is a side effect of commensalism. The traits passed down by gain and loss of genes allow the recipient bacterium to adapt to a new environment and withstand harsh conditions. This helps pathogenic isolates to have a lot of genetic variation and chromosome plasticity. Genomic instability, in addition to gene transfer, is a phenomenon that has recently attracted a lot of attention. Various bacteria, like *Campylobacter jejuni,* have been found to have genomic instability aided in colonization (Ridley *et al.,* 2008). The importance of genomic instability in virulence is increasing as organisms must adapt to their surroundings, and such instability plays a key role. The resiliency of phylogenetic groupings indicates that important events such as ancestral selective sweeps linked to beneficial mutations can represent *E. coli* evolution. Metabolic pathways seem to evolve in phylogenetic order, with *E. coli* strains belonging to the same group sharing more pathways. Several *E. coli* pathovars have explicitly evolved in parallel at varying periods, mostly through the acquisition of mobile genetic elements bearing genes related to specific pathogenic behaviours (Chaudhuri *et al.,* 2012).

2.4. *Escherichia coli*: Pathovars

E. coli strain, K-12 has been the most studied organism in laboratories around the world and represent a classical model. *E. coli* normally exist as a commensal in the lower gut of warm-blooded animals but also exists as a diarrhoeal pathogen as well as an extra-intestinal pathogen (Kaper *et al.,* 2004). The study of *E. coli* genomic diversity was described by DNA hybridization and serotyping studies of the flagellar (H) antigen, somatic (O) antigen, and the capsular (K) antigen and pre-dates the availability of nucleotide sequence data (Chaudhuri *et al.,* 2012). At first, all *E. coli* strains were thought to be harmless commensals, but in the 1940s affiliation of *E. coli* strains with outbreaks of infantile diarrhoea (Bray, 1945) established the pathogenic form of *E. coli* and this specific serotypes of *E. coli*, collectively alluded to as the pathovar enteropathogenic *E. coli* (EPEC). Since then different pathovars of *E. coli* were identified (Nataro *et al.,* 1998; Kaper *et al.,* 2004) pathogenic for humans and animals and their mechanisms of pathogenesis have been extensively studied. Gastroenteritis, diarrhoeal disease, sepsis/meningitis and urinary tract infections (UTIs) are some of the common clinical syndromes that could result from infection with either of these pathogenic variants (***Figure 2.2***). There are six well-characterized diarrhoeagenic pathovars of *E. coli,* which are enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enterohaemorrhagic *E. coli* (EHEC), diffusely adherent *E. coli* (DAEC) and enteroaggregative *E. coli* (EAEC). Urinary tract infections (UTIs) and sepsis/meningitis are responsible for extraintestinal infections and are called ExPEC. Uropathogenic *E. coli* (UPEC) causes urinary tract infections (UTIs), and neonatal meningitis *E. coli* (NMEC) causes meningitis and sepsis (Kaper *et al.,* 2004). While these pathogenic variants frequently target the same host machinery, each pathovar has its array of techniques for altering and exploiting host cells. The mechanism of pathogenesis of these diarrhoeagenic *E. coli* are discussed in brief:

2.4.1. Enteropathogenic *Escherichia coli* (EPEC): Large episodes of infantile diarrhoea in the United Kingdom prompted Bray to establish the pathogenic forms of *E. coli* in 1945. Children with diarrhoea were shown to have a set of serologically different *E. coli* strains called Enteropathogenic *Escherichia coli* and was the first *E. coli* strain

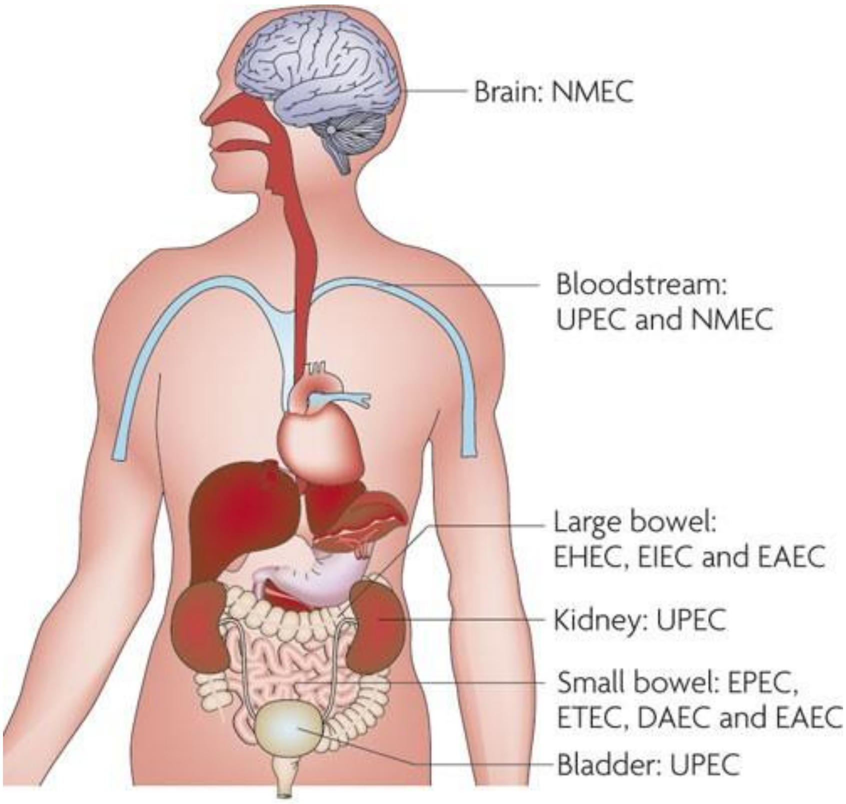

Figure 2.2: Colonization sites of Pathogenic *Escherichia coli* in the human body. Enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), and diffusely adherent *E. coli* (DAEC) colonise the small intestine and cause diarrhoea, whereas enterohaemorrhagic *E. coli* (EHEC) and enteroinvasive *E. coli* (EIEC) colonise the large intestine and cause disease; enteroaggegrative *E. coli* (E Uropathogenic *E. coli* (UPEC) enters the urinary tract and goes to the bladder, causing cystitis and, if untreated, pyelonephritis. Both UPEC and neonatal meningitis *E. coli* (NMEC) can cause septicaemia, and NMEC can pass the blood–brain barrier and cause meningitis in the central nervous system. **Source:** Croxen *et al., Nature Reviews Microbiology.* 2010.

attributed to a human illness. On epithelial cells, the attaching and effacing (A/E) histopathology can be seen at ultrastructural levels which is the hallmark of EPEC infections and lead to profuse watery, sometimes bloody diarrhoea. These observations of lesion formation were critical in establishing adhesion as a key factor in EPEC pathogenesis (***Figure 2.3***) (Deborah *et al.,* 2005).

2.4.2. Enterohaemorrhagic *Escherichia coli* (EHEC): Global epidemics of bloody diarrhoea and hemolytic uremic syndrome (HUS), particularly in developed countries, are caused by the human pathogen Enterohemorrhagic *Escherichia coli*

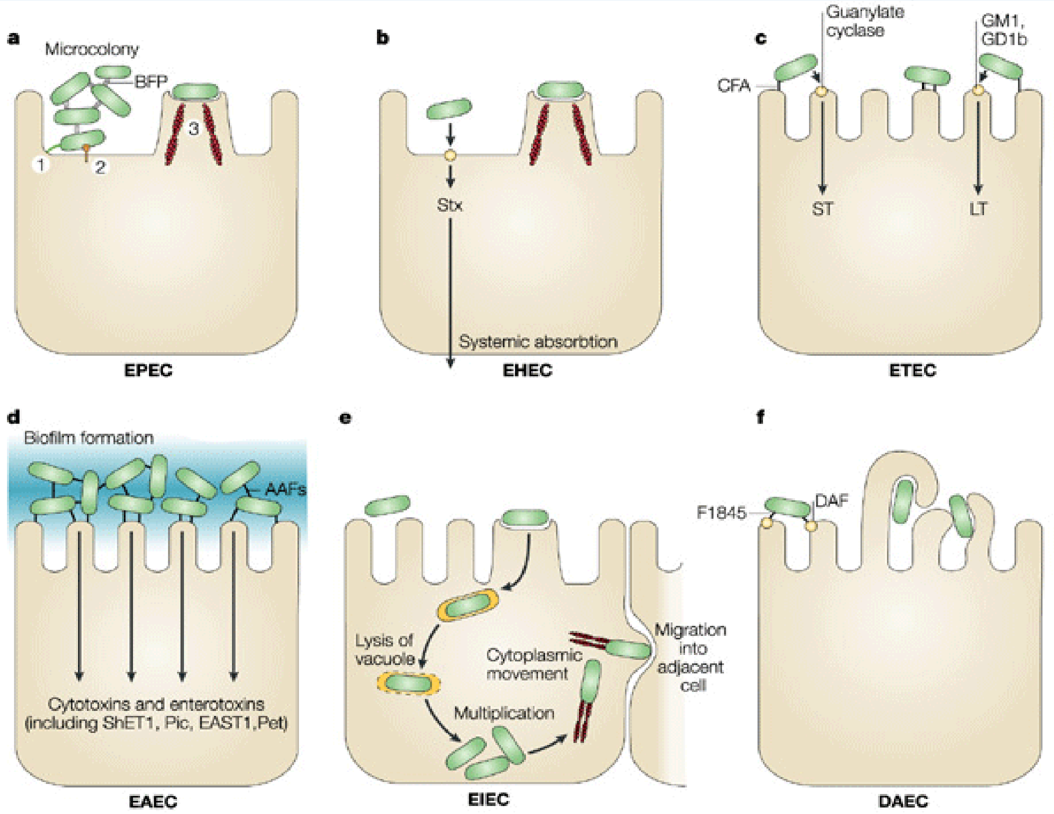

Figure 2.3: Mechanism of pathogenesis of Pathovars of *E. coli.* Each of the six identified types of diarrhoeagenic *E. coli* interacts with eukaryotic cells in a different way. Each category's interaction with a typical target cell is depicted schematically here. These descriptions are mostly based on *in vitro* research and may not accurately reflect what happens in infected persons. **a** | EPEC bind to small bowel enterocytes but damage the natural microvillar architecture, resulting in the adhering and effacing lesion. An inflammatory reaction and diarrhoea accompany cytoskeletal derangements. 1. Initial adhesion, 2. Type III secretion-mediated protein translocation, and 3. Pedestal formation. **b** | EHEC causes the attaching and effacing lesion in the colon as well. The synthesis of Shiga toxin (Stx), whose systemic absorption can lead to potentially fatal consequences, is a distinguishing hallmark of EHEC. **c** | Similarly, ETEC bind to small bowel enterocytes and secrete heat-labile (LT) and/or heat-stable (ST) enterotoxins, which cause watery diarrhoea. **d** | EAEC forms a thick biofilm on small and large intestinal epithelia and produces secretory enterotoxins and cytotoxins. **e** | EIEC invades the colonic epithelial cell, lyses the phagosome, and nucleates actin microfilaments to proceed through the cell. The bacteria could propagate laterally through the epithelium through direct cell-to-cell dissemination or by exiting and re-entering the baso-lateral plasma membrane. **f** | DAEC causes small bowel enterocytes to produce long finger-like cellular projections that wrap around the bacteria, eliciting a distinct signal transduction response. AAF stands for aggregative adherence fimbriae; BFP stands for bundle-forming pilus; CFA stands for colonization factor antigen; DAF stands for decay-accelerating factor; EAST1 stands for enteroaggregative *E. coli* ST1; LT stands for heat-labile enterotoxin; ShET1 stands for Shigella enterotoxin 1; ST stands for heat-stable enterotoxin.**Source:** Kaper *et al., Nature Reviews Microbiology*. 2004.

(EHEC). It's also known as Verocytoxin-producing *Escherichia coli* (VTEC) and Shiga-toxin-producing *Escherichia coli* (STEC). Cattle are the main reservoir, but it is asymptomatic in adult cattle. EHEC creates attaching and effacing (A/E) lesions on the mucosal epithelium at the recto-anal junction (RAJ), allowing it to colonise. In A/E lesions, microvilli are destroyed, bacteria are closely connected to the cell, and polymerized actin accumulates beneath the site of bacterial attachment, forming a pedestal-like structure that cups individual bacteria (***Figure 2.3***). Hemorrhagic colitis (bloody diarrhoea) is a consequence of EHEC infection, and subsequent complications can lead to potentially fatal hemolytic uremic syndrome. (Sperandio *et al.,* 2012).

2.4.3. Enterotoxigenic *Escherichia coli* (ETEC): Enterotoxigenic *Escherichia coli* isolates cause acute diarrhoeal infections in humans and animals globally, and in developing countries, they are accountable for a high rate of infant mortality. They are also a prevalent cause of diarrhoea among travellers from developed countries who travel to tropical or subtropical areas of the globe. ETEC is primarily distinguished by two virulence characteristics: colonization factors, which aid in small intestinal colonization, and enterotoxins, LT and ST, which produce a net secretion of electrolytes and water into the gut lumen and, as a result, cause severe watery diarrhoea. (***Figure 2.3***) (Guth, 2000).

2.4.4. Enteroinvasive *Escherichia coli* (EIEC): Enteroinvasive *Escherichia coli* belong to a group of intracellular pathogenic organisms that may invade colon epithelial cells, propagate within them, and transfer between neighbouring cells in a manner similar to Shigella (***Figure 2.3***). It is characterized by bloody, mucoid diarrhoea and Shigella like dysentery. EIEC strains have a pINV plasmid that is similar to Shigella's and can exhibit Shigella-like invasive behaviour. The molecular machinery required for the bacterium's invasion, survival, and diffusion within the host is encoded by the pINV plasmid. This pathovar is different from the others in that it is obligate intracellular bacteria that lack flagella and adherence factors (Pasqua *et al.,* 2017).

2.4.5. Enteroaggregative *Escherichia coli* (EAEC): In infants and adults in developing and developed countries, enteroaggregative *Escherichia coli* is linked to acute or persistent watery diarrhoea with mucus or blood. After ETEC, it is the second most prevalent cause of traveller's diarrhoea. EAEC is defined by aggregative adherence (AA) to HEp-2 cells and the production of biofilms on the intestinal mucosa with a "stacked-brick" adherence phenotype, which is linked to the presence of a 60 MDa plasmid (pAA) (***Figure 2.3***) (Kaur *et al.,* 2010).

2.4.6. Diffusely adherent *Escherichia coli* (DAEC): Afa/Dr adhesins, which contain Afa, Dr, and F1845 adhesins encoded by the *afa*, *dra* and *daa* operons respectively, are capable of producing diarrhoea in diffusely adherent *Escherichia coli*. (***Figure 2.3***). DAEC diffusely adhered to HeLa and HEp-2 cells. DAEC also causes recurring urinary tract infections (UTIs) in adults (Servin, 2014).

2.4.7. Extraintestinal Pathogenic *Escherichia coli* (ExPEC): Many virulent strains of pathogenic *E. coli* cause extraintestinal infections and are associated with immunodeficiency or immunocompromised hosts. ExPEC bacteria are classified as APEC (avian pathogenic *Escherichia coli*), UPEC (UroPathogenic *Escherichia coli*, UTI), and NMEC (non-pathogenic *Escherichia coli*) (neonatal meningitis-causing *E. coli*). The most frequent ExPEC infection is urinary tract infection (UTI), which is commonly diagnosed in young women and can become recurring, as well as in elderly patients after catheterization. NMEC is associated with meningitis in newborn infants and has a high mortality rate (Smith *et al.,* 2007).

2.4.8. Hybrid Diarrhoeagenic *Escherichia coli*:

EPEC/ETEC: The occurrence of hybrid strain of EPEC/ETEC possessing both *eae* and *elt* virulent genes of EPEC and ETEC, respectively had been reported in a child with acute diarrhoea (Dutta *et al.,* 2015).

STEC/ETEC: According to a study in Finland, hybrid STEC/ETEC strain harbouring genes encoding both Shiga toxins 1and/or 2 along with ETEC specific heat-stable (ST) enterotoxin had STEC genes accounting for 1% of human STEC and 14% of animal and environmental STEC genes. (Nyholm *et al.,* 2015).

2.5. Enterotoxigenic *Escherichia coli* (ETEC)

Enterotoxigenic *Escherichia coli* (ETEC) is an *Escherichia coli* pathovar. It is one of the most common bacterial causes of diarrhoea in underdeveloped countries, particularly among children under the age of five. It is the most prevalent cause of diarrhoea in travellers. The presence of enterotoxins, heat-labile (LT) and/or heat-stable (HT), and colonization factors (CFs) used for attachment to host intestinal cells are the key characteristics of ETEC. Both the enterotoxins as well as colonization factors (CFs) are plasmids encoded and these virulence determinants are known as classical ones. However, ETEC pathogenesis is also aided by several chromosomally encoded factors such as EatA, EtpA, Tia, TibA and LeoA which are called non-classical.

2.5.1. Discovery of ETEC:

In 1956, S.N. De and his colleagues reported for the first time cholera-like illness with *E. coli* strains isolated from children and adults after frequent failure in isolating *Vibrio cholera* from patients with signs and symptoms of typical cholera was convinced that the faeces yielded a pure culture of *Bacterium coli* behaving in the same way as *V. cholera* and confirmed it by rabbit ileal loop (RIL) experiments (***Figure 2.4***) (De *et al.,* 1956). Unfortunately, S.N. De did not isolate these *E. coli*. Later in 1968, R.B. Sack confirmed these findings when he found *E. coli* concentrations in stool samples of patients admitted with cholera-like symptoms which strongly suggested *E. coli* as the etiological agent of this syndrome. By using the RIL assay, he was able to confirm that not only active cultures but also culture filtrates, produced ileal loop inflammation and swelling. (***Figure 2.4***). In 1967, veterinary scientists examined *E. coli* samples from young animals with diarrhoea using the ileal loop model in pigs, dogs, and rabbits, particularly pigs and they identified the enterotoxins LT and ST and colonization factors (Sack, R. 2011). This suggests that ETEC strains from humans and animals are closely related and their studies go side by side.

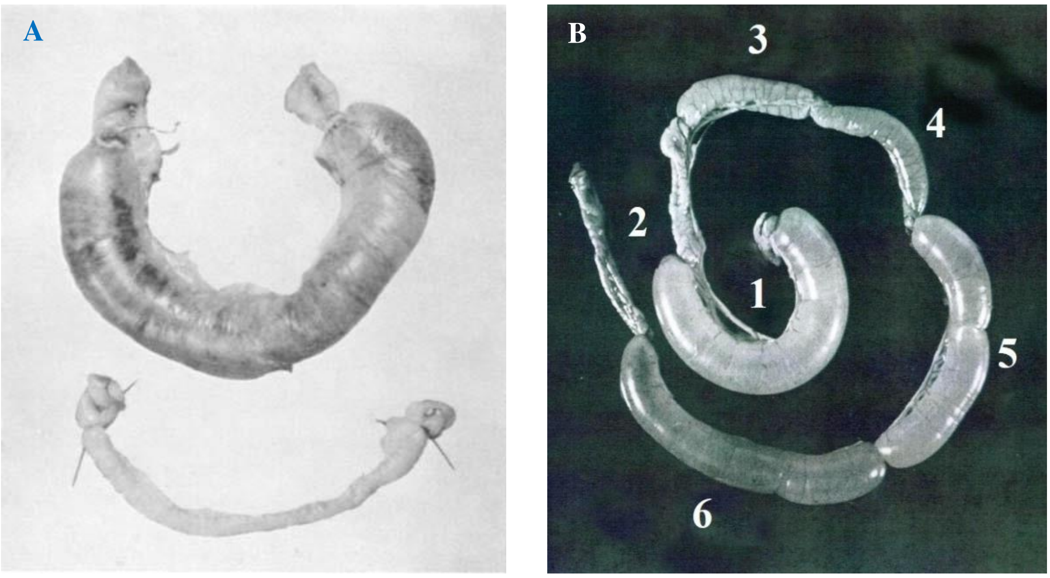

Figure 2.4: Rabbit ileal loop Assay for detection of ETEC. A. Loops of rabbit small intestine isolated by ligatures and injected with *Bacterium coli* cultures at pH 8.4 after 24 hours of injection. The upper loop was injected with pathogenic *Bacterium. coli*, whereas the bottom loop was injected with non-pathogenic *Bacterium coli*. **Source:** De *et al., Journal of Pathology and Bacteriology*. 1956.
B. RIL assay titrations of *E. coli* culture filtrates. Positive (*V. cholerae* enterotoxin) and negative (saline) controls are represented by loops 1 and 2, respectively. The findings obtained with increasing dilutions of the ETEC culture filtrates are shown in Loops 6, 5, 4, and 3.**Source:** Sack R.B., *The Indian Journal of Medical Research*. 2011.

2.5.2. **Pathogenesis of ETEC:**

Like all other gastrointestinal pathogens, ETEC must establish itself by attaching to the intestinal epithelium thereby multiplying and evading the host defence and ultimately causing damage to the host. The strategy of pathogenesis is as follows: (1) Adhere to host cells (2) Multiplication (3) Evasion of host and (4) Damage to hosts (Mims, *et al.,* 2001). Upon entering the small intestine the bacteria use proteinaceous appendages known as surface colonization factors, to establish adherence. After attachment and colonization of bacteria, heat-labile or heat-stable enterotoxins or both enterotoxins are produced which affected the intestinal epithelial cells thereby causing secretory diarrhoea. These colonization factors and enterotoxins are together called classical virulence factors and

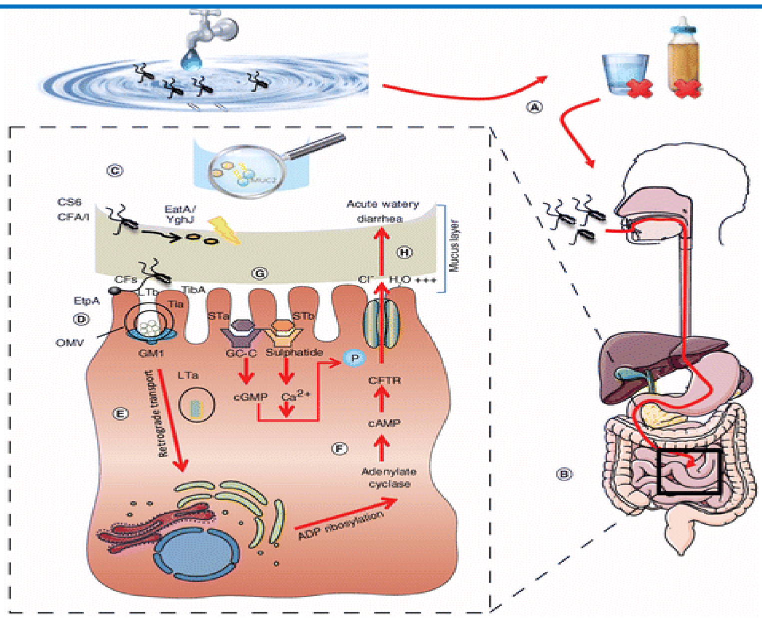

Figure 2.5: Physiopathology of enterotoxigenic *Escherichia coli* infections. The figure resumes the main features of enterotoxigenic *E. coli* physiopathology. (A) Contaminated waters mostly resulting from a lack of sanitation (drinking water and weaning foods) are the most common sources of ETEC infections. (B) ETEC must survive acidic stomach conditions before colonising the small intestine after consumption. (C) Thanks to CFs (such a CS6 or CFA/I), ETEC will be able to adhere both to the mucus layer and enterocytes. Mucin-degrading enzymes EatA and YghJ will favor the adhesion of bacteria to enterocytes and later on the toxin access to their receptors. Non fimbrial adhesins (TibA, Tia and EtpA) also mediate the initial host/pathogen interactions. (D) ETEC adhesion promotes the secretion of LT and/or ST toxins. LTb subunits in OMV bind GM1 monoganglioside at the surface of the enterocyte. (E) The LTa subunit is endocytosed and joins the cytosol via retrograde transport. (F) LTa ADP-rybosylates the GSα subunit of adenylate cyclase leading to an increase in cAMP. (G) STa binds to the GC-C receptor and STb to sulphatide resulting in an increase in cGMP and Ca2+, respectively. **Source:** Roussel *et al., Future Microbiology*. 2017.

both are plasmid-encoded (Mirhoseini *et al.,* 2018). Other virulence factors that contribute to disease have been discovered in addition to the classical virulence factors called non-classical virulence factors (Del Canto *et al.,* 2011). To summarize, ETEC's molecular pathogenesis can be described as all of the virulence traits required for the organization of LT and ST commuting to their respective receptors on the intestinal epithelium (***Figure 2.5***).

2.5.3. Virulence factors of ETEC:

The virulence factors of ETEC consists of classical virulence factors which include enterotoxins and colonization factors and non-classical virulence factors.

2.5.3.1. Enterotoxins:

The enterotoxins heat-labile toxin (LT) and heat-stable toxin (ST) cause diarrhoea by ETEC strains. Strains can express LT alone, ST alone, or both the enterotoxins. Again the enterotoxins are subcategorized into different variants (***Table 2.1***).

Heat-labile enterotoxin:

Heat-labile toxin (LT) is a heterohexameric protein of 85 kDa and is closely related cholera toxin. It is composed of a pentameric B subunit arranged in a ring each of 11.6 kDa, and a single A subunit of 28 kDa (***Figure 2.6***). The A subunit has a disulfide bridge connecting two domains: A1, the functional toxin protein, and A2, the helical part that links the subunit to the B pentamer. (Guth, 2000; Fleckenstein *et al.,* 2010). The LT is encoded by the *eltAB* operon and forms a classic AB_5 toxin. The detrimental A subunit is internalised after the B subunit binds to the host receptor GM1, resulting in a rise in intracellular cyclic AMP (cAMP), dysregulation of the cystic fibrosis conductance regulator (CFTR), and ultimately leakage of water and ions from the epithelium (Sears *et al.,* 1996). Prior to assembly in the periplasm of the bacterial envelope, the toxin subunits are synthesised separately as precursor polypeptides and translocated through bacterial cytoplasmic membranes.

The formation of a disulfide bond in the B subunit is a necessity for the assembly of B-subunit pentamers. The A subunit plays no part in the pentamerization of B-subunits. Association of the A subunit with an intermediate B-subunit assembly is part of the holotoxin assembly pathway. The A subunit will boost the pentamerization of the B subunit, most likely by stabilising an assembly intermediate. The creation of the native holotoxin complex arose from the production of B pentamers followed by association with an A subunit. (Hardy *et al.,* 1988).

$$\mathbf{5B \rightarrow B_5 + A \rightarrow AB_5}$$

Distinct ETEC strains express different allele variations of the LT toxin, which is polymorphic. The LTs are classified into two groups: LT-I and LT-II (Joffré *et al.,* 2016).

Table 2.1: List of different variants of enterotoxins of ETEC.

Enterotoxins	Variants	Encoding Gene	Location of Genes	Host Specificity	Receptor
Heat-labile enterotoxin (LT)	LTIh	*eltAB*	plasmid	humans	GM1a
	LTIp	*eltAB*	plasmid	piglets	GM1a
	LTIIa	*eltAB*	chromosome, prophages	water-buffalo, humans	GD1b
	LTIIB	*eltAB*	chromosome, prophages	unknown	GD1a
	LTIIc	*eltAB*	chromosome, prophages	humans, calves	GM1a
Heat-stable enterotoxin (STa)	STp	*estA1, estA5, estA6*	plasmids	piglets, calves, humans	GC-C
	STh	*estA2, estA3/4, estA7*	plasmids	humans	GC-C
Heat-stable enterotoxin (STb)	STb	*estB*	plasmids	post-weaning pigs	sulfatide
	STbH12N	*estBC34A*	plasmids	post-weaning pigs	sulfatide

GM = Monosialotetrahexosylganglioside; GD = Disialoganglioside; GC-C = guanylate cyclase.
Source: Wang *et al., Toxins*. 2019.

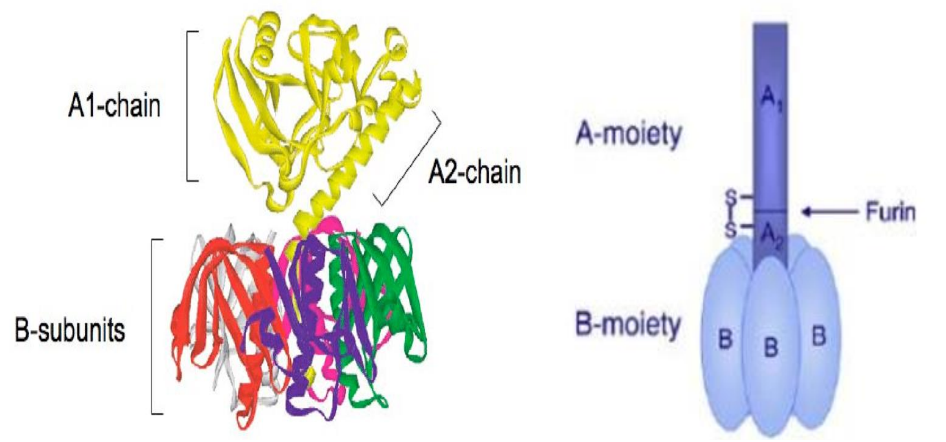

Figure 2.6: Structure of *E. coli* LT toxins A- and B-subunit, LT contain five identical subunits binding to each other, making up a pentameric structure. **Source:** Mirhoseini *et al., Microbial Pathogenesis*. 2018.

Heat-labile enterotoxin I (LT-I): Enterotoxin LT-I is a diarrhoea-causing enterotoxin that has been isolated from humans (LT-Ih) and pigs (LT-Ip) strains and also enhances ETEC and other pathogen adherence to the intestinal epithelium (Wang *et al.,* 2019). LT-I toxin is mainly found in human ETEC and the toxin is neutralizable by anti-cholera toxin sera. The A and B subunits of the LT-I enterotoxin share approximately 80% protein sequence identity with cholera toxin (CT). LT-I and CT both have the same holotoxin structure, principal receptor identity, enzymatic activity, and activity in animal experimentation.

Given the fact that some infections caused by LT-I-producing ETEC strains are as serious as cholera, most LT-I-producing ETEC infections are weaker and last shorter than *V. cholerae* expressing CT. The subunit A1 (21.8 kDa), which has ADP-ribosyltransferase activity, and the A2 (5.4 kDa) subunit, which joins the A1 and B subunits, make up the A fragment. By ADP-ribosylating the alpha subunit of the GTP-binding protein Gs, LT-I stimulates adenylate cyclase activity. ADP-ribosylation inhibits the alpha subunit of Gs' intrinsic GTPase activity, resulting in constitutive adenylate cyclase activation. Elevated intracellular cAMP stimulates the chloride receptor of the cystic fibrosis transmembrane regulator (CFTR), causing electrolyte and water secretion and diarrhoea. (***Figure 2.7***) (Sears *et al.,* 1996).

Heat-labile enterotoxin II (LT-II): The toxin is mainly found in animals and quite rarely found in ETEC infecting humans. Anti-cholera toxin sera is unable to neutralize it. The LT-II variants, unlike the plasmid-encoded LT-I, are encoded by chromosome and prophages. The A subunits of LT-II contain about 57% similarity with the A subunits of LT-I and CT, but the B subunits of LT-I and CT have essentially no homology, which is consistent with the toxins' different ganglioside-binding specificities. LT-IIa and LT-IIb are two distinct members of the LT-II family. The anticipated A and B subunit sequences had 71 and 66% similarity, respectively. By activating adenylate cyclase via the GTP-dependent route, LT-II raises intracellular cyclic AMP levels, just like LT-I (Sears *et al.,* 1996). A new type of LT-II variant had been isolated from the avian host and named as LT-IIc. Antisera that recognise LT-IIa or LT-IIb, but not anti-CT antiserum, partially neutralised the cytotoxicity of LT-IIc. CT, LT-IIa, and LT-IIb were found to be less cytotoxic than LT-IIc (Nawar *et al.,* 2010).

These two LT variations, LT1 and LT2, are especially prevalent and have been linked to clonal ETEC lineages that express various colonization factors. The LT1 variant

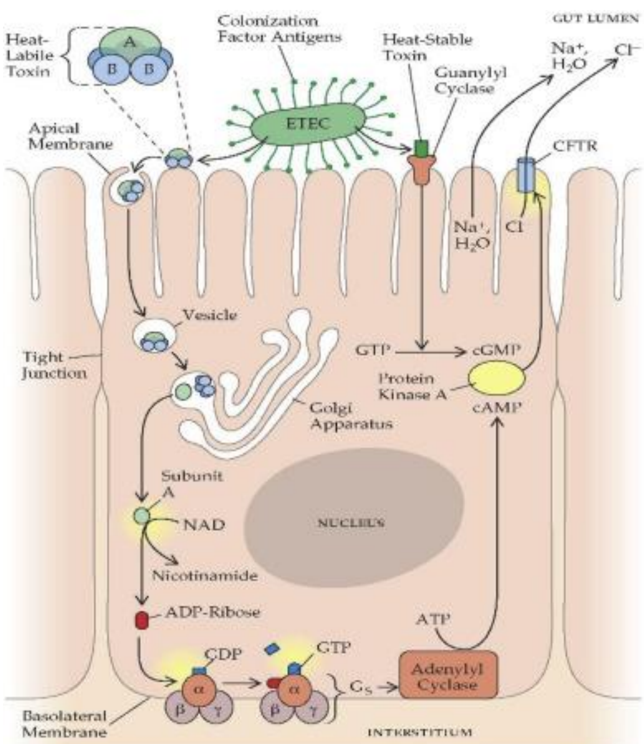

Figure 2.7: Mechanisms of pathogenesis by enterotoxigenic *E. coli*. Pathogenesis of ETEC bacteria invasion involves two steps which are intestinal colonization, followed by elaboration of diarrhoeagenic enterotoxins. Activation of adenylyl and guanylate cyclase lead to formation of cAMP and cGMP, stimulates water and electrolyte secretion by intestinal endothelial cells. **Source:** Mirhoseini *et al., Microbial Pathogenesis*. 2018.

is found in 2 discrete clonal lineages and is associated with CS1+CS3, and CS2+CS3 expressing ETEC. Despite their virulence, they do not generate or secrete large quantities of LT toxin. The LT2 variant, on the other hand, is found in ETEC strains that express CFA/I, CS5+CS6, and generate and secrete significantly more LT toxin. ETEC lineages that have expanded worldwide contain LT1 and LT2 variants, despite differences in toxin output, indicating that these variants are essential for ETEC virulence (Joffré *et al.,* 2016). Wild-type ETEC strains recovered from human hosts have a substantially higher natural diversity of LTs than previously anticipated, which could affect the strains' pathogenesis and disease epidemiology (Lasaro *et al.,* 2008).

Heat-stable enterotoxin:

ETEC secretes small cysteine-rich peptides heat-stable toxins of twenty amino acid residues. STa and STb are two structurally, functionally, and antigenically different kinds. The plasmid-bound *estA* (STI) and *estB* (STII) genes encode these peptides.

Heat-stable enterotoxin a (STa): In ETEC associated with diarrhoea in humans, neonatal piglets, and calves, STa is more important. STa is methanol soluble and protease-resistant. STa is even further subdivided into two categories depending on the host species: STh and STp; which were originally isolated from ETEC strains isolated from human and swine, respectively.

Only human ETEC strains produce STh, while ETEC isolates from porcine, bovine, and human sources widely possess STp. (Wang *et al.,* 2019).

STa is generated as a 72-amino-acid precursor that is processed into a 53-amino-acid peptide by signal peptidase 1 and translocated to the periplasm, where three intramolecular disulfide bonds crucial to toxin activity are formed by DsbA. This peptide is proteolytically cleaved outside the cell before being secreted, resulting in physiologically active STa of 18 or 19 amino acids (***Figure 2.8***) (Sears *et al.,* 1996). The

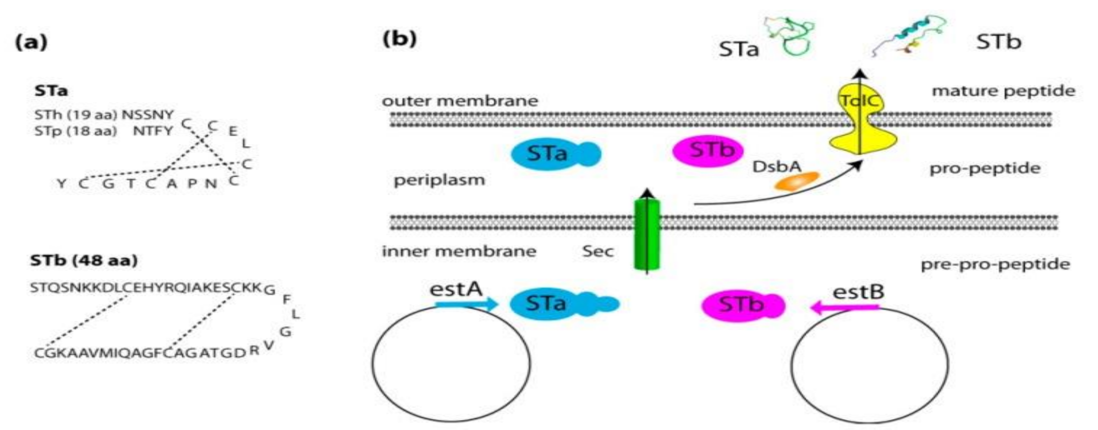

Figure 2.8: ETEC's secretion of heat-stable enterotoxins. (a) The disulfide bonds are represented by the dashed lines in the heat-stable enterotoxin (ST) peptides and the sequences of mature STa and STb peptides. **(b)** STa and STb synthesis and secretion DsbA: Disulfide oxidoreductase; Sec: Secretory pathway. **Source:** Wang *et al.*, *Toxins*. 2019.

heat-stable enterotoxin STa binds to the homotrimeric guanylate cyclase C (GC-C) receptor in the intestine's brush border and activates its intracellular catalytic domain, causing GTP hydrolysis and intracellular cyclic GMP (cGMP) aggregation. The cGMP-dependent protein kinase II is activated as a result of the increased cGMP levels (PKGII). Furthermore, cGMP has been shown to inhibit phosphodiesterase 3 (PDE3), resulting in cAMP-dependent protein kinase A activation (PKA). The cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel is phosphorylated and opened by activated PKGII and PKA, causing Cl and HCO_3 release into the intestinal lumen. Protein Kinase A also prevents Na+ reabsorption by phosphorylating the sodium/hydrogen exchanger 3 (NHE3) (***Figure 2.9***) (Nair *et al.,* 1998; Vaandrager *et al.,* 2002; Wang *et al.,* 2019).

Heat-stable enterotoxin b (STb): STb is mostly associated with animals and particularly in post-weaning pigs. STb is methanol insoluble and protease-sensitive. Enterotoxin STb is highly conserved in ETEC isolates globally in contrast to STa (Wang *et al.,* 2019). In the periplasmic space, the STb gene (*estB*) generates a 71-amino-acid precursor protein with four cysteines that is proteolytically converted into a mature 48-amino-acid protein with two disulfide linkages. This trypsin-sensitive protein, unlike *E. coli* STa, is secreted extracellularly without further processing (Sears *et al.,* 1996). The enterotoxin STb was discovered to interact directly with sulfatide on the surface of intestinal epithelial cells in the pig jejunum. This connection activates pertussis toxin-sensitive GTP-binding regulatory protein (Gi3), which causes calcium to flow into the cell via a receptor-dependent ligand-gated Ca2+ channel (Labrie *et al.,* 2001). The elevated intracellular Ca2+ concentration in response to STb is implicated in the activation of calmodulin-dependent protein kinase II (CaMKII) by the Ca2+-calmodulin

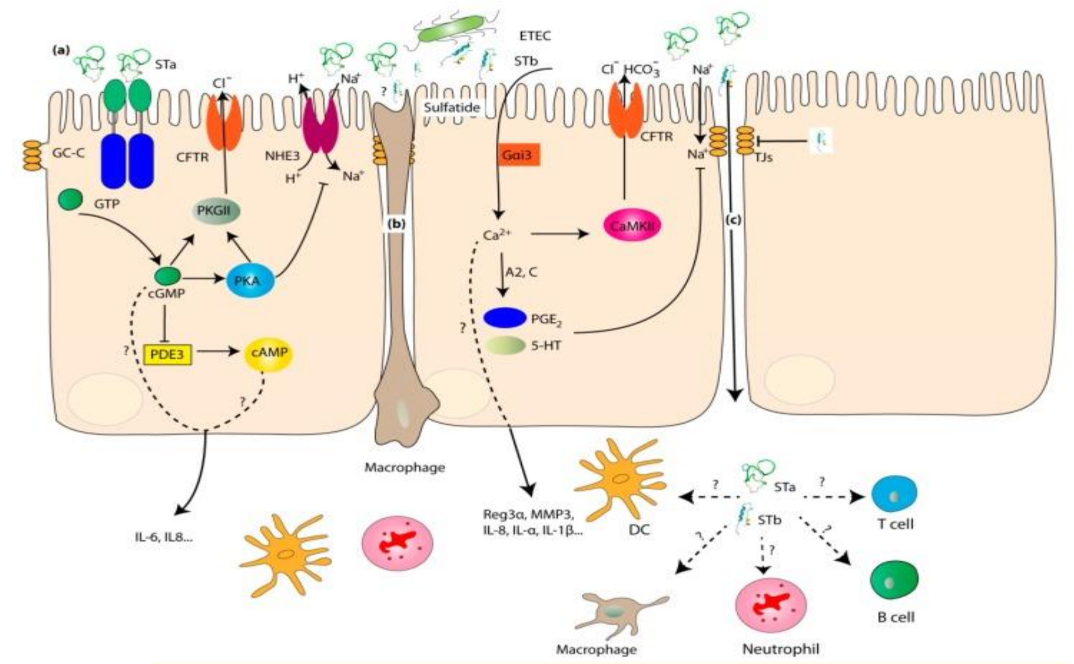

Figure 2.9: The effects of heat-stable enterotoxins on enterocytes and the gut immune system are depicted in this diagram. (a) STs' effects on apical membranes**; (b)** STs' effects on transepithelial dendrites; **(c)** STs' paracellular trafficking**. CFTR:** Cystic fibrosis transmembrane conductance regulator**; GC-C:** Guanylate cyclase C**; PKA:** cAMP-dependent protein kinase; **NHE3**: Na+/H+ exchanger**; PDE3:** cGMP-inhibitable phosphodiesterase 3; **PKGII:** cGMP-dependent protein kinase II; **Gi3:** GTP-binding regulatory protein that is sensitive to pertussis toxin**; A2**: Phospholipases A2; **C**: Phospholipases C; **PGE2**: Prostaglandin E2; **5-HT**: 5-hydroxytryptamine; **CaMKII**: Calmodulin-dependent protein kinase II; **MMP1:** Matrix metallopeptidase 1; **TJs**: Tight junctions. **Source:** Wang *et al.*, *Toxins*. 2019.

pathway, as well as the activation of CFTR by protein kinase C (PKC), resulting in fluid buildup in the intestine (***Figure 2.9***) (Dreyfus *et al.,* 1993; Sears *et al.,* 1996).

2.5.3.2. Colonization Factors:

ETEC pathogenesis is characterised by the adhesion or colonization of the small intestine and, as a result, the release of enterotoxins that cause severe diarrhoea. Adhesins on the surface of bacteria serve a key role in the colonization of the intestine, allowing bacteria to endure intestinal peristalsis and establish infection. The bulk of adhesins are pili (fimbriae) or pilus-related polymeric molecules, although there are a few adhesins that are simple outer membrane proteins that do not form macromolecular structures and are not associated with pili (Gaastra *et al.,* 1996; Madhavan *et al.,* 2015). Colonization factors are the aggregate name for these diverse proteinaceous surface structures (CFs). Based on their antigenic specificity, these CFs are referred to as colonization factor antigens (CFAs), coli surface antigens (CS), and putative colonization factors (PCFs) (Torres, *et al.,* 2005). ETEC CFs' natural intestinal receptor molecules are unknown.

Glycoconjugates are the molecules that ETEC fimbriae bind to. The oligosaccharide sequences of glycoconjugates on eukaryotic cell membranes are extremely diverse (Gaastra *et al.,* 1996; Madhavan *et al.,* 2015). At present more than 27 different CFs in human ETEC have been discovered and among them, 23 are well characterised best-characterized (***Table 2.2***). The various CFs are linked to a small number of serogroups, with some having only one serotype (Gaastra *et al.,* 1996; Madhavan *et al.,* 2015; von Mentzer *et al.,* 2017). The genes that code for ETEC CFs are frequently encoded by plasmids.

Genes clustered together in operons encode all of the genes needed for the assembly of functional CFs. These DNA fragments are flanked by insertion sequences and transposons and GC concentration and codon usage are lower in these bacteria than in *E. coli*. The chaperone-usher route is responsible for the majority of CF synthesis. Fundamentally, these operons are made up of genes that code for pilins, a chaperone gene, and an usher gene. The CF operons include two or three pilin genes, one of which encodes a major pilin that makes up the majority of the pilus shaft and the others minor pilins that only make up a small percentage of the pilus structure.

In other circumstances, pili are made up of two pilins that are proportionately identical. Fimbrial usher proteins (FUPs) are divided into six phylogenetic clades: α, β,

Table 2.2: List of colonization factors (CFs)

CF	Morphology	Size (nm)	Gene	Encoded protein	Mw (kDa)
CFA/I	Fimbrial	7	*cfaB, cfaE*	CfaB, CfaE	15, 41
CS1	Fimbrial	7	*cooA*	CooA	16.5
CS2	Fimbrial	7	*cotA*	CotA	15.3
CS3	Fibrillae	2-3	*cstH*	CstH	15.1
CS4	Fimbrial	6	*csaB, csaE*	CsaB, CsaE	17, 40
CS5	Helical	5	*csfA, csfD*	CsfA, CsfD	21, 41
CS6	Non- Fimbrial	-	*cssA, cssB*	CssA, CssB	18, 15
CS7	Helical	3-6	*csvA*	CsvA	21.5
CS8 (CFA/III)	Fimbrial	7	*cofA*	CofA	20.5
CS10	Non- Fimbrial	-	-	-	16
CS11	Fibrillae	3	-	-	-
CS12	Fimbrial	7	*cswA, cswF*	CswA, CswF	19, 17.4
CS13	Fibrillae	-	*cshE*	CshE	27
CS14	Fimbrial	7	*csuA1*	CsuA1	15.5
CS15	Non- Fimbrial	-	*nfaA*	NfaA	16.3
CS17	Fimbrial	7	*csbA*	CsbA	17.5
CS18	Fimbrial	7	*fotA*	FotA	25
CS19	Fimbrial	7	*csdA*	CsdA	16
CS20	Fimbrial	7	*csnA*	CsnA	20.8
CS21	Fimbrial	7	*lngA*	LngA	22
CS22	Non- Fimbrial	-	*cseA*	CseA	15.7
CS23	Fimbrial	-	*aalE*	AalE	28
CS30	Fimbrial	7	*csmA, csmF*	csmA, csmF	18.5, 17.2

➢ **Source:** Boylan *et al.*,1956; Knutton *et al.*,1987, 1989; McConnell *et al.*,1989, 1990; Heuzenroeder *et al.*,1990; Aubel *et al.*,1992; Jordi *et al.*,1992; Giron *et al.*,1994; Froelich *et al.*,1994, 1995; Schmidt *et al.*,1995; Taniguchi *et al.*,1995; Gaastra *et al.*,1996; Valvatane *et al.*,1996; Wolf *et al.*,1997; Duthy *et al.*,1999; Pichel *et al.*,2000; Ludi *et al.*,2006; Von Mentzer *et al.*, 2017.

γ, π, κ and σ and out of which α, γ and κ clades comprise the pili of ETEC strains that infect humans (***Figure 2.10***) (Gaastra *et al.*, 1996; Wolf *et al.*, 1997; Madhavan *et al.*, 2015). Antigenicity, molecular weight, the major subunit's N-terminal amino acid sequence, and structural morphology either fimbrial, fibrillar, helical, or nonfimbrial are used to classify CF.

The most dominant CFs around the world are CFA/I, CFA/II and CFA/IV groups. CFA/I is a rigid rod fimbrial structure made up of a single antigenic fimbrial type. CS3

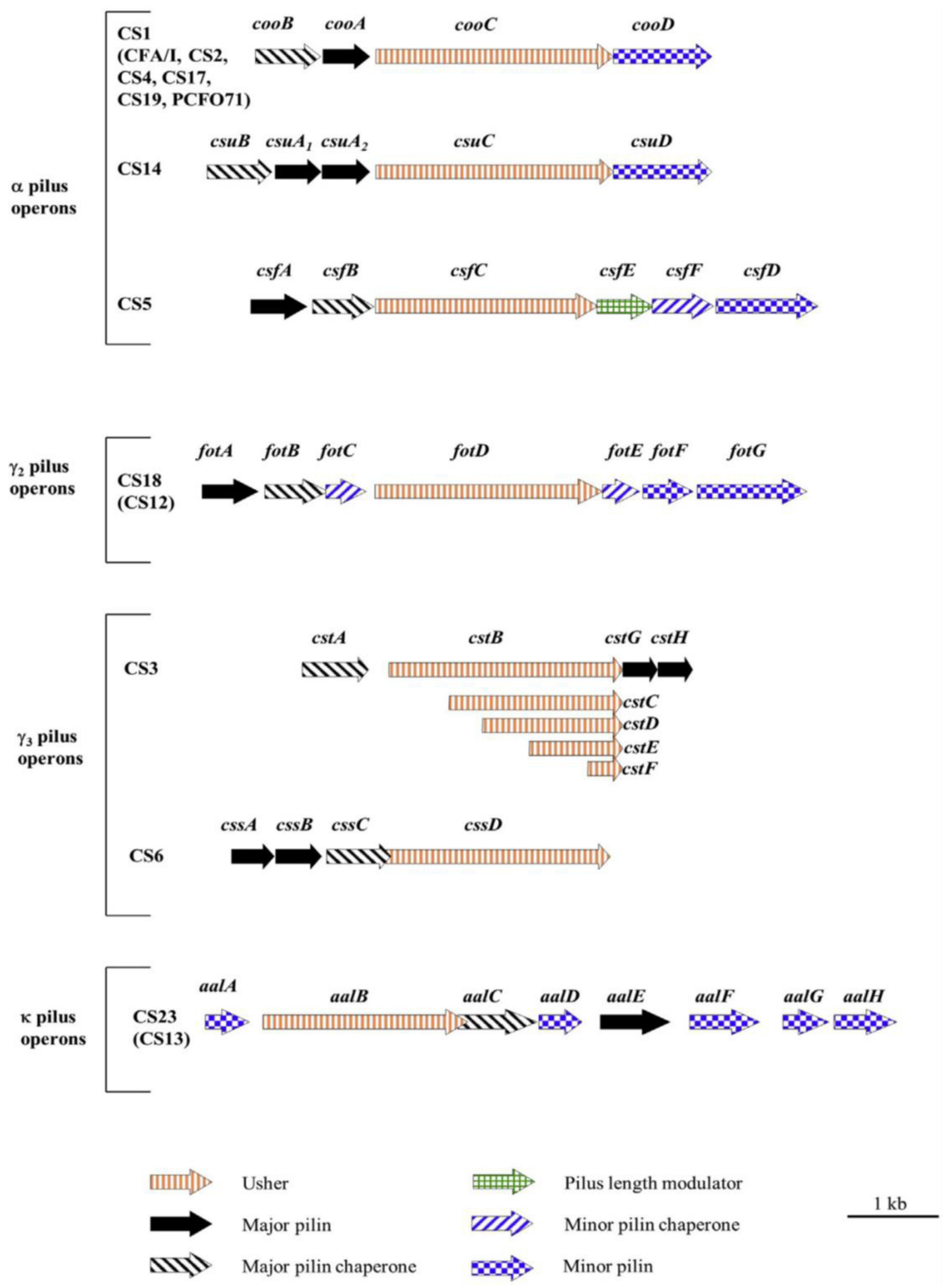

Figure 2.10: Genetic organization of the operons encoding ETEC CFs. Source: Madhavan *et al., Advances in Applied Microbiology*. 2015.

can be found alone or in combination with CS1 or CS2 in CFA/II group. CS6 can be found alone or in combination with CS4 and CS5 in CFA/IV group (Gaastra *et al.,* 1996; Torres *et al.,* 2005). Other less prevalent CFs include CS7, CFA/III (CS8), CS10, CS11, CS12, CS13, CS14, CS15, CS17, CS18, CS19, CS20, CS21, CS22 and CS23. Besides mostly occurring fimbrial CFs some non-fimbrial CFs like CS6, CS10 and CS15 had been reported (Evans *et al.,* 1978; Gaastra *et al.,*, 1996). Nearly 30-50% of ETEC isolates do not have detectable CFs out of the most commonly occurring CFA/1, CS1, CS2, CS3, CS4, CS5, CS6, CS7, CS14, CS17, and CS21, which can be attributed to the lack of CFs, the loss of CF properties on strain culturing, the existence of other adhesins, or the lack of particular techniques for their detection.

2.5.3.3. Non-Classical Virulence Factors of ETEC:

Other pathogenicity-related factors known as "non-classical" virulence factors had been recognised and these include *tibA, tia, etpA, eatA, leo,* etc. (Del Canto *et al.,* 2011).

TibA: ETEC is regarded as non-invasive pathogen however invasion of cultured ileocecal and colonic epithelial cells by ETEC strain H10407 had been reported but does not penetrate deeper tissues via the intestinal epithelium. After invasion, it is unable to propagate within the cell and is restricted to endocytic vacuoles. Two chromosomally encoded genes, *tia* and *tibA*, working independently were found to be responsible for this invasive properties (Elsinghorst *et al.,* 1992). The synthesis of TibA is under *tib* locus. TibA is a glycosylated adhesion factor with a 104 kDA molecular weight that stimulates invasion, bacterial aggregation, and biofilm formation. Only glycosylated TibA allows bacteria to bind to and infiltrate epithelial cells via a specific receptor (Elsinghorst *et al.,* 1992; Elsinghorst *et al.,* 1994; Lindenthal *et al.,* 1999; Sherlock *et al.,* 2005). TibA acts as an adhesion and also as an invasion which suggests the role of *tibA* in ETEC pathogenesis (Lindenthal *et al.,* 2001).

Tia: *tia* is one of the gene responsible for the invasive nature of non-invasive ETEC in intestinal epithelial cell lines. It is chromosomally encoded adhesin on a pathogenicity island of 46-kb which has a lower GC concentration than the rest of the genome (Elsinghorst *et al.,* 1992; Fleckenstein *et al.,* 2000). Tia which is a 25 kDa protein is located to the outer membrane and functions as adhesion as well as invasion and binds a specific receptor on HCT8 cells (Fleckenstein *et al.,* 1996; Mammarappallil *et al.,* 2000).

LeoA: LeoA was discovered during the study of *tia,* where the deletion of one or more elements encoded on the pathogenicity island resulted in significant reduction in LT enterotoxin secretion and in rabbit ileal loop lacked fluid accumulation. *leoA* is responsible for maximal release of LT toxin from the periplasm (Fleckenstein *et al.,* 2000). *leoA* is a GTPase, which is required for maximum LT secretion. The GTP binding domain of *leoA* is also involved in the formation of outer-membrane vesicles (OMVs). The motility of the H10407 ETEC strain is also affected by *leoA* (Brown *et al.,* 2007).

EtpA**:** EtpA was identified as a 170kDa glycosylated exoprotein consisting of three genes, *etpBAC* involving two-partner secretion (TPS) locus on the pCS1 virulence plasmid of ETEC H10407. *etpA* mediates bacterial adherence and is required for optimal colonization of the intestine (Fleckenstein *et al.,* 2006; Roy *et al.,* 2008). *etpA* mimics and binds with highly conserved areas of flagellin, the main subunit of flagella, and these interactions affect flagella-host cell attachment, resulting in adherence and colonization of the intestine (Roy *et al.,* 2009).

EatA: EatA (ETEC autotransporter A) is a 147.7 kDa serine protease autotransporter residing on pCS1 plasmid of ETEC and secreted as 110-kDa protein, passenger protein fragment of EatA (Patel *et al.,* 2004). EatA regulates bacterial adherence by degrading the adhesin EtpA and caused enhanced delivery of heat labile toxin, LT and thereby play a role in ETEC pathogenesis (Patel *et al.,* 2004; Roy *et al.,* 2011). EatA also degrade MUC2, the main component found in the mucosal surface of the small intestine and as a result, by facilitating bacterial access to cell surface receptors, toxin delivery can be increased (Kumar *et al.,* 2014).

YghJ: YghJ is an ETEC metalloprotease released by the type II secretion system (T2SS). It influences intestinal colonization by deteriorating the primary mucins in the small intestine, MUC2 and MUC3, which provide efficient access to small intestinal enterocytes (Luo *et al.,* 2014). YghJ is chromosomally encoded and is conserved among the commensals and the pathogenic variants. It causes tissue hemorrhage and significant fluid accumulation in mouse ileum (Tapader *et al.,* 2017).

2.6. Traveller's diarrhoea:

ETEC as a causative agent of traveller diarrhoea was first reported in 1975 involving five travellers to Mexico, who developed acute watery diarrhoea with ST-only ETEC

(Sack, D. *et al.,* 1975). Study on travellers to Mexico established ETEC as the most common causative pathogen for travellers' diarrhoea with strains expressing either on or both of the enterotoxins (Merson *et al.,* 1976; Morris *et al.,* 1976). ETEC, producing heat-labile and/or heat-stable enterotoxin, as major cause of travellers' diarrhoea was confirmed in the study involving Peace Corps volunteers (PCVs) travelling to Kenya and indicated the presence of ETEC in widely separated geographical areas (Sack, D. *et al.,* 1977).

ETEC emerged as the main causative agent of diarrhoea among USA travellers to developing countries (Gangarosa, 1978). Study from 2010 to 2016 showed that the most common etiological agent isolated from patients with travellers' diarrhoea was ETEC (***Figure 2.11***). ETEC emerge as the dominant pathogen in 42% cases from Latin America and 28% cases in Southeast Asia. The study also showed ETEC expressing ST were most frequent toxin type in Mexico, Guatemala and India (Jiang *et al.,* 2017). Among the 252 travelers with diarrhoea to Mexico, Guatemala, and India, ST-only were most predominant in 58% isolates, 28% produced both LT and ST toxins and 15% were LT-only. CS21 was the most common CF (65%), preceded by CS6 (25%), and CS3 (17%). CS21 was found in 64% of international travellers and 46% of Latin American isolates. CS21 was expressed in 96% of the isolates from India. In comparison to LT-only ETEC, CS21 was discovered more in ST-only and ST/LT ETEC. (Kharat *et al.,* 2017). In the review analysing travellers' diarrhoea from 1973 to 2009, the most common group of agents causing travellers' diarrhoea was identified to be bacterial enteropathogens, with ETEC seems to be the most common causative organism (Shah *et al.,* 2009). The most common toxin type among 52 ETEC isolated from travellers was STh, and the most common colonization factors (CFs) were CS21, CS6, and CS3. EatA and EAST1among the non-classical virulence factors, on the other hand, were frequently found (Rivera *et al.,* 2013). The phenotypes of ETEC strains in travellers differ by country; for example, LT-only ETEC was frequently detected from visitors to Jamaica in 58 % of cases, while LT/ST ETEC was isolated from visitors to India in 45 % of cases. In this analysis, 57 % of the 275 ETEC isolates developed a definite colonization factor (CF). The most common CF was CS6, which was present in 41 % to 52 % of ETEC isolates. ST-only ETEC, on the other hand, was isolated in 51% of visitors to Kenya. Thus, strains circulating in a given nation that primarily infect children and contaminate water and food supplies can be used to determine the type of ETEC afflicting travellers. In Latin

America, ETEC was responsible for 42 percent of diarrhoeal episodes, 36 percent in Africa, and 16 percent in Asia. According to available evidence, ETEC can be the source of 20 to 40% of traveller diarrhoea incidents. Thus, in North Americans and Europeans visiting developing countries, ETEC appears to be the most common cause of traveler's diarrhoea. ETEC is endemic all year, but the rate of occurrence is greatest in the summer, implying that travellers during warm season are more susceptible to ETEC-mediated diarrhoeal illnesses (Qadri *et al.,* 2005).

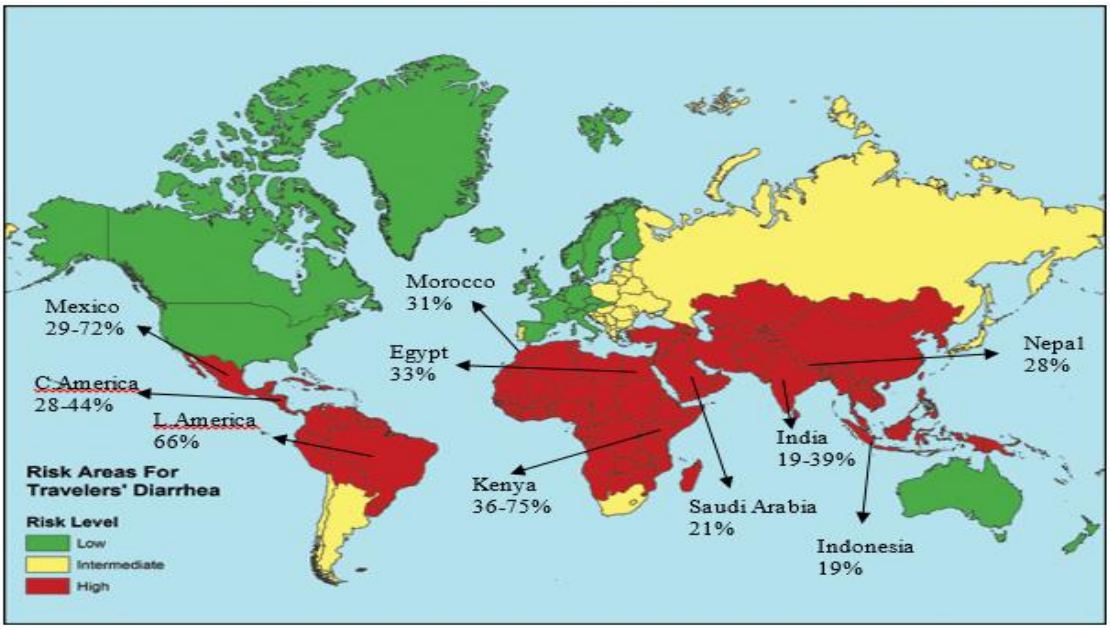

Figure 2.11: Areas of risk for traveler's diarrhoea caused by enterotoxigenic *E. coli*. The world is divided into three risk levels (low, intermediate, and high) of traveler's diarrhoea for travelers from industrialized countries. The percentage occurrence of CS6-expressing ETEC indicated by the arrows. **Source**: Torres *et al., Vaccines for biodefense and emerging and neglected diseases.* 2009.

2.7. Sources of ETEC infection:

ETEC diarrhoea, like other diarrhoeal disorders, can be caused by ingesting contaminated food or water. ETEC is typically a primary cause of diarrhoeal disease in areas with poor drinking water and sanitation (Black *et al.,* 1981). These species have been found in surface waters in developing countries, transmitting the infection to local residents as well as international visitors. ETEC transmission through processed foods has been thoroughly documented. Sack *et al.* discovered in 1977 that 8% of ETEC that produced one or both LT and ST were found in isolates from animal-sourced food in the United States. Episodes of diarrhoea caused by food-borne ETEC were recorded in Sweden and Brazil. *E. coli* infection can also be caused by contaminated weaning food

(Qadri *et al.,* 2005). On multiple occasions, ETEC transmission has been detected on cruise ships (O'Mahony *et al.,* 1986).

2.8. Clinical symptoms of ETEC infection:

In Calcutta, ETEC-caused diarrhoea was initially diagnosed as a cholera-like illness in both adults and children. Since then, numerous studies have shown that ETEC-induced diarrhoea can range in severity from mild to severe. The diarrhoea caused by ETEC is secretory in nature. It starts with a rapid onset of profuse "rice-water" diarrhoea that leads to collapse, small bowel secretory abnormalities, and signs of moderate to severe dehydration. The clinical signs, which include the formation of rice-water stool (without blood or inflammatory cells) and recurring vomiting, extreme dehydration, and shock, are indistinguishable from those of cholera. Furthermore, the physiological anomalies in fluid and electrolyte flow in the small bowel were identical to those reported in cholera, albeit to a lesser extent. The only evident difference was the brevity of the diarrhoea; it usually stopped within 24-30 hours of admission (without antibiotic medication) and was thus substantially less than that found in cholera patients. Intubation investigations revealed a substantial population of *E coli* (10^7-10^9 cfu/mL) in the proximal small bowel during the acute phase of diarrhoea which disappeared from the small bowel during convalescence and these organisms were generally of only one or two serotypes (Sack, R. *et al.,* 1971; 1975). A dry mouth, quick pulse, drowsiness, decreased skin turgor, lowered blood pressure, muscle cramps, and, in the most severe cases, shock ensues from the loss of fluids. Dehydration is classified as mild, moderate, or severe. The disease usually appears 1-3 days after first coming into touch with the pathogen and lasts 3-4 days and it is self-contained; if hydration is maintained, the patients will live with no long-term consequences (Qadri *et al.,* 2005).

2.9. Diagnosis and detection of ETEC:

Because ETEC detection is more difficult than *V. Cholera, Shigella spp., or* rotavirus detection, the actual number of infantile diarrhoeal cases caused by ETEC is frequently deceiving. Because ETEC can only be identified by the enterotoxins it creates, diagnosis requires recognising either LT or ST (Qadri *et al.,* 2005). Over the last few decades, direct identification of ETEC enterotoxins has progressed. Before additional simpler assays could be identified, gold standards for ETEC detection were the physiologic methods, for LT, the rabbit ligated ileal loop model (De *et al.,* 1956); and for ST, the neonatal suckling mouse assay (Dean *et al.,* 1972). Different assays that were employed initially with the use of adrenal cells in tissue culture for detection of enterotoxins were detection of heat-labile *Escherichia coli* enterotoxin (Donta *et al.,* 1974); Heat-labile enterotoxin (LT) activates adenylate cyclase, and morphological changes in cultured Chinese hamster ovary (CHO) cells when exposed to LT enterotoxin (Guerrant *et al.,* 1974). After that GM1 ganglioside enzyme-linked immunosorbent assay (GM1-ELISA) was used for detection of LT (Bäck *et al.,* 1979) and ST (Klipstein *et al.,* 1984). To identify and define ETEC, serotyping (the determination of 'O' serogroups associated with the LPS and H serogroups of the flagella) was utilized (Ørskov *et al.,* 1976). However, it was later discovered that clinical ETEC isolates could belong to a wide range of serotypes, with the serotype profile changing over time (Stoll *et al.,* 1983). Radial immune hemolysis test was used for the detection of LT (Yano *et al.,* 1982). Radiolabeled DNA probes encoding LT and ST was employed for detection of the enterotoxins (Moseley *et al.,* 1980). After that PCR is used for rapid and direct detection of both the LT enterotoxin (Olive, 1989). ELISA was later modified and different ELISA methods are used for the detection of the enterotoxins and CFs (Sjöling *et al.,* 2007). Because LT was highly immunogenic but ST was not, diagnostic tests shifted to LT first. The structural genes for toxins and CFs can be detected using more recently developed DNA probe technologies and PCR, providing them with the advantage of detecting ETEC from samples that have been kept for a long period and may have undergone phenotypic changes. Furthermore, because ETEC CFs are not exposed on bacterial surfaces due to mutations in genes required for surface expression, molecular rather than phenotypic techniques are required in some circumstances to detect them.

2.10. ETEC virulence plasmids:

A plasmid is an extrachromosomal DNA that can replicate without the help of chromosomal DNA. Plasmids are mobile elements that provide host benefits such as antibiotic and heavy metal resistance, toxin and other virulence factors production, hydrocarbon biotransformations, and symbiotic nitrogen fixation. For full pathogenesis, several enteropathogenic bacteria, such as *Shigella, Yersinia, Salmonella*, and *E. coli*, require plasmid-encoded genes (Lim *et al.,* 2010). Despite the large variety of plasmid types seen in *E. coli* strains, plasmids that encode virulence-related traits almost exclusively belong to the IncF incompatibility family (Kaper *et al.,* 2004). Through the acquisition of pathotype-specific properties, the vast majority of these *E. coli* virulence plasmids evolved from a single plasmid backbone type. Some of these pathotype-specific virulence plasmids have a lot of variation, whereas others have a lot of conservation.

ETEC strains have a polyphyletic genetic origin, which means that their genetic history is not preserved. ETEC strains appear to have a plasmid that is more crucial than their chromosomes. These plasmids contain ETEC colonization factors (CFs), toxins, and other virulence factors. It's probable that more ETEC virulence genes will be discovered as more ETEC genomes are completed. The diversity shown here suggests that attempts to find novel virulence determinants and vaccine candidates should concentrate on sequencing the virulence-associated plasmids of a diverse range of isolates (Turner *et al.,* 2006). ETEC strains adhere to the intestinal mucosa using CFs that have a high affinity for the human intestinal epithelia. Human ETEC CFs are normally encoded by a polycistronic operon that comprises the fimbrial subunit genes, chaperones, and ushers and can be plasmid or chromosomal encoded.

On the other hand, the majority of human ETEC CFs are plasmid transcribed and appeared to have been gained horizontally via flanking ISs and transposons. These CFs have gone through a lot of evolutionary changes, which has resulted in a lot of genetic variants. In total, there are more than 27 genetically diverse human ETEC CFs, several of which also have distinct serological characteristics, suggesting that these CFs have the ability to travel and develop quickly (Gaastra *et al.,* 1996). The human ETEC CF-encoding plasmid pCoo was the first to be sequenced in its entirety (Froehlich *et al.,* 2005). To isolate pCoo, researchers employed the human ETEC strain C921b-1, which is known to express CS1 and CS3 (Perez-Casal *et al.,* 1990). In nature, the plasmid was discovered to be cointegrate, with sections that resembled RepI1 plasmid R64 from *Salmonella enterica serovar Typhimurium* and RepFIIA plasmid R100 from *Shigella spp* (***Figure 2.12***) (Johnson *et al.,* 2009). IS100-associated direct repeats separate the composite parts of pCoo, showing that the cointegrate pCoo was formed by recombination between sections of two distinct plasmids (Froehlich *et al.,* 2005).

The plasmid's most remarkable feature was the polycistronic coo operon, which contains four genes encoding the CS1 pilus. Furthermore, EatA, a serine protease autotransporter implicated in ETEC pathogenicity, is connected to pCoo (Patel *et al.,* 2004). The RepI1-like component of pCoo contains the coo genes, while the RepFIIA-like portion contains the *eatA* genes. Clinical CS1 isolates were discovered to all have both cointegrate sections of pCoo, implying that this cointegrate plasmid is stable.

Furthermore, because the two plasmids are flanked by distinct IS elements, they differ slightly inside the CS1-encoding region. As a result, before the integration of RepFIIA backbone components and *eatA* in pCoo, the CS1 operon appears to have been introduced into an ancestral RepI1 plasmid (Froehlich *et al.,* 2005). CFA/I positive ETEC strain H10407 isolated from a patient in Bangladesh, like the majority of ETEC strains, carries several plasmids including several virulence genes. pH10407 95 (95-kb plasmid) encoding CFA/I (Evans *et al.,* 1979), EtpABC (Fleckenstein *et al.,* 2006), and EatA (Patel *et al.,* 2004), and pH10407 66 (66-kb plasmid) encoding a conjugative transfer system and LT are among the plasmids in question.

Although the CFA/I-encoding plasmid pH10407 95 lacks RepI1-related regions, it does have a RepFIIA-like replicon and backbone. The genes for CFA/I and *eatA*, a two-

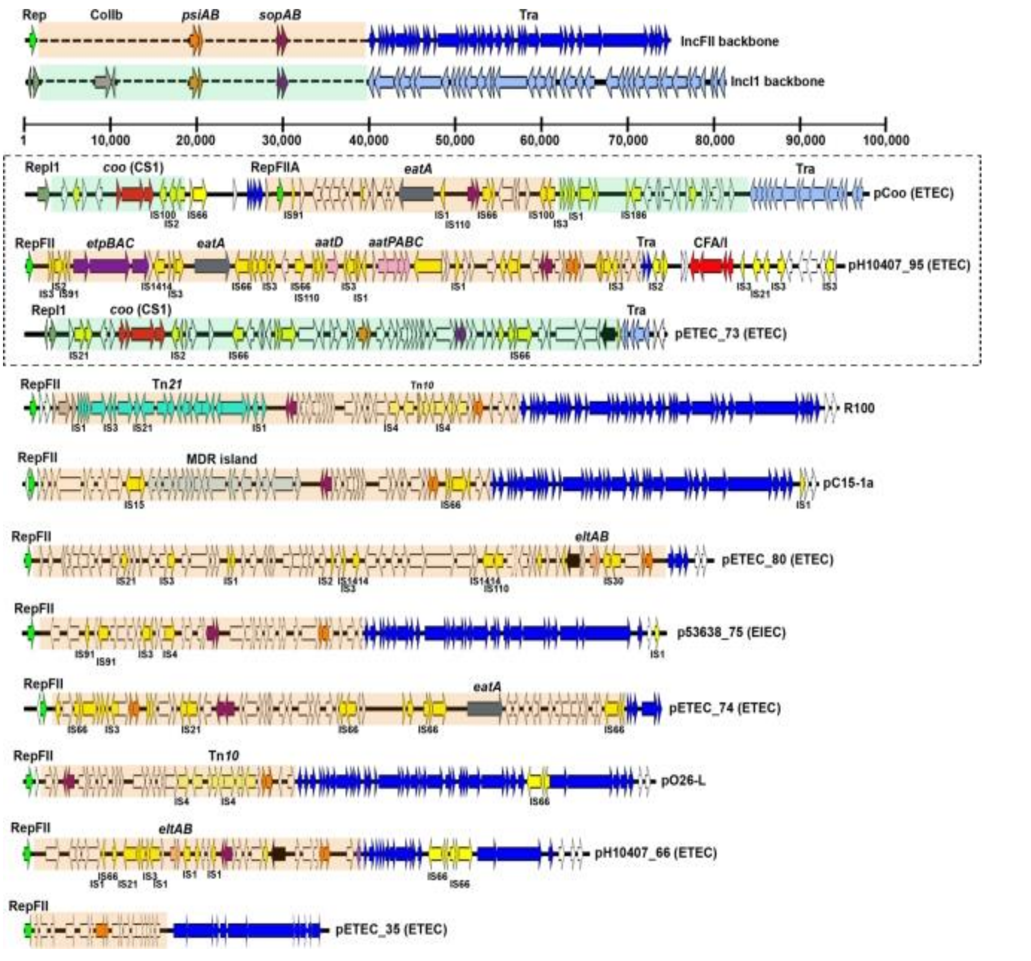

Figure 2.12: *E. coli* virulence plasmids RepFIIA and RepI1 are compared linearly. Basic RepFIIA and RepI1 backbones are highlighted in orange (RepFIIA) and green (RepI1) on the top two maps (RepI1). To demonstrate genetic burden regions, these locations are highlighted across the picture. The scale is based on pairs of numbers. Maps are produced to scale and begin with either RepFIIA or RepI1. ISs are labelled and coloured below each map, as are virulence genes of interest. ETEC plasmids that encode either CS1 or CFA/I are shown by the boxed region. XPlasMap was used to create all of the maps. **Source:** Johnson *et al., Microbiology and Molecular Biology Reviews*. 2009.

partner secretion locus involved in attachment to intestinal epithelia, are found on this plasmid. CFA/ I's existence on pH10407 95, a RepFIIA plasmid, implies that the ETEC CS operons were acquired multiple times on various plasmids. Like pCoo, pH10407 95 has a shorter F transfer region. In fact, CF-encoding plasmids' transfer regions are faulty, making them rely on other plasmids for mobilisation. In addition to the core components and virulence factors mentioned above, many sequenced ETEC plasmids and other sequenced *E. coli* virulence plasmids contain a group II intron-encoded reverse transcriptase/maturase in close proximity to fimbrial operons or other horizontally acquired genetic regions. This region is found in pETEC 73, pETEC 74, pETEC 80, pCoo, pB171, pUTI89, p55989, and the K88_ and K99_ plasmids. E.c.I4 appears to have been absorbed into IS629 and IS911, allowing it to disseminate and transfer between *E. coli* strains. (Johnson *et al.,* 2009).

Only a few human ETEC plasmids have been sequenced, despite the genetic variety of these plasmids in terms of CF types. As a result, we still don't know everything there is to know about the evolution of these plasmids and how it affects virulence. Future comparative genomic research will help us better understand this dynamic and fascinating group of plasmids with implications for human and animal health.

2.11. ETEC and Environmental and Host factors:

When bacteria invade and colonise the gastrointestinal tract, they come into contact with a range of host factors as well as a distinct environment in the intestine, such as bile and sodium bicarbonate, which have been demonstrated to control virulence gene expression in other enteropathogens (Gupta *et al.,* 1997; Abe *et al.,* 2002; Hung *et al.,* 2005; Abuaita *et al.,* 2009). Bile possesses antibacterial characteristics, works as an emulsifier of dietary fats and lipids *in vivo*, and has detergenic qualities (Hofmann, 1999; Begley *et al.,* 2005).

Therefore, optimum expression of colonization factors by the bacteria in the gut environment is important for successful colonization. Stomach pH also plays an important role in colonization. Almost all species require iron as a nutrient. Due to insolubility at physiological pH and sequestration by iron-binding proteins, iron availability inside the human body is limited. Low-iron circumstances are also used by bacteria to sense the gut environment and induce the production of virulence genes (Carpenter *et al.,* 2009).

Iron affects several ETEC virulence variables through the control of the iron-sulfur cluster regulator, IscR. CFA/I is a CF that is controlled by low-iron conditions (Haines *et al.,* 2015). An earlier study has shown specific CFA/I binding is supported by a brush border-associated glycoprotein and that its expression develops as a function of intestinal cell differentiation and is regulated by glucose (Bernet-Camard *et al.,* 1997).

Heat labile toxins (LT) are significantly affected by temperature. LT in a native system is optimally expressed at 37°C but it is modulated when temperature is decreased to 18°C (Trachman *et al.,* 1998). Expression and maturation of curli fimbriae in ETEC was found to be regulated by temperature. Curli was found to be expressed preferentially at low temperatures 30°C than 37°C (Szabo *et al.,* 2005).

The expression of 987P CF in porcine ETEC is modulated by a range of environmental factors. The cAMP-CRP complex, as well as H-NS, control its expression in response to nutritional status, pH, and temperature of the growth medium, according to researchers (Edwards *et al.,* 1997). Water and electrolytes are lost into the intestinal lumen as a result of ETEC infections. Because oral rehydration therapy prevents this, glucose and salt solutions may have an influence on ETEC gene regulatory mechanisms in a CRP and H-NS dependent manner. Salt induces enterotoxin expression in an H-NS dependent way. Furthermore, glucose activates or suppresses toxin gene expression depending on the toxin gene (Haycocks *et al.,* 2015). In addition, there are small proteins in *E. coli* that negatively regulate a large number of unrelated genes.

2.12. Epidemiology of ETEC:

According to WHO and CDC, ETEC causes diarrhoea in developing and developed countries like Argentina, Bangladesh, Brazil, Egypt, India, Japan, Peru, Saudi Arabia, Africa, Spain, Thailand, Latin America, Southeast Asia and United States (***Figure 2.13***). Research on 59 Apache children hospitalised with 64 episodes of severe diarrhoea was the first to link ETEC to children (Sack, R. *et al.,* 1975) and over the last few years, several studies have identified ETEC as a common cause of diarrhoea in infants younger than 5 years of age.

In a study in Bangladesh, ETEC was found in 14 % (n = 662) of the total diarrhoeal cases, with over 70 % of the strains isolated from children below the age of 5 years, with 93 % of those in the 0- to 3-year-old age group. ETEC infection was reported to be the most frequent cause of diarrhoea in children aged 0 to 2 years old in Bangladesh, accounting for 18% of all diarrhoeal cases. Infants and young children have also been found to be more vulnerable in areas with inadequate public health and hygiene (Qadri *et al.,* 2000). According to epidemiological data from India and Bangladesh, ETEC is a cause of adult diarrhoea that is similar to cholera in severity (Qadri *et al.,* 2000). People aged 60 years and above are also affected by ETEC which is detected in a study over the period 1996-2001 in patients admitted in Dhaka Hospital of ICDDR, B (Faruque *et al.,* 2004).

In a group of young children in Guinea-Bissau, ETEC infection was tested in children aged 0 to 2 years, and the rate of primary infections had risen considerably after 3 or 6 months, depending on the type of ETEC that caused the illness (Steinsland *et al.,* 2002). ETEC diarrhoea was found to be 1.5 episodes per child per year in Egyptian infants, accounting for 66% of all first episodes of diarrhoea after birth. The number of episodes per child per year rose from 1.7 in the first 6 months of life to 2.3 in the second 6 months and then decreased (Rao *et al.,* 2003). A 1-Year Prospective Study in a Swedish Clinic for Infectious Diseases conducted to identify enteropathogens in adults ETEC was found in 8% of patients (Svenungsson *et al.,* 2000). The aetiology of acute diarrhoea in 265

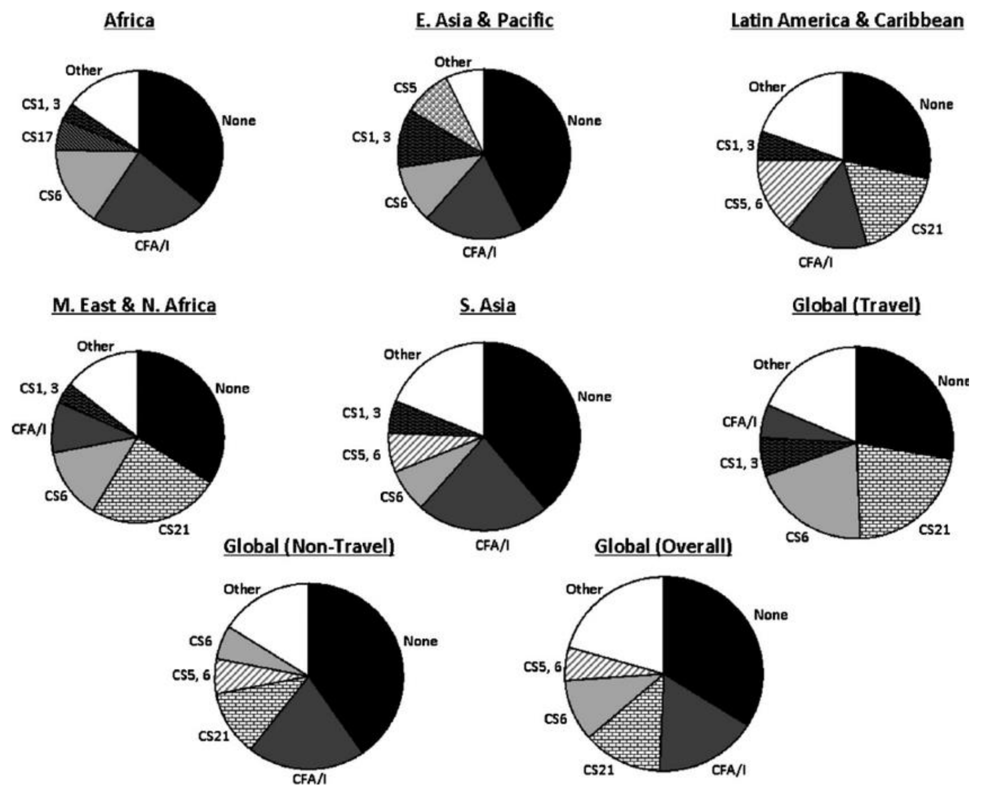

Figure 2.7: Prevalence of ETEC colonization factors by area and travel/non-travel population. In all of the studies that were included, regional estimates were used to describe the distribution of ETEC colonization factors. Estimates from all included studies (i.e., overall) are grouped into global estimates, which are further stratified by travel and non-travel populations. **Source**: Isidean *et al.,Vaccine*. 2011.

Jordanian children under the age of five years was investigated and detected 5.7 % ETEC (Youssef *et al.,* 2000). A study conducted among 489 hospitalized patients in Bali showed the presence of ETEC in 14.9% of the patients and children between 1 and 15 years showed maximum isolation rate (Subekti *et al.,* 2003). ETEC infection is more common in those over the age of 15, and about 25% of ETEC disease occurs in adults.

In a study among Egyptian children under 3 years of age, 933 (27%) ETEC isolates were detected in which 78% was alone and 22% were associated with one or more copathogens (Shaheen *et al.,* 2004). ETEC is a significant diarrhoeal pathogen among children in developing countries. Until the age of five years, an infant born in a developing country is likely to have 0.5 diarrhoeal episodes per year caused by ETEC, after which the annual incidence decreases to 0.1. ETEC caused 280 million episodes of diarrhoea in children under the age of five years per year, and nearly 50 million children in this age group were asymptomatic ETEC carriers.

ETEC was responsible for one out of every seven diarrhoeal episodes in children under the age of one year, and almost a quarter of all diarrhoeal cases in children aged 1-4 years (Wennerås *et al.,* 2004). In developing countries, the frequency of ETEC infections declines between the ages of 5 and 15 years and rises again over 15 years of age, with adults accounting for around 25% of ETEC cases. Patients in this age group (>65 years) also had more extreme dehydration than infants. According to epidemiological data from India and Bangladesh, ETEC is a source of adult diarrhoea that is similar to cholera in severity. It's possible that environmental and immunological factors play a role in declining rates of ETEC infections after childhood and rise in adulthood.

Several studies have shown that ETEC diarrhoea and asymptomatic infections are more common during the summer, implying that visitors to these endemic regions are more likely to contract ETEC infections during the summer (Qadri *et al.,* 2005). ETEC was positive for 18% of stool specimens isolated during the 2004 flood in Dhaka (Qadri *et al.,* 2005). ETEC was the most prevalent pathogen in a cohort of 321 children in Bangladesh who were monitored from birth to two years of age. It was identified in 19.5 % of cases, with an incidence of 0.5 episodes/child/year (Qadri *et al.,* 2007).

ETEC, 32.3 % and 37 %, was the most commonly isolated enteric pathogen from diarrhoeagenic children and from children in the control group respectively in Tunis, Tunisia, among 271 stool specimens that were collected from children (diarrhoeagenic, n = 115 and control, n = 54) and adults (diarrhoeagenic, n = 73 and control, n = 29).

ETEC (12.3 %) was the second most prevalent enteric infection after Salmonella spp. (34.2 %) in adults in the diarrhoeagenic group, whereas ETEC (31 %) was the most common enteric infection in adults in the control group (Al-Gallas *et al.,* 2007). Between September 2000 and August 2003, hospital monitoring in the Nile River Delta detected ETEC in 20.7 % (320/1540) children under the age of five who had diarrhoea and sought hospital treatment being cultured for enteric bacteria (Shaheen *et al.,* 2009). However, there is a scarcity of information on the possible contribution of ETEC to diarrhoeal incidence and mortality in older children, adolescents, and adults. In a review of published works conducted in low and middle-income countries from January 1, 1980, through December 31, 2008, among children 5 years and older, adolescents, and adults, enterotoxigenic *E. coli* (28.1%) was the most common cause of hospitalization (Walker *et al.,* 2010). Between January 2007 and December 2010, 299 out of 3,943 children with diarrhoea who visited five different Bolivian hospitals tested positive for ETEC. Children under the age of 24 months accounted for 259 (86.6%) of the ETEC-positive cases (Gonzales *et al.,* 2013).

Diarrhoeal surveillance studies among the military in Egypt between September 2005 and November 2009, a total of 102 ETEC were identified out of the total of 181 subjects. ETEC was found to be the sole pathogen in 74 (72.5%) and as a mixed infection in 28 (27.5%) of the 102 patients who tested positive for ETEC enterotoxin genes (Nada *et al.,* 2013). ETEC producing heat-stable toxin (ST-ETEC; with or without co-expression of heat-labile enterotoxin) emerged as one of the four main pathogens responsible for diarrhoea in the GEMS study, which was conducted at four sites in Africa and three in Asia for 9439 children with moderate-to-severe diarrhoea and 13,129 control children without diarrhoea. In babies aged 0–11 months, ST-ETEC was connected to an increased risk of case mortality (Kotloff *et al.,* 2013). In a comprehensive analysis of studies on diarrhoea aetiology in hospitalised children under the age of five years to estimate the global burden of diarrhoea mortality by pathogens, ETEC was identified as one of the four pathogens and one of the two bacterial pathogens responsible for more than half of all diarrhoeal fatalities in children worldwide (Lanata *et al.,* 2013).

In an Egyptian birth cohort research, ETEC was revealed to be the most common infection among babies. Over the course of the two-year trial, ETEC was isolated from 18.9% (756/4,001 samples) of diarrhoeal episodes and 13.6% (1,299/9,539 samples) of asymptomatic children's stool samples (Mansour *et al.,* 2014). In 2010, an assessment of Diarrhoeal Illness and Deaths attributable to ETEC among Older Children, Adolescents, and Adults in South Asia and Africa found 42,973 fatalities in Africa and 45,713 deaths in South East Asia. Diarrhoea was responsible for 12.5 % of fatalities in Africa and 19.85 % of fatalities in South Asia among children and adolescents aged 5 to 14 years. Diarrhoea mortality was 6.35 % in Africa and 3.7 % in South Asia among people over the age of 15 years. Outpatients who tested positive for ETEC were predicted to make up 4.6 % of the total. According to these estimates, there are 15.0 million ETEC episodes per year in Africa and 28.7 million ETEC episodes per year in South Asia among people over the age of 5 years. These figures highlight the importance of ETEC as a leading cause of illness and death in older children, teenagers, and adults throughout Africa and South Asia (Lamberti *et al.,* 2014).

Surveillance from January 2007 to December 2012 in Bangladesh detected 11% ETEC diarrhoeal samples out of 15,152 diarrhoeal specimens (Begum *et al.,* 2014).

In a Guatemalan rural community, ETEC was responsible for 26% of severe diarrhoea in children requiring hospitalization and 15% of diarrhoea in the general population (Torres *et al.,* 2015). Over a 10-year period (1997–2007), ETEC was found in 205 (10.6%) of 1941 children with diarrhoea hospitalised with diarrhoea in Blantyre, Malawi. ETEC was found in 37 (7.3%) of 507 children who did not have diarrhoea (Trainor *et al.,* 2016).

Between 2005 and 2009, 1067 ETEC were found out of 8580 stool specimens from diarrhoeal patients at the ICDDR, B Dhaka hospital that were examined for ETEC (Begum *et al.,* 2016). Between 2009 and 2014 in Shenzhen, China, 168 (1.3 %) ETEC strains were identified from 13,324 diarrhoeal outpatients out of which the majority of ETEC-infected patients (82.1%) were between the ages of 20 and 59. Only six of the ETEC-infected patients were youngsters under the age of five years which is a contrasting result with respect to ETEC infection (Li *et al.,* 2017).

2.12.1. Distribution of Virulence factors:

Enterotoxins:

According to a survey that examined ETEC in developing countries, the characteristics of the toxin types and CFs observed on ETEC strains recovered from young children varied among countries where ETEC is endemic. In epidemiological studies, CFs were detected on less than 10% of LT-producing ETEC strains, compared to over 60% of ST- and LT/ST-expressing ETEC strains. The phenotype of the poison did not vary with age. Both the LT and ST phenotypes of ETEC have been linked to diarrhoea in babies in longitudinal investigations. This has been demonstrated in hospital and community-based investigations as well. In hospital-based research, however, ETEC producing both LT and ST or ST alone was relatively more responsible for disease severity (Qadri *et al.,* 2005). A total of 17,205 ETEC isolates were extracted from the 136 investigations that were included in the study. Approximately half of the research (49%) focused on endemic communities, while the remaining 17% focused solely on travel populations. Worldwide 60% of isolates exhibited LT, either alone (27%) or in conjunction with ST (60%) (Isidean *et al.,* 2011). Colonization characteristics in ETEC isolates from children with moderate-to-severe diarrhoea and matched controls were studied in the Global Enteric Multicenter Study (GEMS). Overall, 68.2% of patient isolates (N = 550) were ST-only (N = 291, 36.1%) or LT/ ST (N = 259, 32.1%), genotypes that were significantly associated to MSD in GEMS. In 284 (35.2%) of the 806 cases, STh-only strains were discovered. The remaining 31.8% (N = 256) of case isolates were LT-only (Vidal *et al.,* 2019).

A 2-year study in Bangladesh from September 1996 to August 1998 detected heat-stable toxin (ST) in 49.4 % of the total ETEC isolates, heat-labile toxin (LT) in 25.4 % of the ETEC isolates, and both LT and ST enterotoxins in 25.2 % ETEC isolates. In the hot summer months of May to September, the rate of ETEC isolation rose, then declined in the winter. Among children infected with ETEC as the sole pathogen, there was a trend of significantly more severe disease in children infected with ST-positive or LT-and ST-positive ETEC isolates compared to the severity of disease in children infected with LT only-positive ETEC isolates (Qadri *et al.,* 2000). During a flood in Bangladesh, the toxin profile of ETEC strains revealed that ST-only expressing strains were the most common (67%) followed by strains producing both ST and LT (19%) and LT-only strains (14%) (Qadri *et al.,* 2005). In Bangladesh's urban population, 49% of ETEC strains isolated from diarrhoeal stools expressed ST, 30% expressed LT/ST, and 21% expressed LT exclusively, whereas 46% of ETEC strains isolated from asymptomatic diseases expressed ST, 35% expressed LT, and 19% expressed both LT and ST (Qadri *et al.,* 2007). ETEC isolated from diarrhoeal patients in Bangladesh produced 43% LT/ST toxin, 27% LT, and 30% ST toxin. STp was isolated in 5% of patients among the ST variations (Begum *et al.,* 2014). In this study in Bangladesh, 448 (42%) of the strains produced both LT and ST, with 315 (30%) producing only LT and 304 producing only ST (28%) (Begum *et al.,* 2016).

Among Jakarta residents 177 (72 %) of the 246 ETEC isolates produced ST, 56 (23%) produced LT, and 13 (5 %) produced both ST and LT toxins (Oyofo *et al.,* 2001). ST was identified in 51 (69.9%) of the ETEC strains reported in Bali, Indonesia, whereas LT and ST/LT were discovered in 28.8% and 1.3% of the strains, respectively. Colonization factor Antigens were found in 28.8% of the ETEC strains (Subekti *et al.,* 2003). Heat-stable toxin (ST) was found most frequently (81.5%) in clinical strains from Shenzhen, China, followed by heat-labile toxin (LT) (13.1%) (Li *et al.,* 2017). PCR detection of the LT, ST, and LT+ST toxin genes was found in 77 (49.0%), 49 (31.2%), and 31 (19.7%) of the 157 ETEC cases, respectively, among Israeli recruits working in military field situations (Cohen *et al.,* 2010).

Among 1200 strains representing 1018 ETEC infections in a Guinea-Bissau cohort of young children: 89 with STp-ETEC, 97 with STp/LT ETEC, 168 with STh-ETEC, 89 with STh/LT-ETEC, 565 with LT-ETEC, 5 with STp/STh-ETEC, and 5 with STp/STh/LT-ETEC were found. STp/LT-ETEC (47%) and LT-ETEC (58%) had a larger proportion of CF-negative ETEC than STp-ETEC (16%), STh-ETEC (8%), and STh/LT-ETEC (10%) (Steinsland *et al.,* 2002). In a research in rural Egypt, 61% of the strains only expressed heat-stable enterotoxin (ST), 26% expressed heat-labile enterotoxin (LT) alone, and 12% only expressed both toxins (Shaheen *et al.,* 2004). In 151/320 (47%) of the samples, ETEC was recovered from Egyptian infants with a known CF. LT only (42%; n = 134), ST only (39%; n = 125), or LT/ST (12%; n = 39) were found in ETEC isolates from diarrhoea cases (Shaheen *et al.,* 2009). ST (69.6%; n = 71) was the most common enterotoxin discovered in ETEC isolates from US military soldiers engaging in Operation Bright Star in Egypt from 2005 to 2009, followed by LT (15.7%; n = 16); LTSTh (7.8%; n = 8) and LTSTp (0.9%; n = 1) (Nada *et al.,* 2013).

According to a study comparing ETEC isolates from a birth cohort of children in rural Egypt and asymptomatic children, the distribution of ETEC enterotoxins differed between symptomatic children (44.2 % LT, 38.5% ST, and 17.3 % LT/ST) and asymptomatic children (55.5 % LT, 34.6 % ST, and 9.9 % LT/ST) (Mansour *et al.,* 2014). The morphological and genotypic features of ETEC isolates have never been studied in previous studies of acute diarrhoea in Tunisia. All *E. coli* isolates were screened for ETEC virulence genes *sta* and *elt* using polymerase chain reaction (PCR). Seventy-two percent of ETEC strains (47 of 65) expressed just the *sta* gene, 21.5% (14 of 65) only the *elt* gene, and 6.1% (4 of 65) both genes. The *elt* gene was found in abundance in the outbreak isolates (10 isolates out of 14) (Al-Gallas *et al.,* 2007). In Malawi, STh was the most common toxin type, followed by LT and STp in ETEC isolates from hospitalised children with diarrhoea. Infection with ETEC was most common in infants aged 6–11 months (Trainor *et al.,* 2016).

Over the course of four years, 299 ETEC isolates from Bolivian 5-year-old children with diarrhoea and 55 ETEC isolates from children without diarrhoea (controls) were identified. ETEC strains that only produced heat-labile toxin (LT) or heat-stable toxin (ST) were roughly as common as ETEC strains that produced both toxins (20%). ETEC strains expressing human ST (STh) were found more frequently in children under the age of two, while ETEC strains expressing LT plus STh (LT/STh) were found more frequently in children aged two to five. The toxin profile of the strains had no bearing on the severity of the sickness (Gonzales *et al.,* 2013). Among 40 ETEC clinical isolates from northern Colombia, South America, 21 (52.5%) tested positive for the LT gene, 13 (32.5%) for the ST gene, and 6 (15%) for both ST and LT (Guerra *et al.,* 2014). ETEC strains were detected in 5.3% and 4.3% of diarrhoeal children and healthy children respectively aged 2 to 24 months from Peru. LT was present in 52% and 72 % of ETEC isolates from children with diarrhoea and from controls respectively; 25% and 19% were ST positive respectively; and 23% and 9% were LT positive and ST positive, respectively (Rivera *et al.,* 2010). The most prevalent enterotoxin type among these ETEC identified from occurrences of diarrhoea in Chilean children was STh, which was carried by 65 strains (63.1%) (Del Canto *et al.,* 2011). In Guatemalan children, LT was the predominant enterotoxin followed by ST. LT/ST was the least present. Out of 62 samples positive for ETEC by PCR in a Swedish clinic, 12 (19%) were positive for *estA* only, 28 (45%) for *eltB* only, and 22 (35%) for both genes (Svenungsson *et al.,* 2000).

2.12.2. Classical and non-classical Virulence factors:

Despite the fact that over 27 CFs have been identified in ETEC, only 7 to 8 are more commonly isolated from diarrhoeal stools in developing countries. CFA/I and CS1 to CS6 are the most commonly detected CFs. These CFs are primarily found on ETECs that produce ST or both LT and ST (Qadri *et al.,* 2005). Globally, CFA/I-expressing strains were found in all locations (17%), as were ETEC-expressing CFA/II (9%) and CFA/IV-expressing strains (18%). Toxins and CFs differed significantly between locations and populations. There was considerable inconsistency in CF prevalence estimations between regions and populations. While CFA/I, II, and IV were found in all locations, there was some variation in these and other CFs. Co-expression of CFA/IV component antigens, for example, CS5 and CS6, was not uncommon in Latin America/Caribbean and South Asia (14% and 7%, respectively); however, co-expression of CFA/IV component antigens was not seen in the remaining regions. Both non-travel and travel groups had CFA/I, but the non-travel population had it at a higher rate (20% vs. 6%, respectively).

CS6 (non-travel, 6%; travel, 20%) and CS21 (non-travel, 6%; travel, 20%) showed similar global variability (non-travel, 11%; travel, 22%). In both the total (6%) and non-travel (6%) population estimations, co-expression of CS5 and CS6 antigens was common. The co-expression of CFA/II component antigens CS1 and CS3 was discovered solely in the trip population (6%). Importantly, CS3 alone was the most prevalent phenotype (23%) among CFA/II-expressing ETEC, whereas CS1 and CS3 were the most commonly co-expressed antigens (21%). The most common phenotypes among CFA/IV strains were CS6 alone (30%) or with CS5 (19%). A considerable percentage of ETEC isolates (34%) had no CF reported or identified. There were no discernible variations in the prevalence of CF or toxin over time (Isidean *et al.,* 2011). The Global Enteric Multicenter Study found CFA/I is encoded by 20.4%, CFA/II (i.e., CS3 alone or in combination with CS1 or CS2) is encoded by 14%, and CFA/IV (i.e., CS6 alone or with CS4 or CS5) is encoded by 31.6%. The only significant CF found in LT-only isolates was CS6-alone, which was found in 43 of 256 isolates (16.8%). Only three (1.2%) of the 256 LT-only strains encoded CFA/I or CFA/II (Vidal *et al.,* 2019).

In a 2 year study in Bangladesh from September 1996 to August 1998, approximately 56% of the samples tested positive for one or more of the 12 CFs that were tested. The colonization factor antigen CFA/IV complex's coli surface antigens CS4, CS5 and/or CS6 were the most common (31%), followed by CFA/I (23.5%) and CFA/II's surface antigens CS1, CS2, and CS3 (21%). CS7 (8%), CS14 (PCFO166) (7%), CS12 (PCFO159) (4%), CS17 (3%), and CS8 (CFA/III) (2.7%) were some of the less frequently detected CFs. The ST- or LT-positive ETEC isolates expressed the most common CFs (i.e., CFA/I, CFA/II, and CFA/IV), whereas the bacteria positive for LT alone did not (Qadri *et al.,* 2000). During the 2004 floods in Bangladesh, ETEC was as common as *Vibrio cholera*. ETEC can produce pandemic diarrhoea, just like *V. cholerae* O1. Colonization factors were found in 78% of the strains. The most prevalent phenotype was CFA/I, followed by strains producing CS4 + CS6 or CS5 + CS6, and then others (Qadri *et al.,* 2005). In Bangladesh's urban neighbourhood, about 10% of ETEC strains from diarrhoeal stools (n=242) expressed CFA/I, CS6, or CS5+CS6, while CS6 and CFA/I were common in strains recovered from healthy children (7.0% each on 528 ETEC positive strains). Only CS19-positive ETEC strains were isolated from diarrhoeal cases, while the other CFs were present in the specimens in almost equivalent numbers. In all three ETEC toxin classes, CF-positive ETEC strains were more common in diarrhoeal stools (Qadri *et al.,* 2007). One or more of the 13 evaluated CFs were expressed by 49% (n = 812) of ETEC positive bacteria recovered from diarrhoeal patients in Bangladesh between 2007 and 2012. CS5+CS6 (18%), CFA/I (14%), CS7 (12%), CS6 as well as CS17 (10%) and CS14 (9%) were the most common CF kinds. Between 2007 and 2012, over 51% (n = 845) of the CF negative strains were discovered (Begum *et al.,* 2014). In this investigation in Bangladesh, 523 (49%) of ETEC strains showed one or more CFs; the most common phenotypes during the research period were CFA/I, CS5+CS6, and CS7 (Begum *et al.,* 2016).

CFA testing of 246 ETEC isolates from Jakarta individuals revealed that 21 (8%) had CFA/I, 3 (1%) had CFA/II, 14 (6%) had CFA/IV, and 7 (3%) had PCFO159 and PCFO159 plus CS5. With 201 (82%) of the ETEC strains, no CFAs or PCFs could be found (Oyofo *et al.,* 2001). Out of 73 ETEC isolated in Bali, Indonesia, 8 (10.9%) expressed CFA/I, 3 (4.1%) expressed CS1/CS3, 1 (1.4%) expressed CS5/CS6, 6 (8.2%) expressed CS6, 1 (1.4%) expressed CS17, 1 (1.4%) expressed CS14, and 1 (1.4%) expressed CS5 (Subekti *et al.,* 2003). CFAs were identified in 68% of the isolates from Israeli recruits serving in military field circumstances, with CFAs of the CFA/II group and CS6 being the most common. The CFA/II group, which includes CS1, CS2, and CS3, was expressed by 22% of the isolates. The most common CFA among ETEC isolates was CS6, either alone or in combination with other CFAs, followed by CFAs from the CFA/II family. A CFA was identified in three of the twelve ETEC strains isolated from patients with ETEC-Shigella mixed infections (CS1–CS3, CS12, and CS20) (Cohen *et al.,* 2010). 91 ETEC strains were found to have one or more colonization factors (CFs) in clinical strains from Shenzhen, China (54.2%). CS6 (with or without other CFs) was the most often detected CF (84/91), followed by CS21 (14/91) (Li *et al.,* 2017).

CFs were detected alone or in conjunction with CS3, CS6, CS18, or CS21 in a Guinea-Bissau Cohort of Young Children. With the exception of CS12 and CS18, which appeared in STpLT- and LT-ETEC; CS21 and CFA/I, which showed in STh- and SThLT-ETEC; and CS6, which showed in STp-, STh-, and LT-ETEC 10%), each CF was linked with a specific toxin profile (Steinsland et al., 2002). CFA/I (10%), CS6 (9%), CS14 (6%), and CS1+CS3 (4%) were the most prevalent CF phenotypes found in rural Egypt. Fifty-nine percent of the strains did not express any of the 12 CFs that were included in the test panel (Shaheen *et al.,* 2004). ST CFA/I, ST CS6, ST CS14, and LT and ST CS5 plus CS6 represented 75% of the CFs expressed by ETEC isolates from Egyptian children expressing a detectable CF (Shaheen *et al.,* 2009). From 2005 to 2009, ETEC isolated from US military soldiers engaging in Operation Bright Star in Egypt revealed at least one CF in 67.6% (n = 69) of the isolates. In 46.1% of cases (n = 47), CS6 was shown to be the most common colonization factor. It was found alone (41.1%; n = 42) and in combination with other CFs, such as CS4 (0.3%; n = 3) and CS8 (1.9%; n = 2). In comparison to CS6, CFA/I, CS14, CS2, CS3, CS20, CS22, and were observed with far less frequency (varying from 0.3 % to 3.9%). STh has been identified as a toxin type that lacks a genotypically identifiable CF (63.4 %) (Nada *et al.,* 2013). CFAs were found in 33% of ETEC-related diarrhoea patients. The most typically expressed CFAs with LT-expressing ETEC were CS7 (n=18) and CS17 (n=9); CFA/I (n 48) and CS6 (n=38) with ST-expressing ETEC; and CS3 (n=31), CS6 (n=19), CFA/I (n=12), and CS14 (n=11) with LT/ST-expressing ETEC strains, respectively. In symptomatic children, CFA/I (n=61), CS3 (n=8), CS1+ CS3 (n=24), CS2+CS3 (n=18), CS6 (n=45), CS5+ CS6 (n=11), CS7 (n=25), and CS14 (n=32) were commonly recognised, while CS6 (n=66), CS12 (n=51), CFA/I (n=43), and CS14 (n=20) were frequently isolated from asymptomatic children (Mansour *et al.,* 2014). In the genotypic features of ETEC isolated from acute diarrhoea in Tunisia, CFA/I (44.6%) and CS6 (11%) were the most common colonization factors CFs, with 44.6% of isolates showing no connection with either CFA (Al-Gallas *et al.,* 2007).

In diarrhoeal samples from Bolivian children below 5 years of age, CF-positive isolates were found more frequently than in control samples. CFA/I (16.4%), CS14 (7.4%), CS1+CS3 (6.4%), and CS5+CS6 (6%) were the most prevalent CFs for ETEC isolates from children with diarrhoea (6%). ETECs expressing solely CS6 (3.6%), CS17 (3.3%), CS7 (2.7%), CS13 (2.3%), CS12 (2%), and CS2 + CS3 (1%) were found less frequently. CS14 was found more frequently in STh-only ETEC than in LT-only ETEC, while CS1+CS3 and CS5+CS6 were found more frequently in LT/STh ETEC isolates than in LT ETEC and STh ETEC isolates. *clyA* was found in 92.6% of ETEC isolates from children with diarrhoea, followed by *eatA* (34.8%), *east-1* (31.4%), *tibC* (20.4%), *tia* (12%), and *leoA* (6.7%) (Gonzales *et al.,* 2013). In a study of 40 ETEC clinical isolates from northern Colombia, South America, researchers found CS21 and CFA/I were found in 21 (50%) and 13 (32.5%) isolates, respectively, as the most common CFs. The most prevalent non-classical virulence genes found in more than 60% of the isolates were *eatA,* *irp2,* and *fyuA* (Guerra *et al.,* 2014). CFs were found in 64% of diarrhoeal samples and 37% of control samples of ETEC strains obtained from Peruvian children aged 2 to 24 months. CS6 (14% and 7%, respectively), CS12 (12% and 4%, respectively), and CS1 (9% and 4%, respectively) were the most common CFs from diarrhoeal samples and control samples (Rivera *et al.,* 2010). The most prevalent CF found in Chilean children was CS21, which was carried by 74 strains (71.8%), while 18 (17.5%) of the isolates carrying enterotoxin genes were CF-negative. The non-classical ETEC virulence genes were discovered with varying frequencies among 83 strains (80.6%) in the collection, with the *eatA* (70.9%), *etpA* (74.8%), and *etpB* (62.1%) genes being the most often detected. Among the isolates, 15 (14.6%) isolates were tested negative for NCVF (Del Canto *et al.,* 2011).

The most often expressed CF on community ETEC strains in Guatemala children was CS6, which was discovered in 11 (9.5%) of the strains, followed by CS1+CS3 or CS2+CS3 in 6.9% and CS4+CS6 or CS5+CS6 in 6.0%; other CFs were only discovered in low frequencies, i.e. in 2–4 of the 116 strains. The most common toxin profiles among the hospitalized children were CFA/I, which was identified in 7/22 (31.8%) of the strains, followed by CS1–3 in 2/22 (9%) of the strains, and CS14 and CS17 in 1/22 strains each (both 4.5%). No CFs was detected in 9 (41%) ETEC strains from the hospitalized children (Torres *et al.,* 2015). In total, 25-50% of strains globally are negative for any known CF.

2.13. Mixed Pathogens:

Coinfection with ETEC and other enteric pathogens is common, making it difficult to evaluate whether the symptoms are caused by the ETEC infection and to understand the pathophysiology of the infection. According to a study of patients admitted to the ICDDR, B hospital in Dhaka, mixed infections account for up to 40% of ETEC disease cases, and this number appears to increase with age (Qadri *et al.,* 2000)**.** The most common copathogen was rotavirus, which was followed by *V. cholerae*, *Campylobacter jejuni, Shigella spp.,* and *Salmonella* spp. Rotavirus coinfection was the most common among young children, peaking at 6 to 12 months, while *V. cholerae* was predominantly seen in older children and adults (Faruque *et al.,* 2004). EAEC and *Campylobacter* spp. are prevalent copathogens in travelers (Qadri *et al.,* 2005).

2.14. ETEC infection and malnutrition:

In children under the age of five, ETEC is the second leading cause of death. ETEC is usually the first enteric infection experienced by babies in low-resource nations and endemic areas; practically all children will have at least one ETEC diarrhoeal episode within their first year of life. ETEC infection is linked to malnutrition, growth stunting, and cognitive impairment in children, as it is with other types of diarrhoeal sickness. In the developing world, it is not uncommon for children to have ten severe episodes of diarrhoea in a year, with several of them being caused by ETEC. ETEC infection causes malnutrition and dehydration, resulting in a 15 to 20% decrease in productivity in adults (Qadri *et al.,* 2005). Diarrhoea is a leading cause of death in children with severe malnutrition. A child's body size, which was determined by his age and nutritional state, had a considerable impact on the rate of stool output per kilogramme of body weight. The small intestine of an underweight child is larger in proportion to body weight than that of a normal-weight child, resulting in increased stool losses per kg/body weight. Lower-weight children lost 14-61% more stools than higher-weight children. As a result, children who are small due to age or malnutrition lose a bigger proportion of their overall fluid volume during diarrhoea and are more likely to suffer diarrhoea, which can lead to death if left untreated (Black *et al.,* 1984).

Malnutrition and diarrhoea are two major interconnected health issues affecting pre-school children in developing countries. The incidence of diarrhoea was found to be around 1.6 bouts per child per year, and it was consistent across dietary grades. Children with grade III malnutrition, on the other hand, had a much larger percentage of episodes that resulted in severe dehydration. This group also had more ETEC isolates, indicating a larger bacterial load in the gut. Although pre-existing malnutrition had no effect on diarrhoea incidence, it did appear to have a substantial impact on the severity of the disease (Mathur *et al.,* 1985). Malnutrition was seen in 39% of children who had diarrhoea, regardless of the enteropathogen. Only enterotoxigenic *Escherichia coli (ETEC), Cryptosporidium sp., and Entamoeba histolytica* were considerably more frequent in underweight infants among the enteropathogens discovered (Mondal, D. *et al.,* 2009). Children who are malnourished are more likely to die from diarrhoea. Enteric protozoa, rotavirus, astrovirus, and ETEC were the most common causes of diarrhoea, which occurred 4.69 times per child per year. Malnutrition was found in 16.3% of newborns and 42.4% of infants aged 12 months. Malnourished babies were more likely to contract *Entamoeba histolytica*, *Cryptosporidium*, and ETEC infections, as well as have more severe diarrhoea. Children who were malnourished before the age of 12 months were more likely to have chronic diarrhoea, intestinal barrier dysfunction (Mondal, D. *et al.,* 2012).

ETEC producing heat-stable toxin exhibited a greater relation with moderate-to-severe diarrhoea in children with acute malnutrition, according to the Global Enteric Multicenter Study (Tickell *et al.,* 2020). In underdeveloped nations, micronutrient deficiencies, such as vitamin A and zinc, is extremely frequent, and it worsens the morbidity associated with diarrhoeal diseases although the impact on ETEC diarrhoea morbidity has not been investigated in detail. In Bangladesh, it is estimated that over 40% of children under the age of five suffer from zinc deficiency. For decades, there has been a bidirectional relationship between malnutrition and diarrhoeal morbidity (Qadri *et al.,* 2005).

2.15. Nutritional and Micronutrient Therapy:

Nutritional therapy is an important aspect of the treatment of all childhood diarrhoeas, including those caused by ETEC. Diarrhoea from any source, particularly ETEC, causes a drop in nutritional status and hence stunts growth in children. It has been discovered that including zinc in the treatment of diarrhoea in children results in a shorter length of sickness and a reduction in diarrhoea-related mortality. These investigations were conducted in locations where chronic zinc insufficiency in children is known to exist (Bhandari *et al.,* 2002; Black, 2003). Early in the course of therapy, it is critical to pay attention to supplying sustenance, particularly breast milk. Additional food will aid in catch-up growth during and after the diarrhoeal episode.

2.16. CS6 and CS5:

CS6 is one of the predominant CF among clinical ETEC isolates from developing countries especially in Bangladesh and India (Qadri *et al.,* 2005). It exists either alone or with CS5/CS4. CS6 is present in approximately 30% of ETEC isolated worldwide (Wolf *et al.,* 1997). Co-expression of CS6 and CS5 was found in 14% of Latin America/the Caribbean and 7% of South Asia, respectively. In the instance of CS6, however, global variability was seen in 6% of non-travellers and 20% of travellers. Co-expression of the CS5 and CS6 antigens was estimated to be around 6% in non-travellers and 6% in the total general population (Isidean *et al.,* 2011). EatA is one of the most prevalent non-classical virulence factors among clinical strains as well as in travellers (Del Canto *et al.,* 2011; Rivera *et al.,* 2013; Guerra *et al.,* 2014).

Genetic composition and assembly of CS6:

Coli surface antigen 6 (CS6) is a non-fimbrial component of the colonization factor antigen (CFA) IV complex and is expressed either alone or in conjunction with other colonization factors, CS4 or CS5 on the surface of human ETEC strains. Purified CS6 bonded to mucus from the duodenum and ileum of rabbits, as well as the duodenum, jejunum, and ileum of humans (Helander *et al.,* 1997; Thomas *et al.,* 1985). The CS6 operon is a 4219-base-pair DNA stretch that includes the genes *cssA, cssB, cssC*, and *cssD*. CS6 is unique among CFs in that it has two primary structural subunits, CssA and CssB. CssC and CssD are periplasmic chaperone and molecular usher, respectively, based on sequence homology. A signal sequence is seen in all four genes which is typical of exported proteins (Wolf *et al.,* 1997). CS6 oligomers are made up of two closely linked subunits, CssA and CssB, that form a (CssA-CssB)n complex with a spherical shape and an equal (1:1) stoichiometry (*Figure 2.14*).

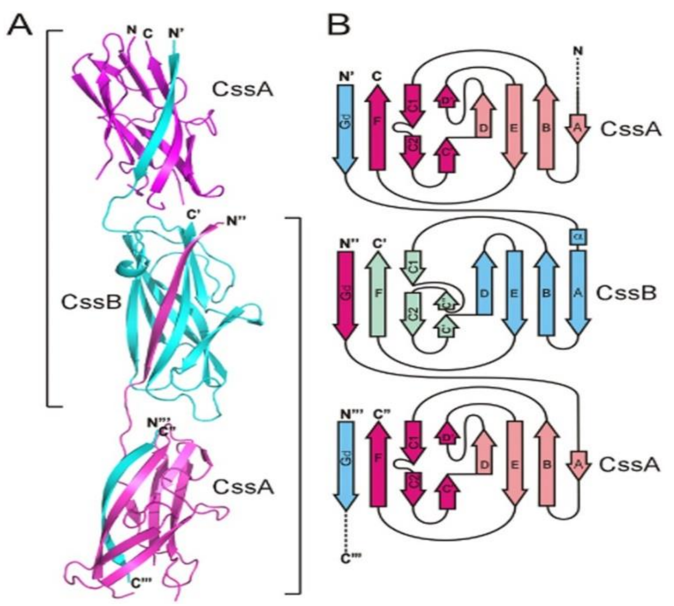

Figure 2.14: Model of the CS6 polymer. A. 3D reconstruction of a CS6 polymer based on the crystal structures of the CssAdsB and CssBdsA subunits. CssA and CssB are shown in magenta and cyan respectively. Brackets indicate two distinct two-subunit repeats in the CS6 fibre. B. Topology diagram of the CS6 fibre. Arrows indicate strands; the single α helix is shown as a rectangle. The strands in the same subunit are shown in slightly different colours to indicate two β-sheets of the β sandwich. **Source:** Roy *et al,. Molecular microbiology*. 2012.

With increasing concentration, the diameter of CS6 oligomers grows proportionately. CssA has low immunogenicity, but CssB has high immunogenicity. Purified CssA has an effective molecular mass of 18.5 kDa due to fatty acid modification, despite the expected molecular mass of 15 kDa. Fibronectin (Fn) has a dose-dependent and saturable interaction with CS6 and as well as CssA (Ghosal *et al.,* 2009; Sabui *et al.,* 2016). For appropriate CS6 expression and cell adhesion, these four proteins are required. CssD recognize the CssA-CssB-CssC complex and transport CssA-CssB to the outer membrane as a colonization factor (***Figure 2.15***) (Wajima *et al.,* 2011). The usher (CssD) is not involved in CS6 assembly or surface expression, but deletion of the chaperon (CssC) considerably lowers CssA levels but not of CssB. Even at very low concentrations, intact CssA is poisonous to the host cell (Tobias *et al.,* 2008). Some, but not all, of the residues in CssA and CssB's N and C termini, appeared to be crucial in intermolecular interactions between these two structural subunits and as well as for chaperone protein CssC (Debnath *et al.,* 2016). The occurrence of point mutations in structural genes, *cssA* and *cssB*, was discovered using sequence analysis which shows

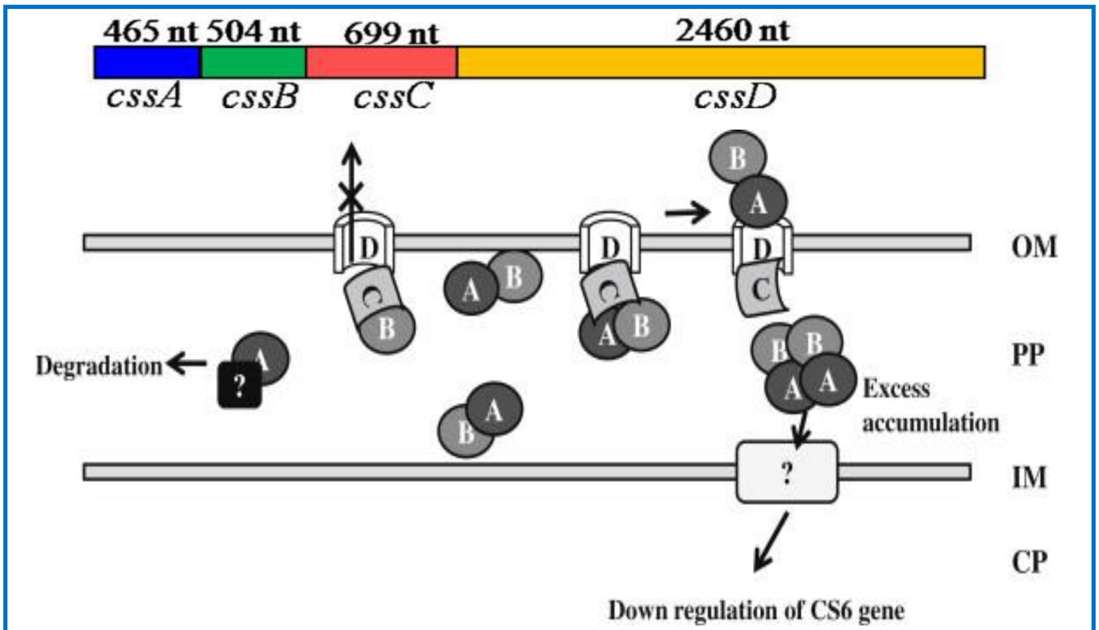

Figure 2.15: Schematic model of CS6 maturation. OM, outer membrane; PP, periplasm; IM, inner membrane; CP, cytoplasm; A, CssA; B, CssB; C, CssC; D, CssD; ?, unknown factors. **Source:** Wajima *et al., Microbial pathogenesis.* 2011.

that CssA has three alleles and CssB has two. Based on these alleles CS6 is subtyped into AIBI, AIIBII, AIIIBI, AIBII, and AIIIBII. CS6 with AI or AIII allelic subtypes had better binding capacity than AII, while BI had better binding capacity than BII.

The AII and BII alleles were found more frequently in controls than cases. In comparison to ETEC isolates expressing AIBI, ETEC isolates expressing AIIBII had lower adhesion to intestinal epithelial cells. Reduced surface-level expression and mucin binding of AIIBII is due to two distinct amino acid changes which resulted in differential adherence between AIBI and AIIBII (Sabui *et al.,* 2010; Debnath *et al.,* 2015).

Genetic composition and assembly of CS5:

ETEC's fimbriae colonization factor CS5 is a 2 nm flexible fibrillar structure, roughly 21.0 kilodaltons (kDa) and is plasmid-encoded (Thomas *et al.,* 1985; McConnell *et al.,* 1988). The CS5 operon encodes for 6 genes and the operon is unique in having two chaperon genes. The major subunit is encoded by *csfA* which is a mature protein of 18.6 kDa (Clark *et al.,* 1992). Five open reading frames (ORFs) downstream of the major subunit gene are *csfB* (25.8 kDa), *csfC* (90.3 kDa), *csfD* (41 kDa), *csfE* (22.8 kDa), and *csfF* (25.5 kDa), all of which are transcribed in the same way as the major subunit and are flanked by a number of insertion sequence sequences (***Figure 2.16***).

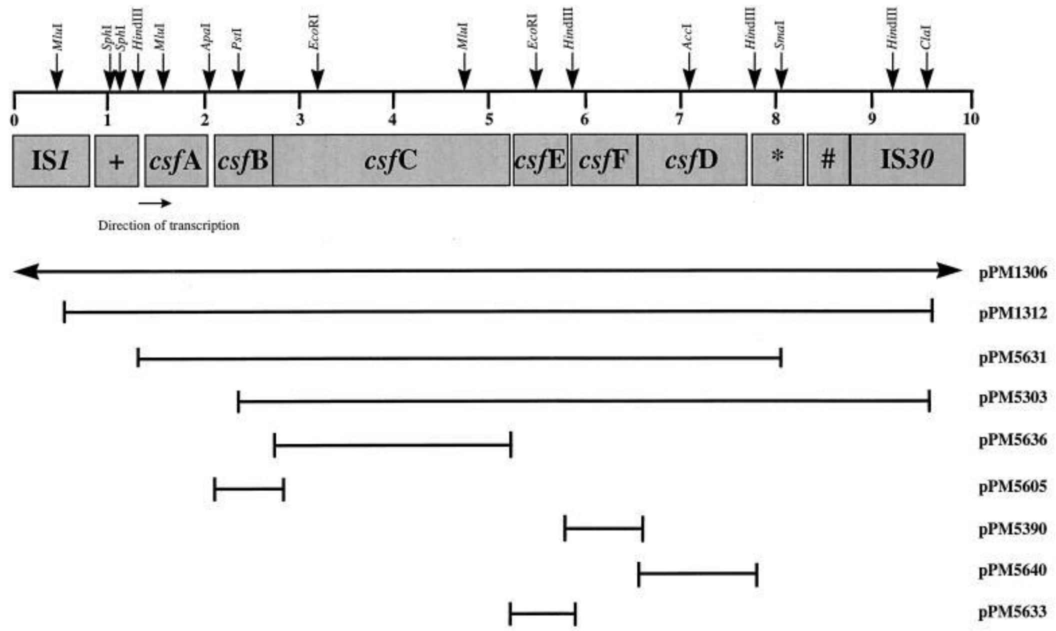

Figure 2.16: Genetic organization of the region required for the biosynthesis of CS5 pili in ETEC O115:H40. ORFs belonging to the CS5 cluster are designated csfA to csfF. The region is flanked by a variety of IS elements. +, region homologous to part of orfB in Tn21; ∗, region homologous to part of IS66; #, region homologous to part of IS911. IS1 and a defective IS30 are indicated. The upper line refers to the position of the genes relative to the sequence length, in kilobases, with relevant restriction sites indicated. Plasmids used in this study and the ORF(s) they encompass are indicated. **Source:** Duthy *et al., Journal of Bacteriology*. 1999.

csfD encodes the minor subunit and is essential for pilus assembly; *csfB* functions as a periplasmic chaperone for the major pilin c*sfA; csfC* functions as usher protein for translocation of CsfA and CsfD pilins across the outer membrane and is required for cell surface CS5 pili expression; *csfE* is important for pilus length regulation and thus determines pilus length; *csfF,* functions as chaperons for the minor pilins CsfD and CsfE. Based on sequence similarity it is suggested that CS5 adhesin is *csfD.* The Sec-dependent pathway is used to transport all Csf proteins through the inner membrane (***Figure 2.17***) (Duthy *et al.,* 1999, 2001, 2002).

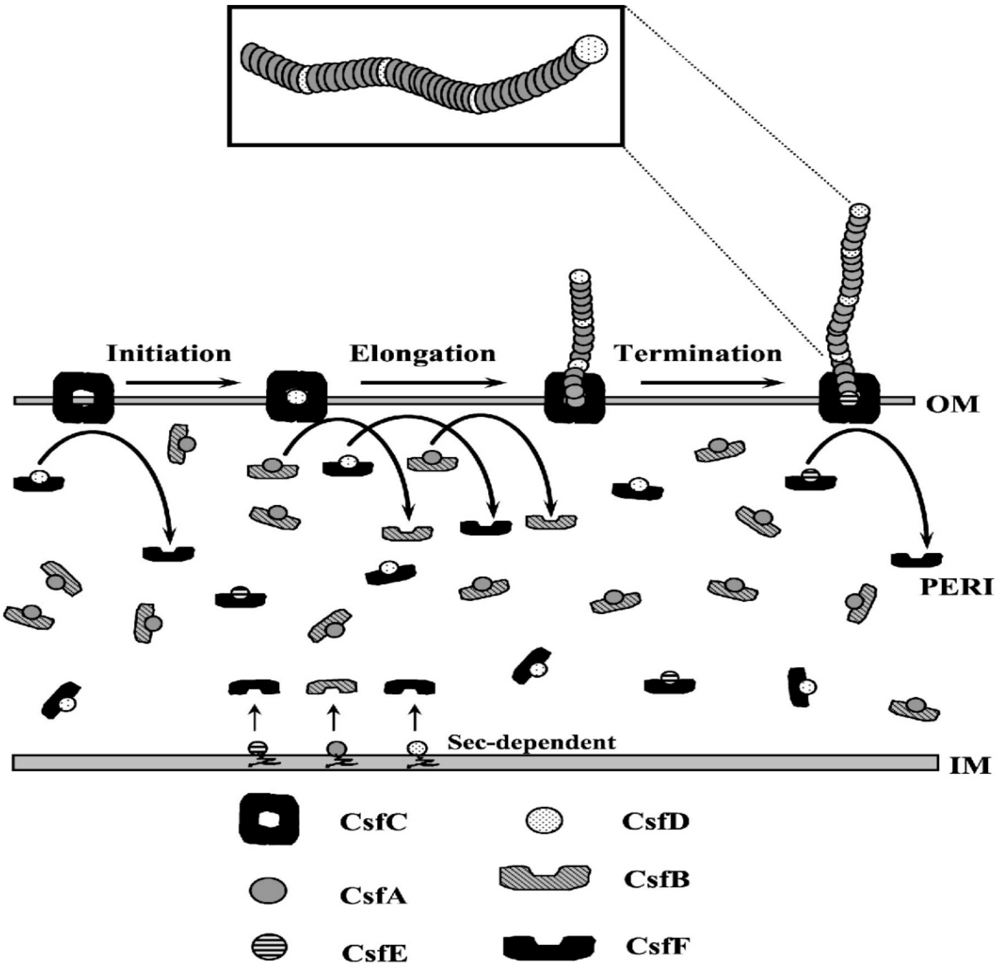

Figure 2.17: Model of CS5 pilus assembly. CS5 pilus biogenesis is hypothesized to be initiated by CsfD-CsfF complexes binding to the outer membrane assembly protein CsfC, resulting in the translocation of CsfD across the outer membrane. This is immediately followed by subsequent rounds of pilus elongation, in which CsfA subunits, delivered via a CsfA-CsfB complex to CsfC, are incorporated. Although the majority of the pilus consists of CsfA, further CsfD-CsfF complexes are also targeted to CsfC. The net result is a flexible CS5 pilus consisting predominantly of CsfA with a minor amount of CsfD. Pilus termination occurs when the rare CsfE-CsfF complex is targeted to CsfC, which is predicted to result in the irreversible incorporation of CsfE with CsfC, thereby preventing further rounds of pi lus elongation. All of the Csf proteins have been assigned patterned structures as shown at the bottom. OM, outer membrane; IM, inner membrane; PERI, periplasm. **Source:** Duthy *et al., Journal of Bacteriology*. 2002.

2.17. Antibiotic treatment and resistance:

Detection of ETEC is a slow process and problematic as coinfection with ETEC and other enteric pathogens are widespread, which can make it difficult to determine whether the symptoms are caused by the ETEC infection and to comprehend the infection's pathophysiology. Mixed infections are common, accounting for up to 40% of all cases.

So, antibiotic treatment is not an integral part of therapy in the treatment of childhood diarrhoea by ETEC (Qadri *et al.,* 2005). Therefore, it has been difficult to study the effect of antibiotics in children with ETEC infection. Antimicrobials, on the other hand, are beneficial in the treatment of travellers' diarrhoea, a diarrhoeal illness with a well-defined clinical symptom and ETEC is recognised as the most common pathogen (Jiang *et al.,* 2002).

When ETEC was originally discovered as the common pathogen for Traveller diarrhoea, it was very susceptible to all antimicrobials, including tetracyclines and trimethoprim-sulfamethoxazole (Sack, B. *et al.,* 1990). Antibiotic resistance developed over time, necessitating the use of newer antimicrobials for the treatment of traveller's diarrhoea. Doxycycline, trimethoprim-sulfamethoxazole, erythromycin, norfloxacin, ciprofloxacin, ofloxacin, azithromycin, and rifamycin are some of the antimicrobials that have been utilised in successful treatment (Sack, B. *et al.,* 1990; Ericsson, 2003).

The study conducted among travellers, including army and service personnel stationed in or visiting ETEC-endemic countries; and endemic population showed the evolution of multidrug-resistant ETEC strains; during 1968 to 1980 period study in Mexico, India, United States of America, Kenya, and Morocco showed that ETEC isolates are sensitive to all antimicrobials; during 1980–1990 single antibiotic resistance appears in Mexico; during 1980–1990 single antibiotic resistance appears in Mexico and Bangladesh; 1990 to 2005 multiple antimicrobial resistance as well as resistance to Ciprofloxacin and fluoroquinolones appears in India, Bangladesh. There is a cause for concern as there is a sharp rise in ciprofloxacin resistance from 1% in 1994 to 1997 to 8% in 2001 to 2004, with the majority of ciprofloxacin-resistant ETEC strains obtained from patients who travelled to India. Travellers to North Africa and Southeast Asia were also identified with ETEC resistance to ciprofloxacin (Qadri *et al.,* 2005).

When antimicrobial resistance in ETEC isolates was compared between the two time periods between 1994 to 1997 and 2001 to 2004, a statistically significant increase in resistance was found. In ETEC the highest levels of resistance were found for trimethoprim/sulfamethoxazole (58.5%), tetracycline (58%) and ampicillin (48%), followed by chloramphenicol (25.5%), nalidixic acid (14%), amoxicillin/ clavulanic acid (9%) and ciprofloxacin (4.5%). ETEC isolates from travellers to North Africa and India had a high percentage of quinolone resistance, which is cause for alarm. Patients with traveller's diarrhoea returning from these geographic locations should use these medicines with caution (Mendez *et al.,* 2009). Due to the frequent use of chemotherapeutic drugs in regions where diarrhoea is endemic, ETEC multidrug resistance is on the rise. Antimicrobial resistance is a global health problem which is due to the widespread and misuse of these antimicrobial agents, especially in endemic areas. To prevent the rise of multidrug resistance strains judicious use of antimicrobial agents are necessary and a search for alternative therapy is the need of this moment.

2.18. Vaccine development against ETEC infection:

Enterotoxigenic *Escherichia coli* (ETEC) is one of the leading bacterial causes of diarrhoea in young children, and also the most common cause of traveller's diarrhoea. Because of the devastating effects of diarrhoeal infections on morbidity and mortality, as well as on nutritional status, particularly among children in endemic locations, an effective vaccination is very desirable. Cholera and diarrhoea caused by ETEC are the two toxin-induced diarrhoeal illnesses that require the most efficient vaccinations. Effective vaccinations could help control the diarrhoeal disease by lowering mortality and morbidity, as well as, ideally, preventing disease transmission (Holmgren *et al.,* 1985). Vaccines that stimulate anti-adhesin immunity in the host to prevent ETEC attachment and colonization, as well as antitoxin immunity to neutralise enterotoxicity, are thought to be the most effective in preventing ETEC diarrhoea (Zhang *et al.,* 2012). Protective immunity against ETEC appears to be mediated by secretory IgA, according to animal research and indirect evidence from clinical trials. Antibodies to the CFs, other surface antigens, and LT; ST, a short peptide, does not elicit neutralising antibodies after natural infection. Thus the concept of a successful ETEC vaccine is feasible as evident from epidemiological studies and results from experimental challenge tests with human volunteers which indicated that ETEC infection resulted in specific immunity against homologous strains. Furthermore, multiple infections with antigenically different ETEC strains appear to result in broad-spectrum ETEC diarrhoea protection (Steinsland *et al.,* 2003; Qadri *et al.,* 2005; Zhang *et al.,* 2012).

There is yet no vaccine that provides widespread protection against ETEC. The difficulties in developing a vaccine can be linked to the wide variety of ETEC isolates found around the world, as well as the different serotypes.

It is generally known that CFs and non-classical virulence factors are geographically variable, with different virulence factors dominating in different parts of the world. As a result, the ETEC vaccination could need to be region-specific (Stoll *et al.,* 1983; Sommerfelt *et al.,* 1996; Qadri *et al.,* 2005). ETEC strains are antigenically diverse displaying different toxins, CFs and non-classical virulence factors. The combination of these factors has to be considered for effective vaccine development and a broad spectrum vaccine may be required owing to so many virulence determinants. Multiple CFs should be included in a broadly protective vaccine to induce anti-adhesin immunity against at least CFA/I and CS1-CS6, which are expressed by the most virulent ETEC strains which may protect against 50 to 80% of ETEC strains in most geographic areas, as well as toxoid antigens to induce anti-toxin immunity against heat-labile and heat-stable toxins. A multivalent toxoid-CF vaccine including an LT toxoid, such as the nontoxic B component LTB or a mutant LT, could give rather broad protection against 80 to 90% of ETEC strains worldwide. By including strains containing CS7, CS12, CS14, and CS17 the possible range of coverage could be expanded to up to 90% of all ETEC strains (Qadri *et al.,* 2005; Svennerholm *et al.,* 2016). Toxin-based, live attenuated, inactivated whole-cell, hybrid, and fimbrial antigen vaccines are among the different types of vaccines that are now being evaluated (***Table 2.3***) (Walker *et al.,* 2007).

2.18.1. Purified CFs and Enterotoxoids:

Various pure CFs have been investigated as oral immunogens, but they have been deemed unsuitable due to their high cost and sensitivity to proteolytic degradation. Purified CFs have been integrated into biodegradable microspheres to protect the fimbriae from degradation in the stomach (Byrd *et al.,* 2005). However, no appreciable protection against later challenges with ETEC expressing the homologous CFs was generated by any formulation of purified CFs, whether immunising with large doses of a mixture of CS1 and CS3 or recombinantly synthesised CS6 (Qadri *et al.,* 2005). The rCTB-CF ETEC vaccine is an oral, inactivated vaccine that expresses a variety of CFs and is complemented with rCTB had been studied extensively (Svennerholm *et al.,* 2016). Both the nontoxic LTB and CTB subunit components are ideal potential antigens for anti-LT immunity since they are immunogenic, stable in the gastrointestinal environment, and capable of binding to the intestinal epithelium. CTB given orally has also been shown to provide considerable (50–70%) protection against ETEC LT illness, however, the protection only lasted 3–6 months.

Based on the existence of distinct epitopes on the LTB molecule, it has been suggested that an LT toxoid may be slightly more effective than CTB in generating protective anti-LT immunity (Svennerholm *et al.,* 2016). The transcutaneous administration of an ETEC vaccination has been investigated as an alternate administration route. *E. coli* CS6 has, for example, been put into patches that are placed to the skin's surface, either alone or in combination with native LT. Immune responses against *E. coli* CS6 and anti-LT responses were generated in around half of the volunteers after they were given *E. coli* CS6 and LT. The evaluation of *E. coli* LT as a potential vaccine following transcutaneous immunisation is now underway (Güereña *et al.,* 2002). Human participants were given three vaccines with active LT in a patch put to the skin surface at 3-week intervals to investigate the ability of transcutaneously administered *E. coli* LT to produce protection against LT-producing ETEC. Transcutaneously administered LT provided highly substantial protection against moderate-to-severe travellers' diarrhoea in American travellers to Mexico, but not against ETEC illness. A phase II trial found that when compared to placebo-controlled travellers to Mexico and Guatemala, the transdermal patch reduced the occurrence of traveller's diarrhoea (Frech *et al.,* 2008). It has also been attempted to transfer ETEC antigens via transgenic plants. Ingestion of LTB expressed in potatoes and corn resulted in the formation of IgG antibodies against LT in human volunteers (Tacket, 2004). Another method involves non-covalently connecting the tip adhesion for CFA/I, CfaE molecule to LTB or CTB to create chimaeras containing ETEC adhesin and enterotoxin B subunit. These chimaeras were discovered to be bifunctional, evoking antitoxin as well as anti-adhesion immune responses (Jobling *et al.,* 2020). The dmLT, a novel mutant LT, is a safe and powerful detoxified enterotoxin that has the ability to act as a mucosal adjuvant for coadministered antigens and elicit anti-LT antibodies without causing negative side effects and could be used in vaccines (Norton *et al.,* 2011). ST has resisted all attempts to develop a non-toxic but immunogenic vaccination until recently. Some recent ideas, on the other hand, appear to have a chance of succeeding. The findings of experimental experiments on ST mutants conducted directly on-site using mutagenesis convinced researchers to develop a safe and immunogenic ST-containing vaccine (Zegeye *et al.,* 2018).

2.18.2. Inactivated Whole-Cell Vaccines:

Using killed ETEC with various forms of CF as immunogenic on the surface of the bacteria is an alternative way in mucosal vaccines. Some inactivated organisms are linked to appropriate LT toxoid i.e., cholera toxin B subunit or LTB. Inactivation is achieved by treating bacteria with colchicine E1 or formalin, which kills the germs but retains their antigenic properties. ETEC CF molecules are more stable than pure CF molecules, retaining their antigenicity and fimbriae structure (Mirhoseini *et al.,* 2018). rCTB–CF ETEC, the most extensively researched ETEC vaccine in clinical trials, is made up of rCTB and inactivated ETEC bacteria expressing CFA/I and CS1–5, as well as several of the most common ETEC O antigens. In a majority (70–90%) of Swedish vaccines, this vaccination was found to be safe and to produce strong IgA immune responses locally in the intestine. The vaccination is highly tolerated in phase I and II studies in Swedish, Bangladeshi, and Egyptian adult volunteers, with 70–100% of vaccines developing mucosal immune responses, or immunological responses in the intestine (gut lavage fluid) or peripheral blood ASCs against distinct vaccine CFs. Furthermore, the vaccination has been found to produce immune responses against CFs and LT in the gut that are comparable to clinical ETEC illness (Svennerholm *et al.,* 2016).

2.18.3. Live-attenuated or oral ETEC Vaccines:

A live vaccination strain expressing CS1 and CS3 fimbriae but lacking the genes for LT and ST provided 75 % protection against challenge with wild-type ETEC expressing equivalent CS factors as well as LT and ST, according to early findings by Levine *et al.* in human volunteers. Following that, non-pathogenic *E. coli*, attenuated *Shigella*, *V. cholerae*, or *Salmonella* expressing different CF components have been produced as possible vaccine candidates, either alone or in combination with an LT toxoid. The rCTB-CF ETEC vaccine is the only vaccination that has been investigated for protective effectiveness in a field trial in young children in places where ETEC is endemic so far. CS2 and CS3 fimbriae, as well as CFA/I, CS2, CS3, and CS4, as well as a detoxified form of human LT, have been expressed in vaccine candidates. Such mutated strains, PTL002 and PTL003, were tested in human volunteers, they were found to be safe and immunogenic when given in a single dose. (Svennerholm *et al.,* 2016). Individual live attenuated ETEC derivatives known as ACE527, which include strains ACAM2025, which expresses colonization factor antigen I (CFA/I) and LTB;

ACAM2022, which expresses CS5, CS6, and LTB; and ACAM2027, which expresses CS1, CS2, CS3, and LTB, have shown potential as ETEC vaccines. Clayton Harro (Johns Hopkins University, Baltimore, MD, USA) conducted phase IIb vaccination challenge trials that revealed a significant reduction in both the incidence and severity of ETEC diarrhoea after the challenge (Harro *et al.,* 2011).

In addition, immunisation with an oral cholera vaccine (Dukoral®, SBL Vaccin, Stockholm, Sweden) containing inactivated whole *Vibrio cholerae* and recombinantly produced CTB (rCTB), which crossreacts immunologically with LTB, has provided statistically significant protection against LT ETEC in both travellers and children in an ETEC-endemic area, as well as any diarrhoea in Spanish travellers (Svennerholm *et al.,* 2016). Recent advancements in vaccination trials and the establishment of an animal model to assess vaccine efficacy, as well as the improvement of a human volunteer-challenge paradigm, have sped up the discovery of an effective vaccine against ETEC diarrhoea. Field studies with big enough sample sizes will aid in the development of a broadly protective ETEC vaccine.

Additional ETEC proteins may have vaccine potential, these novel proteins are categorized as ETEC-specific plasmid-based on EtpA and EatA and Chromosomal-based which may include *Yghj, EaeH* and *Ag* (Fleckenstein *et al.,* 2014; Hosangadi *et al.,* 2018).

Table 2.3: **Summary of the statue and approach to the development of vaccines against ETEC**

Sl.No.	Candidate vaccine	Developer	Progress statue
1	Inactivated tetravalent whole-cell supplemented with LTB–CTB hybrid toxoid; may include dmLT adjuvant (ETVAX)	PATH, Sweden biotechnology laboratory	Phase II clinical
2	Live attenuated cells by genes aroC, omp, omp F with adjuvant dmLT (ACE527)	PATH	Phase II clinical
3	ZH9 attenuated typhoid vaccine expressing LT-ST toxoid (Typhetec)	Prokarium Co.	Preclinical
4	Second generation of attenuated Shigella vaccine expressing CF/CS antigens and LT toxoid	CVD Vaccine Development Center	Preclinical
5	Anti-binding factor subunit vaccine	NMRC	Phase II clinical
6	Anti-binding factor subunit vaccine and toxoid (MEFA)	Kansas University	Preclinical
7	Fusion LT-ST and conjugated LTB and ST	International Consortium of Intestinal Vaccine	Preclinical
8	dmLT	PATH	Phase II clinical
9	Flagellin; EtpA; EatA; EaeH; YghJ	Various*	Preclinical

* Novel toxoids and antigens are being explored by a number of investigators from: Washington University in St. Louis; University of Maryland; University of Virginia; University of Bergen; South Dakota State University; Kansas State University; the Sanger Institute; Johns Hopkins Bloomberg School of Public Health; and Antigen Discovery, Inc.

Source: A. Mirhoseini *et al. Microbial Pathogenesis.* 2018.

Chapter 3

OBJECTIVES

The studies carried out in this thesis work aims to better understand the most prevalent combination of classical and non-classical virulence factors among clinical isolates of ETEC in our settings. Here, we will focus on the prevalent colonization factors in terms of their identification, expression, distribution and regulation during pathogenesis. We will further study the expression of the prevalent colonization factor during pathogenesis taking into account the involvement of other virulence factors and identification of host factors that might control the expression of that prevalent colonization factor during pathogenesis. The objectives are:

Objective I: Detection of common colonization factors and other virulence factors among clinical isolates of ETEC to understand their occurrence.

Objective II: To delineate the expression of identified colonization factors and virulence factors in clinical ETEC.

Objective III: To understand the modulation of prevalent colonization factor and other virulence factors by the host and environmental factors during pathogenesis.

Chapter 4

MATERIALS

&

METHODS

4.1. Bacterial Strains:

The clinical bacterial strains used in this study were from stool samples collected from diarrhoea patients admitted at the Infectious Diseases and Beliaghata General Hospital (ID & BG Hospital, Kolkata) and Dr B.C. Roy Post Graduate Institute of Pediatric Sciences during the period of 2008-2014 under the hospital-based surveillance project of the ICMR-National Institute of Cholera and Enteric Diseases (ICMR-NICED), (limited to first 1000 cases). Strains from the community were also collected from asymptomatic people. A well-characterized clinical ETEC isolate IDH00469 harbouring ST, LT gene will be primarily used as a wild type strain in this study. Other ETEC strains will be used from the strain archive of Bacteriology as and when required. Laboratory strains *E. coli* BL21 (DE3) (Promega) and Top10 (Invitrogen) will be used as control. *JM109* (Promega, USA), *T7 shuffle* were used as required.

4.2. Culture media and conditions:

Stool samples were plated on MacConkey (BD Difco, USA) agar plates and incubated at 37°C for 18 hours for ETEC detection. The lactose-fermenting colonies morphologically resembling *E. coli* were separated and cryo-preserved in Luria Bertani broth containing 15% glycerol and stored at -80°C for future use. The strains were grown routinely on broth and/or agar (BD Difco, USA) plates of LB (BD Difco, USA) or M9 minimal media (BD Difco, USA), at 37°C. These strains were grown in LB media with or without antibiotics at 37°C for overnight to isolate genomic DNA. Routinely agar was used as 1.8%. All media were prepared according to standard methods. All culturing was performed at 37°C. Modifications or changes from the routine methods were made as per the experiment required. Minimal medium, M9, was supplemented with 0.2% glucose as the primary carbon source, glycerol or casamino acids were also used for carbon source.

4.3. Medium Additives:

Crude Bile extract (BactoDifco), taurocholic acid sodium salt hydrate (TCA) (Sigma), sodium cholate hydrate (NaCH) (Sigma), sodium glycocholate hydrate (NaGCH) (Sigma), sodium chenodeoxycholate (NaCDC) (Sigma), and sodium deoxycholate (NaDC) (Sigma); salts of iron, Ferric sulphate (Merck); ferrous chloride (Sigma), Ferric chloride (Sigma), ferric nitrate (Sigma), porcine mucin (Sigma); Sodium Chloride (NaCl) (Merck) were used for the experiments. Appropriate dilutions were

* For media and buffer composition please see appendix

made during experiments by dissolving the powder in DNA/RNA-free water and filter sterilizing it (0.2 µm). Additives were supplemented in M9, LB and CFA broths in different concentrations. Media with no additive added was used as control.

4.4. DNA Extraction:

4.4.1. Boil lysis method- Total genomic DNA was isolated by the boiled template method (Ghosal *et al.,* 2007). In brief, overnight grown ETEC sample was pelleted down by centrifugation for 10 minutes at 16,000 ×*g*, resuspended in TE buffer, boiled for 10 minutes and again centrifuged. The supernatant served as the template for PCR experiments.

4.4.2. Phenol chloroform method (Maniatis, 1982) - Overnight grown ETEC culture was pelleted down by centrifugation for 10 minutes at 10000 ×*g*. The cell pellet was resuspended in lysis buffer and incubated for 1 hour at 37°C. Then an equal volume of phenol/chloroform (1:1) was added and mixed by inverting the tube until the phases are completely mixed. The lysate was centrifuged at 16,000 ×*g* for 15 minutes and the upper aqueous phase was transferred to a new tube. To precipitate the DNA, 3 volumes of cold 100% ethanol was added and mix gently and the tube was kept at -20˚C for 30 minutes and then centrifuged down at 16,000 ×*g* for 15 minutes at 4°C. The supernatant was discarded and the DNA pellet was rinsed with 70% ethanol. After centrifugation, the supernatant was discarded and the DNA pellet was air-dried. The DNA was resuspended in TE buffer.

4.4.3. CTAB method- The CTAB approach was used to isolate whole genomic DNA (William *et al.,* 2012). Briefly, late log phase or early stationary phase cells were pelleted down at 10,000 ×*g* for 10 minutes and the supernatant was discarded. The pellet was resuspended in TE buffer. 10% SDS and Proteinase K (10mg/mL) were added and incubated for 1-3 hours at 60°C. After the cells were lysed (as seen by cleared solution with increased viscosity) 5 M NaCl was added. CTAB (heated to 65°C) was added @ 100µl per 1 mL culture and incubated at 65°C for 1 hour with occasional shaking. Chloroform: Isoamyl alcohol (24:1) @ 500 µl per 1 mL was added and mixed for 20 minutes on ice. Centrifuged at 16,000 ×*g* for 10 minutes at room temperature and the aqueous phase was transferred to a clean microcentrifuge tube to which phenol: chloroform: isoamyl alcohol (25:24:1) @ 500 µl per 1 mL was added and mixed well. Centrifuged at 16,000 ×*g* for 10 min at room temperature. To the aqueous phase 0.6 volume isopropanol (-20°C) was added and incubate at -20°C

* For media and buffer composition please see appendix

for 2 hours. Centrifuged at 10000 ×*g* for 15 minutes at 4°C. The pellet was washed with cold 70% ethanol. Centrifuged at 10000 ×*g* for 5 minutes at 4°C. The supernatant was discarded and the pellet was air-dried at room temperature. Then the pellet was resuspended in DNase free water. Then RNase I @ 10U/µl was added and incubated at 37°C for 1 hour. The enzyme was heat inactivate at 70°C for 15 minutes followed by placing the tube in ice. The DNA was ethanol precipitated and finally, the DNA was eluted in TE buffer and stored at –20˚C.

4.5. Ethanol precipitation:

DNA and PCR products were ethanol precipitated. Briefly, the salt concentration of the DNA sample was adjusted by adding 1/10th volume of 3.2 M sodium acetate, pH 5.2 or an equal volume of 5 M ammonium acetate. Then, 2-2.5 volume 100% ice-cold ethanol was added and mixed well. The mixture was kept at -20˚C for 30 min and spun down at 16,000 ×*g* for 15 minutes at 4˚C. The supernatant was carefully separated and 70% ice-cold ethanol was added and mixed by vortexing. It was again centrifuged for 10 min at 16,000 ×*g* at 4˚C. The supernatant was discarded and the precipitated DNA was air-dried. Finally, the DNA was eluted in TE buffer.

4.6. Primer Designing:

The primer sets for multiplex PCR were created using *Tm* values that were similar. The primers were produced by Integrated DNA Technologies, Inc. in the United States, and the sequences were analysed using the BLAST tool for degeneracy and cross-reactivity with other ETEC genes and CFAs. The BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) was used to check primers for degeneracy and cross-reactivity. Eleven primer sets for colonization factors, two primer sets for toxin genes, and five pairs of primers for non-classical virulence factors were constructed. Using known strains, the specificity of the primer pairs was validated. To confirm the toxins, CFA/I, and all non-classical virulence factors amplicons, ETEC H10407 was employed. Strains used in previous research (Ghosal *et al.,* 2007; Sabui *et al.,* 2010) except for CS21, were used for confirmation of all other traditional colonization variables. Sequencing the first three amplicons from the strains utilised in this study verified the CS21 primer pair. During the confirmation of the primer sets, all amplicons were validated by sequencing on an automated DNA sequencer (ABI 3730 DNA Analyzer; Applied Biosystems) (***Table 9.1; please see appendix***).

 * For media and buffer composition please see appendix

4.7. DNA Amplification and Detection:

Polymerase chain reaction (Ghosal *et al.,* 2007) was done in 0.2 mL PCR tubes with a 20-µl mixture containing the following reagents as described (***Table 4.1***). The reagents for PCR are from Promega, USA. Having confirmed the specificity of each primer set by single PCR, we performed multiplex PCR (Toma *et al.,* 2003) by combining different primer sets in different ratios and tested the control strains in several PCR cycling protocols. For detection of CFA genes, multiplex PCR was performed followed by confirmatory singleplex PCR. PCR was done by using Thermal cycler (Eppendorf, Germany). The volume of primers and water for multiplex PCR were adjusted as per the numbers of primers used in each mixture.

Table 4.1: Composition of PCR reaction mixture

Components for PCR Mixture	Volume
5X Reaction Buffer	4 µl
$MgCl_2$, 25 mM	2 µl
dNTPs, 2.5 mM of each NTP	0.8 µl
Taq Polymerase, 5U/µl	0.2 µl
Forward Primer, 10 µM	0.4 µl
Reverse Primer, 10 µM	0.4 µl
DNA Template	2 µl
Nuclease Free Water	To make up the volume
Final Volume	20 µl

For experiments, the annealing temperature is determined by keeping the temperature 3–5°C lower than the lowest T_m of the primers for PCR amplification (Rychlik *et al.,* 1990). The melting temperature (T_m) of the selected primers were estimated by using the formula: $\boldsymbol{T_m = 4\ (G + C) + 2\ (A + T)}$**.** The final annealing temperature was set by checking the amplicons by gradient PCR. The extension time of PCR depends upon the synthesis rate of DNA polymerase used in the experiments and the length of target DNA. The typical extension time for *Taq* DNA Polymerase was 1 min/kb. The PCR conditions used for amplification for classical genes (***Figure 4.1***) were as follows: initial denaturation at 95°C for 3 minutes, followed by 30 cycles of denaturation at 95°C for 50seconds, annealing at 52°C for 40 seconds and elongation at 72°C for 40 seconds and finally a 5 minutes extension at 72°C. The PCR conditions used for amplification for non-classical genes (*Figure 4.2*) were as follows: initial denaturation at 95°C for 3 minutes, followed by 30 cycles of denaturation at 95°C for 50seconds, annealing at 54°C for 30 seconds and elongation at 72°C for 30 seconds and finally a 5 minutes extension at 72°C. The final step is the storage of the PCR products for an indefinite time at 4°C.

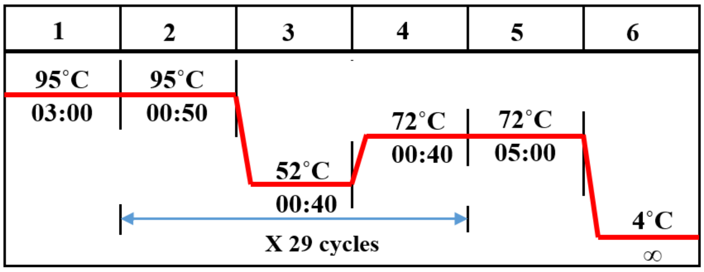

Figure 4.1: Schematic diagram of the PCR conditions used to amplify classical virulence factors genes.

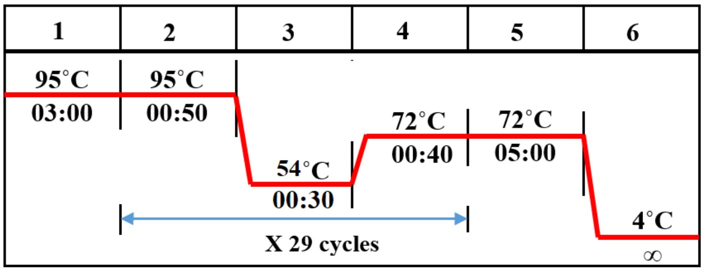

Figure 4.2: Schematic diagram of the PCR conditions used to amplify non classical virulence factors genes.

4.8. Agarose Gel Electrophoresis:

Agarose gel electrophoresis was used to size separate the PCR products to see if they successfully amplified the expected DNA target area. The size of the PCR products was assessed by comparing them to a DNA ladder (Thermo ScientificTM DNA Ladder) that ran alongside the PCR products on the gel. The amplified DNA fragments were resolved by 1.2% agarose gel electrophoresis. Briefly, agarose was heated to dissolve in 1X TAE. The temperature of the suspension was brought to approximately 55°C and poured into a gel caster (AmpliSize; Bio-Rad Laboratories). Electrophoresis was performed and the amplified PCR products were visualized under UV transillumination after staining with ethidium bromide. The images were captured using a gel documentation system (Bio-rad, USA).

4.9. Plasmid DNA Isolation:

Commercial available kit Wizard® plus Minipreps DNA purification system (Promega, USA) was used for plasmid isolation following the manufacturer's protocol.

 * For media and buffer composition please see appendix

In short, 5-10 mL of overnight grown bacterial culture were pelleted by centrifugation at 1,400 x*g* for 10 minutes. In 400µl of cell resuspension solution, the bacterial pellet was resuspended. After that 400µl of cell lysis solution was added which turned the cell suspension clear. After the addition of 400µl of neutralising solution, the lysate was centrifuged at 10,000 x*g* for 15 minutes. A vacuum of 15 inches of Hg was used to pull the resin/lysate fully through the minicolumn after the supernatant was combined with 1mL of resin. After that, the resin was rinsed extensively with wash buffer. Plasmid attached to the membrane was retrieved by centrifugation at 10,000 x*g* for 20 seconds with nuclease-free water.

4.10. Estimation of DNA and RNA concentration:

The concentration of DNA and RNA was quantified by measuring the absorbance at 260 nm and using the following equation (Barbas *et al.,* 2007):

For DNA: $[(A260\times \text{dilution factor}\times 50)/ 1000]\ \mu g\ mL^{-1}$

For RNA: $[(A260\times \text{dilution factor}\times 40)/ 1000]\ \mu g\ mL^{-1}$

Where A_{260} is the absorbance of the diluted DNA at 260 nm. For DNA, A_{260} of 1.0 corresponds to 50µg mL^{-1} concentration and for RNA, A_{260} of 1.0 corresponds to 40µg mL^{-1} concentration. Pure DNA has an OD_{260}/OD_{280} ratio of ~1.8; pure RNA has an OD_{260}/OD_{280} ratio of ~2.0.

4.11. RNA Isolation:

TRIzol™ Reagent (Invitrogen, Thermo Fisher Scientific, USA) method was used to isolate RNA from bacterial cells following the manufacturer's protocol. To isolate RNA, log-phase bacterial cells were pelleted down by centrifugation @ 10,000 x*g* for 10 minutes and the supernatant was discarded. To the pellet, TRIzol™ Reagent was added @ 0.75 mL per 0.25 mL of sample (1×10^7 cells of bacterial origin). The pellet was homogenized by pipetting the lysate. After incubating for 5 minutes at room temperature (RT), 0.2 mL of chloroform per 1 mL of TRIzol™ Reagent was added and mixed well and then incubated for 2-3 minutes. The sample was centrifuged for 15 minutes at 12,000 ×*g* at 4°C and the aqueous phase containing the RNA was transferred to a DNase/ RNase free microcentrifuge tube. Isopropanol was added @ 0.5 mL per 1 mL of TRIzol reagent and incubated for 10 minutes at RT. Total RNA was precipitated by centrifugation at

12,000 ×*g* for 10 minutes at 4°C. The supernatant was discarded and the RNA pellet was resuspended in 75% ethanol @ 1 mL per 1 mL of TRIzol reagent. The sample was vortexed briefly and then centrifuged for 5 minutes at 7500 × *g* at 4°C. The supernatant was discarded and the RNA pellet was air-dried. The pellet was resuspended in 20-50 µl of RNase-free water and incubated in a water bath or heat block set at 55–60°C for 10–15 minutes. The RNA was further treated with DNase and/ or stored at -80°C until further use.

4.12. DNase treatment:

The Ambion® DNA-*free*™ DNase Treatment and Removal Reagents were used to remove contaminating DNA from RNA preparations. 0.1 volume of 10X DNase I Buffer and 1 µl rDNase I were added to the RNA sample and mixed gently. Then it was incubated at 37°C for 20–30 minutes. DNase Inactivation Reagent @ 0.1 volume was added and incubated for 2 minutes with occasional mixing. Then centrifuged at 10,000 ×*g* for 1.5 minutes and the RNA was transferred to a fresh tube. Conventional PCR for *parC* (a housekeeping gene) was used to confirm the removal of genomic DNA.

4.13. cDNA Synthesis (Reverse transcription):

Thermo Scientific Verso cDNA Kit was used to prepare cDNA from 1µg of RNA following the manufacturer's protocol. At first, RNA was heated at 70°C to denature the secondary structures and placed immediately on ice. The reaction mixture was prepared with the components (***Table 4.2***). The components of RT-PCR was then set in a reverse transcription cycling program (***Table 4.3***). RT-PCR products were kept at 4˚C for at least 5 minutes and then stored at -20˚C until further use.

Table 4.2: cDNA Reaction mix preparation

Components	Volume	Final Concentration
5X cDNA synthesis buffer	4 µl	1X
dNTP Mix	2 µl	500 µM each
RNA Primer	1 µl	
RT Enhancer	1 µl	
Verso Enzyme Mix	1 µl	
Template (RNA)	1-5 µl	1 µg
Water, nuclease-free	To 20 µl	
Total volume	20 µl	

* For media and buffer composition please see appendix

Table 4.3: Reverse Transcription Program

Step No.	Step	Temperature	Time	Number of Cycles
1	cDNA synthesis	42 °C	30 minutes	1 cycle
2	Inactivation	95 °C	2 minutes	1 cycle

4.14. Expression study of CS6, CS5 and EatA:

Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) was done for RNA expression study of these virulence factors. First, the RNA was isolated and RT-PCR was done to make their cDNA. Then qRT-PCR was performed to check the expression of the virulence genes using cDNA as templates. Gene-specific forward and reverse primers for the CS5 minor structural subunit gene, *csfD;* CS6 structural subunit gene, *cssB; eatA* passenger domain, and the *E. coli* housekeeping gene, *parC* were designed using PrimerQuest™ Tool of Integrated DNA Technologies, Inc., (IDT), USA and the primers were synthesized by IDT, USA (***Table 9.2; please see appendix***).

The fluorescent dye SYBR Green was used for detection in ABI 7500 Real-Time PCR system (Applied Biosystems). Fold increase or decrease was calculated by the comparative Ct method also referred to as the $2^{(-\Delta\Delta Ct)}$ method (Livak *et al.,* 2001). The housekeeping gene, *parC* was used as an internal control. Negative controls without reverse transcriptase (-RT) were prepared in parallel with the cDNA from the same amount of RNA for all samples. The real-time RT-PCR assays were run in 20µl reactions using a maximum of 100ng cDNA (***Table 4.4***).

Table 4.4: PCR Master Mix for Real-time PCR preparation

Components	Volume	Final Concentration
2X SYBR® Green PCR Master Mix (Applied Biosystems)	10 µl	1X
Forward Primer	1 µl	100 nM
Reverse Primer	1 µl	100 nM
Template	Variable	100 ng
Water	To make 20 µl	
Total	20 µl	

After an initial 10 minutes denaturation step at 95°C, the reactions were subjected to 40 cycles comprising 15 seconds of denaturation at 95°C and 30 seconds of annealing at 58°C and 30 seconds of elongation at 60°C, followed by a dissociation step for melting

* For media and buffer composition please see appendix

temperature (*Tm*) analysis of each amplification product to confirm amplification specificity (***Figure 4.3***).

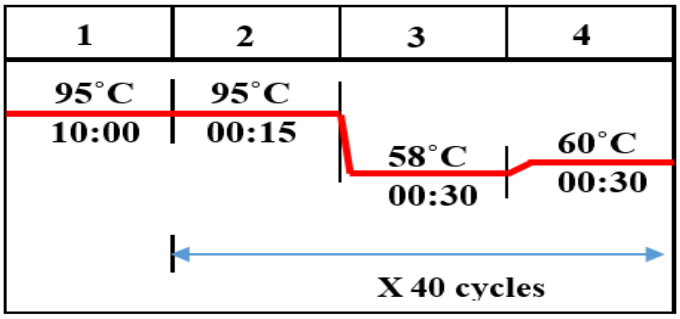

Figure 4.3: Schematic diagram of the qRT-PCR conditions used to amplify genes.

4.15. Cell culture:

Intestinal cell line HT29 was used in the experiments in this study. HT-29 cell line was maintained as per American Type Culture Collection (ATCC) cell culture guidelines. The HT-29 cell line was cultured in Corning T-25 flasks using Dulbecco's Modified Eagle's medium (DMEM) at 37˚C with 5% CO_2 using a HERAcell 150 humidified incubator (Thermo Scientific, USA) to form a monolayer. Complete DMEM media was used for maintaining of cell line and for infection studies incomplete media was used. The spent culture media was replaced with fresh one on alternate days and the monolayer was used after achieving ~90% confluency. The stock of the cultured epithelial cell was prepared by using 95% complete growth media and 5% DMSO (Sigma-Aldrich, USA) and kept in liquid nitrogen.

4.15.1. Monolayer Subculturing: To keep anchorage-dependent cell lines growing in monolayers in exponential development, they must be subcultured at regular intervals. The cells are ready to be subcultured when they are nearing the end of exponential development (approximately 70% to 90% confluent). The used media was discarded and washed with PBS (37˚C) without calcium or magnesium. Then 2 mL to 3 mL of the trypsin-EDTA solution (37˚C) was added and kept at 37˚C for cells to be detached. Then complete growth medium was added to inactivate the trypsin. After centrifugation, the cells are suspended in complete media and dispensed in routine split ratio.

4.15.2. Cell Infection studies: Before the initiation of the experiment, the epithelial cells were serum-starved for 18 hours in an incomplete medium. The confluent cultures in the incomplete medium were infected with varying concentrations of bacteria in

* For media and buffer composition please see appendix

triplicates. Control groups were treated with medium alone and laboratory *E. coli* strains. ETEC was isolated at different time intervals of infection and analysed as per the experiments.

4.15.3. ETEC adherence with cultured epithelial cells by plate count method: In brief, HT-29 epithelial cells were grown up to 80-90% confluent. Bacterial suspension at a concentration of 10^7 cfu mL^{-1} was pelleted down and dissolved in culture medium. This suspension was added to epithelial cells grown on 12-well plates. After 3 hours of incubation at 37°C in 5% CO_2, cells were washed three or four times with PBS to remove unbound bacteria and detached by 0.1% Triton X-100 or by Trypsin-EDTA. Adhered bacteria were counted after serial dilution by plating on MacConkey agar plates for cfu counts.

4.15.4. Quantification of CS6, CS5 and EatA: Quantification of CS6, CS5 and EatA genes upon co-cultured with HT-29 was done by quantitative real-time PCR. For qRT-PCR, ETEC strains were infected with HT-29 and incubated at 37˚C in 5% CO_2 for the required time. After that desired time cells were washed three or four times with PBS to remove unbound bacteria and dissolved by trizol. After that RNA isolation protocol was followed up to qRT-PCR.

4.16. Competent cell preparation:

From overnight grown *E. coli* culture fresh LB broth was inoculated @ 1:100 and grown with shaking at 37˚C till the O.D at A_{600} reaches 0.4-0.5. The culture was kept on ice for 10 minutes. The suspension was then centrifuged at 4,000 x*g* for 15 minutes at 4˚C and the pellet was resuspended in 0.1M ice-cold $CaCl_2$ (filter sterilized). The culture was incubated in ice for 20 minutes and then centrifuged at 4,000 x*g* for 15 minutes at 4˚C. The supernatant was discarded and the pellet was dissolved in 1 mL ice-cold 85% 0.1M $CaCl_2$ and 15% glycerol solution (filter sterilized). Aliquots of 50μl were dispensed in microtubes and preserved at -80˚C.

4.16.1. Ultra Competent *E. coli* Cells preparation (Sambrook *et al.,* 2006):

SOB media was inoculated with *E. coli* cells and grown up to OD_{600} of 0.6-0.8 at 18˚C. After incubating in ice for 10 minutes the cells were centrifuged at 2500 x*g* for 10 minutes at 4˚C. The cells were resuspended in ice-cold TB and incubated in ice for 10 minutes. The cells were centrifuged at 2500 x*g* for 10 minutes at 4˚C and the cells were dissolved in ice-cold TB + 14% DMSO. Ater incubating in ice for 10 minutes the cells

 * For media and buffer composition please see appendix

were dispensed in 50µl aliquots in pre-cooled 1.5 mL micro-centrifuged tubes. The aliquots were stored at -80˚C.

4.16.2. Electrocompetent cells preparation:

Electrocompetent cells were prepared as per New England Biolabs, USA protocol. Fresh inoculum was given in SOB pre-warmed to 37°C from overnight culture and was grown at 37°C with shaking until the cultures reached an OD_{600} of 0.5-0.7. The cultures were placed on ice for 15 minutes and then centrifuged at 5000 x*g* for 10 minutes at 4°C. The cells were completely suspended in ice-cold 10% glycerol while keeping the suspension on ice and again centrifuged. The pellet was again suspended in ice-cold 10% glycerol. After centrifugation at 5000 x*g* for 10 minutes at 4°C, the pellet was dissolved in 1 mL ice-cold 10% glycerol and 100µl was dispensed in cold microcentrifuge tubes on ice. Cells were used for electroporation or kept in -80°C freezer for storage.

4.17. Transformation:

5µl (5ng – 1µg) of plasmid DNA was added to 50µl of competent cell suspension, swirled together, and incubated on ice for 30 minutes. The tubes were given a 45-second heat pulse in a 42 °C water bath before being incubated on ice for 2 minutes. The tube was filled with pre-warmed 0.9mL SOC broth (Invitrogen, USA) and incubated for 1 hour at 37 °C with shaking at 200–250 rpm. Finally, a suitable volume of transformation reactions was plated on LB agar plates with the appropriate antibiotics and incubated overnight at 37 °C. The colonies were examined by PCR the next day to determine which ones were positive.

4.18. PCR product purification:

PCR products either directly or from agarose gel were purified by Wizard® SV Gel and PCR Clean-Up System (Promega, USA) as per the manufacturer's protocol. Shortly, the PCR products were loaded onto the agarose gel and after electrophoresis, the desired bands from the gel were cut and their weights were recorded. To it, Membrane Binding Solution at a ratio of 10µl of solution per 10mg of agarose gel slice was added and incubated at 50–65°C for 10 minutes or until the gel slice was completely dissolved. For PCR Amplified products an equal volume of Membrane Binding Solution was added to the PCR mix. Then the DNA was purified using a vacuum manifold. The dissolved gel mixture or PCR amplified mixture was transferred to the SV Minicolumn and incubated

*For media and buffer composition please see appendix

for 1 minute at room temperature and a vacuum was applied to pull the liquid completely through the SV Minicolumn. The column was washed twice with Membrane Wash Solution previously diluted with 95% ethanol, at first with 700µl and second time with 500µl. The SV Minicolumn assembly was centrifuged for 5 minutes and for 1 minute lid open condition at 16,000 x*g*. Purified DNA was collected by adding Nuclease-Free Water directly to the column followed by centrifugation for 1 minute at 16,000 x*g*. The eluted DNA was stored at 4°C or –20°C as per the usage.

4.19. Generation of isogenic non-polar mutants and ETEC complement strains:

Construction of isogenic non-polar gene-deletion mutants for genes CS5 and EatA were generated in ETEC wild-type strain, IDH00469 by the lambda Red recombination mutagenesis system (Datsenko *et al.,* 2000). The gene to be knocked-out is replaced with the kanamycin resistance gene. The steps are:

Step1: Transformation of wild type ETEC strain IDH00469 with pKD46. The ETEC strain was made competent and transformed with pKD46 plasmid. After overnight growth on LB+ 100µg/mL Ampicillin plates at 30˚C, colonies were checked for pKD46 presence by PCR. Positive colonies were preserved at -80˚C.

Step 2: Substrate DNA designing and generation of fusion linear PCR product for transformation. First 500bp upstream and downstream sequence of the gene to be knockout was searched and primers were designed to get 2 linear fragments from upstream (fragment 1) and downstream (fragment 3) of the gene of interest (GOI) (***Table 9.3; please see appendix***). Then primers were designed to get the kanamycin resistance gene (fragment 2) from pKD13 plasmid including the two FRT sites. Amplification of the 3 regions were done by Phusion™ High-Fidelity DNA Polymerase (Thermo Scientific, USA) and fusion of the three fragments were created to get a single linear PCR fragment having sequence from upstream of GOI followed by Kanamycin resistance gene and sequence from downstream of GOI (fragment 1 fragment 2 fragment 3) (Shevchuk *et al.,* 2004). This fusion fragment was used for transformation. The steps for making the fusion of linear DNA fragments were (***Table 4.3***)**:**

a. Primers I and II were used to amplify fragment 1 and primers V and VI to amplify fragment 3 from ETEC genomic DNA. Using primers III and IV, the kanamycin resistance gene (fragment 2) was amplified from plasmid pKD13.

b. The fragments were gel-purified with extraction from agarose gel using Wizard® SV Gel and PCR Clean-Up System and their concentrations were recorded.

* For media and buffer composition please see appendix

c. Triple reaction step A: No primer fusion. The PCR mix was 4 µl 5X Phusion™ HF Buffer; 0.4 µl 10 mM dNTP mix (Invitrogen); 100ng fragment 1; 100ng fragment 3; 20ng fragment 2; 0.2 µl Phusion™ High–Fidelity DNA Polymerase; water to make up the volume to 20µl. Cycling parameters: initial denaturation 98°C -30 seconds, subsequent steps 98°C -10 seconds, annealing at 50°C -20 s, extension 72°C -1 minute, 10 cycles total, hold at 4°C.

d. Triple reaction step B. Primer fusion. The PCR mix was 4 µl 5X Phusion™ HF Buffer; 0.4 µl 10 mM dNTP mix (Invitrogen); 3µl of unpurified PCR product from step A; 0.2 µl primer I and 0.2 µl primer VI; 0.2 µl Phusion™ High–Fidelity DNA Polymerase; water to make up the volume to 20µl. Cycling parameters: initial denaturation 98°C 30 seconds, subsequent steps 98°C -10 seconds, annealing at 54°C -20 seconds, extension 72°C -1 minute 30 seconds, 30 cycles total, final additional extension 72°C -5 minutes, hold at 4°C.

e. The resulting fused PCR product was analyzed by electrophoresis in 1% agarose.

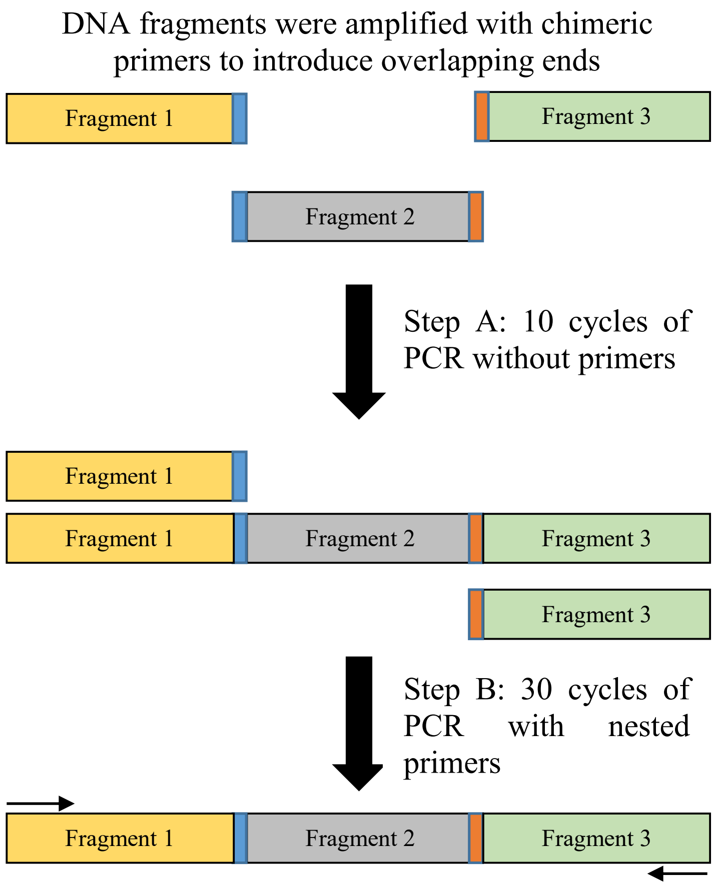

Figure 4.3: Schematic diagram of triple fusion PCR

* For media and buffer composition please see appendix

Step 3: Transformation of the PCR fused fragment by electroporation. The ETEC strain IDH00469 (now maintaining pKD46) were grown at 30˚C in LB + Ampicillin up to O.D_{600} 0.1and to it L-arabinose (Sigma-Aldrich, USA) at a final concentration of 10 mM was added. The cells were again incubated at 30˚C and continued to grow to OD_{600} = 0.4-0.5. The cells were then made electrocompetent and electroporated with linear PCR fragments generated in step 2.

4.19.1. Electroporation Protocol: The electroporation was done in BTX ECM 630 with BTX 1mm electroporation cuvettes. SOC media and LB-antibiotic plates were pre-warmed at 37°C prior to starting the electroporation. The electrocompetent cells were thawed on ice for 10 minutes or freshly made cells were used. Sterile microcentrifuge tubes and electroporation cuvettes were placed on ice. To the cold microcentrifuge tubes, 50µl cells and DNA at varying concentrations were added and mixed by flickering the tubes. For control, pUC19 was added with Milli-Q water and followed for electroporation. The DNA-cell mixture was transferred to the cold cuvette and placed in the electroporation module. The electroporator was set at 1350 V, 5 ms pulse, resistance @ 600 Ohms and capacitance @ 10 µF. Immediately, after pulse, 950 µl SOC was added and transferred to a sterile culture tube. The tube was incubated at 37°C with moderate shaking. Aliquots of the transformed culture were spreaded on LB plates supplemented with kanamycin 30 µg/mL and incubated overnight at 37°C. Next day colonies were screened for gene disruption by PCR and positive colonies were stocked and kept at - 80°C.

Step 4: Removal of the kanamycin cassette. The kanamycin cassette was removed by FLP expression. The transformed colonies were transformed with pCP20 plasmid and grown at 30˚C. Positive colonies were screened by PCR and again incubated overnight at 37 ˚C. Colonies were screened by PCR for the successful removal of the kanamycin marker or alternatively the colonies were screened by streaking on two different LB agar plates, one conditioned with 30µg/mL kanamycin, and the other one without addition of antibiotics. Clones growing only on LB agar plates without kanamycin but no longer on kanamycin supplemented LB agar plates had successfully removed the selection marker by FLP recombination.

4.19.2. Complementation of mutant:

For complementation of the mutants, pBAD-TOPO TA expression vector (Invitrogen) was used and the manufacturer's protocol was followed. PCR primers (***Table***

* For media and buffer composition please see appendix

9.3; please see appendix) were designed and the open reading frame of the gene was PCR amplified by using Platinum® Taq DNA Polymerase (Invitrogen) in a 20µl PCR reaction (***Table 4.5***). The PCR fragment was cloned into pBAD-TOPO TA expression vector and the recombinant vector was transformed into One Shot® TOP10 *E. coli* competent cells. An Insert: vector molar ratio of 3:1 was used in the TOPO® Cloning reaction (***Table 4.6***). The cloned vector was transformed into the mutant strain and the complemented strain was induced by 0.2% arabinose.

Table 4.5: Phusion PCR mix for linear fusion PCR fragments generation

Components for PCR Mixture	Volume
5X Phusion™ HF Buffer	4 µl
10 mM dNTPs	0.4 µl
Forward Primer	0.2 µl
Reverse Primer	0.4 µl
Phusion™ High–Fidelity DNA Polymerase	0.2 µl
DNA Template	2 µl
Nuclease Free Water	To make up the volume
Final Volume	20 µl

Table 4.6: pBAD-TOPO cloning mix

Reagent	Volume
Fresh PCR product	0.5 to 4 µl
Salt solution	1 µl
Sterile water	Add to a final volume of 5 µl
TOPO vector	1 µl
Final volume	6 µl

4.20. DNA sequencing:

For DNA sequencing, at first, the desired gene or gene fragment was amplified by PCR and cloned in pGEM®-T Easy Vectors (Promega, USA). The cloning was done as per the manufacturer's protocol. In short, the ligation reaction was set and incubated overnight at 4°C. The next day, the ligated product was transformed into JM109 High-Efficiency Competent Cells and each transformation culture was plated onto LB/ampicillin/IPTG/X-Gal plates. White colonies positive for the gene of interest was stored and used for plasmid isolation. The pUC/M13 Primers were used to sequence inserts cloned into the pGEM-T easy vectors. The Primer Sequences were Reverse (17mer): 5′-d(CAGGAAACAGCTATGAC)-3′ and Forward (24mer): 5′-d(CGCCAGGGTTTTCCCAGTCACGAC)-3′ from Promega, USA.

Particular primer (either forward or reverse) at a concentration of 3.2 pmole was mixed with 1µl of 10X reaction buffer and 2µl of the reaction mixture (Life technologies, NZ). To make the final volume of 10µl, the volume of water was dependent on the template. Depending on the amplicon size, the amount of template (plasmid) was modified (***Table 4.7***). The PCR conditions for sequencing were Initial denaturation - 96˚C - 2 minutes; 25 cycles at 96˚C -10 seconds, 50˚C – 5 seconds, 60˚C – 4 minutes; and final hold at 4˚C (***Figure 4.4***). After that, the PCR products were ethanol precipitated.

Table 4.7: Amount of template DNA to be used for sequencing PCR

Template size	Amount to be added (ng)
100 – 200 bp	2 – 5
200 – 500 bp	5 – 10
500 – 1000 bp	10 – 20
1 – 2 kb	20 – 50
Greater than 2 kb	40 - 100

The samples were snap-chilled in formamide to destroy secondary structures before loading into an automated ABI Prism 3200 DNA sequencer (Applied Biosystems, USA). The amplicons' identities were validated using the NCBI Blast tool.

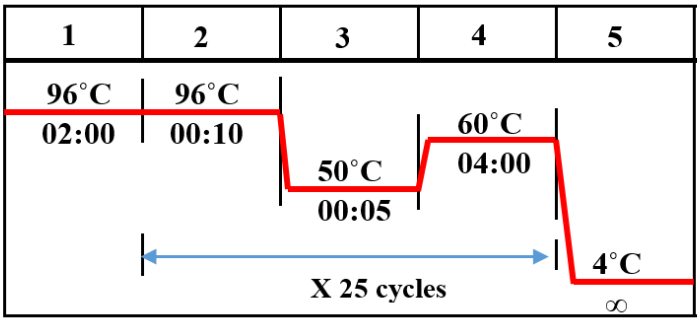

Figure 4.4: Schematic diagram of PCR conditions used for sequencing

4.21. Western blotting:

4.21.1. Protein Extraction: Protein was extracted by the sonication lysis method. First, the stationary phase bacterial culture was centrifuged for 15 minutes at 5000 x*g* and the

 * For media and buffer composition please see appendix

used media was discarded. The pellet was resuspended in Tris-HCl buffer (pH 7.0) @ 1 volume of cell: 4 volumes buffer. The solution was subjected to sonication by keeping the suspension on ice and using a number of short pulses (5-10 seconds) with pauses (10-30 seconds) in ice to re-establish a low temperature. The protein suspension was centrifuged for 15 minutes at 8000 x*g* at 4ºC. The supernatant was used as the protein extract.

4.21.2. Protein precipitation: Trichloroacetic acid (TCA) precipitation method was used to concentrate protein samples before downstream applications such as ELISA or SDS-PAGE. To the protein sample 100% TCA @ 1 volume of TCA: 4 volumes of protein sample was added and incubated on ice for 30 minutes. The samples were centrifuged at 10000 x*g* at 4ºC for 15 minutes. The supernatant was discarded and the pellet was washed with ice-cold acetone to remove any residual TCA. After centrifugation at 10000 x*g* at 4ºC for 15 minutes, the pellet was dried by placing the tubes in a 95°C heat block for 5-10 minutes to drive off acetone.

4.21.3. Protein concentration estimation by modified Lowry method (Lowry *et al.,* 1951)**:** To estimate the amount of protein in a sample we performed Lowry's method. The following reagents were used for estimating the protein concentration:

Reagent A: 2% Na_2CO_3 in 0.1 M NaOH
Reagent B: 1% Na-K tartrate in H_2O + 0.1% SDS
Reagent C: 0.5% $CuSO_4$. $5H_2O$ in H_2O
Reagent D: 48 mL of Reagent A + 1 mL of Reagent B + 1 mL of Reagent C
At room temperature, 200µl of protein sample was incubated with 600µl of reagent D for 15 minutes. The mixture was then re-incubated at 37˚C for 30 minutes with 1 N Folin reagent. At 660 nm, the OD was measured. The protein concentration was calculated using the equation: concentration = OD/0.0026, which was obtained from a BSA standard curve.

4.21.4. Polyacrylamide gel electrophoresis: The protein samples were separated by SDS-PAGE (sodium dodecyl sulphate–polyacrylamide gel electrophoresis) to estimate the purity of the protein samples. The gel percentage was determined based on the protein sample size range and the appropriate separating/resolving gel containing the desired concentration of acrylamide, using the values in Table (***Table 4.8***), was prepared. The SDS-PAGE was run in a continuous buffer system and was carried

 * For media and buffer composition please see appendix

out using vertical gels. Before loading, the sample was heated for 5 minutes at 100°C with loading dye, and the run was performed at 110 V in a gel running system (ATTO, Japan). After electrophoresis, the gel can be stained with Coomassie Brilliant Blue R-250 to see the separated proteins, or it can be processed further for western blotting. The proteins emerged as discrete bands within the gel after staining.

Table 4.8: Composition of SDS-PAGE gel

Reagents	Resolving Gel (pH 8.8)						Stacking Gel (pH 8.8)
	5%	7.5%	10%	12.5%	15%	20%	
Solution A (mL)	1.7	2.5	3.3	4.2	0.45	5	0.45
Solution B (mL)	2.5	2.5	2.5	2.5	2.5	2.5	-
Solution C (mL)	-	-	-	-	-	-	0.75
H_2O (mL)	5.8	5	4.2	3.33	2.5	0.8	1.8
APS (µl)	23	23	23	23	23	23	10
TEMED (µl)	5.5	5.5	5.5	5.5	5.5	5.5	7

4.21.5. Coomassie staining (Meyer *et al.,* 1965)**:** After SDS-PAGE, the gel was fixed for 30 minutes in 50% methanol (MeOH) and 5% glacial acetic acid (AcOH), and then stained for 1 hour at room temperature with a staining solution (50% MeOH, 50% AcOH, and 0.1% coomassie blue). The gel was destained numerous times with 5% MeOH and 7.5% AcOH to visualise the protein bands.

4.21.6. Western Blotting (Towbin *et al.,*1979)**:** For western blotting, the electrophoretic transfer of proteins from SDS-PAGE to nitrocellulose or PVDF membrane was done. The membrane, gel, and filter pads were incubated in transfer buffer (48 mM Tris; 39 mM Glycine; 0.05% SDS; 20% MeOH) for 30 minutes at room temperature to transfer the proteins from polyacrylamide gel to nitrocellulose membrane, and then the transfer cassette was assembled for semi-dry transfer as shown (***Figure 4.4***). The run lasted 1 hour and was carried out at 200 mA. To block non-specific sites, the membrane was incubated in blocking buffer overnight at 4 °C after the transfer. The membrane was cleaned three times with TBS-T the next day and the membrane was dipped in blocking buffer in which the primary antibody with the appropriate dilution was added and was incubated for 2 hours at room temperature. After washing thrice, the secondary antibody was added to TBS-T that had been labelled

 * For media and buffer composition please see appendix

with alkaline phosphatase (AP) or horseradish peroxidase (HRP). NBT (nitro blue tetrazolium)/BCIP (5-bromo-4-chloro-3-indolyl phosphate) were used as substrates for AP-conjugated antibody, and luminal and peroxide were employed as substrates for HRP conjugated antibody. For alkaline phosphate, NBT solution (7.5 mg NBT dissolved in 175 µl DMF, dimethylformamide, and 75 µl water) and BCIP (3.75 mg BCIP dissolved in 250 µl DMF) were freshly mixed with 24.5 mL bicarbonate buffer (210 mg Na_2CO_3 and 5.07 mg $MgCl_2$, pH 9.8 in 25 mL water) and incubated with the membrane in the dark until the desired bands appeared. HRP conjugated antibody detection was done by chemiluminescence which was recorded by chemi doc imaging system, Biorad, USA.

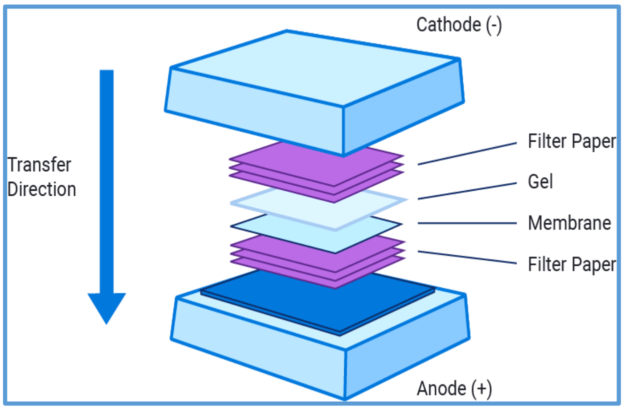

Figure 4.4: Semi dry transfer set up

4.22. Promoter construct:

Promoter regions were cloned in a no promoter β-galactosidase reporter gene containing plasmid, pTL61T (Sigma). The predicted promoter regions were PCR amplified with chimeric primers having restriction enzyme sites at their 5'ends. After amplification with the designed primers (***Table 9.4; please see appendix***) the PCR products were purified and ligated into pGEM-T Easy Vector for Blue-White screening overnight at 4°C. The ligated product was transformed in JM109 High-Efficiency Competent cells (Promega, USA). Positive colonies were screened by PCR and plasmid was isolated. The plasmid (insert) and the vector pTL61T were cleaved with 2 Restriction Enzymes following double digestion as per New England Biolabs (UK) protocol (***Table 4.9***). The enzymes used here were NEB enzymes (XmaI/ XbaI), and buffer 4 for double digestion. The reaction was incubated for either at least 1 hour or overnight at 37˚C. The digested products were purified from a 0.7% agarose gel with the QIAquick gel extraction kit. Ligation between the insert and vector was set up (***Table 4.10***) at a ratio of 6 (insert): 1

* For media and buffer composition please see appendix

(vector) and incubated overnight at 4°C. The following day, the ligated product was transformed into chemically competent *E. coli* Top10 cells. Positive colonies were screened from LB + ampicillin 100µg/mL agar plates by PCR. The fused plasmid was then isolated and transformed into desired chemically competent clinical ETEC strain. Negative control was also prepared where only pTL61T plasmid without the insert was transformed in the same as above chemically competent clinical ETEC strain. The β-galactosidase assay was done to measure the promoter activity and units were measured as Miller units. The deletion of the promoter regions was also prepared.

Table 4.9: Double digestion restriction reaction set up

Reagents	Volume
Buffer (10X)	5µl
BSA (100X)	0.5µl
Enzyme A	2µl
Enzyme B	2µl
DNA product	Xµl
Nuclease-free water to a final volume of	50µl

Table 4.10: Ligation reaction assembling set up

Reagents	Quantity
Vector DNA	17ng
Insert DNA	100ng
Ligase 10X Buffer	1µl
T4 DNA Ligase (Weiss units)	0.1–1 U
Nuclease-Free Water to a final volume of	10µl

4.23. β-galactosidase Reporter Assay (Miller *et al.*, 1984):

Beta-Galactosidase Assay (also known as the "Miller Assay") were used to determine the promoter region of the bacteria by detecting their LacZ activity. Transformed ETEC strains were incubated in LB overnight at 37˚C and the next day, bacteria were transferred to fresh M9 media with/without desired concentrations of additives and were grown to mid-log phase.

Preparation of Cells: The cells were incubated on ice for 20 minutes to stop their growth. The cells were pelleted down by centrifuging for 10 minutes at 5000 x*g* at 4˚C.

* For media and buffer composition please see appendix

The cells were resuspended in the same volume of chilled Z-buffer and the $O.D_{600}$ of the suspension was measured (blank against Z-buffer). The cells were diluted in Z-buffer with 0.5 mL cells + 0.5 mL Z-buffer to a final volume of 1 mL. The cells were permeabilized by adding 100µl chloroform and 50µl 0.1% SDS (w/v) and equilibrated in a 28˚C water bath for 5 minutes.

Assay: The reaction was started by adding 200µl o-nitrophenyl-β-D-galactoside (ONPG) and mixed well by vortexing. The time of addition was recorded. The cells were incubated at 28˚C for sufficient yellow colour to be developed. After that, the reaction was stopped by adding 500µl of 1M Na_2CO_3. The time of addition was recorded. The suspension was centrifuged at 16000 x*g* for 5 minutes to remove debris and chloroform and O.D was measured at 420 nm and 550 nm.

Calculation of units of activity: The β-galactosidase activity was expressed as Miller units.

Miller Units = $1000 \times [(OD_{420} - 1.75 \times OD_{550})] / (T \times V \times OD_{600})$

OD_{420} and OD_{550} are read from the reaction mixture.

OD_{600} reflects cell density in the washed cell suspension.

T = time of the reaction in minutes.

V = volume of culture used in the assay in mL.

The units give the change in A_{420}/min/mL of cells/OD_{600}.

4.24. Cloning PCR Products with pGEM®-T Easy Vectors:

The pGEM®-T Easy Vector Systems were purchased from Promega, USA and their protocol was followed for cloning of PCR products. Briefly, the ligation reaction was set up and mixed by pipetting (*Table 4.11*). The reaction was set for 1 hour at room temperature or alternatively, the reaction was incubated overnight at 4°C.

Table 4.11: Reaction set up for cloning in pGEM®-T Easy Vector Systems

Reagents	Standard Reaction	Positive Control	Background Control
2X Ligation Buffer	5µl	5µl	5µl
pGEM®-T Easy Vector	1µl	1µl	1µl
PCR Product	Xµl	-	-
Control Insert DNA	-	2µl	-
T4 DNA Ligase (3 Weiss units/µl)	1µl	1µl	1µl
Deionized water to a final volume of	10µl	10µl	10µl

* For media and buffer composition please see appendix

4.25. Purification of CS6 and CS5 from clinical ETEC isolates:

CS6 and CS5 were purified from the wild type ETEC strain by heat saline extraction method (Heuzenroeder *et al.,* 1989; Wolf *et al.,* 1989; Ghosal *et al.,* 2009). For CS6 isolation, bacteria were cultured in CFA broth overnight at 37 °C and suspended in 0.85% NaCl. The suspensions were then centrifuged at 8,000 x*g* for 20 minutes after being incubated at 60°C for 20 minutes with moderate shaking at 20 rpm. The heat saline extract was then precipitated with ammonium sulphate. It was first precipitated with 35% ammonium sulphate to remove impurities and then precipitated with a 60% ammonium sulphate saturation.

For CS5 isolation, bacteria were cultured in CFA broth overnight at 37 °C and suspended in phosphate buffer, pH 7.5. The suspension was subjected to heat shock at 56°C for 20 min. This cell suspension was then centrifuged (3000 x*g* for 10 min) to remove cells. The supernatant was then precipitated with 5% ammonium sulphate.

After centrifugation, the pellet of both CS6 and CS5 were dissolved in buffer A (20 mM Tris-HCl, pH 6.8). The protein sample was dialyzed in the same buffer overnight and placed onto a 3-mL Q-Sepharose FF column (GE Healthcare, Sweden) that had been pre-equilibrated in buffer A. A linear gradient of buffer B (20mM Tris-HCl and 1M NaCl, pH 6.8) was used to elute the protein attached to the anion exchange matrix at a flow rate of 1 mL min^{-1}. The fractions containing the desired proteins were combined and dialyzed in buffer A at 4 °C overnight. Step elution in buffer B concentrated the dialyzed material on a 1mL UNO Q1column (Bio-Rad, USA). Peak fractions were fed at a flow rate of 0.5 mL min^{-1} onto a Superdex 200 (GE Healthcare, Sweden) gel filtration column that had been previously equilibrated with 0.15 M NaCl in buffer A. The UNO Q1 column was used to concentrate the protein-containing pooled fractions once again.
These proteins were subsequently used for antiserum production.

4.26. Purification of EatA:

The passenger domain of EatA was purified by cloning. Using primers a 1,483-bp region of the *eatA* gene encoding a portion of the passenger domain was amplified by PCR (***Table 9.5; please see appendix***). The resulting amplicon was cloned into pET-22b(+), using the appropriate sites to create an expression plasmid, of pET-22b(+) + *eatA* that encodes a six-histidine tag-*eatA* fusion protein. To ensure that the insert of *eatA* was in-frame with the polyhistidine tag DNA sequencing of pET-22b(+) + *eatA* was performed.

 *For media and buffer composition please see appendix

The recombinant plasmid encoding a His-protein was over-expressed in BL21λDE3 cells by 0.5mM IPTG induction for 16 hours at 20°C and purified by nickel affinity chromatography. The affinity bead used was Ni-NTA Agarose and it was equilibrated in equilibration buffer (25mM Tris, 100mM- NaCl, 7M Urea, pH 7.5). The protein was eluted in elusion buffer (25mM Tris, 100mM NaCl, 7M Urea, 250 mM-Imidazole, pH 7.5) and stored in the same buffer. The eluted recombinant EatA fragment was used for the generation of polyclonal antisera.

4.27. Antiserum preparation:

Samples of both *cssB* and *csfD* were run on SDS–15% PAGE gels, and bands corresponding to these proteins were excised from the gel and were each emulsified with 0.5 mL of 1X phosphate-buffered saline (PBS). On day 0, each mouse was immunized intraperitoneally. Periodic boosters were given for 3 weeks. Each mouse was exsanguinated by cardiac puncture, and the serum obtained was stored at -20°C in the presence of 0.02% sodium azide. The specificities of the antiserum were examined by Western blot analysis.

For *eatA* antisera generation, first, the r*eatA* protein was run on 12.5% SDS-PAGE to check for purity and was used to immunize New Zealand White rabbits. Periodic boosters were given for 4 weeks with Freund's Incomplete Adjuvant. The final collection of Anti-Sera was done one week after the 4th Booster dosage of immunogenic injections and the serum obtained was stored at -20°C in the presence of 0.02% sodium azide. Polyclonal rabbit antisera were then examined by western blotting.

4.28. Quantification of CS6, CS5 and EatA:

Protein expression of these genes were detected by indirect ELISA and protein was coated which was detected by their respective anti- polyclonal antibody. The steps were:

1. **Coating antigen to microplate (Nunc-Immuno™ MicroWell™ 96 well solid plates, Sigma Aldrich, USA):** The antigen or the protein was diluted to a final concentration of 20 µg/mL in carbonate buffer. The protein was mixed with coating buffer @ 1:1. This suspension was used to coat the wells of the microtiter plate by pipetting 100 µl of the antigen dilution in the wells of the plate. The plates were incubated for 2 hours at RT or at 4°C overnight.

 * For media and buffer composition please see appendix

2. **Washing of plates:** After the incubation time the coating buffer was removed and the plates were washed thrice by PBS.
3. **Blocking:** The coated wells were blocked by adding 300µl blocking buffer and incubated for 2 hours at RT or at 4°C overnight. After incubation, the wells were washed three times with PBS-T.
4. **Incubation with primary and secondary antibody:** 100 µl of the diluted primary antibody was added to each well and incubated for 2 hours at room temperature. The wells were again washed thrice with PBS-T. 100 µl of HRP-conjugated anti-rabbit IgG secondary antibody or HRP-conjugated anti-mice IgG secondary antibody (Sigma-Aldrich, USA), diluted at the optimal concentration (according to the manufacturer) in blocking buffer was added to the wells and incubated for 1-2 hrs at room temperature. The wells were washed thrice with PBS-T.
5. **Detection:** For detection, 3, 3′, 5, 5′- tetramethylbenzidine (TMB) (BD OptEIA, USA) was used as a substrate and incubated for 15-30 minutes for optimum colour development. An equal volume of stopping solution was added and the optical density was measured at 450 nm by an ELISA reader (Bio-Rad, USA).

4.29. Quantification of STh and LT:

LT was quantified by a GM1-based ELISA assay, using an anti-LTB Ab and GM1 as coating antigen. STh was detected by indirect ELISA which was detected by Anti-ST MAb.

4.29.1. GM1-ELISA for LT: Ganglioside-GM1-ELISA was performed to detect LT (Sjöling *et al.,* 2007). ELISA microtiter (NUNC, USA) wells were coated with GM1 (0.5µg/mL in 0.06M sodium carbonate-bicarbonate buffer, pH 9.6) and incubated overnight at 4˚C. Wells were washed thrice with PBS-T and blocked with blocking buffer for 2 hours. After washing with PBS-T, 100µl of the diluted primary antibody was added to each well and incubated for 2 hours at room temperature. The wells were again washed thrice with PBS-T and 100µl of HRP-conjugated secondary antibody (Sigma-Aldrich, USA), diluted at the optimal concentration in blocking buffer was added to the wells and incubated for 1-2 hours at room temperature. After washing TMB was added and incubated for optimum colour development at RT. Stopping solution was added and O.D was measured at 450nm. The standard curve

* For media and buffer composition please see appendix

was generated using 2 fold serial dilution of purified LT and this curve was used to get unknown LT concentration.

4.29.2. Indirect ELISA for ST: For the evaluation of ST production (Rocha *et al.,* 2013), microplates (NUNC, USA) were coated with supernatant of the desired strain after required growth time diluted in 0.05 M sodium carbonate-bicarbonate buffer, pH 9.6 at 4 °C for overnight. At each step, plates were washed with PBS-T. Plates were then blocked with blocking buffer for 2 hours. 100µl of the diluted primary antibody was added to each well and incubated for 2 hours at room temperature. The antigen-antibody reaction was detected by the addition of HRP-conjugated secondary antibody (Sigma-Aldrich, USA) for 1 hour followed by the addition of TMB. O.D at 450nm was measured after stopping the reaction after colour development.

4.30. Quantification of surface expression of CS6 by ELISA:

The phenotypic expression of CS6 on the bacterial surface was quantified by ELISA assays as per Elder *et al.*, 1982. A 10^7 cfu/mL mid-log phase bacterial suspension was incubated with 0.15% SDS, washed with ethanol, and heated for 5 minutes at 100 °C. After being processed, the bacterial pellet was diluted in 100µl carbonate buffer (pH 8.2) and used for coating the ELISA plate to measure total CS6.To quantitate the surface expression, 100 µl of processed ETEC was added to the wells used for coating and incubated overnight at 4˚C. Unbound bacteria were decanted and wells were washed thrice with PBS followed by blocking in 5% (w/v) non-fat skimmed milk in PBS. After washing, the bound CS6 was determined by anti-*cssB* polyclonal antibody as the primary antibody followed by HRP-conjugated secondary antibody. Using a standard curve obtained with purified CS6 the amount of protein was determined.

4.31. Animal experiments:

4.31.1. Ethics statement

All animal experiments were conducted following the standard operating procedure as outlined by the committee for the purpose of supervision and control experiments on animals (CPCSEA), Government of India. The animal experiment protocol was approved by the Institutional Animal Ethics Committee of ICMR-National Institute of Cholera and Enteric Diseases (NICED/CPCSEA/68/GO/(25/294)/2016-IAEC/NSC-1).

 *For media and buffer composition please see appendix

4.31.2. Rabbit ileal loop assay

The rabbit ileal loop assay (De *et al.,* 1956) was performed in young New Zealand white rabbits. Rabbit ileal loop surgery was performed on 6-week-old New Zealand rabbits (male, weighing 1-2 kg) who had been fasted for 48 hours previous to the procedure. Intravenously, the general anaesthetic Ketamine-50 (35 mg kg^{-1} body weight) and the local anaesthetic Xylocaine 2% (5 mg kg-1 body weight) were given. The intestine was removed through a mid-line incision in the abdomen, and 10 cm intestinal loops with 2 cm interloops on both sides were created. The desired strains of ETEC were grown in LB or CFA broth and in each loop, 10^7 cfu/mL of bacterial cells were added. The PBS-inoculated loop was used as a negative control. After 18 hours of incubation, the fluid accumulation (FA) ratio was measured. Recovered ETEC was used for gene expression studies as required by qRT-PCR. Toxin levels were detected by ELISA in the accumulated fluid. Adherence assay was performed and tissue samples were collected for histological analysis.

4.31.3. Mouse virulence assay

Mouse infection study was performed in 4 to 5 days old BALB/c mice and weighed 3.5 ± 1.0 g (Baselski *et al.,* 1977). Six hours before inoculation the infant mice were separated from their mothers. The mice were inoculated orally with 50µl of 10^7 cfu mL^{-1} of the appropriate ETEC strain and incubated for 18 hours. PBS fed mice were used as the negative control. After incubation, the intestines were removed and FA ratio was measured and adherence assay was performed. The experiments involved 4-5 mice for each bacterial strain.

4.31.4. Fluid accumulation (FA) ratio calculation

The fluid accumulation (FA) assay was done in both rabbit and suckling mice. FA ratio in rabbit was calculated by the formula:

FA ratio = Fluid accumulation (in mL) / length of the loop (in cm)

In suckling mice, after incubation, the mice were sacrificed and their entire body weight was determined. Their entire stomach and intestines were removed, blotted dry on absorbent paper and weight of the stomach plus intestine were recorded. Then FA ratio was calculated by the formula:

FA ratio = intestinal weight/ [(whole body weight)-(intestinal weight)]

 *For media and buffer composition please see appendix

4.31.5. Adherence assay

Adherence assay was performed in both rabbits and mice. The adhered bacteria were counted by homogenizing the intestinal section (for rabbit) or the whole intestine (for mice) in 1 mL PBS (taken as crude) after washing the intestinal section with PBS. Then serial dilution was done and plated onto MacConkey agar to enumerate the viable bound ETEC.

4.31.6. Histological studies

For histological analysis, tissue samples (2 cm in length) from rabbit ileal loop experiments were taken and put in 10% neutral-buffered formalin. Tissues were paraffin-embedded and processed according to conventional procedures. Using a Leica rotary microtome, 3 to 4 µm thick sections were cut, stained with hematoxylin and eosin, and viewed under a light microscope. Photographs were obtained with a Leica DMLB microscope (Solms, Germany) fitted with a digital image system at various magnifications.

4.32. Statistical analysis:

All the experiments were conducted at least in triplicates wherever applicable. All data will be expressed as mean±SE (standard error). Data obtained from all the experiments were analyzed by ANOVA with post hoc evaluations. A level of significance was considered at *P* value < 0.05.

4.33. Safety Statement:

All experiments were conducted following the standard operating procedure as outlined by the Department of Biotechnology, Government of India Handbook and was approved by the Institutional Bio-Safety Committee (IBSC), ICMR- National Institute of Cholera and Enteric Diseases, Kolkata, India. (Approval No. NICED/BS/NSC- 004/2016 dated 27/04/2016).

4.34. Whole Genome Sequencing and Plasmid Sequencing:

4.34.1. Isolation, Qualitative and quantitative analysis of DNA: Genomic DNA and plasmid DNA were isolated from samples by Xcelgen bacterial and Plasmid DNA kit. The quality of both DNA was checked on 0.8% agarose gel (loaded 3µl). The gel was run at 110 V for 30 minutes. 1µl of each sample was used for determining concentration using *Qubit® 2.0 Fluorometer*.

4.34.2. Preparation of library: The paired-end sequencing libraries were prepared using *Truseq Nano DNA Library* prep kit. Plasmid DNA and Genomic DNA was isolated and pooled together in equal concentration before library preparation. The library preparation process was initiated with 200ng input. Pooled DNA was mechanically sheared into smaller fragments by *covaris* followed by a continuous step of end-repair where an 'A' is added to the 3' ends making the DNA fragments ready for adapter ligation. Following this step, platform-specific adapters are ligated to both ends of the DNA fragments. These adapters contain sequences essential for binding dual-barcoded libraries to a flow cell for sequencing, allowing for PCR amplification of adapter-ligated fragments, and binding standard Illumina sequencing primers. To ensure maximum yields from limited amounts of starting material, a high-fidelity amplification step was performed using HiFi PCR Master Mix.

4.34.3. Quantity and quality check (QC) of the library: The amplified library was analyzed in Bioanalyzer 2100 (Agilent Technologies) using High Sensitivity (HS) DNA chip as per the manufacturer's instructions.

4.34.4. Cluster Generation and Sequencing: After obtaining the Qubit concentration for the library and the mean peak size from the Bioanalyser profile, the library was loaded onto the Illumina platform for cluster generation and sequencing. The next-generation sequencing was performed on the Illumina platform. Paired-End sequencing allows the template fragments to be sequenced in both the forward and reverse directions. The library molecules will bind to complementary adapter oligos on the paired-end flow cell. The adapters are designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand is then used to sequence from the opposite end of the fragment. The libraries were prepared from the samples by '*Truseq Nano DNA Library preparation kit*'. The library was sequenced on '*Illumina platform*' (2 x 150 bp chemistry) to generate ~3 GB data/Sample.

4.35. Bioinformatics:

Different bioinformatics tools were used throughout this study whenever applicable. '*BLAST*' was used to check the degeneracy and cross-reactivity of primers.

 * For media and buffer composition please see appendix

'*ClustalW*' was used for multiple sequence alignment. Softberry Bacterial Promoter analysis tool '*bprom*' was used for promoter region prediction.

4.35.1. Bioinformatics Analysis for identification of Whole Genome (*Figure 4.5*): De novo assembly of high-quality Paired-End (PE) reads was accomplished using *Velvet v1.2.10* and the assembly was optimized at Kmer-85 and Kmer-83. The scaffolds were further gap filled (stretch of 'Ns' were filled with A, T, G, C bases using *PE reads*) the information using *GapCloser v1.12 software*. The refined assembly obtained was considered as the final assembly for annotation purposes. *tRNAscan-SE v2.0* was used for the identification of probable tRNA genes. The *barrnap v0.9 software* package to identify RNAs in the genome. Simple sequence repeats (SSRs) or micro-satellites (SSR) were identified in scaffold sequences of each sample with the *MISA perl script*. The assembled scaffolds obtained were subjected to gene prediction using *Prodigal v2.60*. The predicted genes samples were subjected to similarity search against NCBI's non-redundant (nr) database using the *BLASTP v2.2.30* algorithm with an e-value threshold of 1e-5. Simultaneously all the predicted genes were searched for similarity against *Uniprot, COG* and *Pfam*

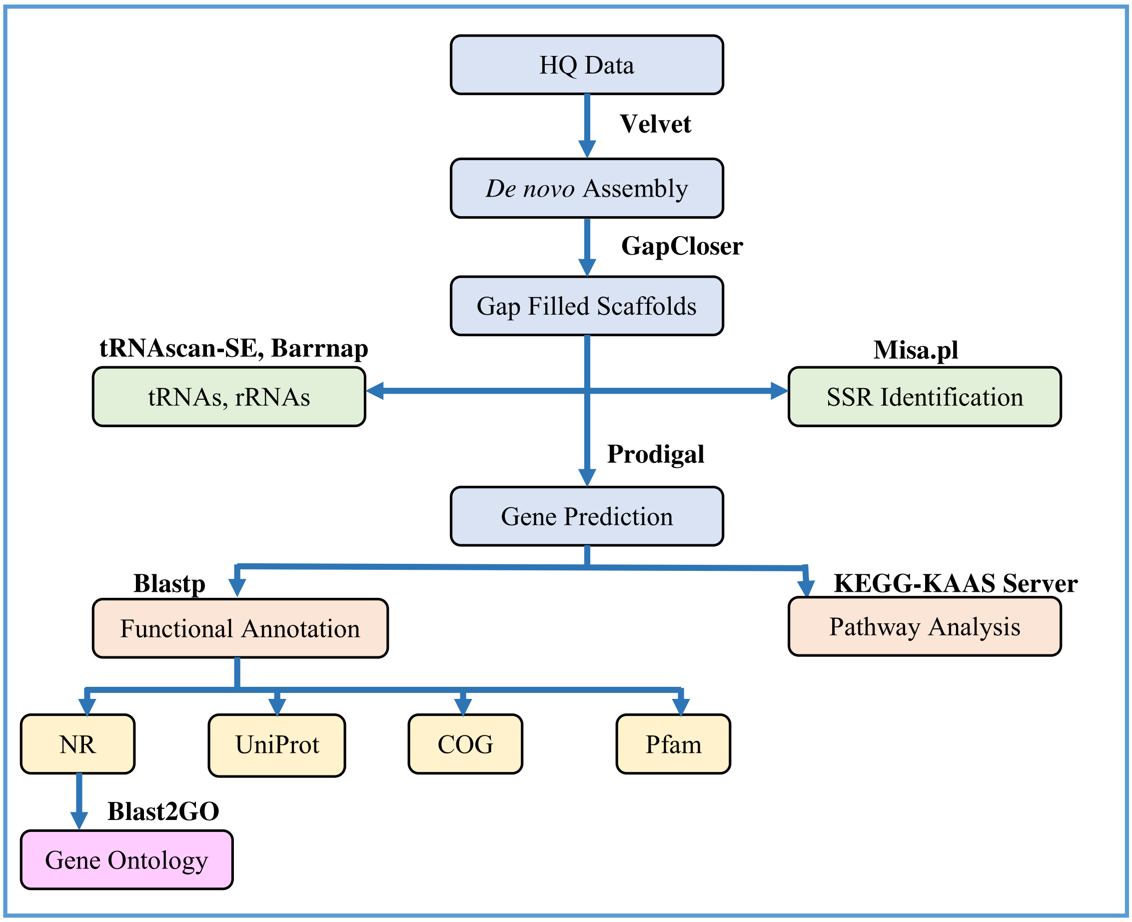

Figure 4.5: Workflow for Bioinformatics analysis of Whole Genome Sequencing

 * For media and buffer composition please see appendix

database using *BLASTP* with an e-value threshold of 1e-5. GO annotation was obtained for NR database annotated proteins using *blast2GO command line v-1.4.1*. Pathway analysis, ortholog assignment and mapping of genes to the biological pathways were performed using *KEGG automatic annotation server (KAAS) vWebserver.*

4.35.2. Bioinformatics Analysis for identification of Plasmid Genome: To achieve plasmid genome from whole-genome sequence (WGS) plasmid *SPAdes v3.13.1* was used, we further mapped samples HQ reads on plasmid genome sequences of *Escherichia coli* using *CLCgenomic workbench v6 gene prediction* using *Prodigal v2.60.* Samples were subjected to similarity search against NCBI's non-redundant (nr) database using the *BLASTP v2.2.30* algorithm with an e-value threshold of 1e-5.

* For media and buffer composition please see appendix

Chapter 5

RESULTS

Objective I:

Detection of common colonization factors and other virulence factors among clinical isolates of ETEC to understand their occurrence.

5.I. Objective I:

The objective was the detection of common colonization factors and other virulence factors among clinical isolates of ETEC to understand their occurrence. As the occurrence of the virulence factors is geographical region based and their prevalence changes from time to time we screened year-wise archived strains of ETEC and the outcome were as follows.

5.I.1. Year-wise distribution of ETEC strains

This study was conducted on ETEC strains isolated between the years 2008-2014. A total of 12243 diarrhoeal samples were analysed. Out of these samples, 350 samples (2.86%) were tested positive for ETEC. During this study period the prevalence of ETEC was 1.33% in 2008; 1.80% in 2009; 1.64% in 2010; 0.25% in 2011; 2.0% in 2012; 0.88% in 2013; and for the year 2014 it was 1.83 %. Out of the 350 ETEC isolates, 171 samples ETEC was the sole diarrhoeal pathogen, here referred to as the 'sole ETEC' and in 179 samples ETEC was present along with other pathogens, here referred to as the 'mixed ETEC'. Thirty-five ETEC strains were also collected from the community which were used as comparators.

5.I.2. Age-wise distribution of ETEC strains

The 350 ETEC strains isolated during this study were categorised in mainly two age groups below the age of 5 years and above or equal to 5 years of age. We reported that 174 (49.71 %) of the 350 ETEC isolates came from children under the age of five, while the remaining 176 (50.28 %) came from those over or equal to the age of five. There were only 17 patients belonging to the age between 5 years and 18 years. Out of 171 total sole ETEC samples, 84 (49.12%) samples were isolated from children below 5 years of age and 87 (50.88%) samples were detected from individuals belonging to ≥ 5 years of age group (***Table 5.I.1***).

In the case of 179 mixed ETEC strains, 90 (50.28%) were such samples belonging to children below 5 years of age and 89 (49.72%) samples were detected from individuals belonging to ≥ 5 years of age group (***Table 5. I.2***).

5.I.3. Distribution of Toxigenic Genes

In case samples the combination of both the toxins LT + ST were present in 61% (n = 212) specimens, followed by ST 25% (n = 89) and LT was 14% (n = 49) (***Figure 5.I.1***).

While for control samples the toxigenic genes distribution showed an opposite result; LT was present in most samples 60% (n = 21); ST in 26% (n = 9) samples; and LT + ST were present in 14% (n = 5) samples (***Figure 5.I.1***). In the case of sole samples, the combination of both the toxins LT + ST were present in 60.82% specimens, followed by ST 23.59% and LT was 15.79%. In the case of mixed samples, the combination of both the toxins LT + ST were present in 60.3% specimens, followed by ST 27.4% and LT was 12.3%.

Table 5. I.1: Distribution of Virulence genes in sole ETEC strains isolated from patients below the age of 5 years

Sole ETEC strains[1]	Age Group	
	< 5 year (n=84)	≥ 5 year (n=87)
***elt* gene**	9	18
***est* gene**	21	19
***est + elt* gene**	54	50
At least one VF gene present[2]	68	56
No VF gene present	16	31
Classical CF[3] present	56	41
Classical CF absent	28	46
Non-classical VF[4] present	54	50
Non-classical VF absent	30	37

[1]ETEC isolated from diarrhoeal patients
[2]Classical colonization factor (CF) genes or non-classical virulence factor (NCVF) genes
[3]Classical CF genes studied here by PCR; please see materials and methods for details
[4]Non-classical VF genes studied here by PCR; please see materials and methods for details

Table 5.I.2: Distribution of Virulence genes in mixed ETEC strains isolated from patients above or equal the age of 5 years

Mixed ETEC strains[1]	Age Group	
	< 5 year (n=90)	≥ 5 year (n=89)
***elt* gene**	11	11
***est* gene**	20	29
***est + elt* gene**	59	49
At least one VF gene present[2]	70	56
No VF gene present	20	33
Classical CF[3] present	54	45
Classical CF absent	36	44
Non-classical VF[4] present	61	40
Non-classical VF absent	29	49

[1]ETEC isolated from diarrhoeal patients
[2]Classical colonization factor (CF) genes or non-classical virulence factor (NCVF) genes
[3]Classical CF genes studied here by PCR; please see materials and methods for details
[4]Non-classical VF genes studied here by PCR; please see materials and methods for details

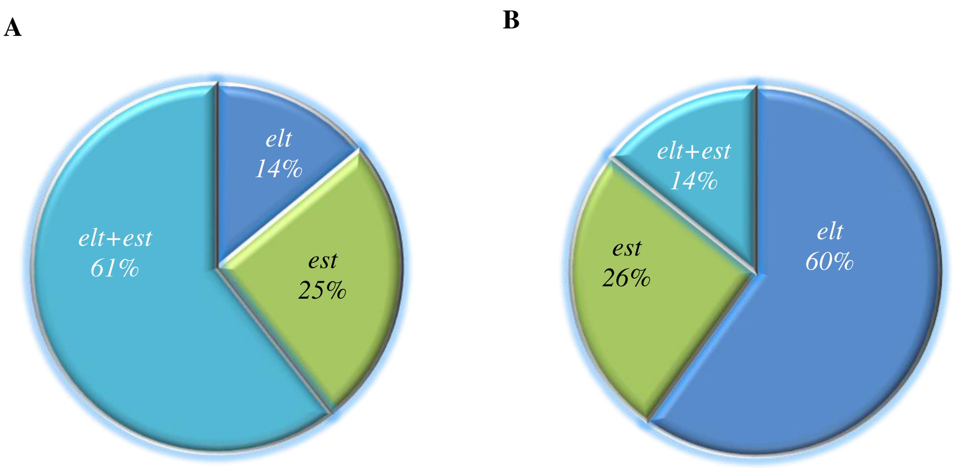

Figure 5.I.1: Graphical representation of distribution of enterotoxigenic genes across the clinical and community based control strains (A – Clinical case samples; B – Community based control samples)

5.I.3.1. Toxin genes identified over time in ETEC isolates

The distribution of toxin genes among the ETEC strains isolated during the period (2008-2014) was studied for their year-wise variation (*Table 5.I.3*). The number of strains harbouring both the enterotoxins (*est* and *elt*) was nearly double that of strains harbouring either of the toxin genes in 2008. The number of *est* positive strains was larger than the number of *est+elt* or *elt* positive strains between 2009 and 2011. In 2012 and 2013, the number of strains harbouring either of the enterotoxin genes was quite low, and *elt+est*

Table 5.I.3: Distribution of Virulence genes throughout the study period

Sole ETEC strains[1]	2008 (n=15)	2009 (n=25)	2010 (n=28)	2011 (n=4)	2012 (n=35)	2013 (n=20)	2014 (n=44)
***elt* gene**	5	6	8	1	0	0	7
***est* gene**	5	8	13	3	4	0	7
***est* + *elt* gene**	5	11	7	0	31	20	30
At least one VF gene present[2]	12	15	24	4	22	14	33
No VF gene present	3	10	4	0	13	6	11
Classical CF[3] present	9	13	20	3	17	10	25
Classical CF absent	6	12	8	1	18	10	19
Non-classical VF[4] present	10	13	18	2	20	11	30
Non-classical VF absent	5	12	10	2	15	9	14

[1]ETEC isolated from diarrhoeal patients
[2]Classical colonization factor (CF) genes or non-classical virulence factor (NCVF) genes
[3]Classical CF genes studied here by PCR; please see materials and methods for details
[4]Non-classical VF genes studied here by PCR; please see materials and methods for details

strains were prevalent. It's worth noting that in 2011, just ten ETEC strains were isolated. In 2014, strains harbouring only *est* and only *elt* toxin reappeared. Since 2012, however, bacteria containing the *est+elt* toxin gene have been identified in greater numbers than those carrying the other toxin genes. In 2012, double toxin genes were reported in 89 % of ETEC strains, nearly 100% in 2013, and 64% of total isolates in 2014. In 2012, no *elt*-positive strains were detected, and just one *elt*-positive strain was found in 2013. Strains harbouring *est* gene were hard to come by in 2012, and none were accessible in 2013.

5.I.3.2. Toxin genes in ETEC isolates based on age

In all age groups, both enterotoxins harbouring ETEC isolates were found in greater numbers than isolates with only either of the enterotoxin. In the age categories <5 years and ≥5 years, the *est+elt* harbouring ETEC strains were 65% and 56%, respectively. When equated to the age group ≥5 years (17%), the number of *elt*-carrying strains in the <5 years age group was lower (11%). For children below 5 years of age, *est* presents in higher sample numbers than *elt*. Although *est*-containing strains were more abundant than *elt*, the proportion of *est*-containing strains was both younger and higher age groups are similar (24% and 27% respectively) (***Table 5.I.1; Table 5.I.2***).

5.I.4. Distribution of classical and non-classical virulence genes

Among the 350 ETEC isolates, there were 250 strains (71%) in which at least one of the tested virulence genes was present. Both the classical and non-classical virulence genes were detected in 151 (43%) ETEC isolates. We detected only classical virulence factors (CF) in 45 (13%) ETEC strains and only non-classical virulence factors (NCVF) in 54 (15%) ETEC strains. There were 100 strains (29%) negative for all the virulence genes that were studied (***Figure 5.I.3***). There were 196 strains (56%) that were positive for the CFs. There were 205 (59%) ETEC strains that contain NCVFs genes.

In the case of sole ETEC strains, there were 47 (27%) strains that had no virulence gene; 77 (45%) strains had both classical and non-classical virulence genes; 27 (16%) strains were negative for classical virulence genes, and 20 strains (12%) were negative for non-classical virulence genes.

Among the mixed strains, 53 (30%) had no virulence gene; 74 (41%) possessed both classical and non-classical virulence genes; 27 (15%) were negative for classical virulence genes, and 25 (14%) were negative for non-classical virulence genes.

For control ETEC isolates there were 21 strains (60%) which possessed at least one virulence determinants out of which both classical and non-classical genes were harboured by 11 strains (31%); only CF genes were harboured by 2 strains (6%); and only NCVF genes were harboured by 8 strains (23%). There were 14 strains (40%) that were negative for the virulence determinants that were investigated in this study (***Figure 5.I.3***).

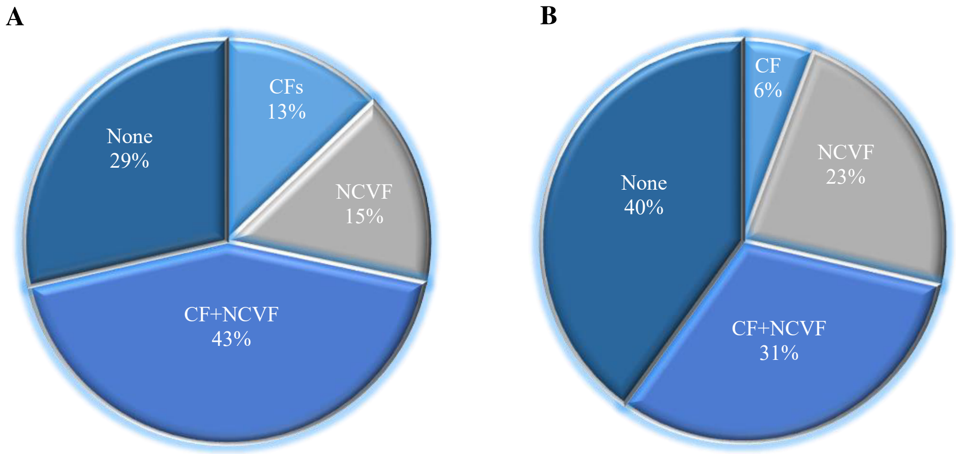

Figure 5.I.3: Graphical representation of distribution of classical and non-classical virulence genes across the clinical and community based control strains (A – Clinical case samples; B – Community based control samples)

5.I.4.1. Classical colonization factor (CF) gene circulation in ETEC isolates

The most common CFs harboured were CS21 (37%) and CS6 (36%) among the 11 CFs examined in this investigation for detection. CS5 (24%) and CFA/I (18%) were two other frequent CF genes. Among the least detected CFs were CS4 (3%) and CFA/III (2%). There were 72 ETEC strains that harboured CS21 and among them only 20 (10%) strains had CS21 as their only CF and other strains had other CF(s) associated with CS21 (27%). CS21 was mainly linked with CS1, CS3, CFA/I, and CS17. Among the 71 CS6 harbouring strains, 23 (12%) strains had CS6 alone and 48 (24%) strains had others CFs along with CS6. CS5 was found in 92% of CS6-positive ETEC strains, indicating that the two genes co-occur more frequently in clinical ETEC isolates. Only 4 ETEC isolates, however, had CS5 genes as their only CF. Only 8 strains carried extra CF genes in addition to CS6+CS5. In our settings, CFA/I was found in 36 (18%) strains.

Among the sole ETEC, the most prevalent CF was CS6 (35.45%) in all strains where CF genes were found, followed by CS21 (32.72%), CS5 (26.36%), and CS3 / CFA/I

(16.36 %). CS4 (3.6%) and CS14 (1.8%) were the least prevalent CFs found, whereas CFA/III was not found in any of the samples (***Figure 5.I.4***). In 20 samples, CS6 was found in conjunction with CS5 (18.18%). In 5 samples, CS21 was the sole virulence gene, whereas, in 29 samples, it was found in conjunction with other conventional and non-classical virulence genes.

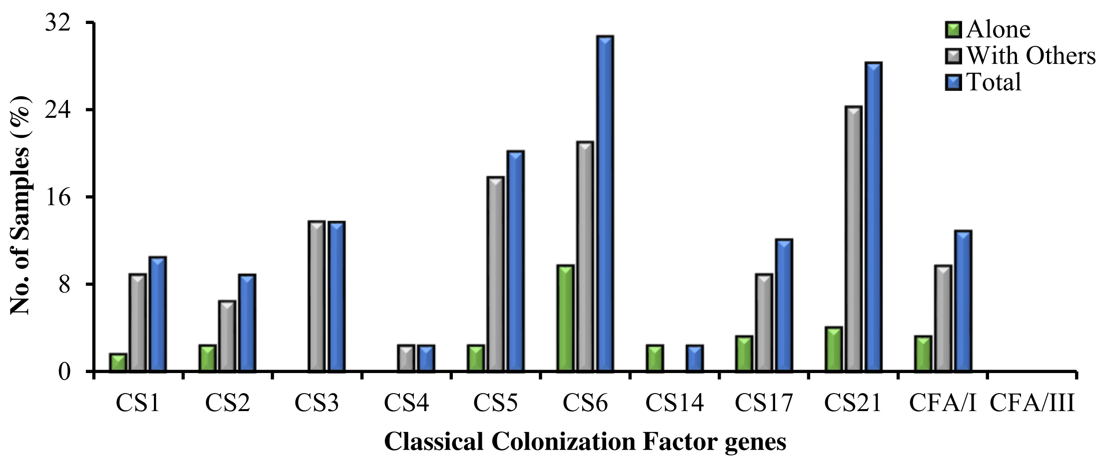

Figure 5.I.4: Classical colonization factors (CFs) allocation amongst Sole ETEC variants

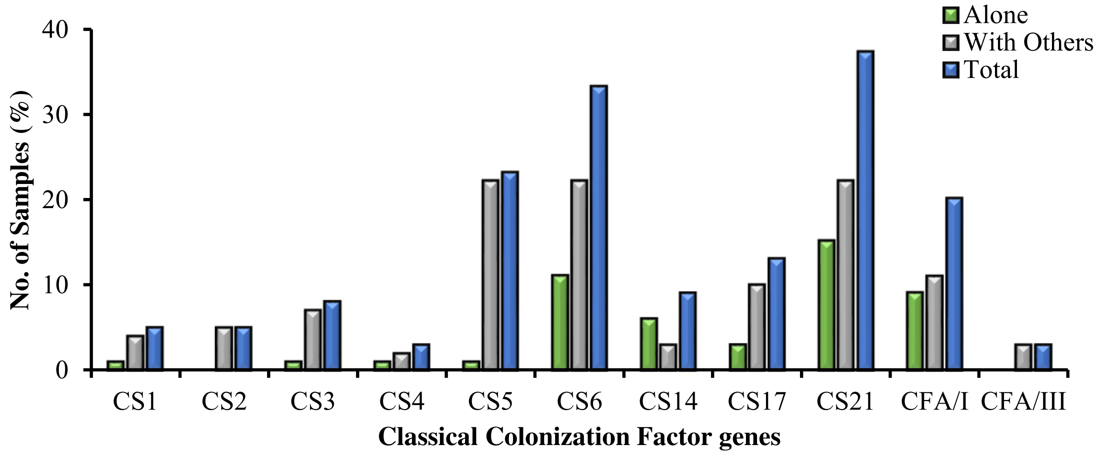

Figure5.I.5: Classical colonization factors (CFs) allocation amongst Mixed ETEC variants

The most prevalent CF was CS21 (37%) in all mixed strains where classical virulence genes were found, followed by CS6 (33%), CS5 (23%), CFA/I (20%), and CS17 (13%). CS14 (1.8%) and CFA/III (3%) were the two least common CFs found (***Figure 5.I.5***). Genes are more likely to be found in combination with other classical or non-classical genes than when they are present alone. In 5 samples, CS21 was the sole virulence gene, whereas, in 29 samples, it was found in conjunction with other virulence genes. In 22 (22%) samples, CS6 was found in conjunction with CS5.

For the control samples, the predominant CFs was CS6 and CS5 and the least predominant were CS4 and CS14. No ETEC were harbouring CS17 and CFA/III (***Figure 5.I.6***).

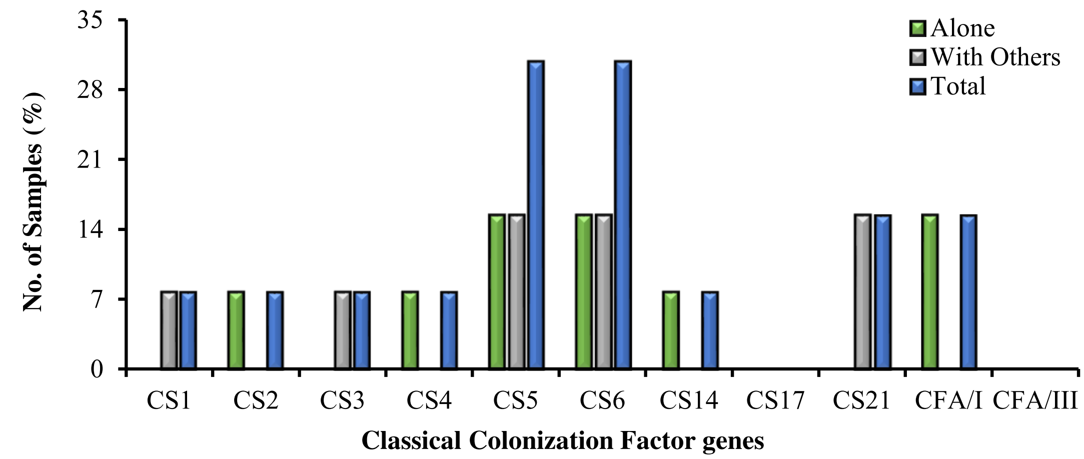

Figure 5.I.6: Classical colonization factors (CFs) allocation amongst control ETEC variants

5.I.4.2. Detection of Colonization Factor genes in ETEC isolates in the different age categories

There were 196 ETEC isolates positive for CFs, out of that 110 (56%) were isolated from children under the age of five, and 86 (44%) were from patients above or equal to the age of five. The CS21 (35%) and CS6 (34%) carrying strains were diagnosed more

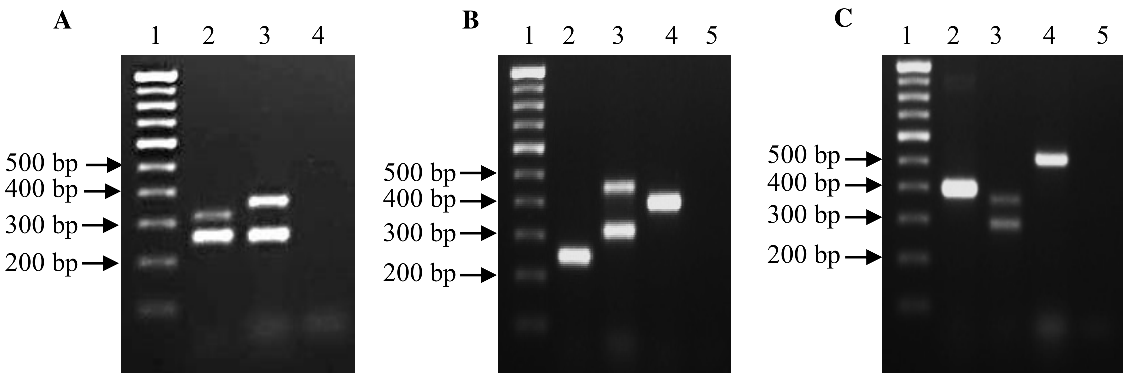

Figure 5.I.7: Representative picture of multiplex PCR assay results for detection of CFs from ETEC strains.
A – Lanes: 1 – DNA marker (100 bp); 2 – CS1 (324 bp) and CS3 (264 bp); 3 – CS2 (368 bp) and CS3 (264 bp); 4 – negative control (no template).
B – Lanes: 1 – DNA marker (100 bp); 2 – CS4 (250 bp); 3 – CS5 (453 bp) and CS6 (321 bp); 4 – CS21 (400bp); 5 – negative control (no template).
C – Lanes: 1 – DNA marker (100 bp); 2 – CS14 (394 bp); 3 – CFA/I (364 bp) and CS17 (290 bp); 4 – CFA/III (438 bp); 5 – negative control (no template).

than other colonization factor genes in the age range of 5 years. CS6 (40%) and CS21 (40%) were also often discovered in the ≥5-year-old age group. With 25% of the younger age group and 23% of the older age group having CS5, it was the third most frequent CF. In both age groups, low levels of CS4, CS14, and CFA/III were found. CFA/I (13%) was relatively high in the age category of <5 years than in the age category of ≥5 years (7%) (***Table 5.I.1; Table 5.I.2***). The dependent variables were determined to be age and CF (P = 0.001).

5.I.4.3. Detection of Colonization Factor genes in ETEC isolates during the time period

CS6 was diagnosed in around 41% of ETEC strains in 2008, 35% of ETEC strains in 2009, 32% of ETEC strains in 2010, 57% of ETEC strains in 2011, 24% of ETEC strains in 2012, 29% of ETEC strains in 2013, and 49% of ETEC strains in 2014. In the years 2010, 2012, and 2013, CS21 was discovered more frequently (34%, 43% and 50% strains respectively). One sample in 2008 and two isolates in 2012 were found to contain the infrequently observed colonization factor gene CFA/III. In 2012, four CS4-containing strains were identified, with one each in 2009 and 2014. From 2012 onwards, CS6 was found less frequently, whereas CS21 was discovered more frequently. In previous years of research, however, we found more CS6-containing strains than CS21-containing bacteria.

5.I.4.4. Non-classical virulence factor (NCVF) gene circulation in ETEC strains

NCVFs were present in 205 (59%) ETEC strains. Among the NCVF positive strains, *eatA* (65%) and *etpA* (51%) were the most commonly detected virulence factors. Following these, *tibA* (18%) and *tia* (16%) were the most common VF genes. In our study, the *leoA* gene was the least frequently found NCVF gene (3%). In comparison to other non-classical virulence factors, *eatA*-positive ETEC bacteria were mostly identified in samples with at least one classical CF.

In 27% of CF-negative ETEC strains *etpA* was present. For sole ETEC strains where non-classical virulence genes were present, *eatA* (67.47%) was the predominant virulence gene; followed by *etpA* (46.34%). The least common non-classical gene was *leoA* (3.25%) (***Figure 5.A.8***). As for the mixed ETEC isolates the predominant non-classical virulence genes was *eatA* (64%); followed by *etpA* (57%). *leoA* was the least

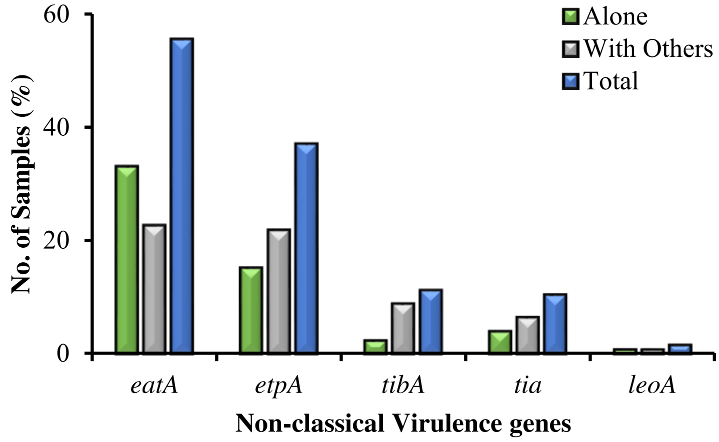

Figure 5.I.8: Non-classical virulence factors (NCVFs) allocation amongst Sole ETEC variants

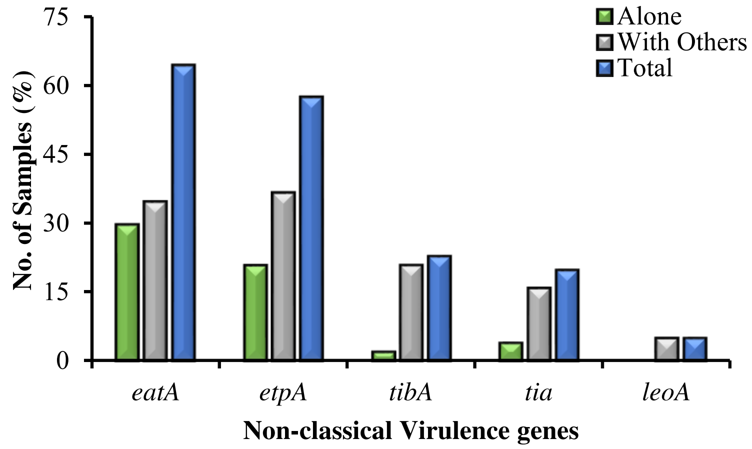

Figure 5.I.9: Non-classical virulence factors (NCVFs) allocation amongst Mixed ETEC variants.

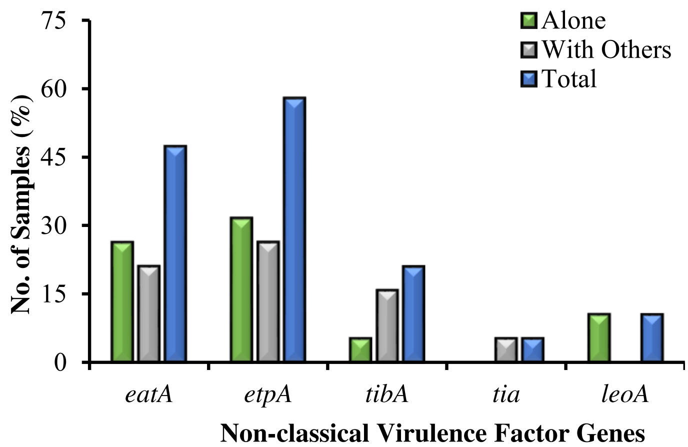

Figure 5.I.10: Non-classical virulence factors (NCVFs) allocation amongst control ETEC variants

common non-classical gene (5%) (***Figure 5.A.9***). Here genes are more in combination than present in alone. Among the comparator control strains, 19 samples contained the

NCVF gene. Among the NCVF present strains *etpA* (58%) was the most frequently occurred NCVF, followed by *eatA* (47%). Here the least detected NCVF was *tia* (5%). The other NCVF genes detected were *tibA* (21%) and *leoA* (11%) (***Figure 5.I.10***). There was no significance in distribution between NCVF and clinical strains.

5.I.4.5. Detection of Non-classical Virulence Factor genes in ETEC strains in different age

From children belonging to the age group < 5 years, 115 (56%) ETEC strains were identified, while ETEC samples from the age group ≥5 years accounted for 90 (44%) cases. In all ages (***Figure 5.I.12***), *eatA* was the most common NCVF, followed by *etpA*, with no age group prediction. *tibA* was found in higher numbers in strains isolated from children under the age of five. In both age categories, the distribution of *tia* was

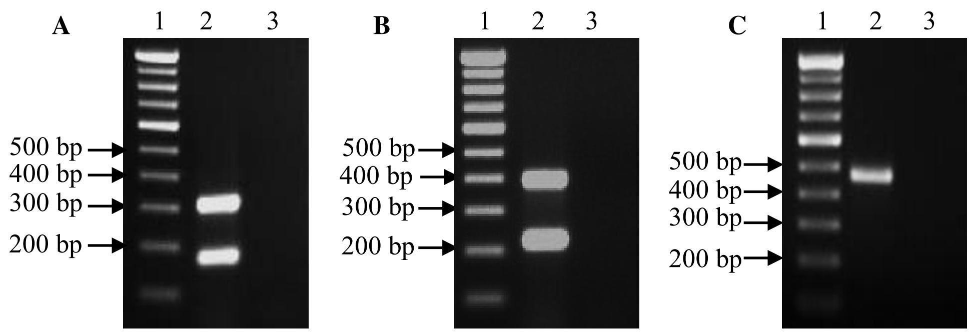

Figure 5.I.11: Representative picture of multiplex PCR assay results for detection of NCVFs from ETEC strains.
A – Lanes: 1 – DNA marker (100 bp); 2 –*tia* (172 bp) and *leoA* (315 bp); 3 – Negative control (no template)
B – Lanes: 1 – DNA marker (100 bp); 2 – *etpA* (221 bp) and *tibA* (396 bp); 3 – Negative control (no template)
C – Lanes: 1 – DNA marker (100 bp); 2 – *eatA* (465 bp); 3 – Negative control (no template)

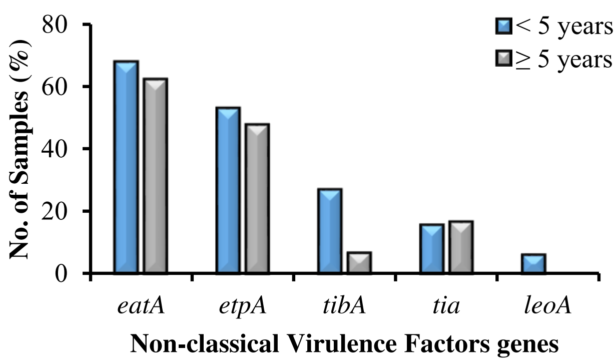

Figure 5.I.12: Distribution of Non-classical Virulence factors (NCVFs) in different age category

essentially the same. All 7 *leo*-containing ETEC strains were obtained from children under the age of five.

5.I.4.6. Detection of Non-classical Virulence Factor genes in ETEC strains during the time period

During the study period between 2008 to 2014 *eatA* was the predominant NCVF followed by *etpA*. In 2013 and 2014, the numbers of *eatA* and *etpA* were found to be about identical. *LeoA* was detected from 2011 onwards and was the least detected NCVF in all the years. *Tia* and *TibA* were present more or less in similar numbers throughout the study period except in 2011 where no *tibA* was detected.

5.I.5. Virulence factor genes in association with toxin genes

In 350 ETEC clinical isolates, the distribution of CF genes and non-classical VF genes was examined in addition to the toxin genes. The clinical strains and the classical colonization factors were shown to be associated (P = 0.033), while the NCVF genes were not (P = 0.624).

In 23 (12%) out of the 196 clinical ETEC strains with CF genes had the *elt* gene, 45 (23 %) had the *est* gene, and 128 (65%) contained both the enterotoxin genes. Thirteen of the 35 asymptomatic control ETEC strains had the CF genes, out of that 5 (38%) had the *est* gene, 4 (31%) had the *elt* gene, and 4 (31%) had both the enterotoxin genes. There exists a correlation between clinical strains harbouring the classical virulence genes and toxin genes (P = 0.031) as well as the colonization factor genes (P = 0.031). In clinical isolates, however, there was no affiliation between enterotoxins and CF genes (P = 0.114).

In the 205 NCVF-containing ETEC strains, the *elt* gene was found in 21 (10%) strains, the *est* gene in 58 (28%) strains, and the *elt+est* toxin genes in 126 (62%) strains. In the control, ETEC strains there were 19 strains in which NCVF genes were present. Amongst these 19 strains, *elt* gene was harboured by 10 strains (53%); *est* gene was harboured by 5 (26%) strains, and *elt+est* gene was harboured by 4 strains (21%). An affiliation was present between clinical strains and enterotoxin genes (P = 0.036), but none was between clinical strains and NCVF genes (P =0.624). In clinical isolates, however, there was a correlation between the enterotoxin genes and NCVF genes (P = 0.036). In addition, in clinical isolates, an association was found between CFs and NCVFs (P < 0.01).

Among the 100 strains where no virulence factors were found *elt* gene was present in 18 (18%) isolates, the *est* gene was present in 19 (19%) isolates and *elt+est* toxin genes were present in 63 (63%) isolates.

5.I.6. Genotypic profile of frequently circulating ETEC strains during the period of study

The profile of virulence factors among the most common circulating strain was *est+elt*+CS6+CS5+*eatA* (11%), followed by *est+etpA* (5%), *elt*+CS6 (3%), *elt+est+etpA* (3%), and *est+elt*+CS21+CS3+*etpA* (2%). ETEC samples that contained CS21 were also widespread, however these strains came in a wide range of combinations.

5.I.7. Association of different Diarrhoea type

We classified the diarrhoea associated with hospitalized patients into 5 categories – watery stool, loose stool, stool with blood, stool with mucus and stool with both blood and mucus. In both the sole and mixed ETEC strains watery stools were present in the maximum number of patients, 52.05% for sole and 59.78% for mixed strains followed by loose stools in both the cases; 36.26% for sole and 35.19% for mixed strains. For sole strains, there was no ETEC isolated from patients having stool with blood diarrhoea whereas only 1 patient was from the mixed strains. There was very few reported occurrences of stool with mucus and/or blood for diarrhoea occurred by both sole and mixed ETEC isolates.

In 196 samples where patients had watery stool, *elt* detected in 28 (14%) isolates, *est* detected in 55 (28%) isolates and *elt+est* detected in 113 (58%) isolates. Among the 125 samples which were isolated from patients having loose stool 19 (15%) harboured

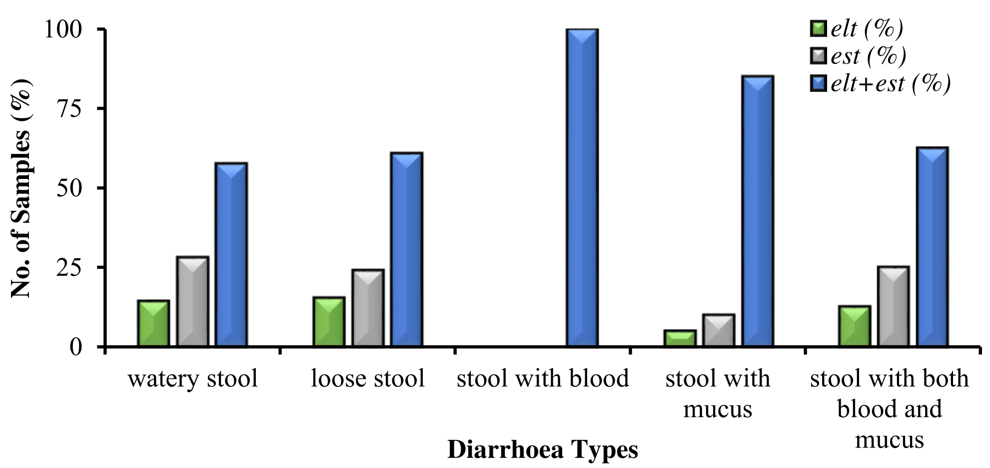

Figure 5.I.13: Distribution of toxin genes across different types of diarrhoea

elt, 30 (24%) harboured *est* and 76 (61%) harboured *elt+est* as their toxin genes (***Figure 5.I. 13***).

When we analysed the diarrhoea type with respect to age we found that patients belonging to below 5 years age category mainly had loose stool (58%) whereas patients belonging to ≥5 years age category mainly had watery stool (81%) (***Figure 5.I.14***). All the cases of stool with blood and stool with mucus were from children belonging to the below 5 years age group. Patients suffering from stool with both blood and mucus belonged to both age groups.

For 196 strains isolated from patients having watery stool, there were 60 (31%) strains that do not harbour any of the tested virulence genes. Among the CFs, the most common was CS6 (22%), followed by CS21 (18%), CS5 (13%) and CFA/I (7%). Among the NCVF genes, the most common was *eatA* (38%) followed by *etpA* (29%). There were 125 strains from patients with loose stool, out of which 26% strains were negative for all the tested virulence genes. Among the CFs, the most common was CS21 (22%), followed by CS6 (17%), CS5 (14%) and CFA/I (14%). Among the NCVF genes, the most common was *eatA* (38%) followed by *etpA* (30%).

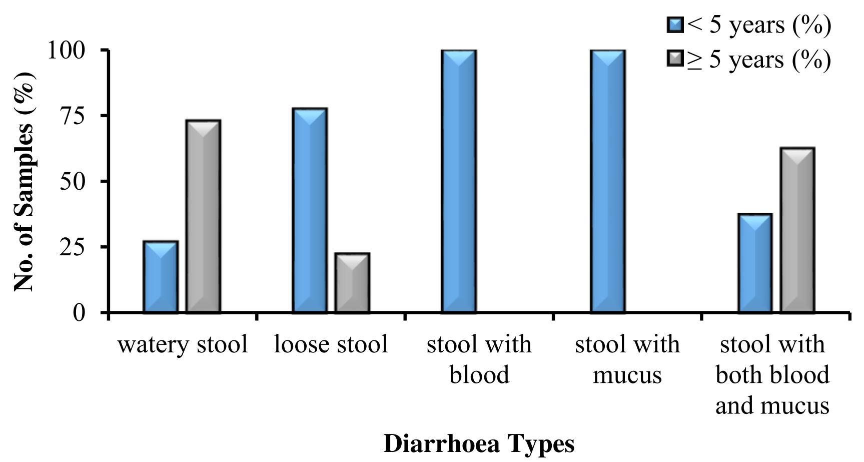

Figure 5.I.14: Different types of diarrhoea in different age groups

Objective II:

To delineate the expression of identified colonization factors and virulence factors in clinical ETEC.

5.II. Objective II:

The objective was to delineate the expression of identified colonization factors and virulence factors in clinical ETEC. All the present virulence factors do not express their respective RNA and to study their expression in lab conditions this objective was done and the outcomes were as follows.

5.II.1. Expression pattern of Sole ETEC strains

The strains that were positive for at least one of the virulence genes i.e. CFs and NCVFs, were checked for their expression of the virulence genes among the sole ETEC isolates. Among the 97 strains which had classical genes present, 81 strains have expressed their respective classical genes. Sixteen strains lacked the ability to express their classical genes. CS6 was expressed in 34 of the 38 strains that have it. CS21 was expressed in 27 of the 35 strains, while CS5 was expressed in 23 of the 25 strains (***Figure 5.II.1***). Non-classical virulence factor genes were expressed in 75 out of the 104 ETEC

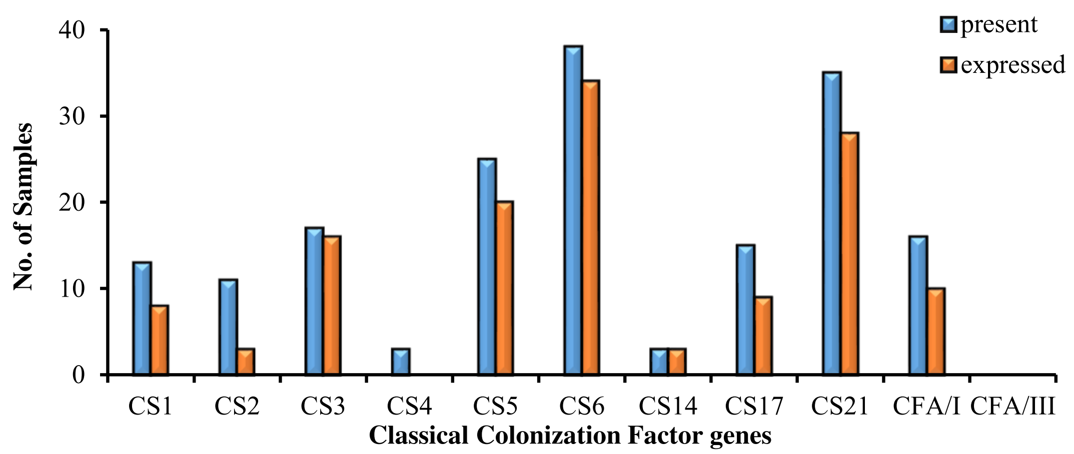

Figure 5.II.1: Distribution of the Presence and Expression of Classical colonization factors (CFs) among the sole clinical strains

isolates, while the genes were not expressed in 29 strains. EatA was expressed in 40 strains out of 69 strains, while EtpA was expressed in 40 strains out of 44 strains (***Figure 5.II.2***). There were 3 strains in which neither of the virulence factors was expressed in the expression profile.

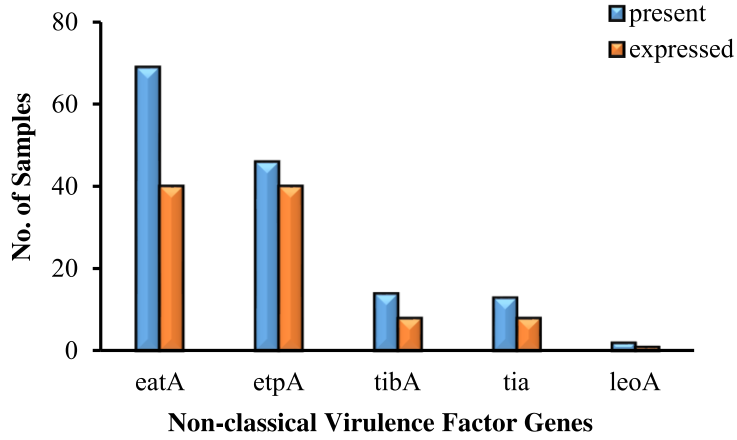

Figure 5.II.2: Distribution of the Presence and Expression of Non-classical Virulence Factors (NCVFs) among the sole clinical strains.

5.II.2. Expression pattern of Mixed ETEC strains

In the mixed strains 67 strains expressed their virulence genes out of 99 strains which were positive for classical colonization factors genes, while 32 strains did not express their CF genes. CS6 was expressed in 26 of the 33 strains that have it. CS21 was expressed in 18 of the 37 strains, while CS5 was expressed in 18 of the 23strains (***Figure 5.II.3***). Out of the 101 strains positive for NCVF genes, 60 strains expressed their respective genes and 41 strains do not express their NCVF genes. EatA was expressed in 40 strains out of 65 strains, while EtpA was expressed in 29 strains out of 58 strains (***Figure 5.II.4***). There were 8 strains in which neither of the virulence factors was expressed in the expression profile.

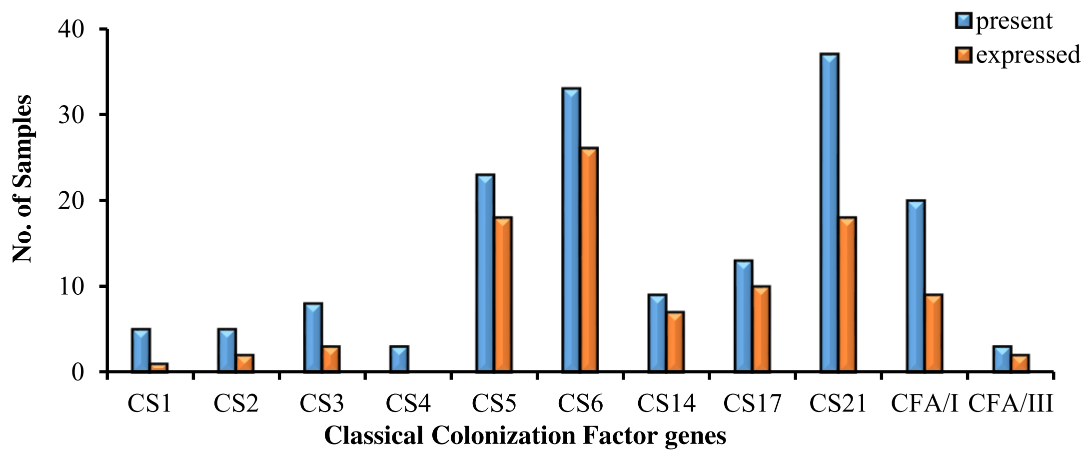

Figure5.II.3: Distribution of the Presence and Expression of Classical colonization factors (CFs) among the mixed clinical strains

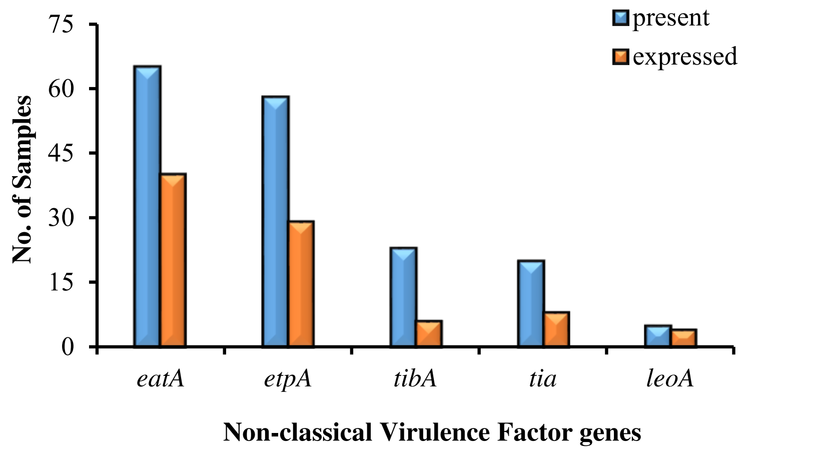

Figure 5.II.4: Distribution of the Presence and Expression of Non-classical Virulence Factors (NCVFs) among the mixed clinical strains.

5.II.3. Expression pattern of Control ETEC strains

In the community control strains, 11 of the 13 classical colonization factors were expressed, while 12 of the 19 non-classical genes were expressed. All the strains harbouring CS6, CS5 and CS21 expressed their genes (***Figure 5.II.5***). EatA was expressed in 7 strains out of 9 strains and EtpA was expressed in 7 out of 11 strains (***Figure 5.II.6***).

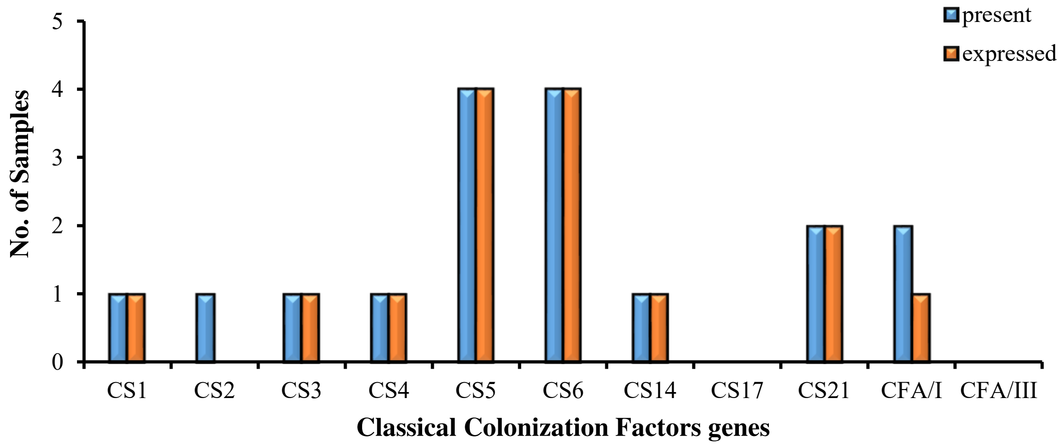

Figure 5.II.5: Distribution of the Presence and Expression of Classical colonization factors (CFs) among the control clinical strains

5.II.4. CS6 expression profile

However, when we looked into the expression of these genes, not all were expressed. Since CS6 was prevalent with respect to gene expression, we focused our study on CS6 (***Table 5.II.1***). When we looked into the expression level of CS6 we found that not all strains expressed CS6 at the same level (***Figure 5.II.7***). So, we have strains that express

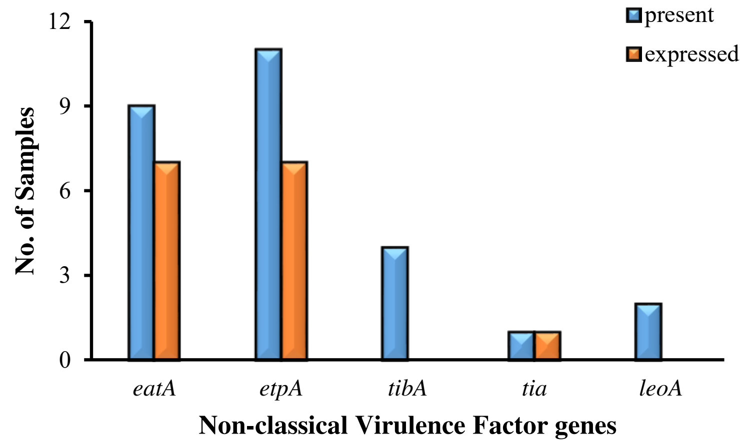

Figure 5.II.6: Distribution of the Presence and Expression of Non-classical Virulence Factors (NCVFs) among the mixed clinical strains

CS6 differentially as well as strains that have CS6 genes but do not express the gene. The strains which don't express their genes were confirmed by SDS-PAGE and ELISA (*Figure5.II.8*). The different media used for culturing does not influence the non-expressive nature of CS6. To see if the lack of CS6 expression and differential expression of CS6 was related to mutations, the promoter region was sequenced. Upon sequencing we found only single-base variations in the promoter region at position -208; -157; -137; -135; -132; -111; -48 (*Figure 5.II.9*).

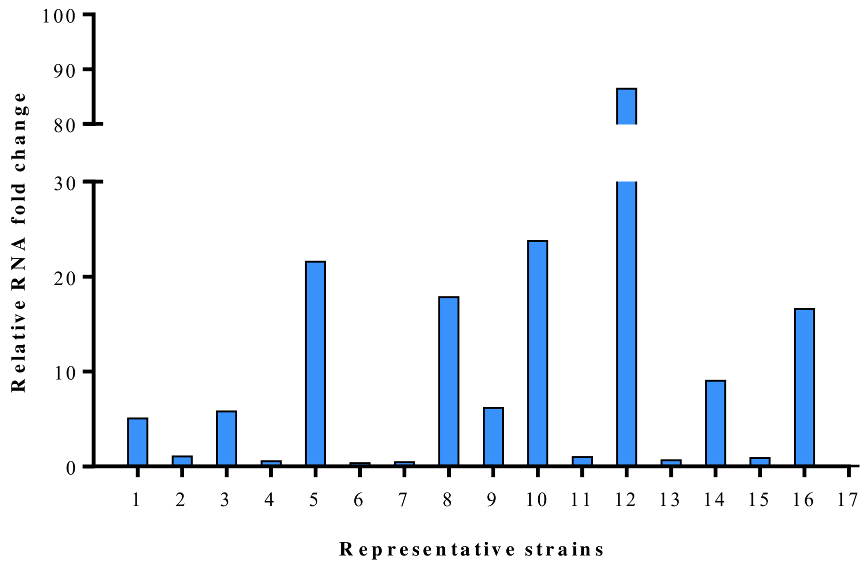

Figure 5.II.7: Differential Expression profile of different representative CS6 strains

Table 5.II.1: Expression profile of CS6 harbouring strains

CF	NCVF	Present	Expressed					
			CS6	CS5	*eatA*	*tia*	*etpA*	Total
CS6	*None*	13	+					8
			—					5
CS6	*eatA*	8	+		+			5
			+		—			3
CS6	*tia*	1	—			—		0
CS6	*etpA*	1	—				+	1
CS6 + CS5	*None*	2	+	+				1
			—	—				1
CS6 + CS5	*eatA*	29	+	+	+			23
			+	+	—			4
			+	—	+			2
CS6 + CS5	*eatA + tia*	3	+	+	+	—		2
			+	+	—	+		1
CS6 + CS5	*eatA + etpA*	2	+	+	+		—	1
			+	—	—		+	1
CS1+CS5+CS6+CFA/I	*eatA*	1	+	+	+			1
CS1 + CS3 + CS5 + CS6 + CS21	*eatA + etpA*	1	+	+	—		—	1
CS1 + CS3 + CS6	*etpA*	1	+	+			+	1
CS5 + CS6 + CFA/I	*eatA*	2	+	+	+			2
CS6 + CS4	*eatA*	1	+		+			1
CS6+CS5+CS17	*eatA + etpA+tia*	3	+	+	+	+	+	2
			—	—	—	—	—	1
CS6 + CS17	*eatA + etpA*	1	+		+		+	1
CS6 + CFA/1	*tia*	1	+			+		1
CS6 + CS5 + CS21	*eatA*	1	+	—	+			1

'+' indicates expression of that gene
'—' indicates no expression of that gene

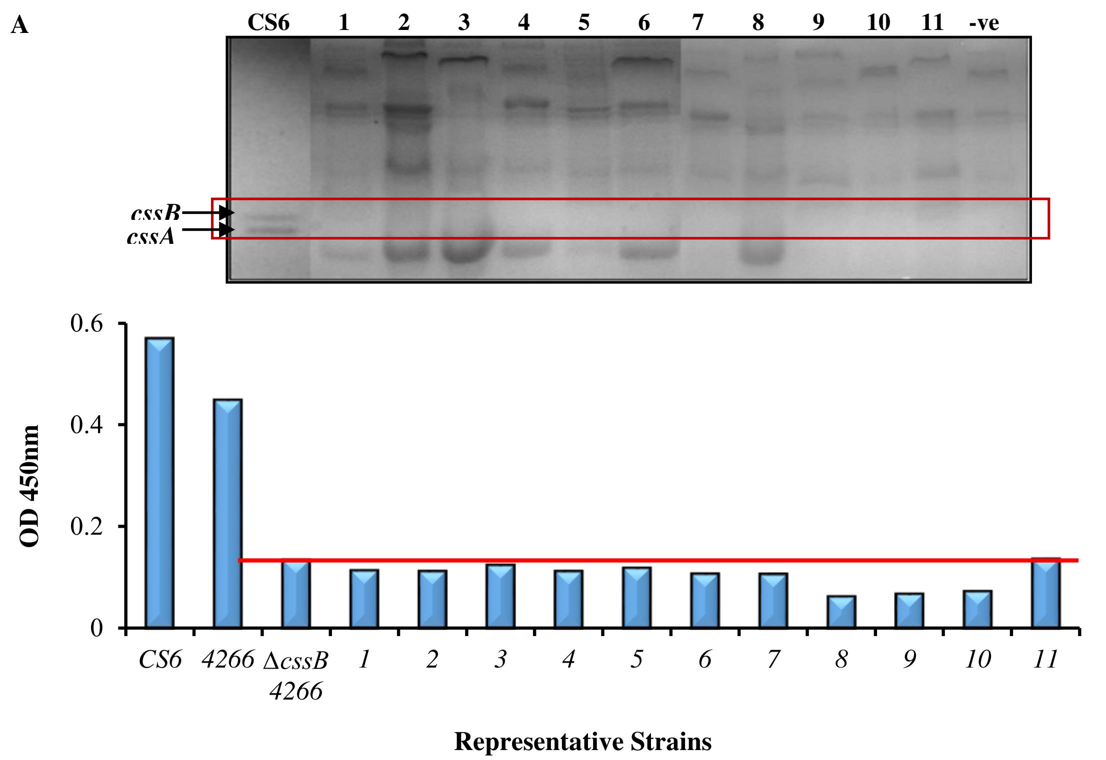

Figure 5.II.8: Confirmation of non-expression of CS6 strains.
A) by SDS-PAGE: SDS-PAGE stained with Coomassie Blue showing *cssA* and *cssB* expression in purified CS6. Non-expressive CS6 showed no protein band corresponding to the expected molecular weight of *cssA* and *cssB*
B) by ELISA: Comparison of *cssB* expression level of non-expressive CS6 strains with the purified CS6 and ETEC 4266 by ELISA
1 – 11 – representative strains
CS6 – Purified CS6
-ve – BSA was used as protein sample
4266 – ETEC 4266 standard lab strain known to express and produce CS6
Δ*cssB* 4266 – ETEC 4266 strain in which *cssB* gene had been knock out.

5.II.5. Effect of different additives:

When we further looked into the effect of host factors on a representative strain of ETEC which does not express CS6 we found that CS6 had expressed when supplemented. In M9 media (supplemented with 0.2% glycerol as background carbon source) the strain harbouring CS6 do not express CS6 but when the media was supplemented with different additives the strain expressed CS6 as quantified by qRT-PCR (quantitative real-time PCR) (*Figure 5.II.10*). In presence of 75mM NaCl, the CS6 expression was 9 fold upregulated while in 0.05mM $FeSO_4$ [iron (II) salts] the expression was increased by 3 fold and in presence of 1% glucose the expression was 8 fold higher as compared to only M9 media.

```
Strain 4266    CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCAACAAAGC
Strain 1_      CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCAACAAAGC
Strain 2_      CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCAACAAAGC
Strain 3_      CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCAACAAAGC
Strain 4_      CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCAACAAAGC
Strain 5_      CAACCCGCTACGGGGGAGACTTACCCGAAATCTGATTTATTCTACAAAGC
               ****************************************** *******

Strain 4266    CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAATTCCATC
Strain 1_      CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAATTCCATC
Strain 2_      CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAATTCCATC
Strain 3_      CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAATTCCATC
Strain 4_      CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAATTCCATC
Strain 5_      CTCACGGCTGGGGAATCTCTCTCCGGCAGCCTTCAGGAAGAAACTCCATC
               ******************************************* ******

Strain 4266    AGACGGCTGCTTAAAAAAGAACAAATGGTAGTGTCCGCTATTGCCAGTAC
Strain 1_      AGACGGCTGCTTAAAAAAGAACAAATGGTAGTGTCCGCTATTGCCAGTAC
Strain 2_      AGACGGCTGCTTAAAAAAGAACAAATGGTAGTGTCCGCTATTGCCAGTAC
Strain 3_      AGACGGCTGCTTAAAAAAGAACAAATGGTAGTGTCCGCTATTGCCAGTAC
Strain 4_      AGACGGCTGCTTAAAAAAGAACAAATGGTAGTGTCCGCTATTGCCAGTAC
Strain 5_      AGACGGCTGCTTAGAGAAAAACAAATGGTAGTGTCCGCTGTTGCCAGTAC
               ************* * ** ******************** **********

Strain 4266    ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
Strain 1_      ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
Strain 2_      ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
Strain 3_      ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
Strain 4_      ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
Strain 5_      ACCTCACTCACCAATAAAAGCGTCAATACGGTGCTCCGTTGACACATTAC
               **************************************************

Strain 4266    GAATGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
Strain 1_      GAATGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
Strain 2_      GAATGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
Strain 3_      GAATGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
Strain 4_      GAATGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
Strain 5_      GATTGTTATGTATACAATAAAAATGATTATAGCAATAGTAATGGTGTTAT
               ** ***********************************************
```

Figure 5.II.9: Alignment of Promoter DNA sequences of different CS6 harbouring strains by CLUSTAL 2.1 multiple sequence alignment. The alignment showed the presence of conserved regions by '*' symbol and non-conserved regions by 'blank'. 'Strain 4266' and 'Strain 1 to Strain 4' were different CS6 expressing strains while 'Strain 5' represent non-expressive CS6 strain. Yellow highlighted part indicate the changes in the sequence.

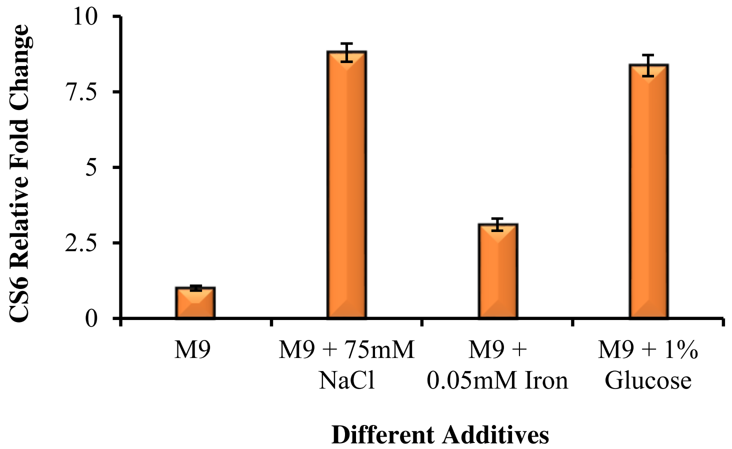

Figure 5.II.10: Expression of CS6 in CS6-expression negative ETEC with different additives

5.II.6. Expressive versus non-expressive CS6 strain by Whole Genome Sequencing

We compared the expressive CS6 strain (ETEC 4266) with the non-expressive CS6 strain (IDH01470) by whole genome sequencing using Illumina Platform. *Truseq* Nano DNA Library preparation kit was used to make libraries from samples IDH01470 and ETEC 4266. For the samples IDH01470 and ETEC 4266, the average library size is 445 bp and 561 bp, respectively. On an *Illumina* platform (2 x 150 bp chemistry), the library was sequenced, providing roughly 3 GB of data per sample. *Velvet v1.2.10* was used for de novo assembly of high-quality PE reads from samples IDH01470 and ETEC 4266, and the assembly was optimised at *Kmer-85* and *Kmer-83* for sample IDH01470 and ETEC 4266, respectively. Using PE readings and GapCloser v1.12 software, the scaffolds were gap-filled (stretch of 'Ns' were filled with A, T, G, and C bases). The amended assembly was regarded as the final assembly for annotation purposes. An *in-house perl script* was used to calculate the final assembly pieces.

Identification of RNAs: *tRNAscan-SE* was used for the identification of probable tRNA genes. Predicted tRNAs for IDH01470 were 76 and ETEC 4266 were 70. The *barrnap 0.9* software package to identify rRNAs in the genome of the IDH01470 and ETEC 4266 samples. As a result, 6 rRNAs were identified from the IDH01470 sample, out of which two 5S_rRNA, one 16S_rRNA and two 23S_rRNA were present, whereas 5 rRNAs were identified from ETEC 4266 sample, out of which two 5S_rRNA, two 16S_rRNA and one 23S_rRNA were present.

SSR (Simple Sequence Repeats) Identification: The *MISA perl script* was used to find SSR in each sample's scaffold sequences. The total number of identified SSRs for IDH01470 and ETEC 4266 were 7 and 5 respectively.

Gene prediction: Using Prodigal v2.60, the constructed scaffolds from both samples (IDH01470 and ETEC 4266) were submitted to gene prediction, yielding 4,537 and 4,562 gene sequences from IDH01470 and ETEC 4266, respectively (***Table 5.II.2***). The projected 4,537 and 4,562 genes from IDH01470 and ETEC 4266 samples, respectively, were compared to NCBI's non-redundant (nr) database using the BLASTP method and with an e-value threshold of 1e-5. For IDH01470 out of a total of 4537 proteins, there were 4513 proteins with hits and 24 proteins with no hits. For ETEC 4266 out of total 4562 proteins, there were 4536 number of proteins with hits and 26 number of proteins with no hits. For both IDH01470 and ETEC 4266 samples, the top-hit species distribution revealed that the majority of the hits were against the species *Escherichia coli*, followed by *Shigella sonnei*. Simultaneously all the predicted 4,537 and 4,562 genes from IDH01470 and ETEC 4266 samples respectively were compared for similarity against Uniprot, COG and Pfam database using BLASTP with an e-value threshold of 1e-5. Results for similarity search against all three databases are shown in table (***Table 5.II.2***). Results for gene distribution are represented in form of COG functional categories (***Figure 5.II.11; Figure 5.II.12***) and in form of Venn diagram (***Figure 5.II.13; Figure 5.II.14***).

Table 5.II.2: Blast statistics against Uniprot, COG and Pfam database for IDH01470 and ETEC 4266 sample

Sample	Total proteins	Database	No. of proteins with Hits	No. of proteins with No Hits
IDH01470	4,537	*UniProt*	4,253	284
		Pfam	3,370	1,167
		COG	3,726	811
ETEC 4266	4,562	*UniProt*	4,286	276
		Pfam	3,367	1,195
		COG	3,738	824

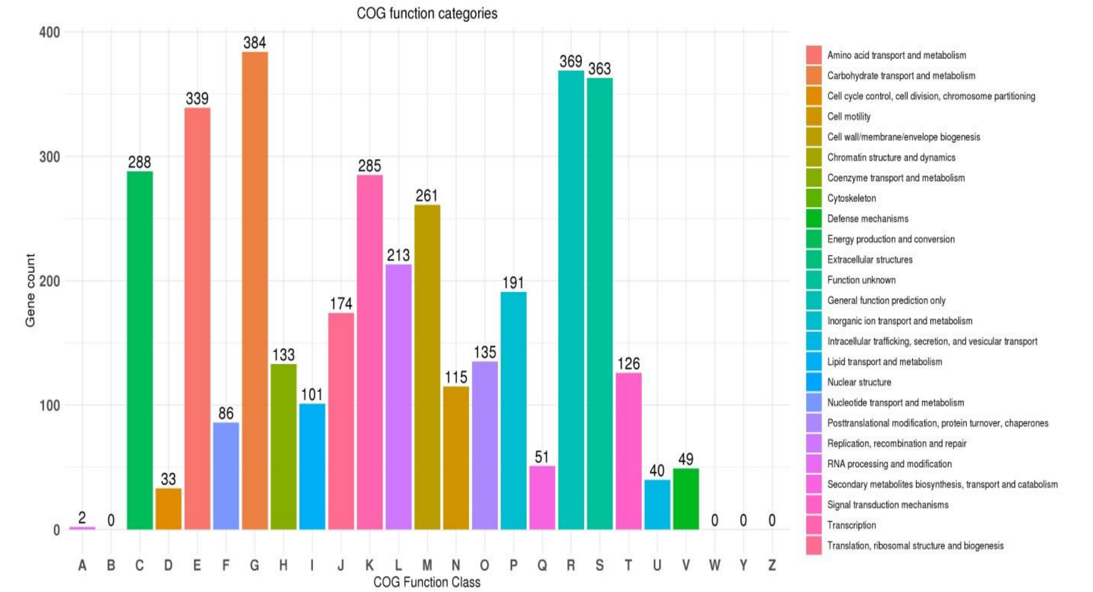

Figure 5.II.11: COG functional classification of ETEC 4266 sample

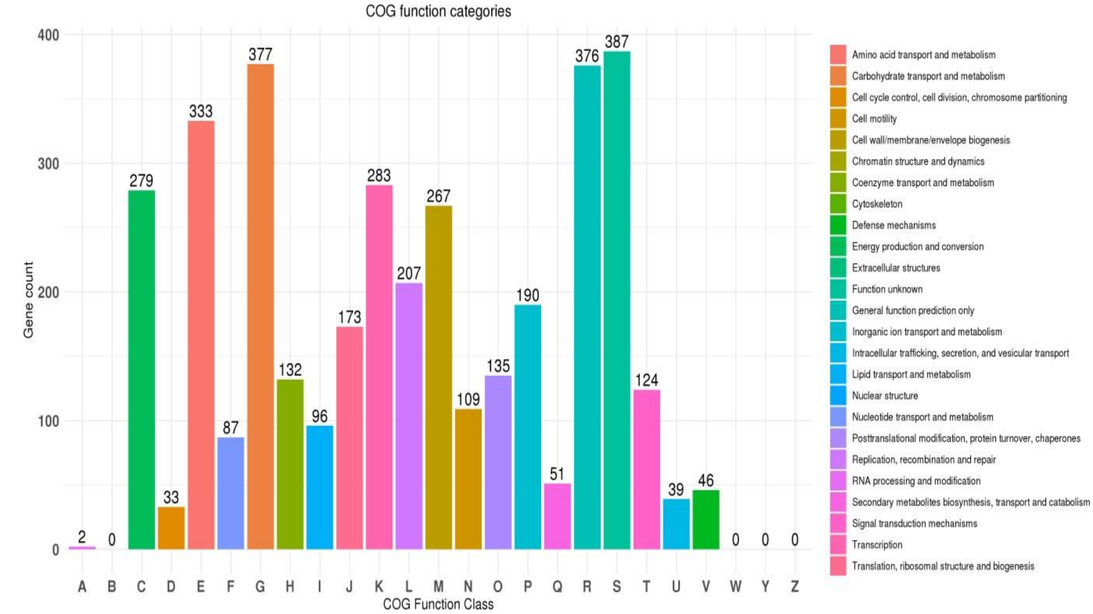

Figure 5.II.12: COG functional classification of IDH01470 sample

The GO distribution was carried out using the Blast2GO command-line tool. GO terms were assigned to a total of 2,625 and 2,753 genes in IDH01470 and ETEC 4266 samples respectively. Wherein 2,223 genes were assigned to Biological Process, 2,203 genes to Cellular Component and 2,203 genes to Molecular functions for IDH01470 sample (***Figure 5.II.15***). For ETEC 4266 sample 2,323 genes were assigned to Biological

Process, 1,672 genes to Cellular Component and 2,318 genes to Molecular functions (***Figure 5.II.16***).

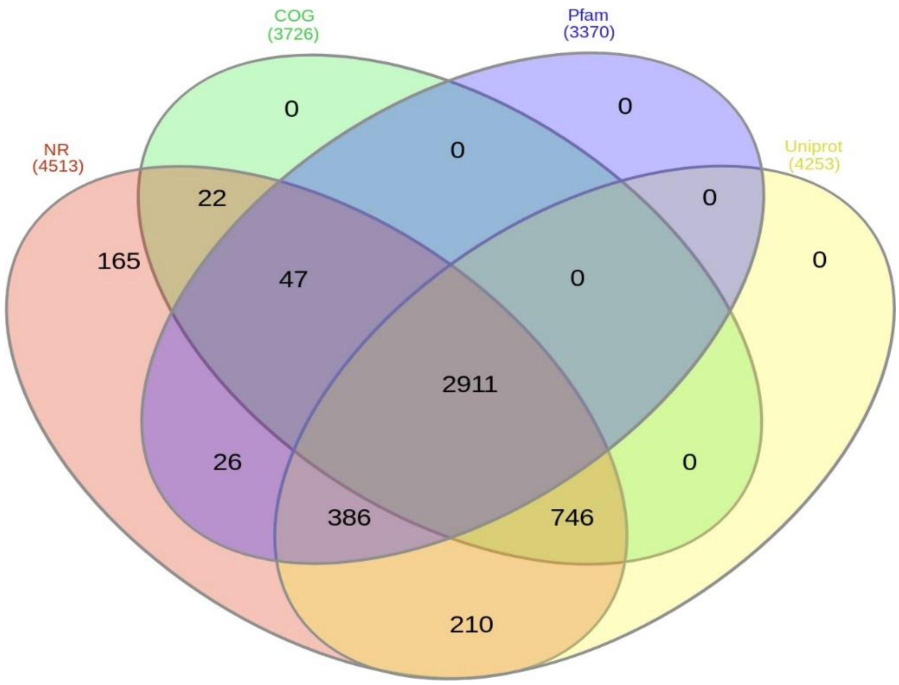

Figure 5.II.13: Venn diagram for IDHO1470 sample, annotated proteins count in different databases.

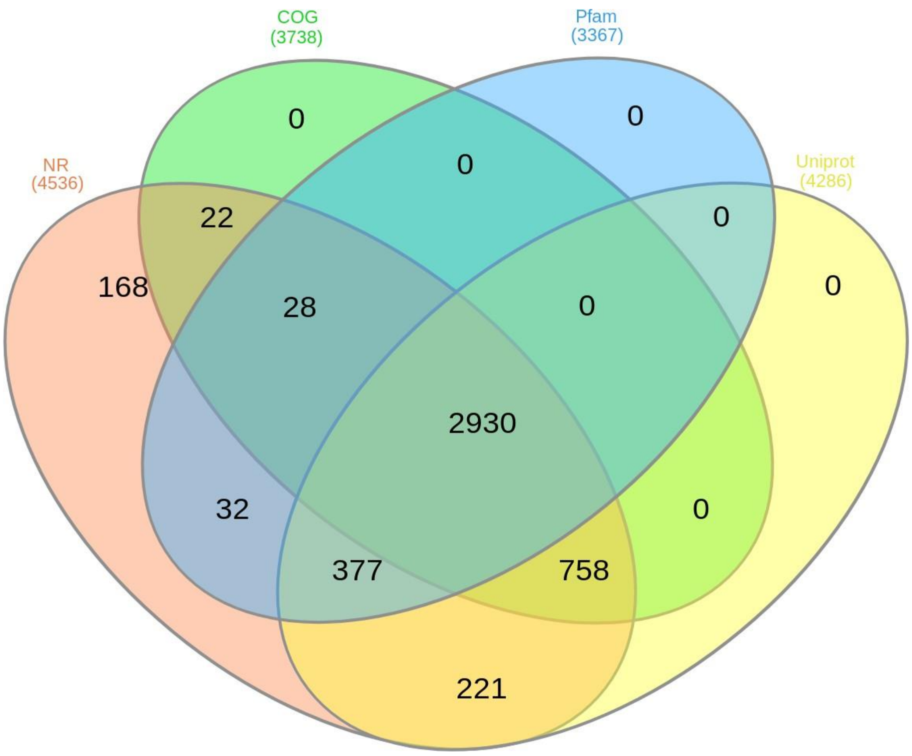

Figure 5.II.14: Venn diagram for ETEC 4266 sample, annotated proteins count in different databases.

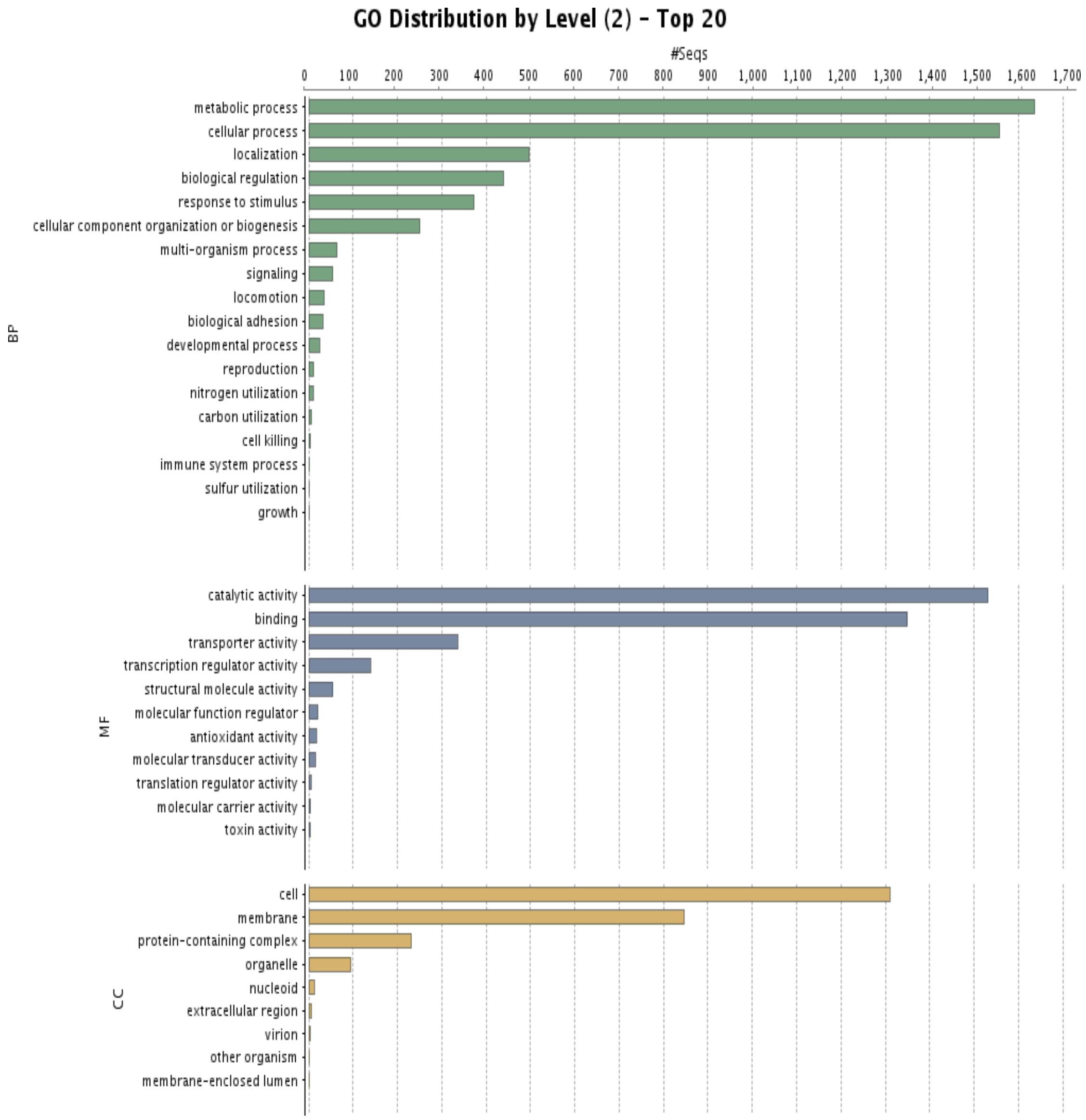

Figure 5.II.15: GO distribution for IDH01470 sample

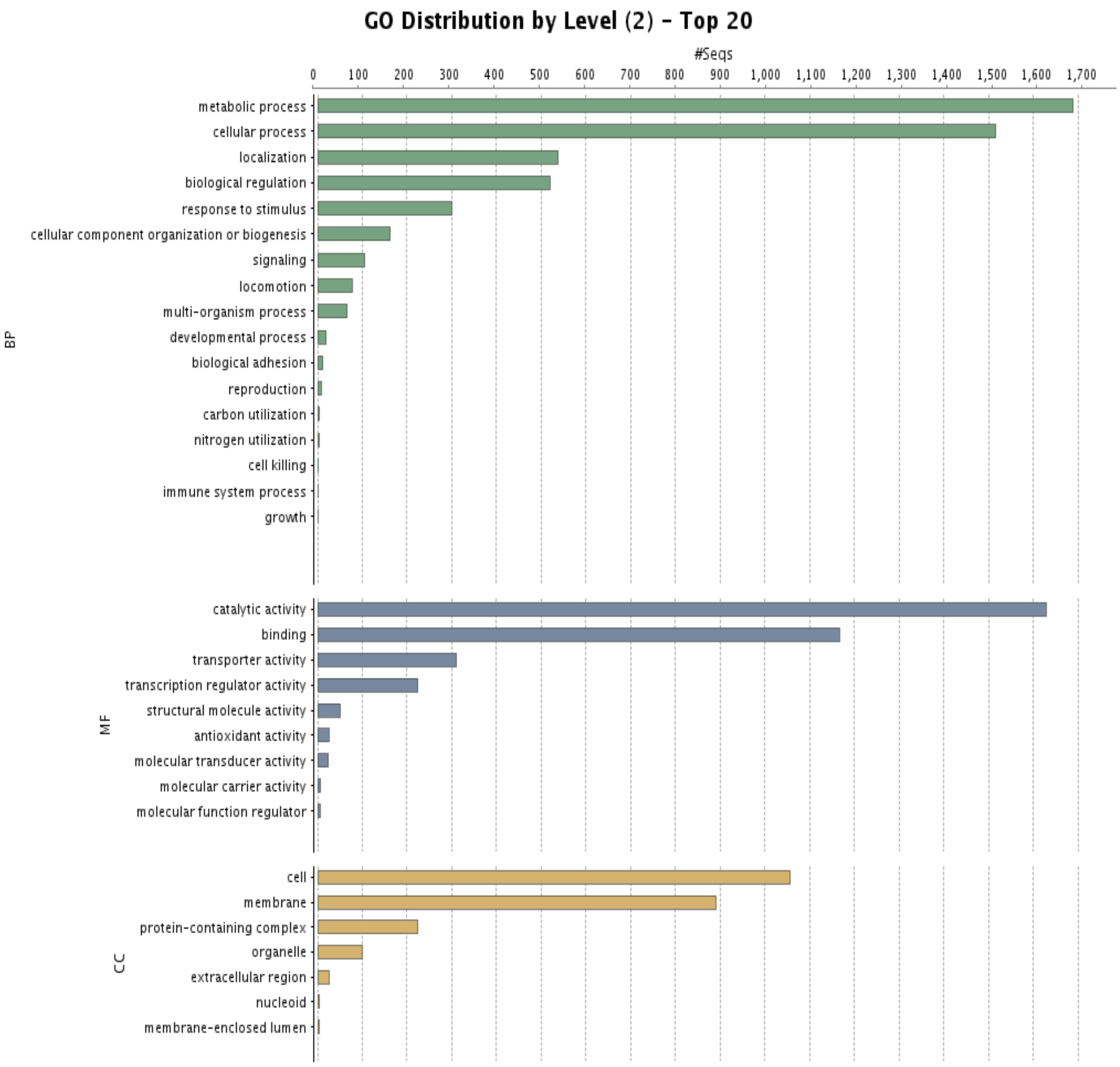

Figure 5.II.16: GO distribution for ETEC 4266 sample

Pathway Analysis: The *KEGG automatic annotation server (KAAS)* was used to perform pathway analysis, ortholog allocation and gene mapping to the biological pathways. Using *BLASTP* with a threshold bit-score value of 60 (default), all predicted 4,537 and 4,562 genes sequences from IDH01470 and ETEC 4266 samples were compared against the KEGG database. The distribution of genes attributed to various pathways is shown in Table (***Table 5.II.3***)

Table 5.II.3: Pathway Distribution in from IDH01470 and ETEC 4266 samples

Level1	Level2	IDH01470	ETEC4266
Metabolism	Carbohydrate metabolism	492	507
	Energy metabolism	189	189
	Lipid metabolism	85	85
	Nucleotide metabolism	128	127
	Amino acid metabolism	275	276
	Metabolism of other amino acids	84	87
	Glycan biosynthesis and metabolism	68	67
	Metabolism of cofactors and vitamins	187	189
	Metabolism of terpenoids and polyketides	33	33
	Biosynthesis of other secondary metabolites	43	43
	Xenobiotics biodegradation and metabolism	79	79
	Not included in regular maps	1	1
Genetic Information Processing	Transcription	4	4
	Translation	86	86
	Folding, sorting and degradation	53	53
	Replication and repair	89	88
Environmental Information Processing	Membrane transport	272	266
	Signal transduction	173	172
Cellular Processes	Transport and catabolism	8	8
	Cell motility	60	60
	Cell growth and death	19	19
	Cellular community - prokaryotes	169	165
Organismal Systems	Environmental adaptation	2	6
Human Diseases	Infectious disease: bacterial	32	33
	Drug resistance: antimicrobial	61	61

Circos plot: Scaffolds, CDS, tRNAs and rRNA predicted were used for the representation of IDH01470 genome (***Figure 5.II.17***) and ETEC 4266 genome (***Figure 5.II.18***) in circular form.

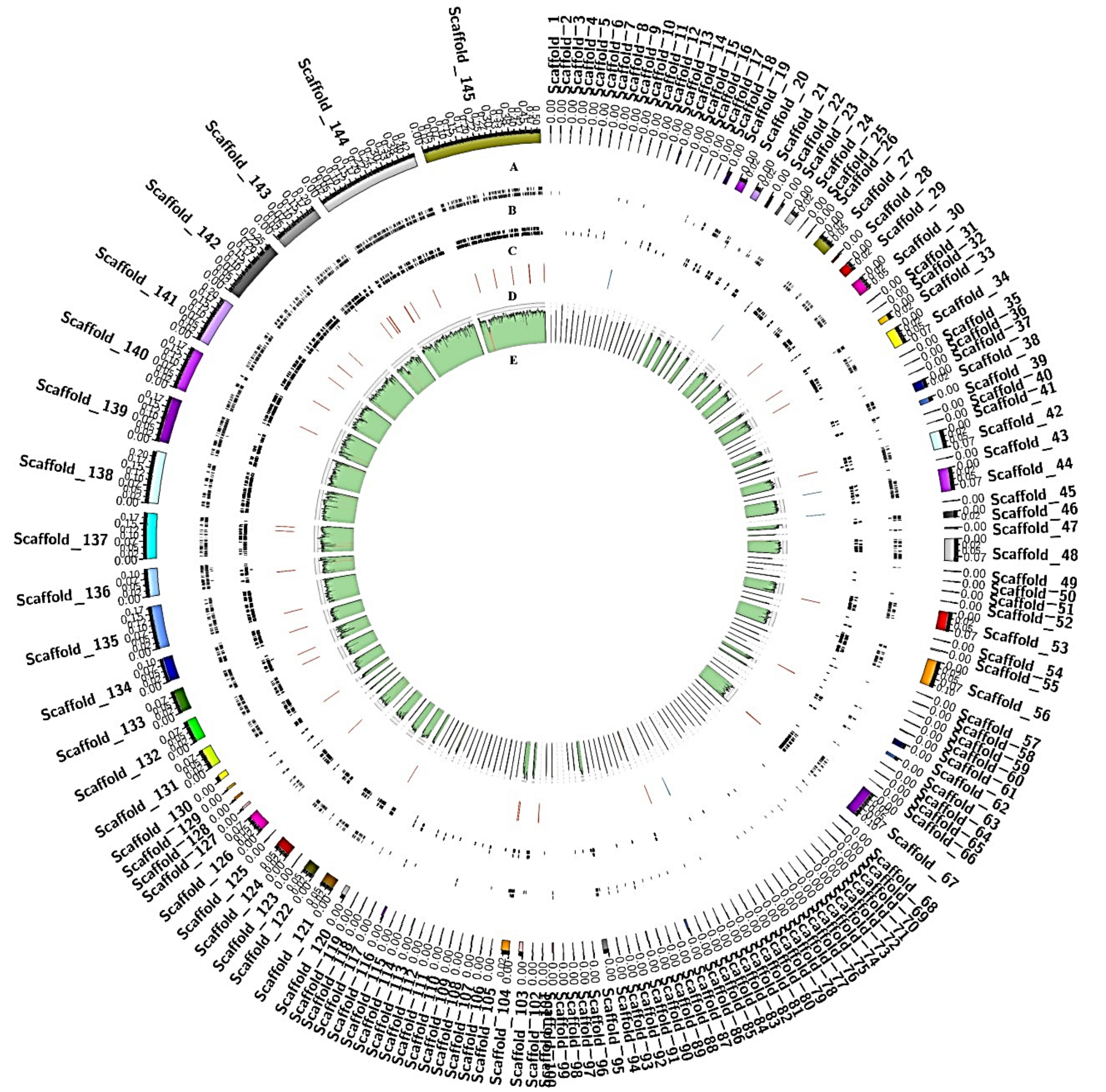

Figure 5.II.17: The assembled 145 scaffolds of the IDH01470 sample are represented in the outer circle A and circle B represents CDS predicted in the positive frame. Circle C represents CDS predicted in the negative frame, circle D represents tRNA(represented in red tiles) , blue tiles rRNA. The inner E circle represents distribution of GC .The red shade in the E circle represents regions of low GC contents i.e; less than 35% whereas green shade represent GC contents higher than 35%

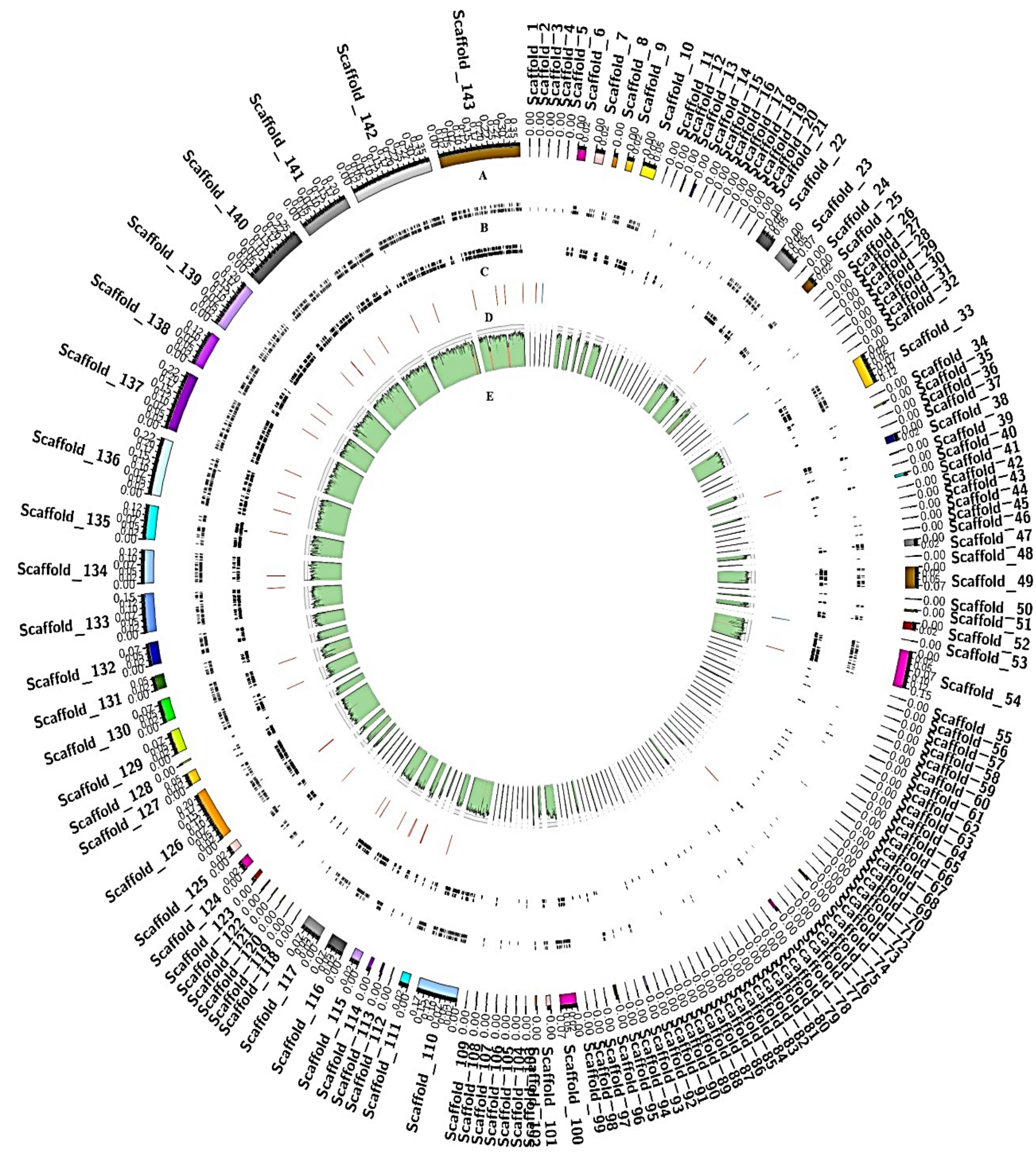

Figure 5.II.18: The assembled 143 scaffolds of the ETEC 4266 are represented in the outer circle A and circle B represents CDS predicted in the positive frame. Circle C represents CDS predicted in the negative frame, circle D represents tRNA(represented in red tiles) , blue tiles rRNA. The inner E circle represents distribution of GC. The red shade in the E circle represents regions of low GC contents i.e; less than 35% whereas green shade represent GC contents higher than 35%

Identification of Plasmid Genome: Plasmid were assembled using *plasmidSPAdes.* A total of 60 and 175 scaffold were obtained after assembly from which a total of 221 and 478 genes were identified using *prodigal* from IDH01470 and ETEC 4266 samples respectively.

Comparison of the expressive CS6 strain (ETEC 4266) with the non-expressive CS6 strain (IDH01470) revealed that ETEC 4266 has more predicted genes than IDH01470. Chromosomal analysis revealed that the genes of both samples are closely related. However, plasmid analysis showed a huge difference between the numbers of predicted genes among the samples as ETEC 4266 has 478 genes while IDH01470 has 221 genes. Since the genes for colonization and virulence are present on the plasmids for ETEC the difference between the number of genes may hold the key for virulence although the only detectable colonization factor for IDH01470 is non-expressive.

Objective III:

To understand the modulation of prevalent colonization factor and other virulence factors by the host and environmental factors during pathogenesis.

5.III. Objective III:

The objective was to understand the modulation of the prevalent colonization factor and other virulence factors by the host and environmental factors during pathogenesis. To understand the modulation of prevalent colonization factor and other virulence factors by the host and environmental factors during pathogenesis we focused our study on CS6, CS5 and *eatA* since this is the most prevalent combination of CF (classical colonization factor) and NCVF (non-classical virulence factors) in terms of gene presence as well as gene expression. Different host factors were supplemented as additives in M9, LB and CFA broths in different concentrations as indicated. Media with no bile salt added was used as control. The effect of different factors on the expression of CS6, CS5 and *eatA* were quantified by quantitative real-time PCR and protein level expression were quantified by ELISA by comparing with the standard curve. The toxins were quantified by ELISA and the values were calculated from the standard curve. The outcomes were as follows.

5.III.1. Effect of different host factors on the expression of virulence genes: Classical colonization factor, CS6 and CS5; Non-classical virulence factor, *eatA*; and enterotoxins, LT and ST.

A. Crude Bile

CS6 expression was mostly independent with the addition of crude bile at different concentrations in the media (*Figure 5.III.1*). In M9 broth at a concentration of 0.05% bile the expression of CS6 was upregulated by 1.8 fold but at other concentrations of bile, CS6 was downregulated. However, CS5 expression was significant in presence of bile

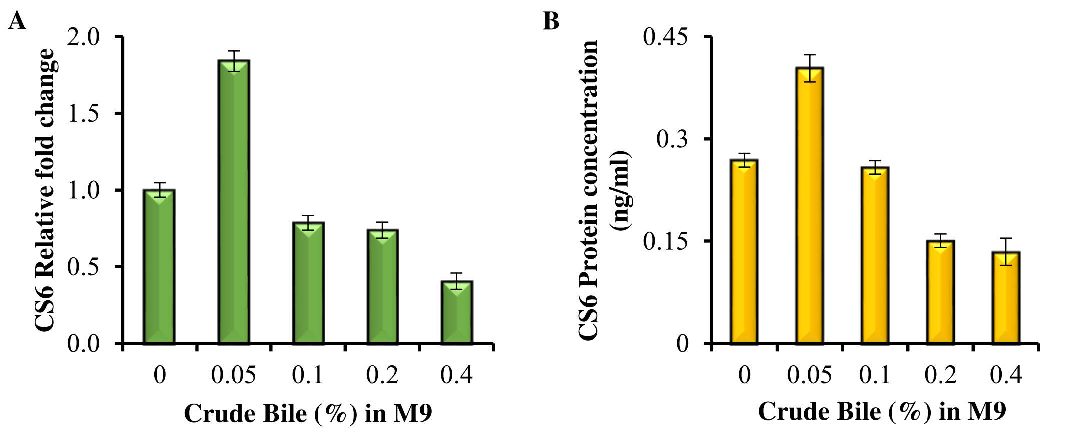

Figure 5.III.1: Effect of crude bile on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

having an increase in relative fold expression by 80 fold when compared with the control (*Figure 5.III.2*). EatA expression was not significantly affected in presence of bile (*Figure 5.III.3*). The amount of LT secretion was decreased in presence of bile salts. The secretion of ST was not affected by the addition of bile.

As observed in RNA expression, in protein level CS6 production was also independent in presence of crude bile salt in M9 broth. At the concentration of 0.05% crude bile, the production of CS6 was upregulated but at other concentrations of bile, CS6 was downregulated. CS5 production was more in presence of crude bile, maximum at 0.05% crude bile concentration. EatA production was not significantly affected when bile is present. There was up-regulated production of EatA at 0.05%, 0.1% and 0.2 % concentration of crude bile.

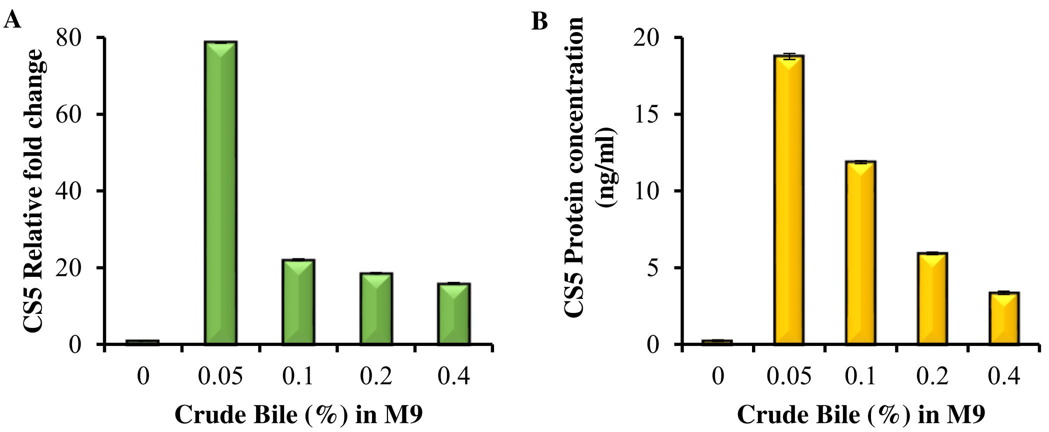

Figure 5.III.2: Effect of crude bile on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

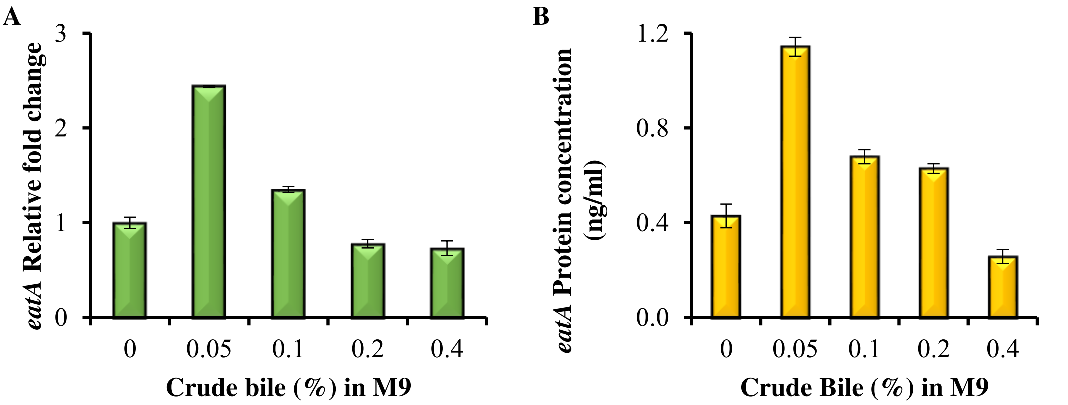

Figure 5.III.3: Effect of crude bile on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

B. Sodium cholate hydrate (NaCH)

CS6 expression was not significantly affected, just upregulated by 1.5 fold ($P< 0.05$) in presence of 0.05% Sodium cholate hydrate (NaCH) salt, a component of bile, in M9 media (***Figure 5.III.4***). CS5 expression was 2 fold at 0.05% but with an increase in NaCH concentration, the expression level decreased (***Figure 5.III.5***). EatA expressed was increased by 2 fold at 0.05% NaCH concentration (***Figure 5.III.6***). LT secretion was lowered. ST secretion was more in the 0.05% NaCH bile concentration in M9 media.

CS6 protein production was not affected significantly in presence of Sodium cholate hydrate salt, which coincides with the RNA expression pattern. At a concentration of 0.05%, the production of CS6 was upregulated but at other concentrations of bile, CS6 was downregulated which was not so significant. CS5 was presence of NaCH was

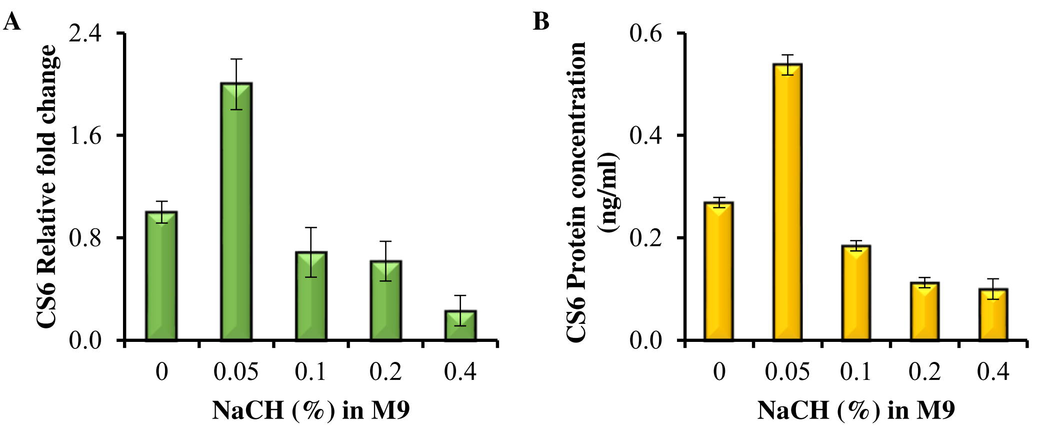

Figure 5.C.4: Effect of NaCH on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

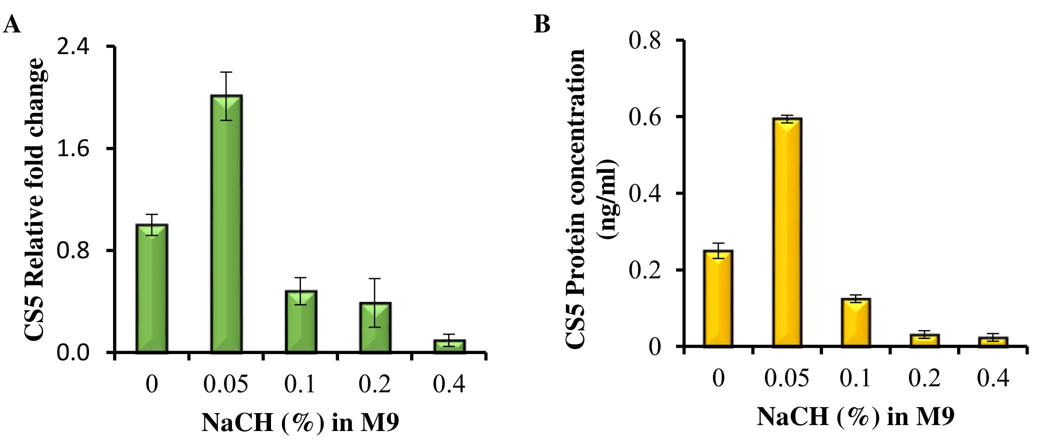

Figure 5.C.5: Effect of NaCH on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

produced more than control only in 0.05%, while in more concentration of NaCH it was downregulated. EatA in presence of NaCH was produced more than control only in 0.05%.

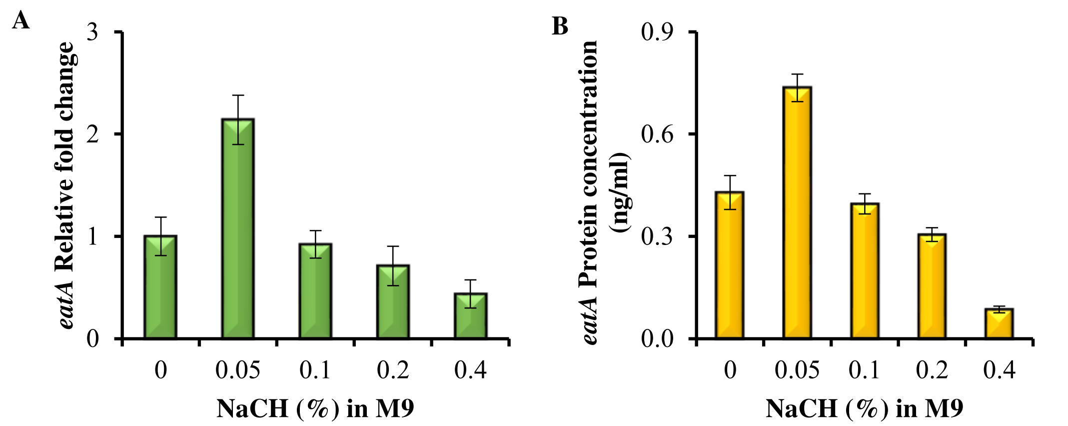

Figure 5.III.6: Effect of NaCH on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

C. Sodium Deoxycholate (NaDC)

In M9 broth at a concentration of 0.05% Sodium Deoxycholate (NaDC), a component of bile, the expression of CS6 was up-regulated by 3 fold (***Figure 5.III.7***). CS5 was expressed most (60 fold) at 0.2% (***Figure 5.III.8***). EatA was maximally expressed by 5.5 fold at 0.2 % (***Figure 5.III.9***). LT secretion was induced in presence of NaDC. ST secretion was decreased with an increase in NaDC concentration.

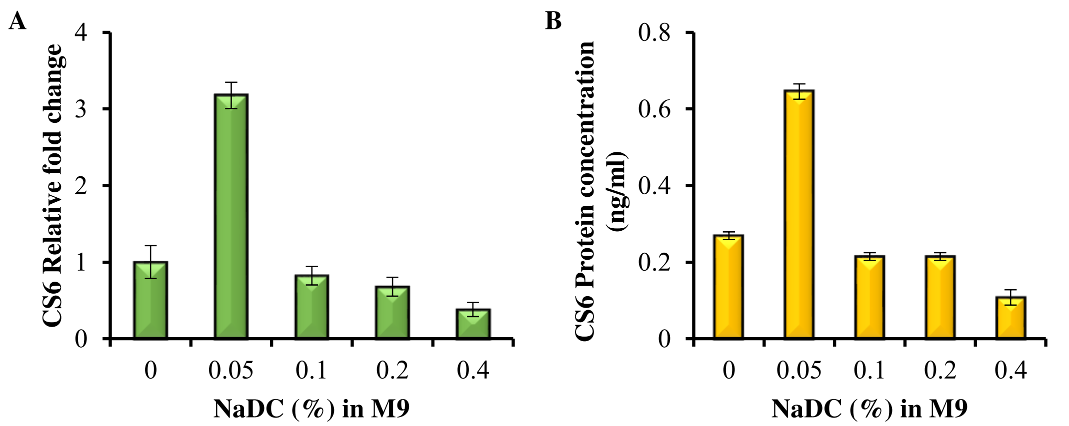

Figure 5.III.7: Effect of NaDC on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

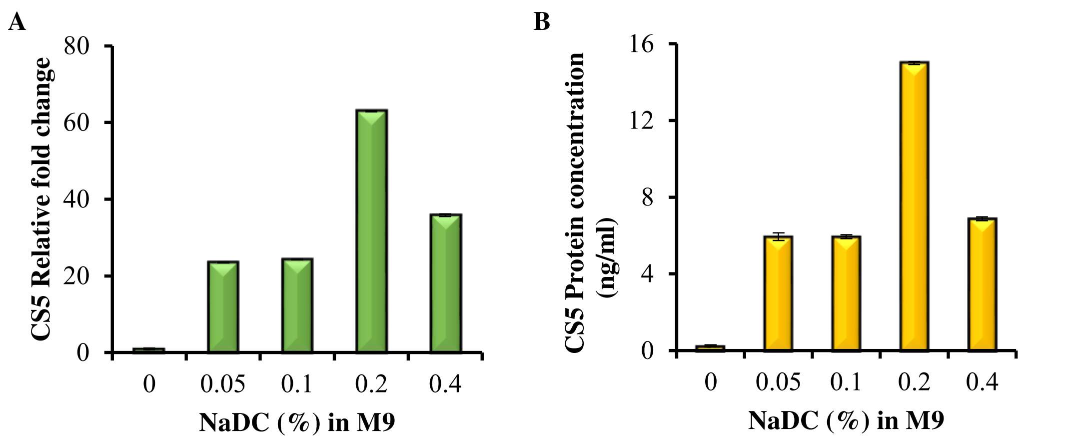

Figure 5.III.8: Effect of NaDC on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

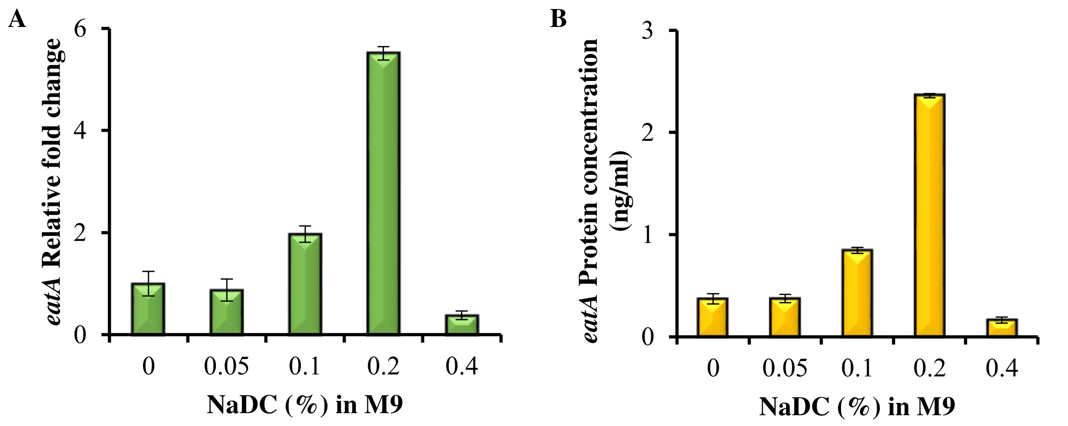

Figure 5.III.9: Effect of NaDC on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

CS6 production was increased in presence of NaDC at 0.05%, while at other concentrations CS6 was not affected significantly. CS5 in presence of NaDC was produced more than control in all the concentrations and maximum expression was at 0.2%, followed at 0.4% NaDC. EatA in presence of NaDC was produced more than control only in 0.2 %, but with an increase in NaDC concentration, the protein level decreases.

D. Sodium glycocholate hydrate (NaGCH)

CS6 expression was upregulated by 2.6 fold in presence of Sodium glycocholate hydrate (NaGCH) salt, a component of bile, at 0.1% (***Figure 5.III.10***). CS5 was expressed most at 0.05 % (***Figure 5.III.11***) and EatA at 0.1% (***Figure 5.III.12***). LT secretion was decreased with an increase in the concentration of NaGCH. The secretion of ST was more in the 0.05% NaGCH concentration in M9 media.

CS6 protein production was upregulated maximally in presence of 0.1% sodium glycocholate hydrate salt. With an increase in NaGCH concentration the production level decreases. CS5 protein in presence of NaGCH was produced more than control in all concentrations with the maximum at 0.05 %. EatA protein in presence of 0.1% NaGCH is upregulated most. With an increase in NaCDC concentration the production level decreases.

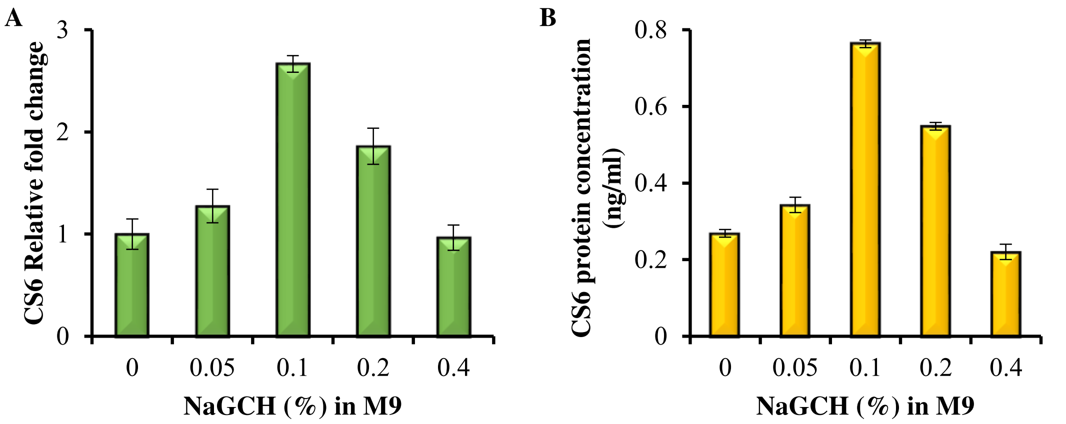

Figure 5.III.10: Effect of NaGCH on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

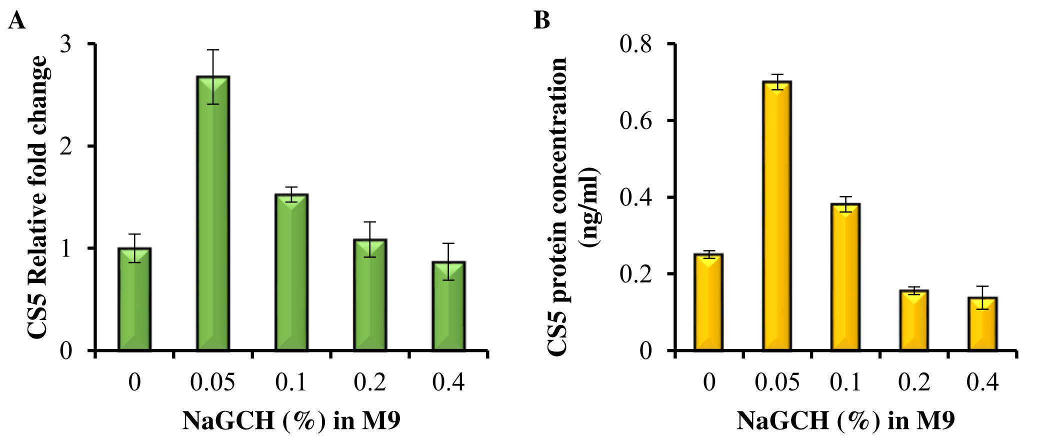

Figure 5.III.11: Effect of NaGCH on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

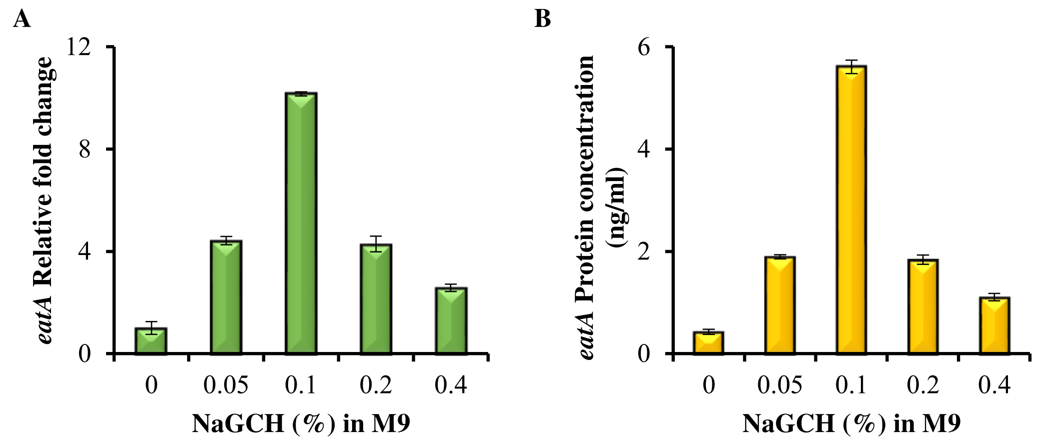

Figure 5.III.12: Effect of NaGCH on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

E. Taurocholic acid sodium salt hydrate (TCA)

CS6 expression was upregulated in presence of Taurocholic acid sodium salt hydrate (TCA) salt, a component of bile, when added in the media. At a concentration of 0.1% TCA, the expression of CS6 was maximally up-regulated by 5 fold (***Figure 5.III.13***). CS5 and EatA expression were upregulated by 7 fold in presence of 0.1% TCA (***Figure 5.III.14***) (***Figure 5.III.15***). LT secretion was increased in presence of TCA salt whereas ST secretion was decreased with an increase in TCA concentration.

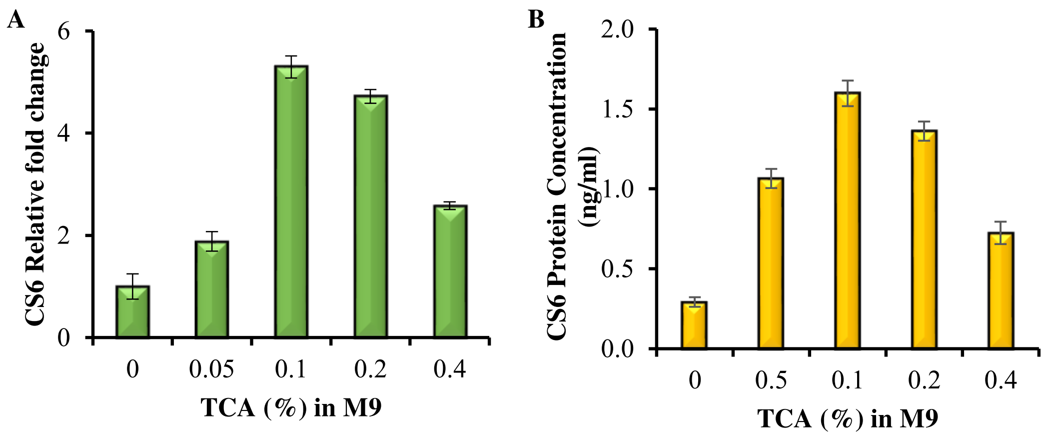

Figure 5.III.13: Effect of TCA on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

CS6 protein was upregulated most in presence of TCA at a concentration of 0.1% but at other concentrations, CS6 production was downregulated with an increase in TCA concentrations. CS5 protein production was more when TCA is present. In all the test concentrations the production was more than the control set but maximum production was observed in 0.1% TCA. EatA in presence of TCA was produced more than control only in 0.1% and 0.2 %; most at 0.1% TCA.

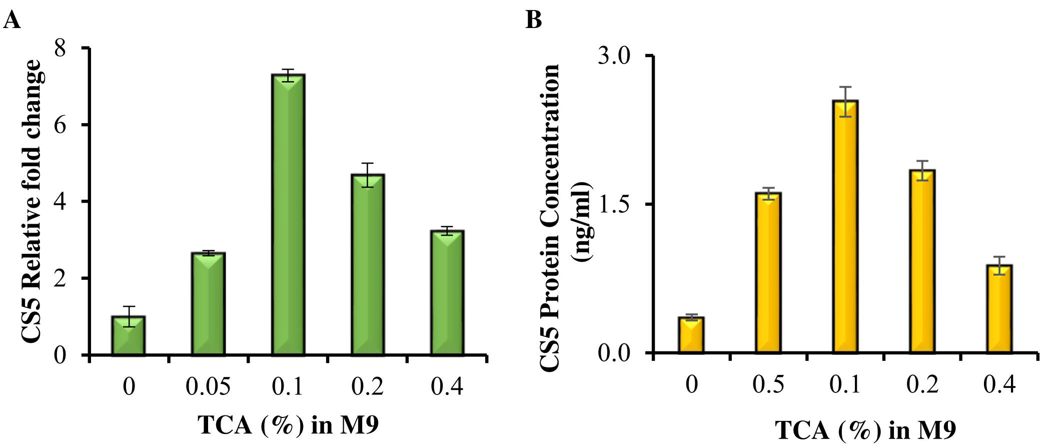

Figure 5.III.14: Effect of TCA on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

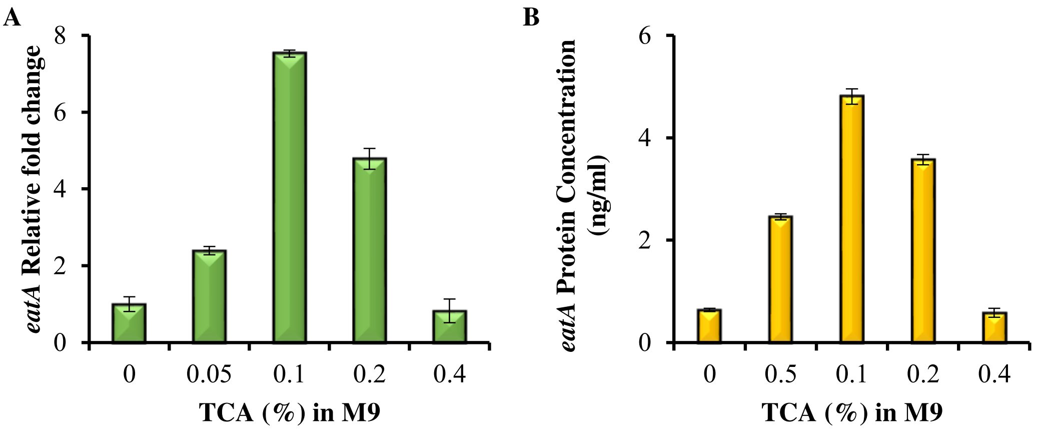

Figure 5.III.15: Effect of TCA on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

F. Sodium chenodeoxycholate (NaCDC)

CS6 expression was upregulated in presence of Sodium chenodeoxycholate (NaCDC) salt, a component of bile. At concentration 0.1 % NaCDC, the expression of CS6 was upregulated by 40 fold (***Figure 5.III.16***), CS5 expression by 27 fold (***Figure 5.III.17***) and EatA expression by 19 fold (***Figure 5.III.18***), respectively. LT secretion was increased at a lower concentration of NaCDC. The secretion of ST was almost uniform in all the conditions of NaCDC.

CS6 production was upregulated in presence of NaCDC at a concentration of 0.1 %. CS5 in presence of NaCDC was produced more than control in all concentrations with the maximum increase at 0.1%. EatA in presence of NaCDC was produced most at 0.1%.

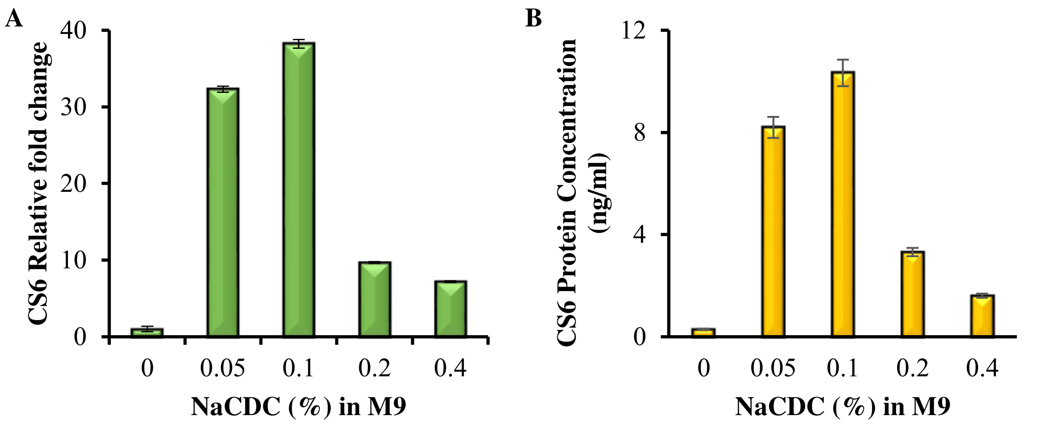

Figure 5.III.16: Effect of NaCDC on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

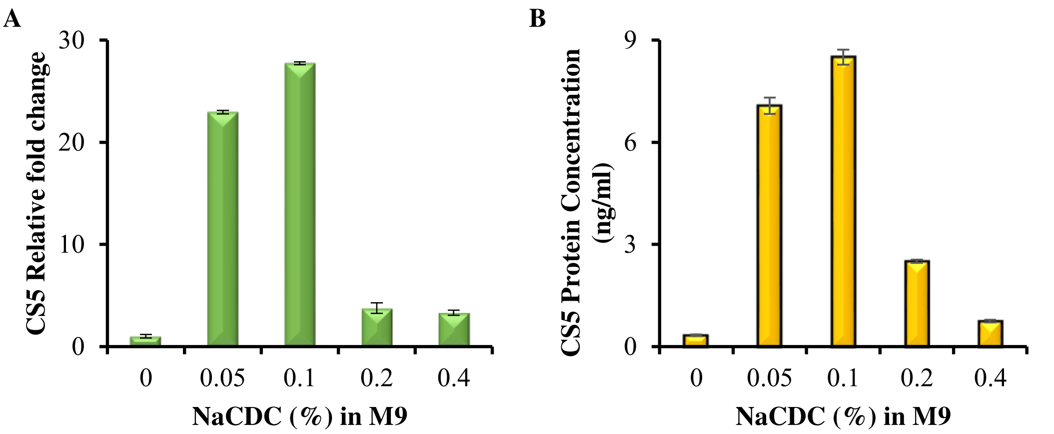

Figure 5.III.17: Effect of NaCDC on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

G. Mucin

CS6 expression was increased by 2 fold in presence of 0.0125% mucin but in other test concentrations, the expression level was same as in the control set (*Figure 5.III.19*). CS5 expression was increased in presence of mucin by 4 fold (*Figure 5.III.20*). EatA expression was not significantly affected by the addition of mucin having a fold increase of just 2 fold at 0.0125% mucin (*Figure 5.III.21*). LT secretion was inhibited in presence of mucin and ST secretion is not much affected.

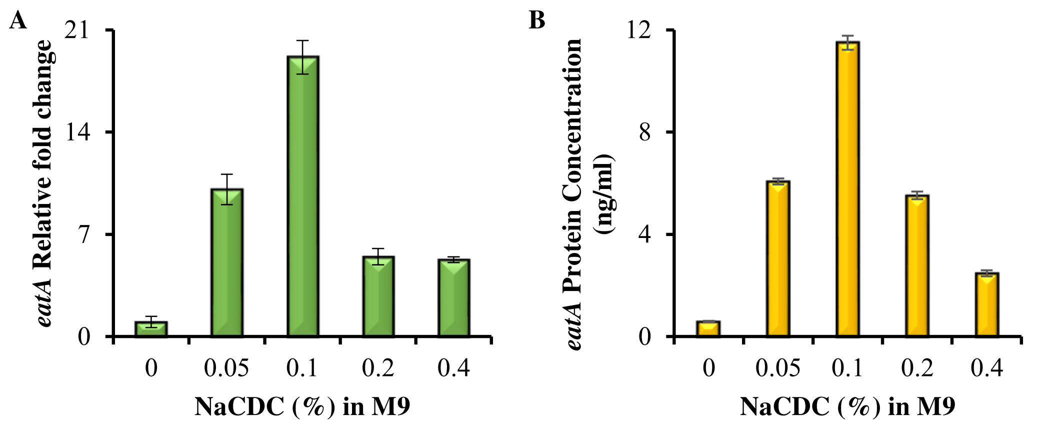

Figure 5.III.18: Effect of NaCDC on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

CS6 production was increased in presence of mucin at 0.0125% but not so significant. But in other test concentrations, the production level was inhibited. CS5 production was increased in presence of 0.025% mucin but in other test concentrations, there was a gradual down regulation of CS5 production as the concentration of mucin increased. EatA production increased in presence of 0.0125% mucin in M9 media. There was gradual downregulation of EatA production as the concentration of mucin increased.

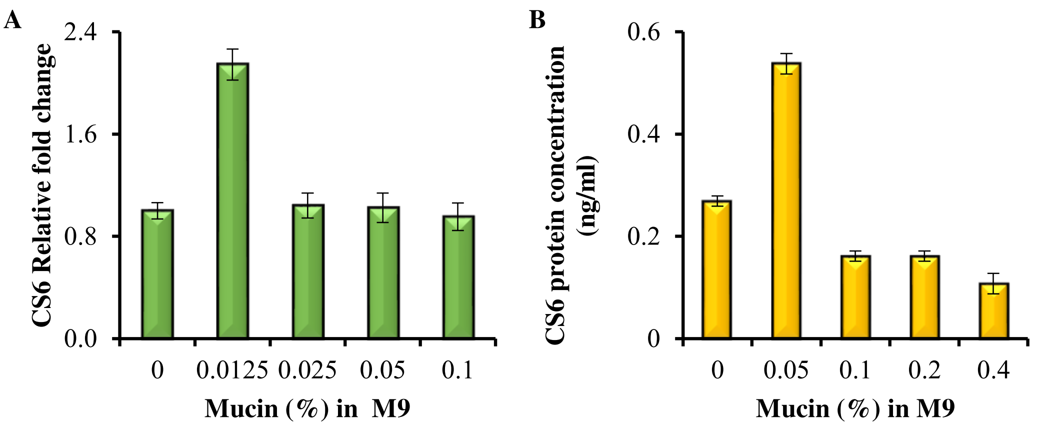

Figure 5.III.19: Effect of mucin on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

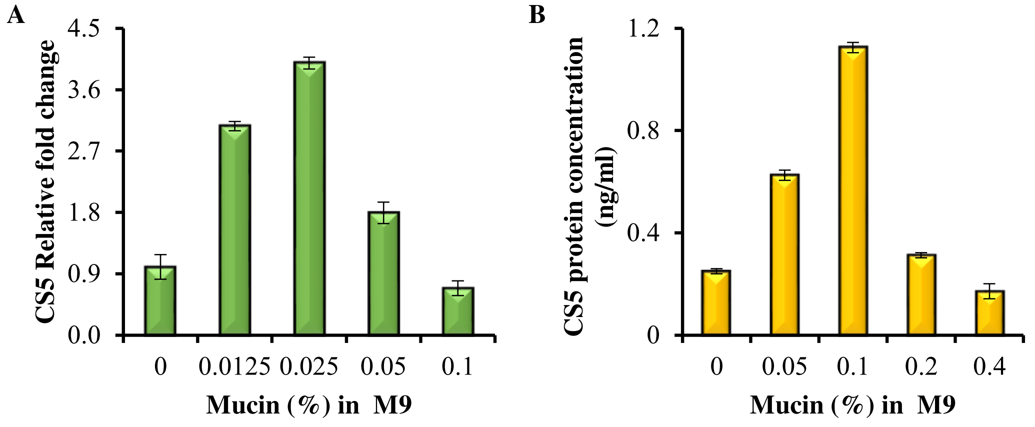

Figure 5.III.20: Effect of mucin on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

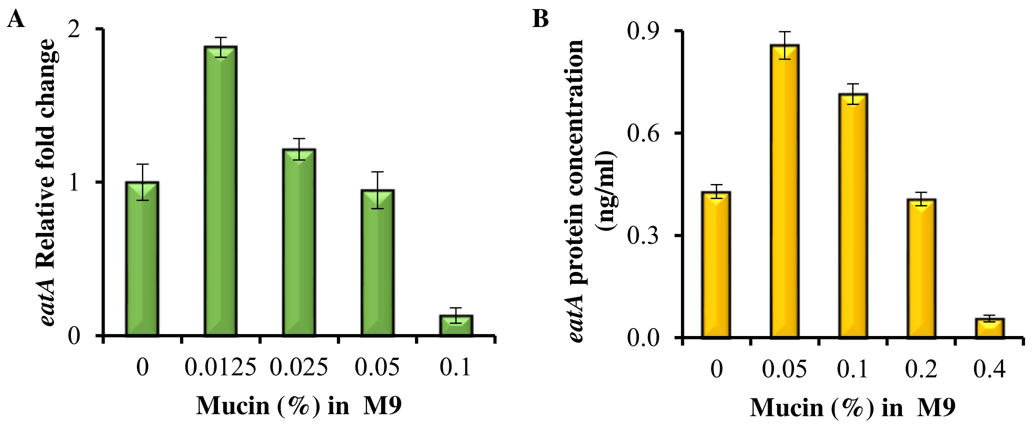

Figure 5.III.21: Effect of mucin on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

H. Iron Salt

CS6 expression was upregulated by 50 fold at 0.2mM concentration of Iron salts (FeSO4; Iron (II) sulfate) with respect to untreated ETEC (***Figure 5.III.22***). Similarly, 30 fold upregulation was seen for CS5 expression at 0.2mM concentration of iron salt (***Figure 5.III.23***). EatA expression was 40 fold increased (***Figure 5.III.24***). LT and ST secretion decreased with an increase in iron concentration.

CS6 production was most at 0.2mM concentration of iron salts and above or below that concentration the production decreases. Similarly, CS5 production was maximum at 0.2mM iron salts. EatA production was upregulated most at 0.2mM iron concentration.

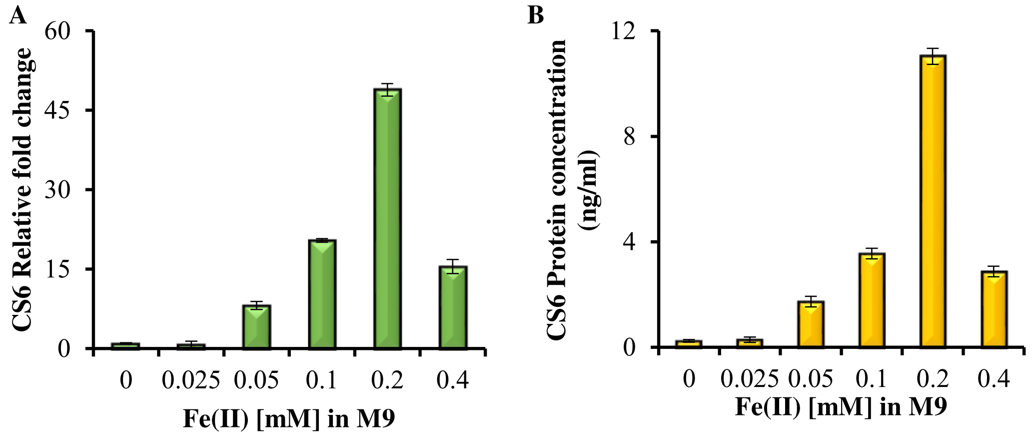

Figure 5.III.22: Effect of Fe (II) on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

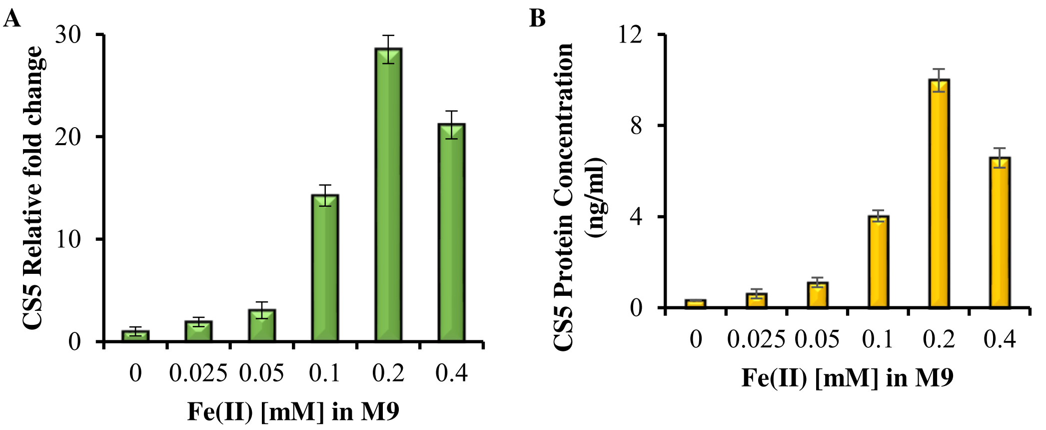

Figure 5.III.23: Effect of Fe (II) on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

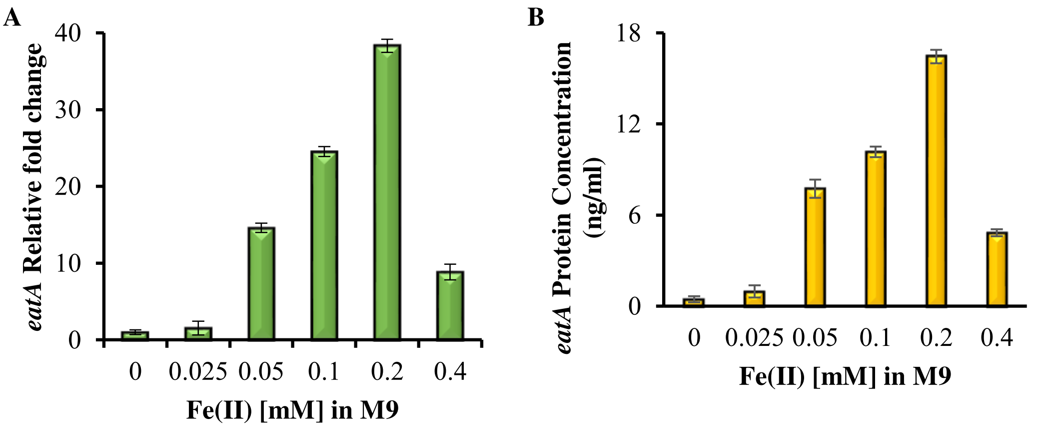

Figure 5.III.24: Effect of Fe (II) on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

I. Glucose

CS6 expression in presence of 1% Glucose concentration was just about 2 fold (***Figure 5.III.25***). CS5 expression was 3.2 fold (***Figure 5.III.26***) and EatA expression remained similar at 1% glucose concentration (***Figure 5.III.27***). All the virulence genes were downregulated at 0.25% and 0.5% glucose.

CS6 protein was produced most at 1% glucose. CS5 and EatA were produced most at 1% glucose concentration.

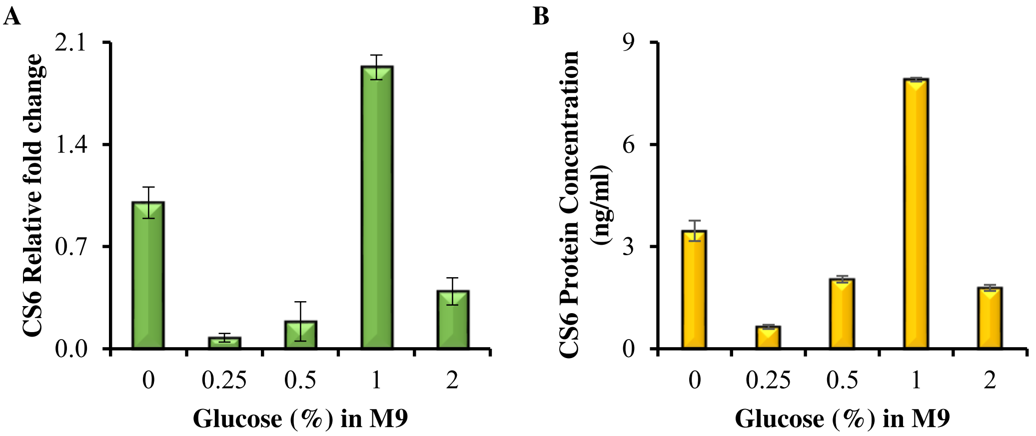

Figure 5.III.25: Effect of glucose on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3). [Glycerol was used as background carbon source]

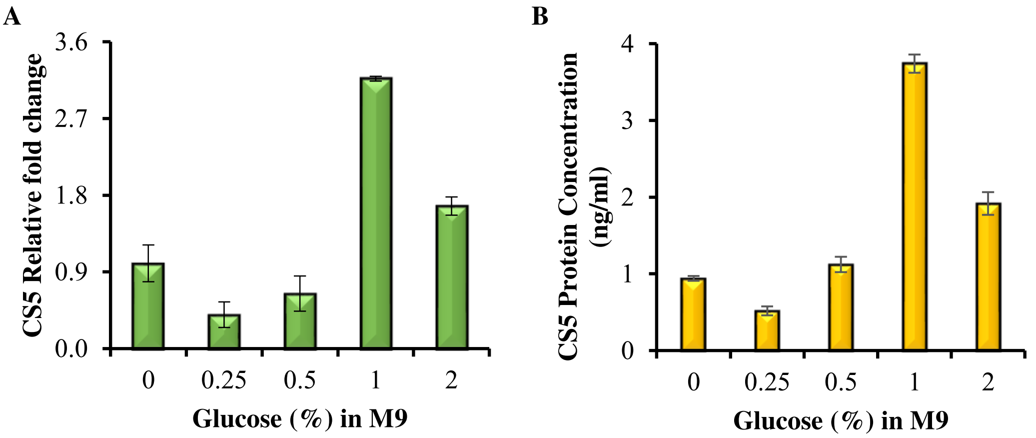

Figure 5.III.26: Effect of glucose on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3). [Glycerol was used as background carbon source]

LT secretion was most at 0.25% glucose concentration in all the tested media. At higher glucose concentrations the LT secretion is lowered. ST secretion is downregulated at all concentrations of glucose.

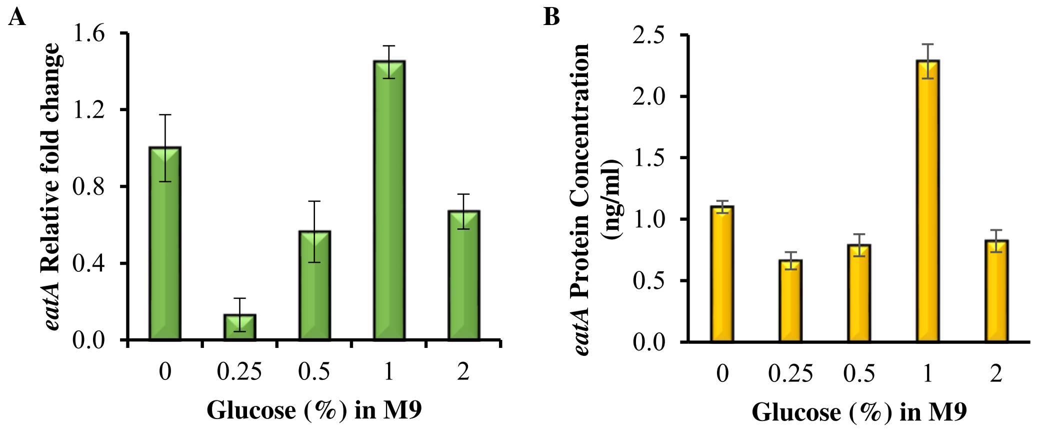

Figure 5.III.27: Effect of glucose on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3). [Glycerol was used as background carbon source]

5.III.2. Effect of different host factors on the expression of virulence genes: Classical colonization factor, CS6 and CS5; Non-classical virulence factor, *eatA*; and enterotoxins, LT and ST.

A. pH

CS6 expression was higher at pH 6 and pH 9 (***Figure 5.III.28***). CS5 (***Figure 5.III.29***) and *eatA* (***Figure 5.III.30***) expression were most at pH 6. Maximum secretion of LT and ST toxin was at pH 8.

CS6 production increased at pH 6 and decreased at alkaline pH which leaps at pH 9. CS5 production was most at pH 6 in all the media and downregulated at pH 8 and 9. But there was a leap in the production at pH 9 than pH 8. EatA production was most at pH 6 in all the media and downregulated at pH 8 and 9.

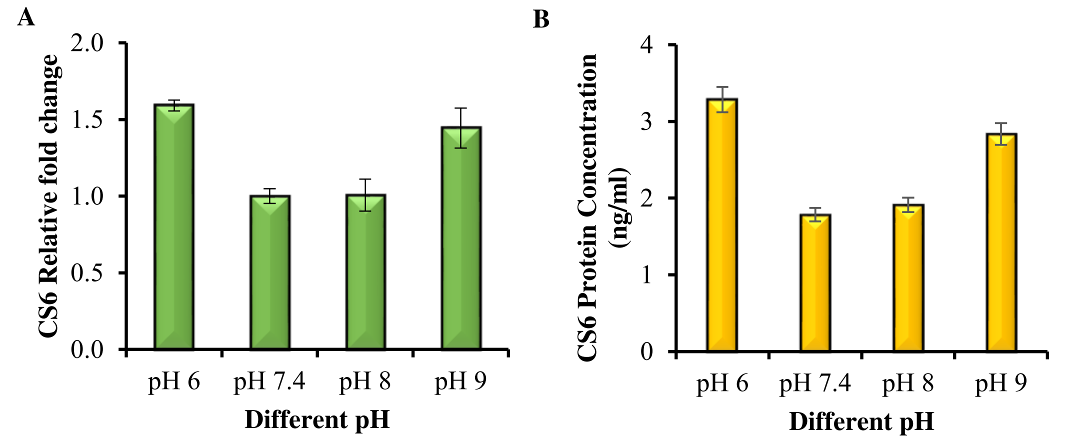

Figure 5.III.28: Effect of different pH on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

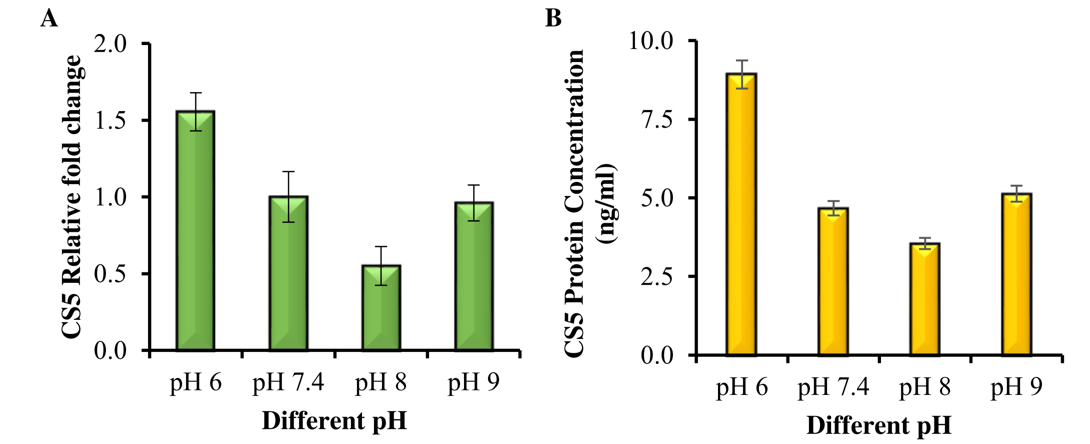

Figure 5.III.29: Effect of different pH on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

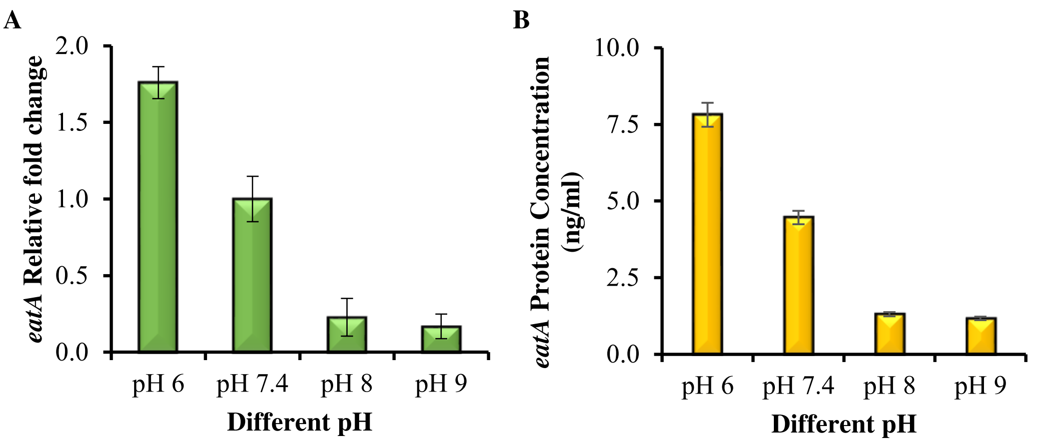

Figure 5.III.30: Effect of different pH on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

B. Temperature

The expression of CS6 was most at 37ºC, an increase of 60 fold when compared with respect to 16 ºC (***Figure 5.III.31***). A similar trend was observed for CS5 (***Figure 5.III.32***) and EatA (***Figure 5.III.33***) expression. LT and ST secretion was also maximum at 37ºC.

The protein production of CS6, CS5 and EatA were maximum at 37ºC.

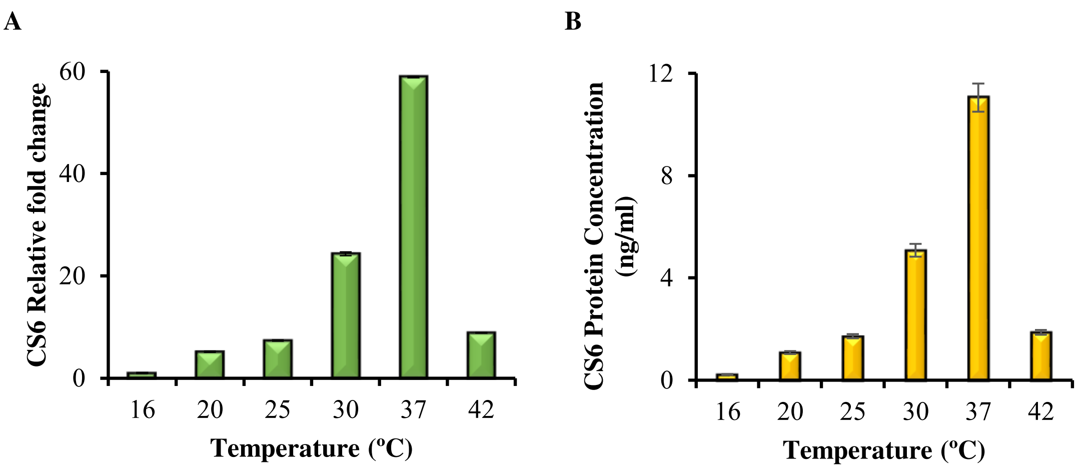

Figure 5.III.31: Effect of different temperature on CS6 expression. The effect on CS6 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

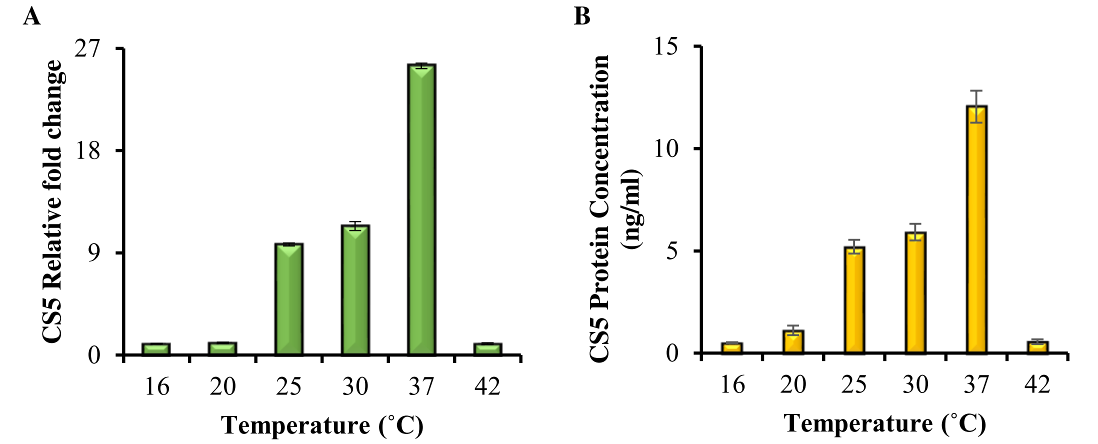

Figure 5.III.32: Effect of different temperature on CS5 expression. The effect on CS5 expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

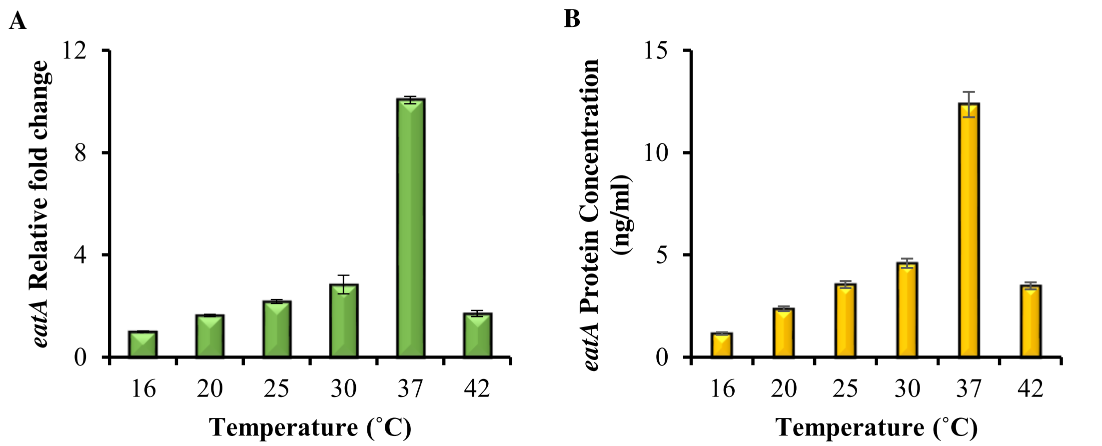

Figure 5.III.33: Effect of different temperature on *eatA* expression. The effect on *eatA* expression was studied at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

5.III.3. Quantitative analysis of Classical colonization factor, CS6 and CS5; Non-classical virulence factor, *eatA*; and enterotoxins, LT and ST expression at different stages of pathogenesis in tissue culture and in the animal model

A. Tissue Culture

The intestinal epithelial HT29 cells were cultured in DMEM supplemented with 10% FBS following standard technique at 37˚C in a humidified CO_2 incubator under 5% CO_2. The HT29 cells at 80% confluency in serum-free medium were infected with varying concentrations of bacteria (10^5 to 10^8 cfu/mL) in triplicates. Control groups were treated with medium alone and laboratory *E. coli* strain, DH5α. ETEC was isolated at different time intervals of infection and analysed by qRT-PCR for expression of CS6, CS5, *eatA* and toxins. After 8 hours the HT-29 cells start to dislodge from the flask surface and at 18 hours the tissues completely dislodge from the surface and do not remain adhered.

For CS6 expression (***Figure 5.III.34***) no significant change could be observed at 2-hour intervals at all the given concentrations. But with the increase in the time interval, the transcription of CS6 significantly increases at concentrations of 10^7 and 10^8 cfu/mL of bacteria. While for the other two concentrations significant changes were measured at 18 hour time interval. At 18 hour interval, the highest two concentrations were same and CS6 expression had reached its maximum.

For CS5 expression (***Figure 5.III.35***) in RNA, no significant change can be observed up to 4 hour interval at all the given concentrations. But with the increase in the time

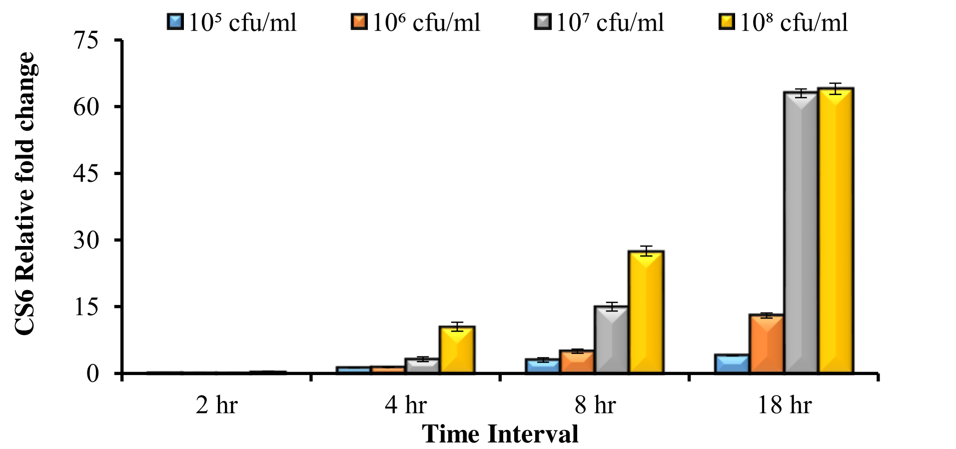

Figure 5.III.34: CS6 expression in contact with HT-29 cells at different time point

interval, the transcription of CS5 significantly increased at all the concentrations of bacteria. A significant increase in RNA level was seen at the 18-hour interval.

For *eatA* expression (*Figure 5.III.36*) in RNA significant change can be observed from 2-hour interval at all the given concentrations. But with the increase in the time interval, the transcription of *eatA* significantly increased at all the concentrations of bacteria. A significant increase in RNA level was seen at 18-hour interval in all the concentrations of bacteria.

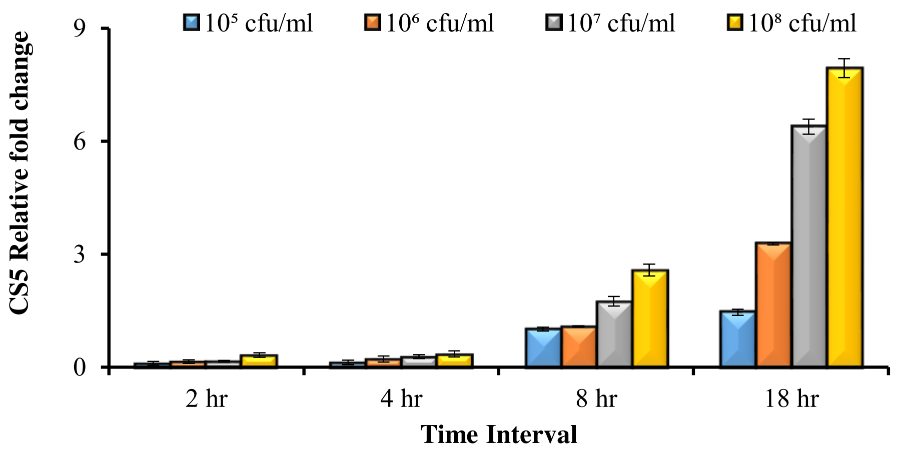

Figure 5.III.35: CS5 expression in contact with HT-29 cells at different time point

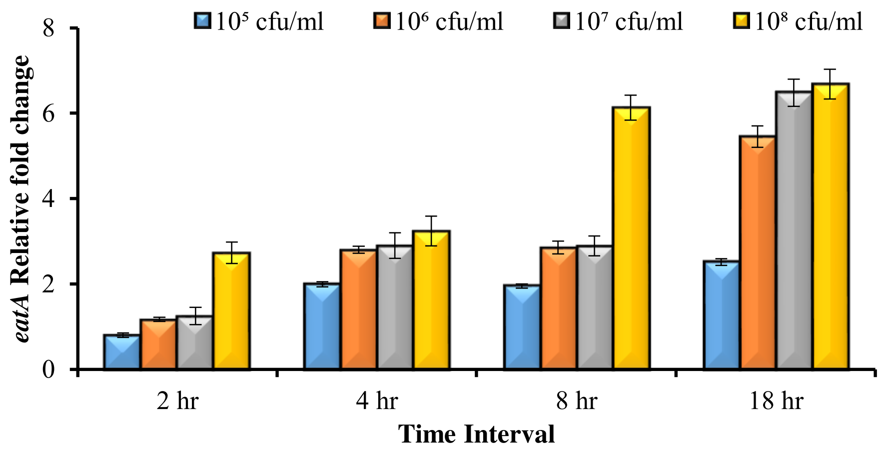

Figure 5.III.36: *eatA* expression in contact with HT-29 cells at different time point

Toxins were most secreted at the highest concentration. LT secretion (***Figure 5.III.37***) is most at the highest concentration of bacteria from the start. In others concentrations, LT secretion reaches pick after 8 hours. For ST secretion (***Figure 5.III.38***) all the concentrations nearly have the same level of St throughout after 2 hours. LT is secreted at a higher level than ST.

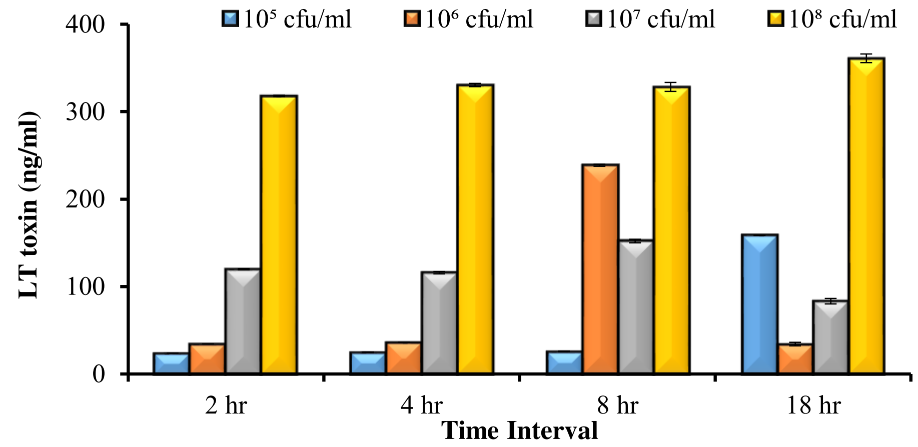

Figure 5.III.37: Expression of LT in contact with HT-29 cells at different time interval

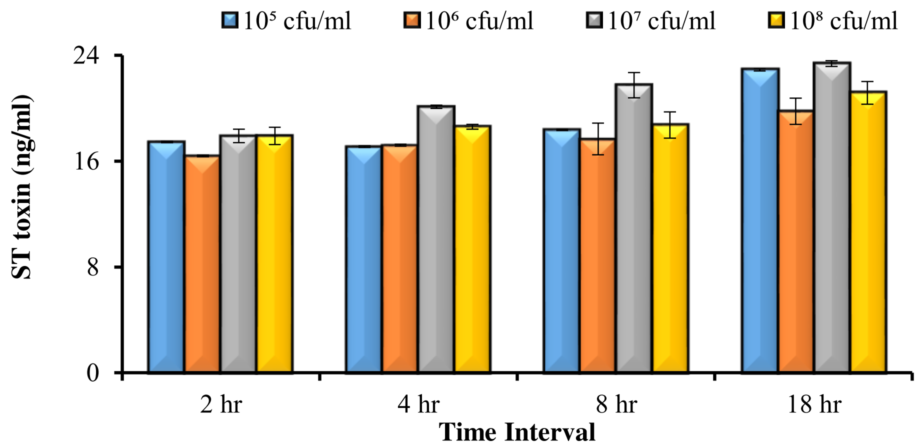

Figure 5.III.38: Expression of ST in contact with HT-29 cells at different time interval

B. Animal Model: Rabbit Ileal Loop

Ileal loop assay (***Figure 5.III.39***) was done and the recovered bacteria from the loop were used for qRT-PCR for transcription analysis of CS6, CS5 and EatA. The accumulated fluid was used to measure the toxin levels by ELISA. The bacterial concentration used were 10^5 cfu/ mL, 10^6 cfu/ mL, 10^7 cfu/ mL and 10^8 cfu/ mL.

In a rabbit model, lower cell density did not produce fluid. At higher cell density, there was an increased fluid accumulation ratio (FA) (***Figure 5.III.40***).

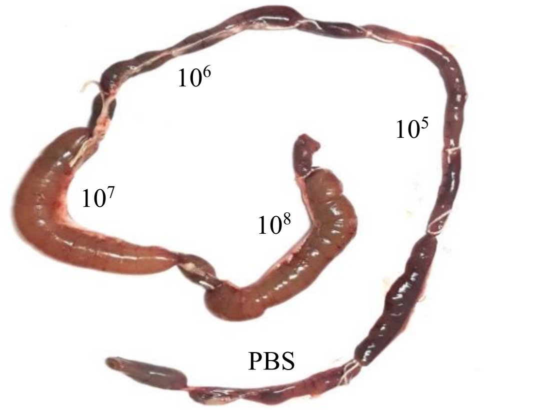

Figure 5.III.39: Rabbit Ileal Loop Assay with different concentration of bacteria (PBS loop was used as control)

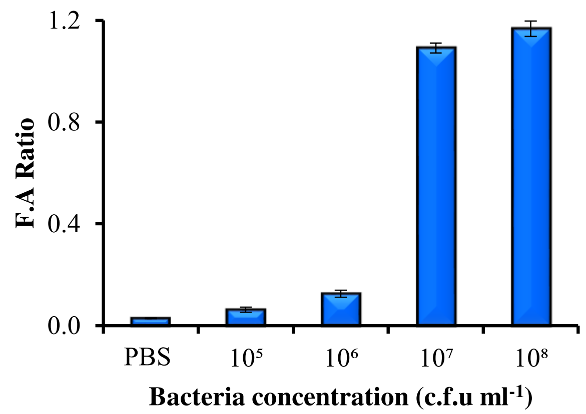

Figure 5.III.40: Fluid accumulation ratio in Rabbit ileal loop by different concentration of bacteria.

All the genes expression was higher under this condition. The expression of the 3 genes was most at 10^8 cfu/mL and lowest at 10^5 cfu/mL (***Figure 5.III.41; Figure 5.III.42; Figure 5.III.43***). Toxins were most secreted at the highest concentration. With the increase in the concentration of bacteria the toxin levels were also increased.

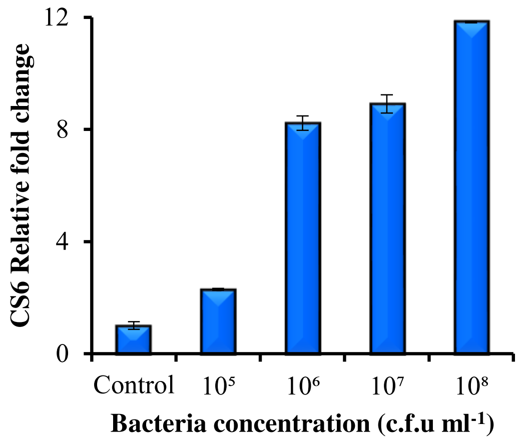

Figure 5.III.41: Expression of CS6 from Rabbit ileum at different concentrations of ETEC

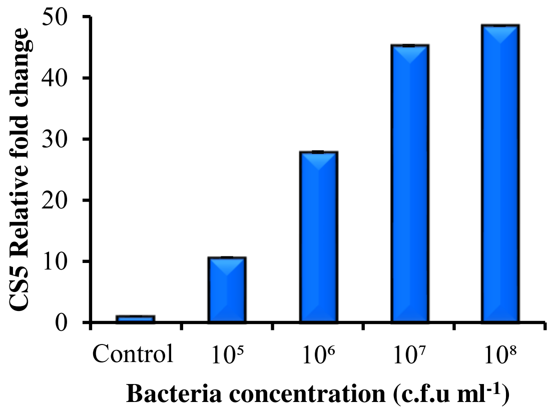

Figure 5.III.42: Expression of CS5 from Rabbit ileum at different concentrations of ETEC

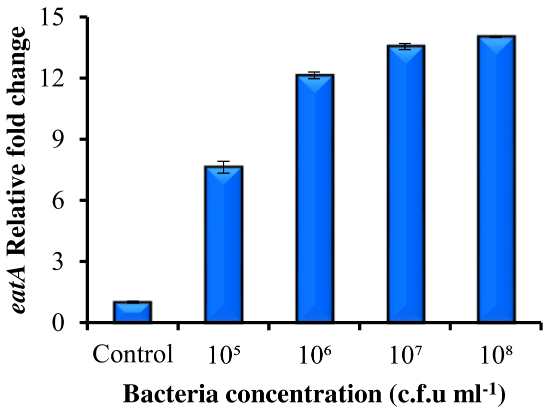

Figure 5.III.43: Expression of *eatA* from Rabbit ileum at different concentrations of ETEC

Histological observation:

The tissues were collected from the ileal loops and histological changes were observed at different concentrations of ETEC after staining with haematoxylin/eosin. The ileum (Rabbit) tissues were observed under the microscope in 10X magnification. There was almost no histologic damage in infection with the lowest cell density (10^5 cfu/mL) but with increasing density of bacterial infection histologic damage was more and more prominent in each step, in comparison with PBS treated loop (***Figure 5.III.44***).

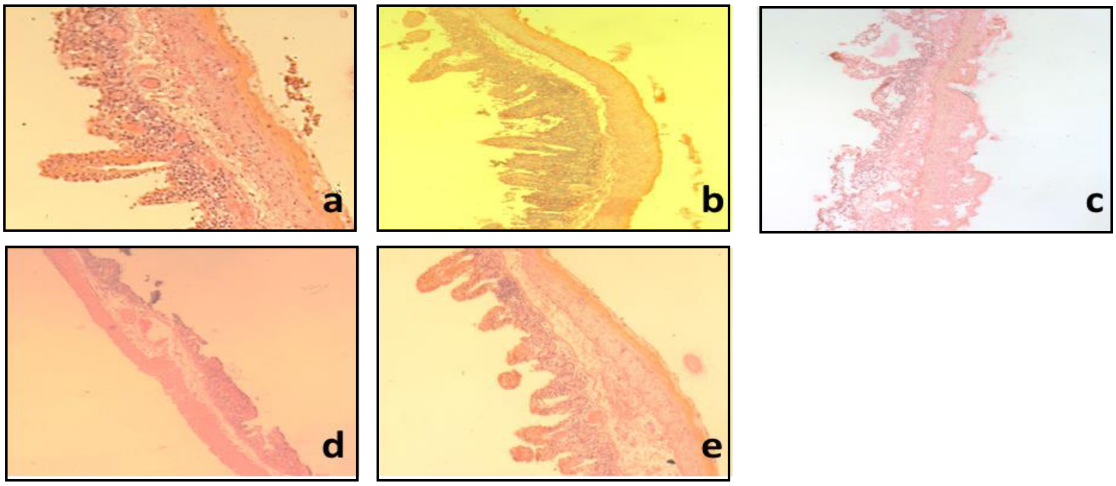

Figure 5.III.44: Histological images of Rabbit ileum at different concentration of ETEC. Development of histologic damage (Indicated with arrow) in rabbit small intestine (ileum) with increasing cell density from 10^5 cfu/mL **(a)**, 10^6 cfu/mL **(b)**, 10^7 cfu/mL **(c)**, and 10^8 cfu/mL **(d)** in comparison with control **(e)** (PBS treated) can be observed.

5.III.4. Promoter Characterization

The promoter of CS6, CS5 and *eatA* were predicted using bioinformatics ‘softberry bacterial promoter prediction.’ Upon prediction, we cloned that part in a promoter-less plasmid, pTL61T and check for promoter activity. For negative control, the empty plasmid was introduced into strains. Deletion of the promoter was also prepared. The promoter constructs were transformed into ETEC test strains and relative β-Galactosidase activity expressed in arbitrary units was used to analyse promoter activity.

CS6 promoter: The result of the prediction for CS6 Promoter (***Figure 5.III.45***): *-10 box* gtatacaat; *-35 box* ttgaca; oligonucleotides from known TF binding sites: *arcA*: aataaaaa; *argR*: ataaaaat; *fur*: aaaatgat. We cloned the predicted promoter region upstream and some portion of *cssA* gene of CS6 from -573 to +25 bases in pTL61T plasmid and transformed them into test strains. Indeed this region was the promoter region as after cloning the β-galactosidase activity of the plasmid was observed by miller assay (***Figure 5.III.46***). The effect of different host and environmental factors on the promoter was checked to find the optimum concentration for maximum promoter activity which are given in the table (***Table 5.III.1***).

A series of regions upstream of CS6 were also deleted and studied for promoter activity after deletion to locate the key region for promoter activity. When upstream bases from -573bp to -350bp were deleted, promoter activity remained the same as without deletion. But after the deletion of the upstream region -350bp to -255bp, promoter expression decreases drastically to 26% from 99% (***Figure 5.III.47***). Further deletion also

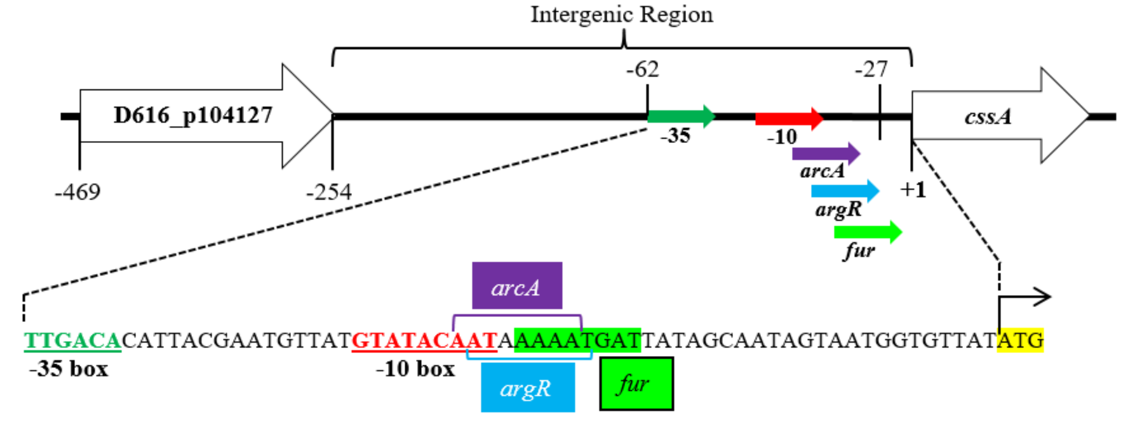

Figure 5.III.45: Pictorial representation of Predicted CS6 promoter region with oligonucleotides from known TF binding sites are indicated and highlighted.

decreases promoter activity up to a little range. So the region -350bp to -255bp holds the key promoter sequence for the CS6 gene.

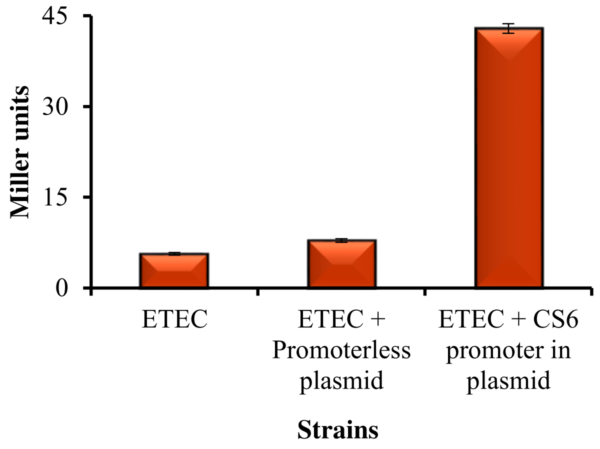

Figure 5.III.46: Expression of predicted CS6 promoter in promoter-less plasmid by β-galactosidase assay

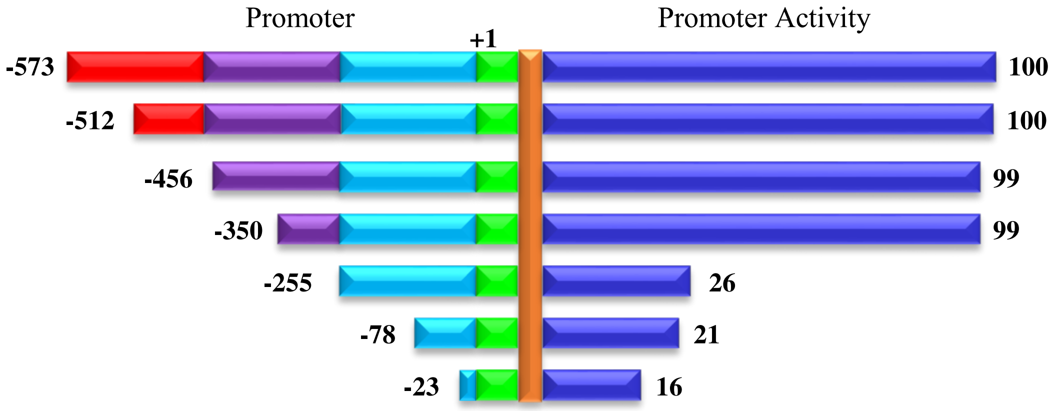

Figure 5.III.47: Effect of promoter deletion on CS6 promoter activity.

CS5 promoter: The result of the prediction for CS5 Promoter (***Figure 5.III.48***): *-10 box* tgttattgt; *-35 box* ttgacg; oligonucleotides from known TF binding sites: *rpoD17*: ttttactt; *farR*: tgtattat. We cloned the predicted promoter region upstream and some portion of *csfA* gene of CS5 from -207 to +25 bases in pTL61T plasmid and transformed into test strains. Indeed this region was the promoter region as after cloning the β-galactosidase activity of the plasmid was observed by miller assay (***Figure 5.III.49***). The effect of different host and environmental factors on the promoter was checked to find the optimum concentration for maximum promoter activity which are given in the table (***Table 5.III.1***).

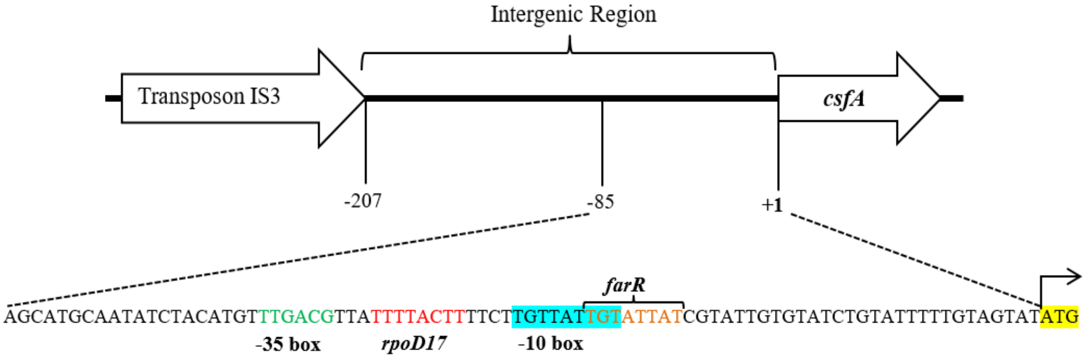

Figure 5.III.48: Pictorial representation of Predicted CS5 promoter region with oligonucleotides from known TF binding sites are indicated and highlighted.

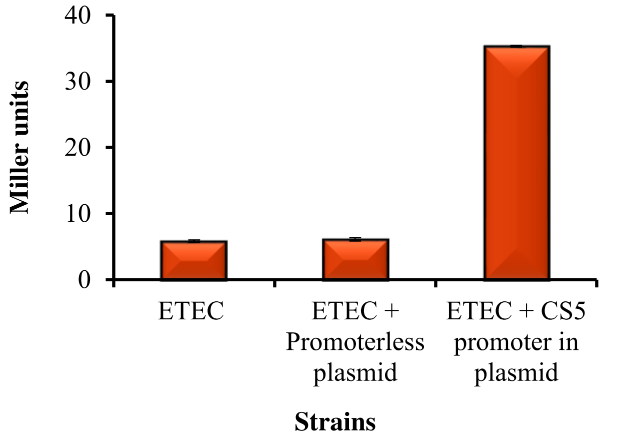

Figure 5.III.49: Expression of predicted CS5 promoter in promoter-less plasmid by β-galactosidase assay

A series of regions upstream of CS5 were also deleted and studied for promoter activity after deletion to locate the key region for promoter activity. When upstream bases from -573bp to -85bp were deleted, promoter activity remained the same as without

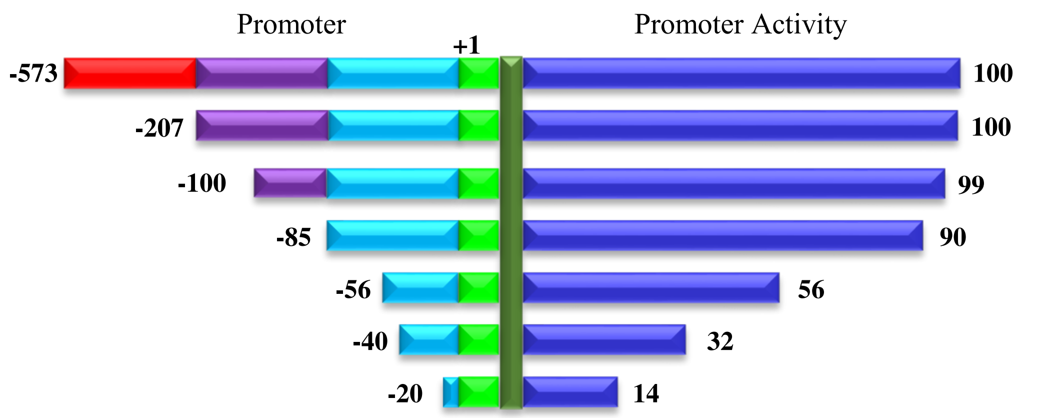

Figure 5.III.50: Effect of promoter deletion on CS5 promoter activity.

deletion. But after the deletion of the upstream region -85bp to -56bp, promoter expression decreases drastically to 56% from 90% (***Figure 5.III.50***). Further deletion also decreases promoter activity up to a little range. So the region -85bp to -85bp holds the key promoter sequence for the CS5 gene.

eatA* promoter:** The result of the prediction for *eatA* Promoter (Figure 5.III.51***): *-10 box* tgctaaaat; *-35 box* ttaaaa; oligonucleotides from known TF binding sites: *cpxR*: taaaaaga; *arcA*: aataaaaa; *metJ*: gctaaaat; *lexA*: ataaataa; *tyrR*: taaataaa; *crp*: aataaatt; *argR*: aattcaat. We cloned the predicted promoter region upstream and some portion of *eatA* gene from -286 to +25 bases in pTL61T plasmid and transformed them into test strains. Indeed this region was the promoter region as after cloning the β-galactosidase activity of the plasmid was observed by miller assay (***Figure 5.III.52***). The effect of

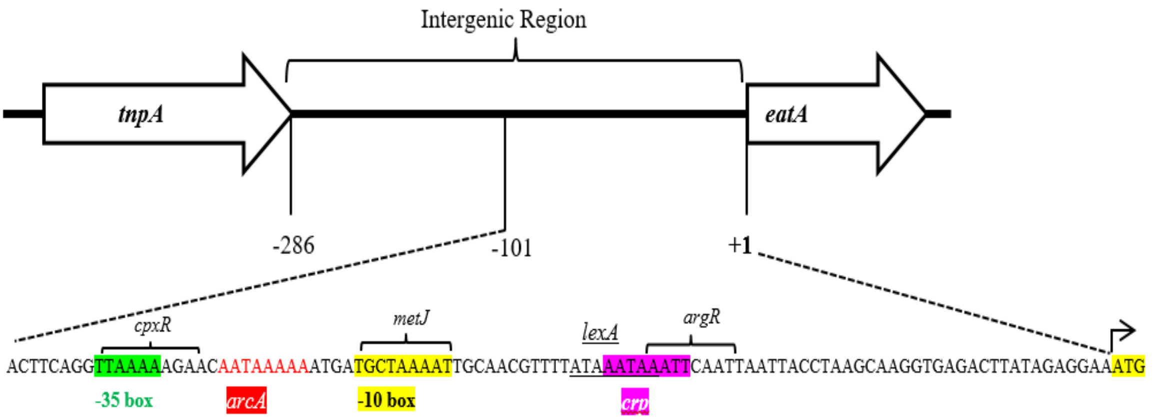

Figure 5.III.51: Pictorial representation of Predicted *eatA* promoter region with oligonucleotides from known TF binding sites are indicated and highlighted.

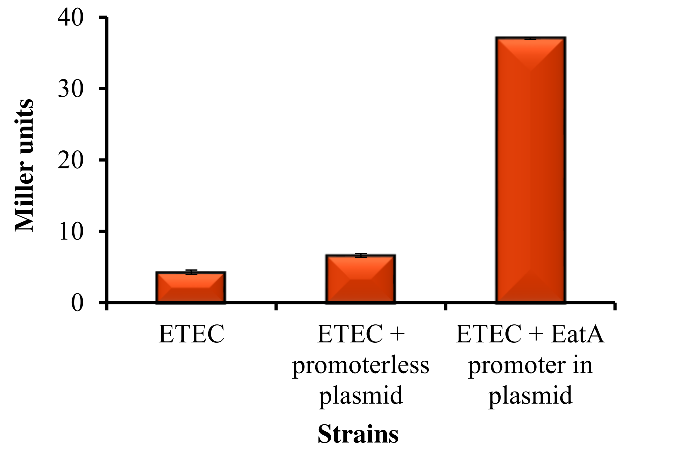

Figure 5.III.52: Expression of predicted *eatA* promoter in promoter-less plasmid by β-galactosidase assay

different host and environmental factors on the promoter was checked to find the optimum concentration for maximum promoter activity which are given in the table (***Table 5.III.1***).

A series of regions to the upstream of *eatA* were also deleted and studied for promoter activity after deletion to locate the key region for promoter activity. When upstream bases from -286bp to -82bp were deleted, promoter activity remained the same as without deletion. But after the deletion of the upstream region -82bp to -60bp, promoter expression decreased drastically to 30% from 85% (***Figure 5.III.53***). Further deletion also decreases promoter activity up to a little range. So the region -82bp to -60bp holds the key promoter sequence for *eatA* gene.

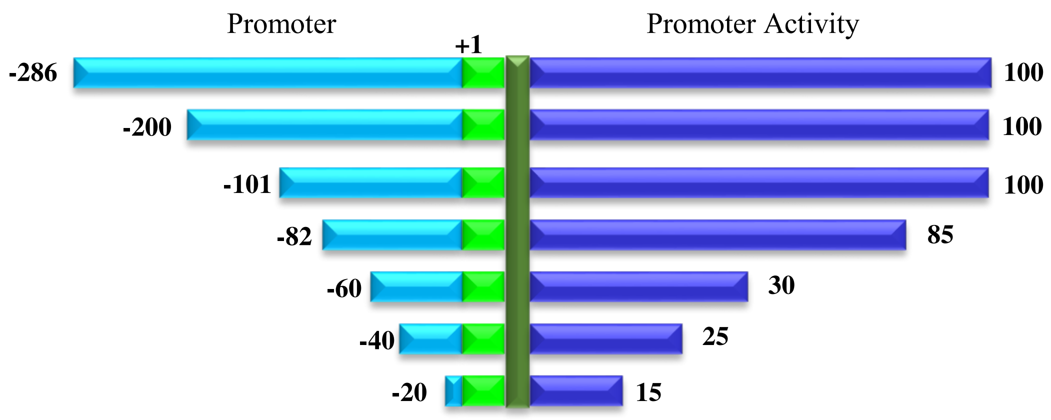

Figure 5.III.53: Effect of promoter deletion on *eatA* promoter activity.

Table 5.III.1: **Concentration of additives at highest promoter activity**

Components	CS6	CS5	EatA
Crude bile	N/A	0.05 %	N/A
Sodium cholate hydrate	0.05 %	0.05 %	0.05 %
Sodium deoxy cholate	N/A	0.2 %	0.2 %
Sodium tauro deoxy cholate	0.1 %	0.1 %	0.1 %
Sodium cheno deoxy cholate	N/A	0.1 %	0.1 %
Taurocholic acid sodium salt	0.1 %	0.1 %	N/A
Iron (II) $FeSO_4$	0.2 mM	0.2 mM	0.2 mM
Glucose	1 %	N/A	1 %
Mucin	N/A	N/A	N/A
pH	6	6	7.4
Temperature	37˚ C	37˚ C	37˚ C

5.III.5. Effect of iron salts on prevalent colonization factor, CS6

As iron salts was emerged as the main inducible component which showed the maximum impact on the prevalent colonization factor, CS6 we investigated further into the effects of iron salts on CS6.

A. Surface expression of CS6 was induced in presence of Iron: The surface expression of CS6 increased gradually in a concentration-dependent manner in presence of iron salt however it decreased when the concentration was 0.4mM and beyond (***Figure 5.III.54***). According to the findings, the optimum dose for the highest CS6 expression was 0.2 mM iron (II).

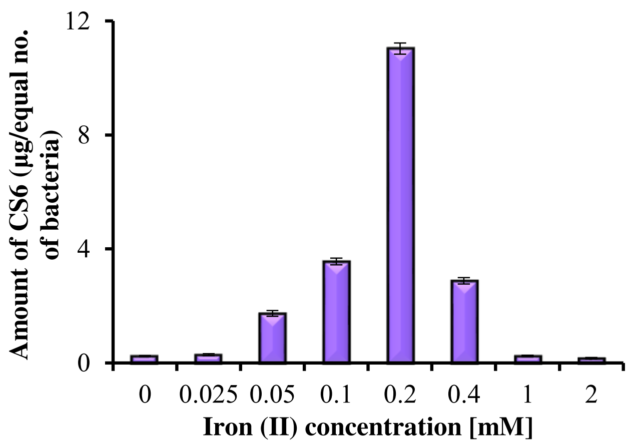

Figure 5.III.54: Phenotypic expression of CS6 on ETEC surface in presence of iron.

B. Transcription of CS6 under other iron compounds: Iron mainly occurs in either of two oxidation states in biological systems as and Fe (III) or ferrous and ferric, respectively. Fe (II) is favoured by low pH and low oxygen concentrations. Ferric (Fe III) iron is the predominant form of iron under aerobic conditions. The basis of many redox reactions in cells is the ability of iron to be interconverted between these two states. In this study, we also investigated the expression of CS6 in ferric iron compounds as ETEC has to survive in the outer environment outside the host and outside environment is aerobic, where the predominant form of iron in ferric. CS6 expression was also enhanced in the presence of Fe (III) compounds. In both forms ferrous (***Figure 5.III.55***) and ferric (***Figure 5.III.56; Figure 5.III.57***), iron induces the expression of CS6 in the *in vitro* culture conditions.

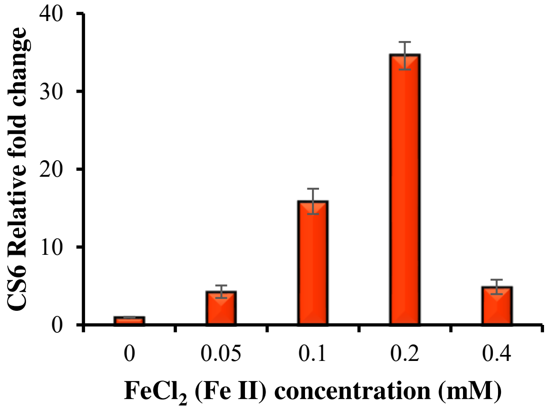

Figure 5.III.55: Effect of $FeCl_2$ on CS6 expression was studied at RNA level by quantitative RT-PCR. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

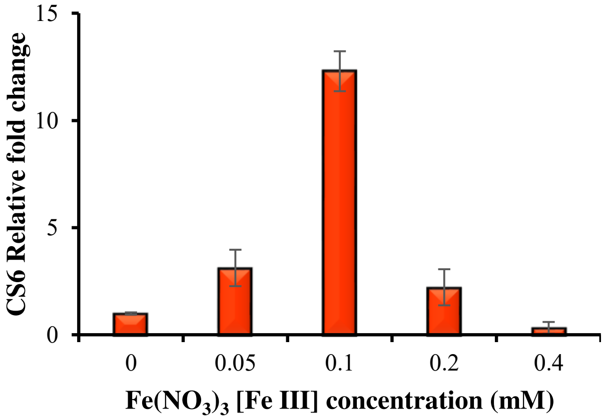

Figure 5.III.56: Effect of Fe(NO3)3 on CS6 expression was studied at RNA level by quantitative RT-PCR. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

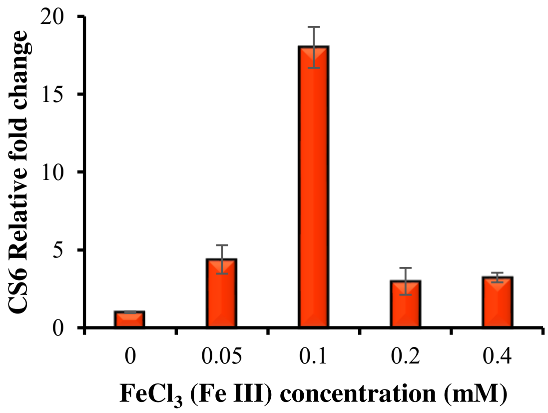

Figure 5.III.57: Effect of $FeCl_3$ on CS6 expression was studied at RNA level by quantitative RT-PCR. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

C. **Transcription of CS6 under iron chelation was reduced:** Upon supplementation of iron chelator along with iron salt at 0.2mM in a concentration-gradient manner, showed that CS6 expression decreased. At 2mM and 4mM chelator, deferoxamine (Def) (***Figure 5.III.58***) and 2,2' dipyridyl (Dip) (***Figure 5.III.59***) concentration, the expression of CS6 was downregulated.

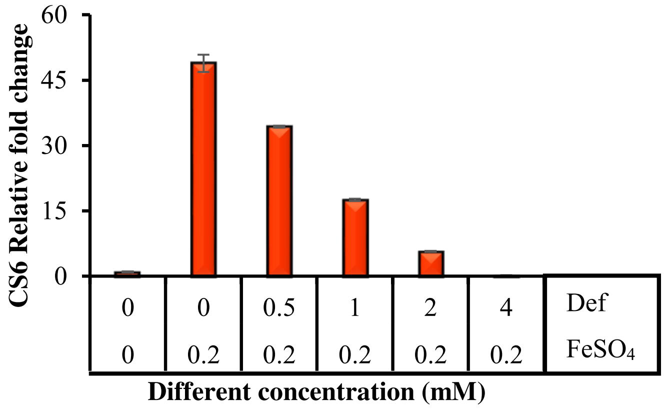

Figure 5.III.58: Effect on CS6 expression of ETEC in presence of iron chelation compound, deferoxamine (Def) in dose-dependent concentration.

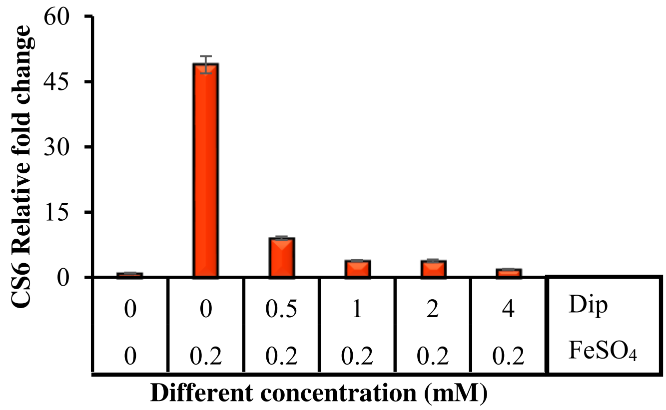

Figure 5.III.59: Effect on CS6 expression of ETEC in presence of iron chelation compound, 2,2' dipyridyl (Dip) in dose-dependent concentration.

D. **Adhesion of ETEC to HT-29 in presence of iron was enhanced:** For ETEC, adhesion to the epithelial layer of the small intestine is a prerequisite for pathogenesis. For evaluating bacterial attachment, co-culturing HT-29 cells with CS6 harbouring ETEC clinical isolates in the presence of 0.2mM Fe(II) was done.

We found that iron significantly enhanced the ability of ETEC cells to attach to HT-29 cells compared to untreated ETEC.

E. **Iron enhanced the growth of ETEC in presence of iron:** Iron in a dose-dependent manner enhanced the growth of ETEC *in vitro*. It was found that up to 0.2mM concentration of iron the growth increased but beyond that concentration the growth was retarded (***Figure 5.III.60***).

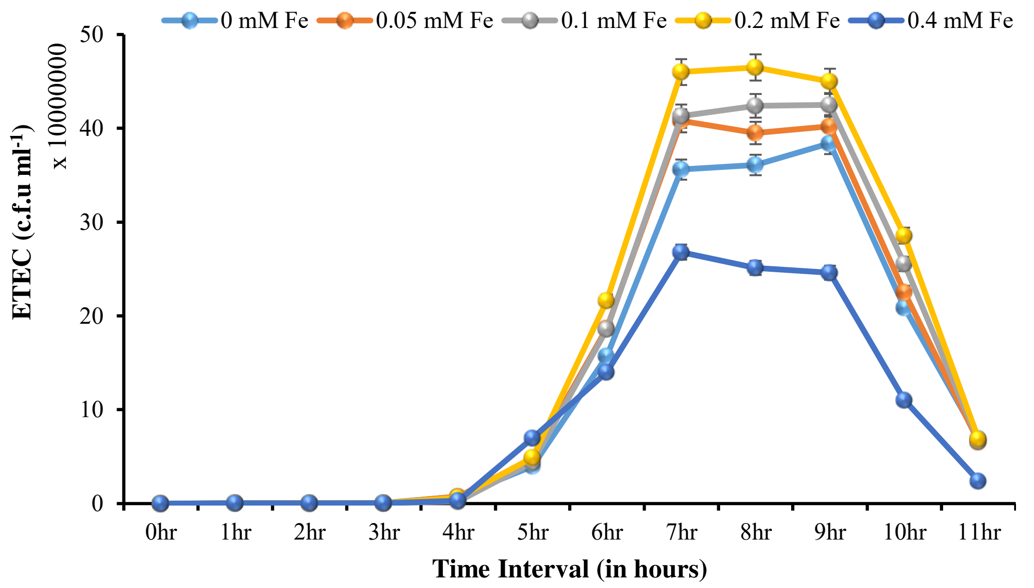

Figure 5.III.60: Effect of iron salt ($FeSO_4$) on ETEC growth as measured by colony counting method

F. **Animal Experiments:**

In vivo adhesion of bacterial strains were also evaluated in animal models.

Rabbit ileal loop assay: In young New Zealand white rabbits, the Rabbit ileal loop assay (***Figure 5.III.61***) was performed. ETEC strains grown in the presence and absence of varying concentrations of iron salts were inoculated in rabbit ileum. As a negative control, a PBS-inoculated loop was employed.

The FA ratio was calculated after 18 hours of incubation, and intestinal loop sections were washed three times in PBS, homogenised, and serially diluted in PBS. Introducing these bacterial cultures on MacConkey agar plates yielded the adhering bacterial count. When ETEC was grown in presence of 0.2 mM iron salts then adherence increased by 2 fold whereas in presence of 2 mM iron the growth retarded by half in comparison to the untreated condition (***Figure 5.III.62***).

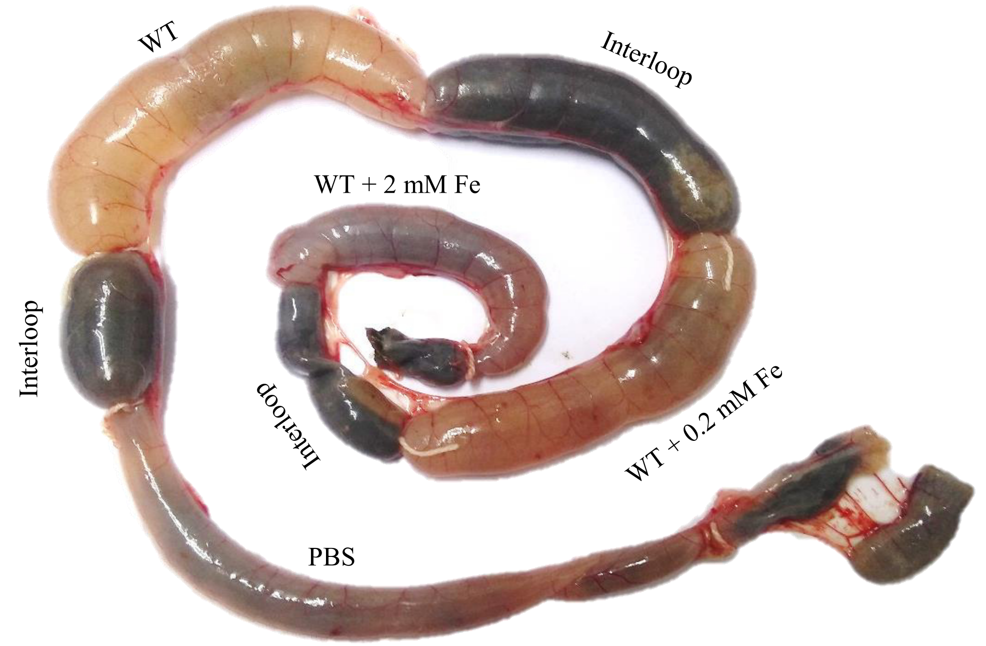

Figure 5.III.61: Rabbit ileal loop assay to evaluate enterotoxigenic activity of ETEC at different iron concentrations. Rabbit ileal loops were inoculated with 10^7 cfu of ETEC strain (WT) and incubated for 18 h. PBS strain was used as control

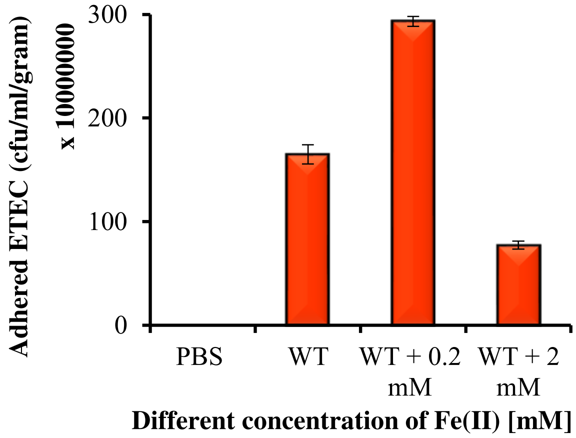

Figure 5.III.62: Effect of iron on ETEC adherence or colonization to the rabbit ileal epithelium by ligated rabbit ileal loop assay.

Fluid accumulation was highest in the loop that had 0.2mM iron supplementation followed by the loop that had no iron supplementation and lowest in the loop that had 2mM iron supplementation (***Figure 5.III.63***).

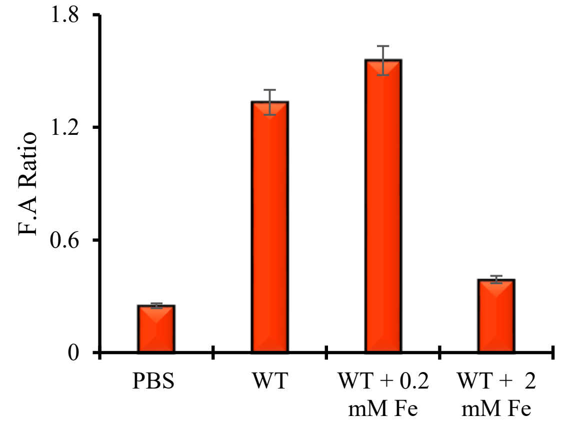

Figure 5.III.63: Analysis of fluid accumulation of WT ETEC strain in presence of different concentrations of iron in Rabbit ileal loop assay. The results are reported as fluid accumulation (FA) (in millilitres) per loop length (in centimeters). The mean ± SEM are shown; n = 3.

Suckling Mice assay: In the suckling mice assay the adherence was more when iron was supplemented as compared to no iron supplementation (***Figure 5.III.64***). It was most in 0.2mM iron addition condition. Fluid accumulation was highest in the wild type strain loop that had no iron supplementation and lowest in the loop that had 2mM iron supplementation (***Figure 5.III.65***).

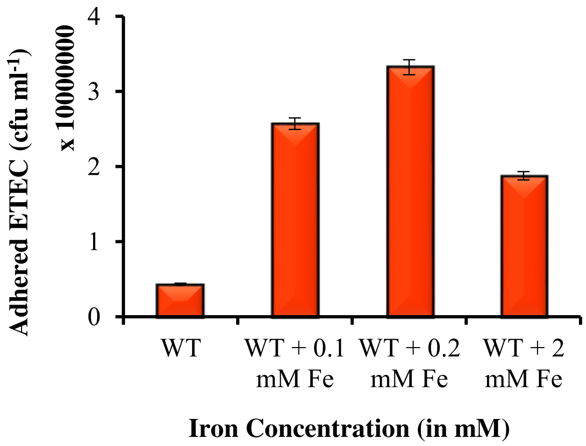

Figure 5.III.64: Mice colonization assay to demonstrate the effect of iron on ETEC adherence or colonization to the mice intestinal epithelium by suckling mice assay.

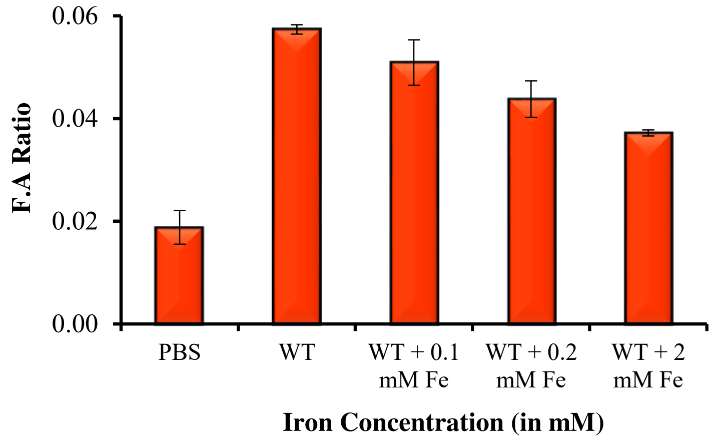

Figure 5.III.65: Analysis of fluid accumulation of WT ETEC strain in presence of different concentrations of iron in suckling mice. Shown are mean ± SEM; n = 3.

5.III.6. Effect of iron salts on the expression of CS6 in a dysregulated condition of CS5 and *eatA* and effect on toxin genes (*elt* and *est*)

CS6 being one of the prevalent colonization factor in this region and it co-occurs with CS5 and *eatA*, it would be interesting to find out does these virulence factors affect the expression of CS6. We knockout CS5 and *eatA* gene by λ-Red recombination method from a strain that harboured CS6, CS5 and *eatA* and prepare the following mutant strains – 1. ΔCS5CS6; 2. Δ*eatA*CS6; 3. ΔCS5Δ*eatA*CS6 and tested the effect of iron salts that was previously enhancing the expression and production of CS6 by qRT-PCR. For the experiments, HT 29 cells were infected with ETEC wild type (WT) strains (strain harbouring CS6, CS5 and EatA) and mutants at 10^7 cfu/mL and different additives. We compare the conditions at which we get the maximum expression to the condition closely resembling the normal human physiological concentrations. CS6 protein production was also measured. We also checked the effect by tissue culture and animal model.

A. Toxin expression

Expression of both the toxins does not get affected by the dysregulated condition of CS6 (***Figure 5.III.66; Figure 5.III.67***). The toxins and the CFs have independent expression.

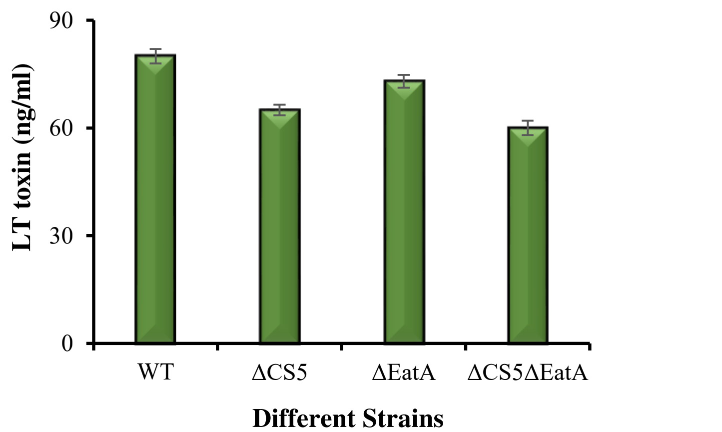

Figure 5.III.66: Expression of LT enterotoxin in contact with HT-29 cells in different mutant strains and wild-type strain.

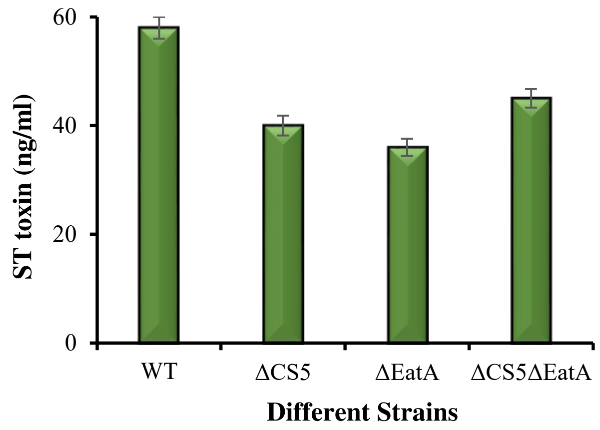

Figure 5.III.67: Expression of ST enterotoxin in contact with HT-29 cells in different mutant strains and wild-type strain.

B. Iron Salt

CS6 expression is up-regulated in wild type and all mutant conditions when 0.2mM of Iron Salt was added to the media. In ΔCS5 (CS6+*eatA*) condition expression of CS6 is down-regulated by 3fold, in Δ*eatA* (CS6+CS5) condition CS6 expression down-regulated by 1.34 fold and in double mutant condition or ΔCS5Δ*eatA* (CS6 only) condition CS6 expression is the lowest showing approximately 4 fold decrease (***Figure 5.III.68***). Here the comparison is done with the WT strain. This pattern of change in CS6 expression is also seen in the case of protein lysate analysis. Here also wild type strain shows the highest expression of protein and lowest in double mutant condition.

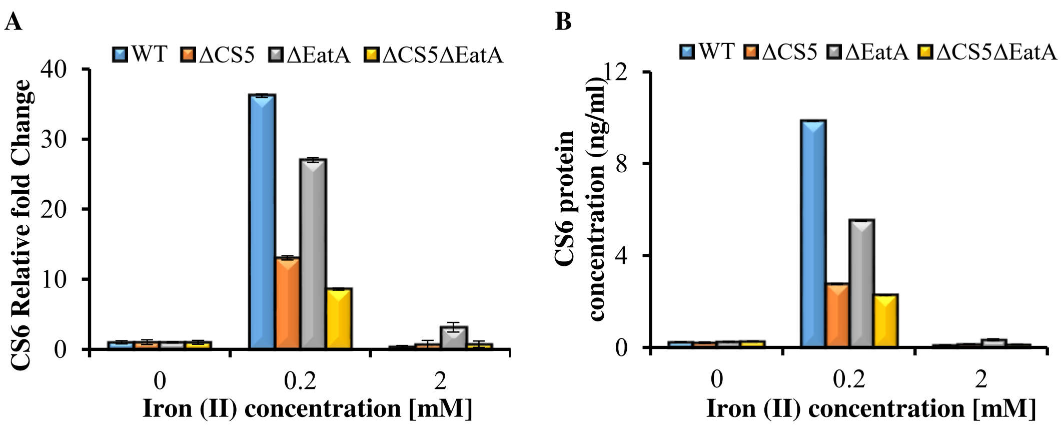

Figure 5.III.68: Effect of Fe (II) on CS6 expression in WT (wild-type) strain and different mutant strains. The effect on CS6 expression was studied in HT-29 tissue culture at (A) RNA level by quantitative RT-PCR, (B) protein level by ELISA. The data was provided as mean ± SEM after each experiment was conducted three times (n=3).

C. *In vivo* analysis

Mice Colonization Assay: BALB/c mice, aged 4 to 5 days old were orally administered with wild type ETEC and different mutant variants.

Adherence was most in the wild type strain that harboured CS6, CS5 and EatA followed by ΔCS5 mutant and Δ*eatA* mutant. The least adherence was observed in the double mutant strain, Δ*eatA* ΔCS5 variant (***Figure 5.III.69***).

Wild type strain showed the highest fluid accumulation as well as the highest cfu/mL count. The double mutant variant showed the least number of cfu/mL (***Figure 5.III.70***).

LT and ST toxin expression were most in the WT strain. However all the mutant variants did not had much difference in their toxin profile (***Figure 5.III.71***, ***Figure 5.III.72***).

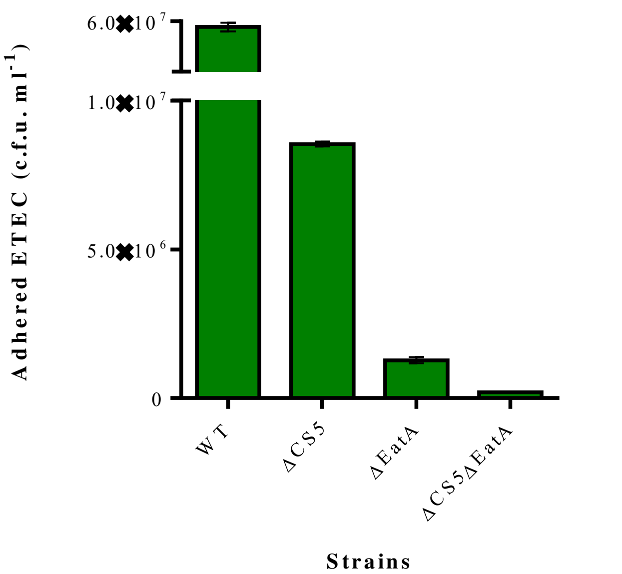

Figure 5.III.69: Adherence or colonization of different strains, WT and mutant variants to the rabbit ileal epithelium by suckling mice colonization assay.

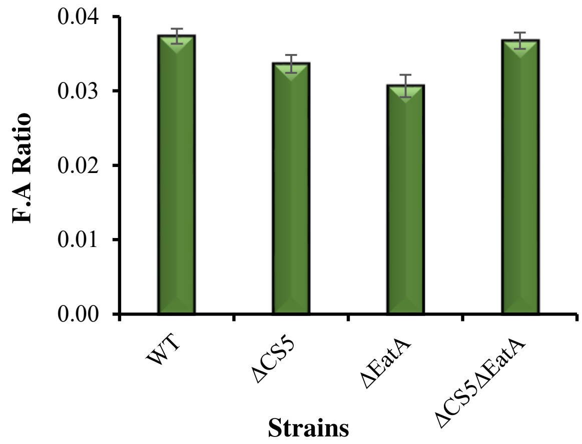

Figure 5.III.70: Analysis of fluid accumulation of WT and mutant ETEC strain in Mice colonization assay. Shown are mean ± SEM; n = 3.

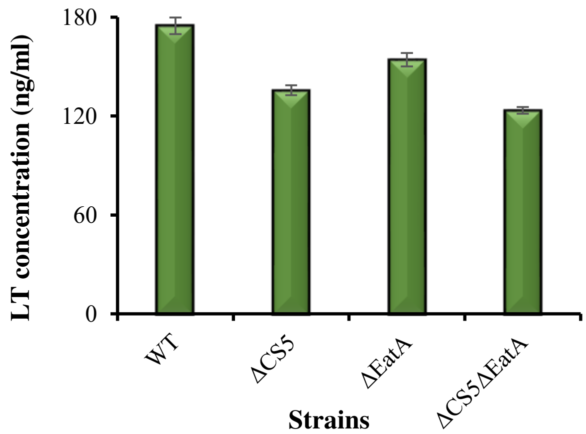

Figure 5.III.71: Expression of LT enterotoxin from suckling mice colonization assay in wild-type strain and different mutant strains.

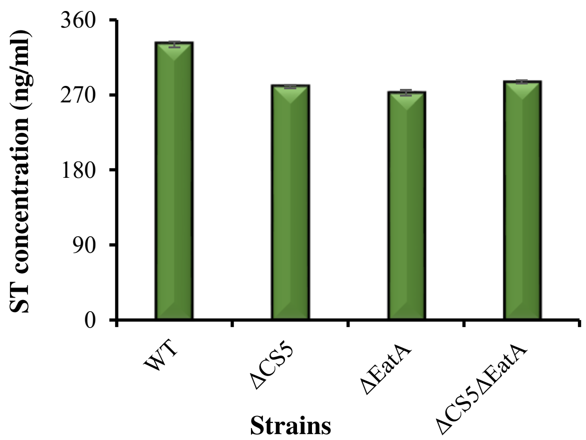

Figure 5.III.72: Expression of ST enterotoxin from suckling mice colonization assay in wild-type strain and different mutant strains.

Chapter 6

DISCUSSION

One of the most common enteric pathogens worldwide is Enterotoxigenic *Escherichia coli* (ETEC), causing significant death and morbidity, especially in low- and middle-income nations (Anderson IV *et al.,* 2019) and also in travellers to these areas. ETEC is among the 4 pathogens responsible for moderate-to-severe diarrhoea among sub-Saharan Africa and South Asia children population aged 0–59 months (Kotloff *et al.,* 2013). According to WHO reports annual deaths due to ETEC is attributable to 280-400 million diarrhoeal episodes in children below the age of 5 years and 74,100 deaths annually. Despite the fact that mortality has decreased in recent decades, morbidity has not (Hosangadi *et al.,* 2018).

To lower the ETEC burden, drinkable water and sanitation, as well as effective clinical care for acute disease, are essential. These treatments, on the other hand, are not always feasible and may not provide enough long-term coverage. Vaccination may be the most budget-friendly and fundamental approach to primary prevention. However, because multiple virulence factors are important vaccine candidates and are vital for ETEC pathogenesis, more evidence is needed to speed up investment and enhance decision-making for ETEC vaccinations. There is currently no vaccination available that offers broad protection against ETEC. The difficulties in developing a vaccine can be linked to the wide variety of ETEC isolates found around the world, as well as the different serotypes.

It is generally known that CFs and non-classical virulence factors are geographically variable, with different virulence factors dominating in different parts of the world. As a result, the ETEC vaccination could need to be region-specific. ETEC vaccine candidate probably should contain enterotoxins, prevalent colonization factors and non-classical virulence factors (Zhang *et al.,* 2012; Svennerholm *et al.,* 2016). Because ETEC has numerous significant virulence characteristics, a multivalent vaccine design could give broad protection against a variety of ETEC strains. The use of multivalent chimeric proteins in the development of vaccines could also be a viable option. In both circumstances, it's critical to understand the most common virulence factors so that the most significant ones can be targeted for a broad spectrum of protection. As a result, a thorough epidemiological examination is required to learn about the most frequently distributed strain and virulence genes in this area.

In order to find the most prevalent virulence determinants circulating in this region, ETEC strains were isolated from diarrhoeal patients during the period 2008 to 2014 and toxigenic genes, colonization factor genes, and non-classical virulence genes were all screened. According to this study, the detection rate of ETEC has been 1.5% across the years, with the highest detection rate in 2014 (1.8 %) and the lowest detection rate in 2011 (0.25%). The frequency of ETEC in this region is substantially lower than in previous reports (Dutta *et al.,* 2013), but much higher than in other parts of India (Shetty *et al.,* 2012).

Though ETEC is most commonly related with children under the age of five, it also impacts adults, which was also studied in relation to ETEC. Studies related to ETEC strains are often done from children below the age of 5 years as evident from GEMS study conducted in children <59 months in years in Asia and Africa (Kotloff *et al.,* 2013). According to a Minnesota research, the average age of patients was 44 years, out of those samples 15% of cases were under the age of 18 (Buuck *et al.,* 2020). ETEC is mostly discovered in children under the age of five and children aged five to fourteen, as previously reported from our region (Dutta *et al.,* 2013); however, our findings revealed that the age groups most affected were those under the age of five and those over the age of eighteen, which is consistent with further studies of ETEC isolates during 2015-2019 from our lab (Mondal, I. *et al.,* 2021) and with the findings of a Shanghai surveillance research (Tang *et al.,* 2015). The type of diarrhoea is substantially connected to age; severe diarrhoea is associated with age groups <5 years, whereas no diarrhoea is associated with age groups ≥5 years (6.8% versus 77%; $p<0.005$).

We found that in both the sole and mixed clinical ETEC strains *elt+est* are the most prevalent toxin profile, followed by *est*-alone strains, while *elt*-alone strains are the least common in clinical isolates. This profile pattern was also evident from studies in this region (Dutta *et al.,* 2013) and further studies in our lab conducted on ETEC isolated from patients during 2015-2019 (Mondal, I. *et al.,* 2021). Researchers in Bangladesh have noticed this pattern as well, indicating a departure from their previous reporting (Begum *et al.,* 2014). Our findings, however, contradict the global preponderance of toxin genes, in which the presence of LT is substantially more than that of ST, either alone or in combination with LT (ST+LT), despite the fact that both toxins are prevalent in this geographical location (Isidean *et al.,* 2011). According to Asia and Africa's Global Enteric Multicenter Study (GEMS), 68% of ETEC strains were either ST alone or LT/ST (Kotloff *et al.,* 2013; Vidal *et al.,* 2019). In comparison studies in Bangladesh (1996-1998) (Qadri *et al.,* 2000) and Shanghai (2012-2013) (Tang *et al.,* 2015) had found ST as the most prevalent toxin type and ST+LT as the least dominant one. We discovered that the pattern is reversed in our community-based comparator strains, with ST+LT being the least prevalent toxin and LT being the most widespread. This enterotoxin distribution profile is comparable to those seen in an Egyptian cohort study (Mansour *et al.,* 2014). While the trend in Northern Colombia (Guerra *et al.,* 2014) and Guatemala (Torres *et al.,* 2015) was LT > ST > ST+LT. It should be noted that comparing the comparator control strains to the archival strains is one of the study's limitations, and a study with a greater number of paired strains should be evaluated to give a more accurate image.

In 56% of the clinical strains examined in this investigation, at least one colonization factor was found. In Bangladesh, 56% of the isolates tested positive for the CFs (Qadri *et al.,* 2000). Only 33% of ETEC isolates from a birth cohort of infants in rural Egypt were linked to any of the CFAs that were examined (Mansour *et al.,* 2014). In 53.3 % of ETEC strains isolated from children with or without diarrhoea, CF genes were found, and 21.1% in controls from Teresina/PI, Brazil (Nunes *et al.,* 2011). In Peruvian children with ETEC, CFs were detected in 64% of diarrhoeal samples and 37% of control samples (Rivera *et al.,* 2010). In Bangladesh, ETEC isolates tested positive for at least one CF pathogen in 49% of cases (Begum *et al.,* 2014). According to a study of Nicaraguan youngsters, At least 50% of ETEC strains were found to be positive for one of the screened CF (Vilchez *et al.,* 2014). In Shenzhen, China, one or more CF strains were found in 54% of ETEC samples (Li *et al.,* 2017). Further studies in our lab with the inclusion of 24 CFs found that 91% of strains had at least one CF associated with it (Mondal, I. *et al.,* 2021).

Among the colonization factors, CS6 and CS21 were the predominant ones. According to a research on the ETEC population in Nepal, the most common CFs were CS21 (62.6%) and CS6 (30.2%), in accord with our findings (Margulieux *et al.,* 2018). In previous studies from our region, CS6 emerged as the major circulating CF (Ghosal *et al.,* 2007; Sabui *et al.,* 2012; Dutta *et al.,* 2013) which was also the major CF in Guatemala, Egypt, Zambia and China (Mansour *et al.,* 2014; Torres *et al.,* 2015; Li *et al.,* 2017; Simuyandi *et al.,* 2019). Reports from Bangladesh showed that there was a shift of predominance of CF from CS21 to CS6 (Begum *et al.,* 2014). In our lab study conducted on ETEC isolates from 2015-2019 had CS6 as mostly present CF, instead of CS21, followed by CFA/I (Mondal, I. *et al.,* 2021). CS6 was also the prevalent CF detected from ETEC cases for traveller's diarrhoea (Rivera *et al.,* 2013). In North Colombia, CS21 was the common CF (Guerra *et al.,* 2014) and in traveller's coming back to Spain (Rivera *et al.,* 2013). CS21 is highly expressed in isolates from India and

Latin America, according to a recent study (Kharat *et al.,* 2017). In ETEC isolated from Chilean newborns, CS21 was likewise shown to be the most common CF (Montero *et al.,* 2017). The colonization factor antigen group IV (CFA/IV) is made up of CS4, CS5, and CS6 and CS6 is known to be expressed in conjunction with or in addition to CS4 or CS5 (Gaastra *et al.,* 1996; Wolf *et al.,* 1997). We found a strong co-occurrence of CS6 in combination with CS5 however we got only one strain harbouring CS6 along with CS6. Co-occurrence of CS6 and CS5 was previously reported from our region (Dutta *et al.,* 2013; Ghosal *et al.,* 2007; Sabui *et al.,* 2012). From our study, there was no significant co-occurrence of CS21 with any other CFs which contradicts the report of the strong association of CS21 with CFA/I (Guerra *et al.,* 2014). The most common classical colonization factors in the control samples were CS5 and CS6. There was an association between CS6 and CS5 in the control samples.

We only looked at the most common non-classical VF genes. In 59% of the clinical strains, at least one of these NCVFs was discovered. EatA was the most common among them in clinical strains, followed by *etpA*. When no CF genes were detected, the *leoA* and *tia* genes were found in the least frequency and in the majority of instances. These findings were in line with those from Chile, Colombia, and Guatemala (Del Canto *et al.,* 2011; Guerra *et al.,* 2014; Torres *et al.,* 2015). Among the community control strains, *etpA* was the prevalent NCVF followed by *eatA*. An association had been observed between *eatA* and *etpA* as both of them co-occurred.

However, *eatA* was found to be associated with colonization factors, suggesting that EatA may play a role in pathogen adherence. EatA and colonization factors were also associated in Chilean ETEC isolates in a similar pattern (Del Canto *et al.,* 2011). There was no link between these parameters in ETEC strains obtained from clinical samples and asymptomatic people. A research in Bolivia found a similar outcome (Gonzales *et al.,* 2013). There was no evidence of a significant link between EtpA and colonization factors. EtpA interacts with highly conserved regions of flagellin, which promotes adhesion and improves intestinal adherence and colonization by intestinal cells (Roy *et al.,* 2009).

Among the clinical strains, 29% isolates lack both the classical CFs and NCVFs. The screened virulence genes were found to be negative in 40% of the community specimens. ETEC isolates that were negative for CFs and non-classical virulence factors had been reported all over the world. Only 56% of the isolates during a 2 years period in Bangladesh were positive for the screened CFs (Qadri *et al.,* 2000), whereas, during the period 2007-2012, 51% of the isolates were negative (Begum *et al.,* 2014). Classical and non-classical adhesins were found to be negative in 16% of ETEC strains in Chile (Del Canto *et al.,* 2011). From strains identified in Thailand, 41% of the strains tested negative for any CFs (Puiprom *et al.,* 2010). Only 25% of the isolates from North Colombia were negative for CF (Guerra *et al.,* 2014). In 33.3% of ETEC isolates from Iranian toddlers no CFs were detected (Nazarian *et al.,* 2014). Similarly in Shenzen, China, 46% of strains lacked any evaluated CFs (Li *et al.,* 2017). In ETEC strains obtained from travellers, 31% of the strains lacked any of the evaluated CFs (Rivera *et al.,* 2013).

A large percentage of ETEC isolates from indigenous youngsters and international visitors to a rural Guatemalan hamlet were negative for CFs (Torres *et al.,* 2015). These findings suggested that there could be more colonization determinants that have yet to be discovered. New adhesion factors like CS23 (Del Canto *et al.,* 2012) and, more recently, CS30 (von Mentzer *et al.,* 2017) have been discovered after studies using ETEC strains that did not test positive for any of the recognised CFs. The finding of structures related to CU-pili of the γ2 and κ-families and π-CU pili are new inside ETEC adhesin after genomic investigation of ETEC strains with no known CFs with the goal of discovering new CU pili loci (Montero *et al.,* 2017). These CU-pili are the latest addition to the ETEC adhesion family. However, the inclusion of these CFs, as well as other classical and non-classical virulence factors, that were not included in this study may limit the study. By including these CFs, the number of CF negative strains may be reduced which is evident from the study in our lab with the inclusion of CFs resulting in only 3% strains negative for any virulence factors (Mondal, I. *et al.,* 2021).

During the years 2008-2011, the distribution of toxin genes was consistent; however, between 2012 and 2014, there was a large increase in the number of *est+elt*-containing isolates. On the contrary, we detected year-to-year changes in the prevalence of CS6 and CS21, with no discernible trend. In the instance of CFA/I, there was also some fluctuation. In the annual ETEC isolates, however, there was no fluctuation in the distribution of *eatA* and *etpA*. As a result, these variants may be tracked for vaccine development.

Similar to a previous report from our region (Dutta *et al.,* 2013), sole ETEC isolates were associated with watery diarrhoea. CS6 had been mainly linked to watery diarrhoea, either alone or in combination with other CFs. EatA, the predominant non-classical virulence gene, was mainly linked to watery type diarrhoea either alone or in combination with other virulence genes.

However, when we looked into the expression of these genes, for both the sole and mixed clinical strains we found that all the genes were not expressed. However, for control isolates we found most of the genes to be expressed. Since CS6 was prevalent with respect to gene expression, we focused our study on CS6. When we looked into the expression level of CS6 we find that not all strains express CS6 at the same level. So, we have strains that express CS6 differentially as well as strains that have CS6 genes but do not express the gene. The GEMS study reported that among the CS6-only isolates only 38.5% were positive for phenotypic expression by dot-blot (Vidal *et al.,* 2019). CS6 is expressed constitutively on the bacterial surface, but its phenotypic expression is dependent on the presence of a functional chaperone component (IV), which protects structural subunits from degradation in the periplasm during transit to the bacterial surface (Wolf *et al.,* 1997). One possible explanation for the lack of expression is that the *in vitro* growth conditions used did not trigger the controlled production of isolates expressing specific CFs, as evidenced by the fact that when we supplemented media with different host factors, the genotypically positive strain's non-expressive nature became expressive. Single nucleotide polymorphisms, minor alterations in structural or chaperone genes, or lower copy number plasmids may all allow PCR amplification but diminish or abolish CF expression in isolates that are PCR positive but expression negative. This notion is supported by the fact that a single point mutation in the untranslated region upstream of the chaperone encoding gene caused CS6 phenotypic expression to be negative in CS6 genotypically positive ETEC strains. The function of this mutation is unknown, however, it could be involved in the chaperone subunit's ribosome binding and subsequent translation (Nicklasson *et al.,* 2008; Tobias *et al.,* 2008). The ETEC virulence regulator *rns* had been discovered to positively regulate the expression of CS14 pili, which is also crucial for the expression of CS1 or CS2 colonization factor antigens as well as for adhesion (Caron*et al.,* 1989; Bodero *et al.,* 2016). CFA/I pilin expression is also controlled by an *rns*-like regulator gene (Caron, *et al.,* 1990). CFA/I, CS1, CS2, CS4, and CS5 expression is influenced by transcriptional regulators such CfaD (also known as CfaR) and Rns, both of which belong to the AraC family of transcriptional regulators (Caron *et al.,* 1990; Hodson *et al.,* 2017). When we sequenced the promoter region we found mutations in the promoter sequence and probably these mutations were regulating the transcriptional modulators which in turn were responsible for the expression negative nature of CS6 strains.

ETEC colonise the small intestine and induce disease in humans and other hosts under a variety of environmental situations. This organism's virulence qualities over commensal *E. coli* are due to its capacity to overcome various environments and cope with various stimuli.

Understanding how virulence factors are regulated at the infection site where the disease develops is crucial to understanding the pathophysiology of ETEC. Most ETEC virulence variables are poorly understood, particularly during human intestinal infection. During the infection process, one of the goals of this thesis was to investigate how different host factors in the human gut, as well as environmental factors, influence the expression of the ETEC enterotoxins ST and LT, common colonization factors CS5 and CS6, and common non-classical virulence factor *eatA*.

ETEC-caused diarrhoea appears to necessitate colonization of the human gut by various colonization factors. The significance of different host factors in the human gastrointestinal tract, as well as environmental factors as possible regulators of these genes, were studied in this study by quantifying expression levels of the major ETEC colonization factors CS5, CS6, and the non-classical colonization factor EatA, *in vitro* and *in vivo*. Temperature, bile, glucose, glutamine, and iron concentrations, as well as proximity to epithelial cells, all affect the expression of CFs (Wolf *et al.,* 1989; Haines *et al.,* 2015; Haines *et al.,* 2015; Hodson *et al.,* 2017). However, despite the fact that certain growth circumstances, like temperature, affect CS6 expression (Wolf *et al.,* 1989), no unique positive regulator for CS6 has been discovered (Favre *et al.,* 2006; Nicklasson *et al.,* 2012). The modulation of the common colonization factors CS5, CS6, and EatA was the topic of this research, with a special emphasis on CS6, which is becoming more widely recognised around the world particularly in this region. The healthy gut microbiota and the gut microenvironment were shaped by host immunity, metabolism, and environmental variables (Sekirov *et al.,* 2010). As a result, the output of virulence factors and the programmes associated with virulence must be precisely coordinated in response to various environmental conditions for the disease to progress properly. We show here that interactions between ETEC and host factors cause considerable changes in the ETEC transcriptome, which could be accompanied by significant changes in the architecture of these organisms at the pathogen-host interface, which is crucial for pathogenesis. These studies showed that the expression of CS6, CS5 and EatA was enhanced in the presence of iron and some bile components which are present in the human intestine. In HT-29 and Rabbit ileal loop assay, it is shown that 10^7

cfu/mL is the minimum inoculum necessary for pathogenesis. In HT-29 tissue culture, it is shown that 18 hours is the time that is needed for optimum expression of virulence factors. Expression studies at the protein level were also in accord with the RNA expression.

When ETEC bacteria enter the small intestine's duodenum, they are introduced to bile secreted by the gall bladder; the amount of bile salts in the bile is reduced when absorbed and redelivered to the liver via the small intestine, and this plays a role in antimicrobial defence (Begley *et al.,* 2005). *In vivo*, bile works as an emulsifier of dietary fats and lipids and has detergenic effects (Hofmann, 1999). Therefore, the optimum expression of colonization factors by the bacteria in the gut environment is important for successful colonization. Crude bile seems to have no significant effect on CS6 but individual components like NaDC cause significant upregulation. The presence of bile salts, which are abundant in the colon and are known to be bactericidal, up-regulated the expression of *estA* and *eltA* enterotoxin genes and down-regulated colonization factors CS1 and CS3 in E24377A ETEC strain, according to an RNA sequencing study (Sahl *et al.,* 2012). Crude bile, sodium deoxycholate and sodium glycocholate hydrate (components of bile) induced CS5 which is the same as previous findings where they were found to be a potent inducer of colonization factor CS5 (Nicklasson *et al.,* 2012; Joffre *et al.,* 2019). However, studies on transcription regulation on CS6 were not addressed. CS6 is bile independent but some components may trigger its expression.

EatA expression is not significantly up-regulated when crude bile is present but individual bile salts, NaGCH, TCA, NaCDC significantly upregulate EatA expression in M9 media. LT secretion is downregulated in presence of crude bile; whereas ST secretion is upregulated which is also shown in earlier reports (Isidean *et al.,* 2011; Sahl *et al.,* 2012; Joffre *et al.,* 2019). Expression of virulence factors can be induced by bile salts in ETEC and CF expressions showed variable modulation in presence of bile, some are activated while others are unaffected or dissuade by bile components (Sjöling *et al.,* 2007). These data suggest that CS6 expression depends on the composition of bile in the gut and the different components can trigger the variable response of CS6 but the expression may also be dependent on the type of ETEC strain and its response in presence of bile components.

Mucins' oligosaccharides and glycoconjugates may impact pathogenic microorganisms' virulence-related capacities and can be used by certain ETEC strains as environmental cues (Pacheco *et al.,* 2015). ETEC expresses several colonization factors that enable them to bind mucins and colonize the small intestine. Intestinal mucin is a crucial host defence mechanism that prevents commensal and pathogenic organisms from interacting with enterocytes, preserving the mucosa's integrity. Furthermore, due to the continual replacement of the mucus layer, neutralised gastric mucins are almost always present in the small intestine. Indeed, we discovered that mucin has little effect on the expression levels of CS5, CS6, and EatA. Contradictory to our results, the previous study on CFA/I showed porcine mucin is a positive inducer (Haines *et al.,* 2015). Toxins were inhibited in the presence of mucin which coincides with the results from previous work (Haines *et al.,* 2015).

Glucose is the preferred carbon source for prokaryotes. Glucose concentration in the upper parts of the small intestine is high but slowly decrease lower down the intestine due to the intensive absorption of glucose by glucose transporters on the epithelial (Brückner *et al.,* 2002). Exposure of ETEC to glucose at 1% had a positive effect on CS6 and associated genes expression but the response was repressed at higher concentrations. Earlier works showed that glucose induces expression of CFA/I (Sahl *et al.,* 2012; Haines *et al.,* 2015) at lower concentrations but inhibited at higher concentrations (Karjalainen *et al.,* 1991; Ares *et al.,* 2019) which is the same as our results. A previous study found that a brush border-associated glycoprotein supports specific CFA/I binding and that its expression develops as a consequence of intestinal cell differentiation and is regulated by glucose (Bernet-Camard *et al.,* 1997). Surprisingly, while glucose has no effect on CS1 or CS3 gene expression in strain E24377A (Sahl *et al.,* 2012), it reduces both CS1 and CS3 surface expression in strain PB-176 (Evans *et al.,* 1991). The presence of glucose in the growth medium upregulated LT and downregulates ST which also established the previous findings (Sahl *et al.,* 2012; Haycocks *et al.,* 2015; Sahl *et al.,* 2015).

Stomach pH also plays an important role in colonization. ETEC must be able to grow in the pH range of 4.5 and pH 9 for successful colonization. Over this wide pH range, ETEC preserves its nucleic acid stability and successful presentation of colonization factors (Gonzales-Siles *et al.,* 2016). Expression of CS6 and CS5 were most at acidic pH 6 which is present at the upper part of the stomach and at alkaline pH 9 present in the duodenum while EatA had most expression at pH 6. This pH may be an indication used by the ETEC to sense the entry points at different sections of the gastrointestinal tract. However, changing circumstances can impact bacteria's growth rate and survivability, and increased secretion at pH 9 could be attributed to membrane leakage, as alkaline pH stresses *E. coli* membranes. The secretion of LT is maximal at pH 8 which confirms previous results (Gonzales *et al.,* 2013) whereas ST secretion is not pH-dependent as shown in an earlier study (Johnson *et al.,* 1978).

Exposure to body temperature is likely one of the most impactful changes, as pathogenic bacteria's virulence factors are frequently induced in response to 37°C (Gonzales-Siles *et al.,* 2016). CF expression is usually thermoregulated, CFs are induced at 37°C but not expressed at ambient temperatures(Gaastra *et al.,* 1996) but not below 22°C, although exceptions have been reported. CS6 and the related genes are maximally expressed at physiological temperature, at 37°C. Both toxins have an optimal expression at 37°C. LT in a native system is optimally expressed at 37°C but it is modulated when the temperature is decreased to 18°C (Trachman *et al.,* 1998). The expression and maturation of curli fimbriae in ETEC were found to be regulated by temperature. Curli was found to be expressed preferentially at low temperatures 30°C than 37°C (Szabo *et al.,* 2005).

A crucial step in the pathogenesis of enterovirulent bacteria is the ability to associate effectively with the numerous cell phenotypes lining the host intestinal barrier. Pathogenic *E. coli* strains may encounter HT29 and other human intestinal epithelial cells that exemplify a cell type that mimics *in vivo* conditions (Kesty *et al.,* 2004; Liévin-Le *et al.,* 2013). As the incubation time with the HT-29 grew, the intensity of expression associated with virulence factors increased. These results show that the expression of the virulence factors depends on the starting inoculum concentration and correlates with the time of incubation of the pathogen with the HT-29 cells.

CS6 expression increases *in vivo* as compared to *in vitro* which is true in all respect as rabbit ileal loop is a more close resemblance to the human gut than *in vitro* laboratory conditions. Fluid accumulation was highest in 10^7 and 10^8 cfu/mL but expression can also be seen at 10^6 cfu/mL concentration. As previous findings, we here also seen expression and fluid accumulation at concentrations of 10^6 to 10^8 colony forming units (cfu) per mL (Sack, R. 2011). In the rabbit ileal loop assay and mice model, we find that the wild type strain is more effective than the mutants. The wild type strain is causing more fluid accumulation and adherence than the mutants. Histological images showed that the major histological damage was done by ΔCS5 variant, ΔCS5ΔEatA variant and wild type strain. The ΔEatA variant could not cause much damage compared to other variants. Although the conditions of a healthy human's normal gut are usually assumed to affect virulence in laboratory experiments, it's vital to remember that the ecology of the small intestine changes during a diarrhoeal episode, which might affect epithelial cells. These conditions can cause a local increase or decrease in ions, osmolarity, and pH, among other things.

From the promoter analysis, it can be said that other than the predicted promoter sequences there are other sequences present upstream of the gene that controls its expression. Detailed analysis of the sequence upstream of the -10 region may suggest the control of CS5, CS6 and CS5 expression through blocking the regions upstream of the gene as transcriptional blockers. Promoter activity under the influence of different host and environmental factors showed optimal promoter activity in accordance with RNA and protein expression concentration. Deletion of the upstream sequences led to the identification of regions other than the predicted ones that are important for promoter activity and thereby the expression of the genes. The regions that were deleted from upstream regions caused a significant reduction of the promoter activity despite the media and supplementation used thereby indicating the activity of that region as important for gene expression.

Iron is an important mineral to almost all species. Iron availability plays a central factor in bacterial pathogenesis initiation, signalling the presence within the small intestine (Muir *et al.,* 1985). We found that the virulence factors CS6, CS5, and EatA had increased expression at 0.2 mM Fe concentration i.e., under iron-rich conditions. But the presence of Fe inhibited toxin secretion. Previous results showed induction of CF under iron starvation conditions (Carpenter *et al.,* 2009) and contrastingly also under iron excess conditions (Haines *et al.,* 2015). In the heterologous ETEC strain E24377A, analyses of virulence gene expression for both *cooA* (CS1) and *cstA* (CS3) show differential regulation of CF and toxins under bile and glucose, as seen in our study (Sahl *et al.,* 2012). CS6 expression decreased with the dysregulated condition. According to the findings, iron functions as a positive regulator of CS6, and CS6 expression is greatly boosted in iron-rich environments. Bacterial infection is reduced in an iron-deficient environment, thus it's reasonable to assume that an iron-rich environment may promote infection by ETEC harbouring CS6. Indeed previous studies that iron can increase the production of virulence factors (Zimmermann *et al.,* 2010; Kortman *et al.,* 2012). In order for invading pathogens to survive in low iron environments, the host reduces iron availability as part of its natural defence. Humans need iron-binding proteins like transferrin and lactoferrin to lower their iron levels. Low iron levels operate as a signal to pathogens to induce virulence, as seen in our experiment, where toxins are produced at low iron concentrations and decrease at high iron concentrations.

Previous research with enterohaemorrhagic *E. coli* Shiga-like toxin I revealed toxin induction in the presence of iron deficiency. To compensate for the lack of iron, pathogens use siderophores. (Andrews *et al.,* 2003). In other CS6 positive clinical strains expressing both LT and ST toxins, we also found an enhanced transcription of CS6 in presence of 0.2mM iron which confirmed the ability of iron to induce the expression of CS6. In contrast, under iron deficiency, however, the synthesis of the CFA/I fimbriae increases and LT secretion is suppressed in the ETEC H10407 prototype strain (Haines *et al.,* 2015). Iron fortification increased the phenotypic expression of CS6 which was also seen in the increased production of flagellin by iron supplementation in extraintestinal pathogenic *Escherichia coli* (Magistro *et al.,* 2017). In trichomonads, iron increased the synthesis and surface placement of adhesins (Alderete *et al.,* 2004).

In contrast to our findings, it has previously been discovered that iron deficiency increases the expression of CS1, CS3, and CFA/I (Haines *et al.,* 2015). Iron chelation reduced the expression of CS6. It was also evident from previous studies in pathogens such as *Staphylococcus aureus, Staphylococcus epidermidis, Escherichia coli, Enterobacter* (Van Asbeck *et al.,* 1983). Against *S. aureus, Plasmodium falciparum* and other organisms, iron chelation therapy had worked successfully (Carver, 2018). In a concentration-dependent way, iron availability increased the adhesion of CS6 harbouring strains to HT-29 tissue culture which suggested the positive role of iron on CS6 of ETEC on attachment to the intestinal epithelial layer. As evidenced by previous investigations, iron increased the attachment of the enteric pathogen *S. typhimurium* but did not impact or even reduced the adhesion of *E. faecalis* and *L. plantarum* (Kortman *et al.,* 2012). Free iron can have a catastrophic effect on natural infection resistance, resulting in the elimination of defensive bactericidal mechanisms in tissue fluids, allowing bacteria to multiply rapidly (Bullen *et al.,* 2005).

In most *in vivo* and *in vitro* situations, the CS5 operon was transcribed at elevated amounts than the CS6 operon, indicating that the microenvironment in diarrhoeal conditions may be more advantageous for the production of the CS5 operon than the CS6 operon. This overexpression, according to theory, could be important for the pathogenesis or dissemination of ETEC infection. More research is needed, however, to determine how ETEC bacteria adapt to their host's environment.

Even though this study employed a standardised commercial bile extract, it was observed that the exact amount of bile produced varied between groups and was dependent on the nutritional state of the host. Similarly, depending on the individual's nutritional state and diet, the relative concentrations of glucose and other factors may vary substantially. The concentrations of nutritional and bile salts in the intestinal lumen and mucosa are also likely to differ. As a result, our ETEC infection model should only be regarded as a speculative representation of the impact of the unique host and environmental variables investigated in this thesis. More research is needed to fully understand the processes of regulation of ST, LT, CS5, CS6 and EatA by the host and environmental conditions outlined in this paper.

When we compared the expression of CS6 of wild type strain harbouring CS5, CS6 and EatA with the mutants' ΔCS5 and/or ΔEatA, CS6 expression in all mutants were down-regulated when compared to wild type conditions. In every experiments analysis of protein expression also coincides with the RNA-fold change data. The WT strain induced greater adhesion and fluid accumulation than the mutants in the rabbit ileal loop model and the suckling mouse colonization assay. In terms of adhesion, fluid accumulation, and histological damage, deletions of CS5 and/or EatA mutants, are less pathogenic. Following the mutants, it can be said that maximal expression of CS6 is dependent on the presence of CS5 and EatA in CS6+CS5+EatA harbouring strains. The deletion of CS5 and/or EatA appears to have rendered the mutant weak or binding-deficient. However, the toxins were independent of the mutants and produced a similar amount of toxins. However, further studies are needed to fully understand the mechanism of CS6 regulation and the role of CS5 and EatA on the regulation of CS6.

Chapter 7

CONCLUSION

This thesis aimed to understand the distribution of most prevalent combination of classical and non-classical virulence factors among clinical isolates of ETEC in our settings and focus on their regulation of expression by different host and environmental factors during pathogenesis. The objective-wise conclusions are

Objective 1 was the detection of common colonization factors and other virulence factors among clinical isolates of ETEC to understand their occurrence. The conclusions are:

- Both the children the adults must be the goal for medications during ETEC infection.
- The value of a prevalence analysis of ETEC according to geographical regions for vaccine development is heightened by the variation in toxin dominance in different regions.
- Based on these findings, it can be stated that ETEC strains in this region have the same pattern of colonization factors and that their connection differs from those of other regions.
- Although the strains had a variety of CFs and non-classical virulence genes, CS6-containing ETEC strains remained the most prevalent. Between CS6+CS5 and *eatA*, as well as the *est* gene alone or in combination with *elt*, a distinct pattern may be seen. Our research is limited by the use of common CFs and non-classical virulence genes.
- These results indicated that there may be more colonization determinants still to be discovered, and the strains that are CF negative in this zone should be examined further to identify novel and relevant CFs that could contribute to the development of an ETEC vaccine.
- CS21, on the other hand, has become a prominent CF. The CS21-containing strains, on the other hand, had a wide range of CFs and NCVF genes associated with it.
- To our understanding, no non-classical virulence gene investigation has ever been conducted in this zone. Our findings suggested that non-classical virulence variables should be included in the ETEC vaccine candidate.
- Based on our study and investigations in other geographical realms, it appears that the development of an ETEC vaccine must be geo-region specific. Antigenically, ETEC isolates exhibit a wide range of toxins, CFs, and non-classical virulence factors. For successful vaccine development, a mixture of these elements must be considered.

- Recognizing the dominating profile and assessing the co-existence of CFs and NCVFs variables may aid in tracking ETEC's global spread and prioritising possible vaccine candidates.

Objective 2 was to delineate the expression of identified colonization factors and virulence factors in clinical ETEC. The conclusions are:

- Not all genetically positive CF and NCVF genes are expressed.
- One of the significant findings from this thesis is that a positive CS6 result obtained by genotypic approaches may not be a reliable indicator of a strain's CS6 expressing capacity. This information must be taken into account while creating an ETEC vaccination because CS6 has become one of the most common ETEC CFs in recent years. For detecting CS6 and other CFs, such investigations frequently use genotypic assays. One benefit of genotypic approaches is that phenotypic features may be lost when strains are stored. However, as demonstrated in this work, genotypic approaches can be deceiving. Vaccines for protection against ETEC strains expressing CS6 are very important to consider. As a result, we propose that both genotypic and phenotypic assessments be included in ETEC CF epidemiological investigations for vaccine development.
- This work integrates a large amount of experimental and epidemiological data into a genetic and phenotypic perspective, and it serves as a prototype for future diagnostic and intervention efforts.
- This thesis emphasises the importance of studying the regulation of ETEC virulence factor phenotypic surface expression and secretion, as well as the variability of ETEC strains and their virulence, in order to gain a better understanding of the pathogenic mechanisms of this heterogeneous enteropathogen as a foundation for various intervention strategies.

Objective 3 was to understand the modulation of prevalent colonization factor and other virulence factors by the host and environmental factors during pathogenesis. The conclusions are:

- Recent studies have indicated that the pathogenicity of ETEC is multifactorial and depends on the presence of additional virulence factors other than ST/ LT enterotoxins and the expression of CFs for successful infection in the small intestine.

- Pathogenesis is the result of a synergistic effect of multiple virulence determinants at different stages of pathogenesis. The ability to control gene expression during the course of an infection is critical for a pathogen's survival in the host. As a result, virulence gene expression responds in a variety of ways to host variables and environmental conditions.
- Results are expected to provide insight into a better understanding of the differential expression of the virulence genes in ETEC and how they relate to pathogenesis. This will advance our knowledge of the virulence mechanisms of ETEC in general. The research outcome would also provide understanding regarding the interactions in the regulatory network of ETEC and virulence gene expression in this important human pathogen during infection.
- Since CS6 is a prevalent CF, an important virulence factor and is an essential vaccine candidate, we have used this as a model system to study the regulation and expression during ETEC pathogenesis.
- The results would further help in elucidating the transcriptional targets and putative mechanisms at the molecular level. This will have direct relevance in the development of therapeutic strategies towards the prevention and treatment of diarrhoeal illnesses by targeting the regulatory networks.
- If bacteria can trigger linked genetic and metabolic pathways, they may have a selective advantage that is required for survival when they encounter signals indicative of the intestinal environment, such as the presence of different compounds in growth media in our experiment, or they may be detrimental to their survival, causing the host to become more protective.
- Based on our findings, ETEC triggered in response to conditions comparable to those seen in the intestinal environment plays a critical role in the initial attachment to the intestinal epithelium, as well as the growth of microcolonies and the release of toxins.
- ETEC has to adapt to the different conditions of the gut that it encounter and has to optimize its virulence mechanisms in a multifactorial process. It is also important to remember that virulence regulation conditions examined in the lab may differ during human infection as we cannot totally mimic the gut microbiome and the physiological conditions.
- We discovered that iron, glucose, and other host and environmental factors affect ETEC pathogenicity in relation to ST, LT, CS5, CS6, and EatA, showing that nutrition and dietary may have an impact on ETEC infection outcomes.

- Our findings reveal that free iron availability increases the ETEC colonization factor CS6, supporting the theory that iron might boost enteric pathogen proliferation and virulence. According to our findings, ETEC can express CS6 in response to iron concentrations.
- The host and environmental factors, besides creating a favourable condition for the growth and survival of the organisms, are also the key regulators of the expression and activity of their genes.
- The minimum $ETEC^{WT}$ inoculum required for infection is 10^7 cfu/mL. 8-hour post-infection, bacteria start showing histological damages in the intestine.
- 0.2 mM iron, bile components like Na-chenodeoxycholate, taurocholic acid causes significant upregulation of CS6 expression.
- However, in presence of 0.2 mM iron, CS6 expression depends on the other two factors. Therefore, CS6 without CS5 and/or EatA showed a dysregulated expression.
- The mutants become less pathogenic in terms of adherence, fluid accumulation and histological damage when compared to $ETEC^{WT}$.
- The minimum region of CS6 promoter lies between -254 and +1. Important region(s) lie between -350 bp to -255 bp upstream in the promoter which might have important elements needed to control CS6 gene expression.
- CS6 expression is dependent on CS5 and EatA. CS6 expression is dependent on the presence of other virulence factors that may assist or act in synergy for the enhanced expression of CS6.
- ETEC has to adapt to the different conditions of the gut that it encounters and has to optimize its virulence mechanisms in a multifactorial process. It is also important to remember that virulence regulation conditions examined in the lab may differ during human infection as we cannot totally mimic the gut microbiome and the physiological conditions.

Significance and Impact:

We advocate that a multivalent vaccine candidate containing LT and ST, CS6, CS5, and EatA positive ETEC strains from the most common clonal complex, which was found in strains isolated from both children and adults in this region over several years of research, be considered for vaccine development research. Molecular details of the regulatory pathway will help in better understanding of the ETEC pathogenesis and this knowledge could be translated for effective and safe drug for disrupting regulators of ETEC pathogenesis.

Future Prospects:

Our research will help in filling up important knowledge gaps in our understanding of ETEC pathogenesis in the intestine at different stages of infection. This knowledge will also help in evaluating the efficacy of live attenuated vaccines. In the future, more research works will be conducted to translate this basic knowledge into the rational design of safe and effective drugs against the regulators to disrupt the pathogenesis. Additionally, this concept may be applicable to a broad range of colonization factors in ETEC as well as other enteric pathogens which require future works on other colonization factors related to this geographical region. With further basic studies into the regulation of CS6 and its biochemical properties will help to elucidate the real organization of CS6 operon and reveal how it functions inside the human gut. These knowledge will generate important aspects of CS6 which will establish it as a potent vaccine candidate.

Chapter 8

REFERENCES

Abe, H., Tatsuno, I., Tobe, T., Okutani, A., & Sasakawa, C. (2002). Bicarbonate ion stimulates the expression of locus of enterocyte effacement-encoded genes in enterohemorrhagic *Escherichia coli* O157: H7. *Infection and Immunity, 70*(**7**), 3500-3509.

Abuaita, B. H., & Withey, J. H. (2009). Bicarbonate induces *Vibrio cholerae* virulence geneexpression by enhancing ToxT activity. *Infection and Immunity, 77*(**9**), 4111-4120.

Al-Gallas, N., Abbassi, S. M., Hassan, A. B., & Aissa, R. B. (2007). Genotypic and phenotypic profiles of enterotoxigenic *Escherichia coli* associated with acute diarrhoea in Tunis, Tunisia. *Current Microbiology*, ***55*(1),** 47-55.

Al-Gallas, N., Bahri, O., Bouratbeen, A., Haasen, A. B., & Aissa, R. B. (2007). Etiology ofacute diarrhoea in children and adults in Tunis, Tunisia, with emphasis on diarrhoeagenic *Escherichia coli*: prevalence, phenotyping, and molecular epidemiology. *American Journal of Tropical Medicine and Hygiene*, ***77*(3)**, 571-582.

Alderete, J., Nguyen, J., Mundodi, V., & Lehker, M. (2004). Heme-iron increases levels of AP65-mediated adherence by *Trichomonas vaginalis*. *Microbial Pathogenesis, 36*(**5**), 263-271.

Anderson IV, J. D., Bagamian, K. H., Muhib, F., Amaya, M. P., Laytner, L. A., Wierzba,T., & Rheingans, R. (2019). Burden of enterotoxigenic *Escherichia coli* and shigella non-fatal diarrhoeal infections in 79 low-income and lower middle-income countries: a modelling analysis. *The Lancet Global Health*, ***7*(3)**, e321-e330.

Andrews, S. C., Robinson, A. K., & Rodríguez-Quiñones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiology Reviews, 27*(**2-3**), 215-237.

Ares, M. A., Abundes-Gallegos, J., Panunzi, L. G., Torres, J., Alcántar-Curiel, M. D., Giron, J. A., & De la Cruz, M. A. (2019). The coli surface antigen CS3 of enterotoxigenic *Escherichia coli* is differentially regulated by H-NS, CRP, and CpxRA global regulators. *Frontiers in Microbiology, **10***, 1685.

Bäck, E., Svennerholm, A., Holmgren, J., & Möllby, R. (1979). Evaluation of a ganglioside immunosorbent assay for detection of *Escherichia coli* heat-labile enterotoxin. *Journal of Cinical Microbiology*, ***10*(6),** 791-795.

Barbas, C. F., Burton, D. R., Scott, J. K., & Silverman, G. J. (2007). Quantitation of DNAand RNA. *Cold Spring Harbor Protocols*, ***2007*(11),** pdb-ip47.

Baselski, V., Briggs, R., Parker, C. J. I., & immunity. (1977). Intestinal fluid accumulationinduced by oral challenge with *Vibrio cholerae* or cholera toxin in infant mice. *Infection and Immunity, 15*(**3**), 704-712.

Begley, M., Gahan, C. G., & Hill, C. (2005). The interaction between bacteria and bile. *FEMS Microbiology Reviews, 29*(**4**), 625-651.

Begum, Y. A., Baby, N. I., Faruque, A. S., Jahan, N., Cravioto, A., Svennerholm, A.-

M., & Qadri, F. (2014). Shift in phenotypic characteristics of enterotoxigenic *Escherichia coli* (ETEC) isolated from diarrhoeal patients in Bangladesh. *PLoS Neglected Tropical Diseases, 8*(**7**), e3031.

Begum, Y.A., Talukder, K.A., Azmi, I.J., Shahnaij, M., Sheikh, A., Sharmin, S., Svennerholm, A.M. and Qadri, F. (2016). Resistance pattern and molecular characterization of enterotoxigenic *Escherichia coli* (ETEC) strains isolated in Bangladesh. *PloS One*, ***11***(**7**), e0157415.

Bernet-Camard, M.-F., Duigou, F., Kerneis, S., Coconnier, M.-H., & Servin, A. L. (1997).Glucose up-regulates expression of the differentiation-associated brush border bindingsite for enterotoxigenic *Escherichia coli* colonization factor antigen I in cultured humanenterocyte-like cells. *Infection and Immunity, **65***(**4**), 1299-1306.

Bhandari, N., Bahl, R., Taneja, S., Strand, T., Mølbak, K., Ulvik, R.J., Sommerfelt, H. and Bhan, M.K. (2002). Substantial reduction in severe diarrhoeal morbidity by daily zinc supplementation in young north Indian children. *Pediatrics, **109***(**6**), e86-e86.

Black, R. E., Merson, M. H., Rowe, B., Taylor, P. R., Alim, A. A., Gross, R. J., & Sack, D.A. (1981). Enterotoxigenic *Escherichia coli* diarrhoea: acquired immunity and transmission in an endemic area. *Bulletin of the World Health Organization*, ***59***(**2**), 263.

Black, R. E., Merson, M. H., Eusof, A., Huq, I., Pollard, R. (1984). Nutritional status, body size and severity of diarrhoea associated with rotavirus or enterotoxigenic *Escherichia coli*. *The Journal of Tropical Medicine and Hygiene, **87***(**2**), 83-89.

Black, R. E. (2003). Zinc deficiency, infectious disease and mortality in the developing world. *The Journal of Nutrition*, ***133***(**5**), 1485S-1489S.

Bodero, M. D. R., & Munson, G. P. (2016). The virulence regulator Rns activates the expression of CS14 pili. *Genes, **7***(**12**), 120.

Bray, J. (1945). Isolation of Antigenically Homogeneous Strains of Bact. coli neapolitanum from Summer Diarrhoea of Infants. *The Journal of Pathology and Bacteriology*, ***57***(**2**)**,**239-247.

Brown, E. A., & Hardwidge, P. R. (2007). Biochemical characterization of the enterotoxigenic *Escherichia coli* LeoA protein. *Microbiology, **153***(**11**), 3776-3784.

Brückner, R., & Titgemeyer, F. (2002). Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiology Letters, **209***(**2**), 141-148.

Bullen, J. J., Rogers, H. J., Spalding, P. B., & Ward, C. G. (2005). Iron and infection: the heart of the matter. *FEMS Immunology and Medical Microbiology, **43***(**3**), 325-330.

Buuck, S., Smith, K., Fowler, R. C., Cebelinski, E., Lappi, V., Boxrud, D., & Medus, C. (2020). Epidemiology of Enterotoxigenic *Escherichia coli* infection in Minnesota, 2016–2017. *Epidemiology and Infection*, ***148***.

Byrd, W., De Lorimier, A., Zheng, Z. R., & Cassels, F. J. (2005). Microencapsulated subunitvaccine approach to enterotoxigenic *Escherichia coli* and other mucosal pathogens. *Advanced Drug Delivery Reviews*, ***57*(9)**, 1362-1380.

Caron, J., Coffield, L. M., & Scott, J. R. (1989). A plasmid-encoded regulatory gene, rns, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. *Proceedings of the National Academy of Sciences*, ***86*(3)**, 963-967.

Caron, J., & Scott, J. R. (1990). A rns-like regulatory gene for colonization factor antigen I (CFA/I) that controls expression of CFA/I pilin. *Infection and Immunity*, ***58*(4)**, 874- 878.

Caron, J., Maneval, D. R., Kaper, J. B., & Scott, J. R. (1990). Association of rns homologs with colonization factor antigens in clinical *Escherichia coli* isolates. *Infection and Immunity*, ***58*(10)**, 3442-3444.

Carpenter, B. M., Whitmire, J. M., & Merrell, D. S. (2009). This is not your mother's repressor: the complex role of fur in pathogenesis. *Infection and Immunity,* ***77*(7)**, 2590-2601.

Carver, P. L. (2018). The battle for iron between humans and microbes. *Current Medicinal Chemistry,* ***25*(1)**, 85-96.

Chaudhuri, R. R., & Henderson, I. R. (2012). The evolution of the *Escherichia coli* phylogeny. *Infection, Genetics and Evolution*, ***12*(2)**, 214-226.

Clark, C. A., Heuzenroeder, M. W., & Manning, P. A. (1992). Colonization factor antigen CFA/IV (PCF8775) of human enterotoxigenic *Escherichia coli*: nucleotide sequence of the CS5 determinant. *Infection and Immunity*, ***60*(3)**, 1254-1257.

Cohen, D., Tobias, J., Spungin-Bialik, A., Sela, T., Kayouf, R., Volovik, Y., Yavzori, M. & Ephros, M. (2010). Phenotypic characteristics of enterotoxigenic *Escherichia coli* associated with acute diarrhoea among Israeli young adults. *Foodborne Pathogens and Disease*, ***7*(10)**, 1159-1164.

Croxen, M. A., Law, R. J., Scholz, R., Keeney, K. M., Wlodarska, M., & Finlay, B. B. (2013). Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*, ***26*(4)**, 822-880.

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, ***97*(12)**, 6640-6645.

De, S. N., Bhattacharya, K., & Sarkar, J. K. (1956). A study of the pathogenicity of strainsof Bacterium coli from acute and chronic enteritis. *Journal of Pathology and Bacteriology*, ***71*(1)**, 201-9.

Dean, A. G., Ching, Y. C., Williams, R. G., & Harden, L. B. (1972). Test for

Escherichia coli enterotoxin using infant mice: application in a study of diarrhoea in children in Honolulu. *Journal of Infectious Diseases*, ***125*****(4)**, 407-411.

Debnath, A., Wajima, T., Sabui, S., Hamabata, T., Ramamurthy, T., & Chatterjee, N. S.(2015). Two specific amino acid variations in colonization factor CS6 subtypes of enterotoxigenic *Escherichia coli* results in differential binding and pathogenicity. *Microbiology*, ***161*****(4)**, 865-874.

Debnath, A., Sabui, S., Wajima, T., Hamabata, T., Banerjee, R., & Chatterjee, N. S. (2016). Functional Role of N-and C-Terminal Amino Acids in the Structural Subunits of Colonization Factor CS6 Expressed by Enterotoxigenic *Escherichia coli*. *Journal of Bacteriology*, ***198*****(9)**, 1429-1441.

Deborah Chen, H., & Frankel, G. (2005). Enteropathogenic *Escherichia coli*: unravelling pathogenesis. *FEMS Microbiology Reviews*, ***29*****(1)**, 83-98.

Del Canto, F., Valenzuela, P., Cantero, L., Bronstein, J., Blanco, J. E., Blanco, J., Sommerfelt, H. (2011). Distribution of classical and nonclassical virulence genes in enterotoxigenic *Escherichia coli* isolates from Chilean children and tRNA gene screening for putative insertion sites for genomic islands. *Journal of Clinical Microbiology,* ***49*****(9)**, 3198-3203.

Del Canto, F., Botkin, D. J., Valenzuela, P., Popov, V., Ruiz-Perez, F., Nataro, J. P., Levine, M.M., Stine, O.C., Pop, M., Torres, A.G. & Vidal, R. (2012). Identification of coli surface antigen 23, a novel adhesin of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***80*****(8)**, 2791-2801.

Donta, S. T., Moon, H. W., & Whipp, S. C. (1974). Detection of heat-labile *Escherichia coli*enterotoxin with the use of adrenal cells in tissue culture. *Science*, ***183*****(4122)**, 334-336.

Dreyfus, L. A., Harville, B., Howard, D. E., Shaban, R., Beatty, D. M., & Morris, S. J. (1993). Calcium influx mediated by the *Escherichia coli* heat-stable enterotoxin B (STB). *Proceedings of the National Academy of Sciences*, ***90*****(8)**, 3202-3206.

Duthy, T. G., Staendner, L. H., Manning, P. A., & Heuzenroeder, M. W. (1999). CS5 pilusbiosynthesis genes from enterotoxigenic *Escherichia coli* O115: H40. *Journal of Bacteriology,* ***181*****(18)**, 5847-5851.

Duthy, T. G., Manning, P. A., & Heuzenroeder, M. W. (2001). Characterization of the CsfCand CsfD proteins involved in the biogenesis of CS5 pili from enterotoxigenic *Escherichia coli*. *Microbial Pathogenesis*, ***31*****(3)**, 115-129.

Duthy, T. G., Manning, P. A., & Heuzenroeder, M. W. (2002). Identification and characterization of assembly proteins of CS5 pili from enterotoxigenic *Escherichia coli*. *Journal of Bacteriology,* ***184*****(4)**, 1065-1077.

Dutta, S., Guin, S., Ghosh, S., Pazhani, G. P., Rajendran, K., Bhattacharya, M. K., Takeda, Y., Nair, G.B. & Ramamurthy, T. (2013). Trends in the prevalence of diarrhoeagenic *Escherichia coli* among hospitalized diarrhoeal patients in Kolkata, India.*PLoS One, 8*(**2**), e56068.

Dutta, S., Pazhani, G. P., Nataro, J. P., & Ramamurthy, T. (2015). Heterogenic virulence in a diarrhoeagenic *Escherichia coli*: evidence for an EPEC expressing heat-labile toxinof ETEC. *International Journal of Medical Microbiology*, ***305***(**1**), 47-54.

Edwards, R. A., & Schifferli, D. M. (1997). Differential regulation of *fasA* and *fasH* expression of *Escherichia coli* 987P fimbriae by environmental cues. *Molecular Microbiology, **25***(**4**), 797-809.

Elder, B. L., Boraker, D. K., & Fives-Taylor, P. M. (1982). Whole-bacterial cell enzyme- linked immunosorbent assay for *Streptococcus sanguis* fimbrial antigens. *Journal of Clinical Microbiology, **16***(**1**), 141-144.

Elsinghorst, E. A., & Kopecko, D. J. (1992). Molecular cloning of epithelial cell invasion determinants from enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***60***(**6**), 2409-2417.

Elsinghorst, E. A., & Weitz, J. A. (1994). Epithelial cell invasion and adherence directed bythe enterotoxigenic *Escherichia coli* tib locus is associated with a 104-kilodalton outermembrane protein. *Infection and Immunity*, ***62***(**8**), 3463-3471.

Ericsson, C. D. (2003). Travellers' diarrhoea. *International Journal of Antimicrobial Agents*, ***21***(**2**), 116-124.

Escherich, T. (1988). The intestinal bacteria of the neonate and breast-fed infant. *Clinical Infectious Diseases, **10***(**6**), 1220-1225.

Evans, D. G., Evans Jr, D. J., Tjoa, W. S., & DuPont, H. L. (1978). Detection and characterization of colonization factor of enterotoxigenic *Escherichia coli* isolated from adults with diarrhoea. *Infection and Immunity*, ***19***(**2**), 727-736.

Evans, D. G., Evans Jr, D. J., Clegg, S. T., & Pauley, J. A. (1979). Purification and characterization of the CFA/I antigen of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***25***(**2**), 738-748.

Evans, D. G., Evans, D. J., Karjalainen, T. K., & Lee, C.-H. (1991). Production of colonization factor antigen II of enterotoxigenic *Escherichia coli* is subject to cataboliterepression. *Current Microbiology, **23***(**2**), 71-74.

Faruque, A. S., Malek, M. A., Khan, A. I., Huq, S., Salam, M. A., & Sack, D. A. (2004). Diarrhoea in elderly people: aetiology, and clinical characteristics. *Scandinavian Journal of Infectious Diseases*, ***36***(**3**), 204-208.

Favre, D., Lüdi, S., Stoffel, M., Frey, J., Horn, M. P., Dietrich, G., Spreng, S., & Viret, J.F. (2006). Expression of enterotoxigenic *Escherichia coli* colonization factors in *Vibrio cholerae*. *Vaccine*, ***24***(**20**), 4354-4368.

Fleckenstein, J. M., Kopecko, D. J., Warren, R. L., & Elsinghorst, E. A. (1996).

Molecularcharacterization of the tia invasion locus from enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***64*****(6)**, 2256-2265.

Fleckenstein, J. M., Lindler, L. E., Elsinghorst, E. A., & Dale, J. B. (2000). Identification of a gene within a pathogenicity island of enterotoxigenic *Escherichia coli* H10407 required for maximal secretion of the heat-labile enterotoxin. *Infection and Immunity*, **68(5)**, 2766-2774.

Fleckenstein, J. M., Roy, K., Fischer, J. F., & Burkitt, M. (2006). Identification of a two- partner secretion locus of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***74*****(4)**, 2245-2258.

Fleckenstein, J. M., Hardwidge, P. R., Munson, G. P., Rasko, D. A., Sommerfelt, H., & Steinsland, H. (2010). Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. *Microbes and Infection, **12***(**2**), 89-98.

Fleckenstein, J., Sheikh, A., & Qadri, F. (2014). Novel antigens for enterotoxigenic *Escherichia coli* vaccines. *Expert Review of Vaccines*, ***13*****(5)**, 631-639.

Frech, S.A., DuPont, H.L., Bourgeois, A.L., McKenzie, R., Belkind-Gerson, J., Figueroa, J.F., Okhuysen, P.C., Guerrero, N.H., Martinez-Sandoval, F.G., Meléndez- Romero, J.H. and Jiang, Z.D. (2008). Use of a patch containing heat-labile toxin from *Escherichia coli* against travellers' diarrhoea: a phase II, randomised, double-blind, placebo-controlled field trial. *The Lancet*, ***371*****(9629)**, 2019-2025.

Froehlich, B., Parkhill, J., Sanders, M., Quail, M. A., & Scott, J. R. (2005). The pCoo plasmid of enterotoxigenic *Escherichia coli* is a mosaic cointegrate. *Journal of Bacteriology*, ***187*****(18)**, 6509-6516.

Gaastra, W., & Svennerholm, A.-M. (1996). Colonization factors of human Enterotoxigenic *Escherichia coli* (ETEC). *Trends in Microbiology, **4***(**11**), 444-452.

Gaastra, W., & van der Zeijst, B. A. (1996). Fimbriae of human enterotoxigenic *Escherichia coli* and their possible use as vaccine components. *Reviews in Medical Microbiology*, ***7*****(3)**, 165.

Gangarosa, E. J. (1978). Epidemiology of *Escherichia coli* in the United States. *The Journal of Infectious Diseases*, ***137*****(5)**, 634-638.

Ghosal, A., Bhowmick, R., Nandy, R. K., Ramamurthy, T., & Chatterjee, N. S. (2007). PCR-based identification of common colonization factor antigens of enterotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology, **45***(**9**), 3068-3071.

Ghosal, A., Bhowmick, R., Banerjee, R., Ganguly, S., Yamasaki, S., Ramamurthy, T., & Chatterjee, N. S. (2009). Characterization and studies of the cellular interaction of native colonization factor CS6 purified from a clinical isolate of Enterotoxigenic *Escherichia coli. Infection and Immunity, **77***(5), 2125-2135.

Gonzales, L., Ali, Z. B., Nygren, E., Wang, Z., Karlsson, S., Zhu, B., Quiding-Järbrink, M. & Sjöling, Å. (2013). Alkaline pH is a signal for optimal production and secretion of the heat labile toxin, LT in enterotoxigenic *Escherichia coli* (ETEC). *PLoS One*, ***8*****(9)**, e74069.

Gonzales, L., Sanchez, S., Zambrana, S., Iñiguez, V., Wiklund, G., Svennerholm, A. M., & Sjöling, Å. (2013). Molecular characterization of enterotoxigenic *Escherichia coli* isolates recovered from children with diarrhoea during a 4-year period (2007 to 2010) inBolivia. *Journal of Clinical Microbiology*, ***51*****(4)**, 1219-1225.

Gonzales, L., & Sjöling, Å. (2016). The different ecological niches of enterotoxigenic *Escherichia coli*. *Environmental Microbiology,* ***18*****(3)**, 741-751.

Güereña-Burgueño, F., Hall, E. R., Taylor, D. N., Cassels, F. J., Scott, D. A., Wolf, M. K.,Roberts, Z.J., Nesterova, G.V., Alving, C.R. & Glenn, G. M. (2002). Safety and immunogenicity of a prototype enterotoxigenic *Escherichia coli* vaccine administered transcutaneously. *Infection and Immunity*, ***70*****(4)**, 1874-1880.

Guerra, J. A., Romero-Herazo, Y. C., Arzuza, O., & Gómez-Duarte, O. G. (2014). Phenotypic and genotypic characterization of enterotoxigenic *Escherichia coli* clinical isolates from northern Colombia, South America. *BioMed Research International*, ***2014***.

Guerrant, R. L., Brunton, L. L., Schnaitman, T. C., Rebhun, L. I., & Gilman, A. G. (1974).Cyclic adenosine monophosphate and alteration of Chinese hamster ovary cell morphology: a rapid, sensitive *in vitro* assay for the enterotoxins of *Vibrio cholerae* and *Escherichia coli*. *Infection and Immunity*, ***10*****(2)**, 320-327.

Gupta, S., & Chowdhury, R. (1997). Bile affects production of virulence factors and motilityof *Vibrio cholerae*. *Infection and immunity*, ***65*****(3)**, 1131-1134.

Guth, B. E. C. (2000). Enterotoxigenic *Escherichia coli*-an overview. *Memórias do Instituto Oswaldo Cruz*, ***95***, 95-97.

Haines, S., Arnaud-Barbe, N., Poncet, D., Reverchon, S., Wawrzyniak, J., Nasser, W., &Renauld-Mongénie, G. (2015). IscR regulates synthesis of colonization factor antigenI fimbriae in response to iron starvation in enterotoxigenic *Escherichia coli*. *Journal ofBacteriology,* ***197*****(18)**, 2896-2907.

Haines, S., Gautheron, S., Nasser, W., & Renauld-Mongénie, G. (2015). Identification of novel components influencing colonization factor antigen I expression in enterotoxigenic *Escherichia coli*. *PLoS One,* ***10*****(10)**.

Hardy, S. J., Holmgren, J., Johansson, S., Sanchez, J., & Hirst, T. R. (1988). Coordinated assembly of multisubunit proteins: oligomerization of bacterial enterotoxins *in vivo* and *in vitro*. *Proceedings of the National Academy of Sciences*, ***85*****(19)**, 7109-7113.

Harro, C., Sack, D., Bourgeois, A. L., Walker, R., DeNearing, B., Feller, A.,

Chakraborty,S., Buchwaldt, C. & Darsley, M. J. (2011). A combination vaccine consisting of threelive attenuated enterotoxigenic *Escherichia coli* strains expressing a range of colonization factors and heat-labile toxin subunit B is well tolerated and immunogenic in a placebo-controlled double-blind phase I trial in healthy adults. *Clinical and Vaccine Immunology*, ***18*****(12)**, 2118-2127.

Haycocks, J. R., Sharma, P., Stringer, A. M., Wade, J. T., & Grainger, D. C. (2015). The molecular basis for control of ETEC enterotoxin expression in response to environmentand host. *PLoS Pathogens,* ***11*****(1)**, e1004605.

Helander, A., Hansson, G. C., & Svennerholm, A. M. (1997). Binding of Enterotoxigenic *Escherichia coli* to isolated enterocytes and intestinal mucus. *Microbial Pathogenesis*, ***23*****(6)**, 335-346.

Heuzenroeder, M. W., Neal, B. L., Thomas, C. J., Halter, R., & Manning, P. A. (1989). Characterization and molecular cloning of the PCF8775 CS5 antigen from an enterotoxigenic *Escherichia coli* 0115: H40 isolated in Central Australia. *Molecular Microbiology*, ***3*****(3)**, 303-310.

Hodson, C., Yang, J., Hocking, D. M., Azzopardi, K., Chen, Q., Holien, J. K., Parker, M.W., Tauschek, M. & Robins-Browne, R. M. (2017). Control of virulence gene expression by the master regulator, CfaD, in the prototypical enterotoxigenic *Escherichia coli* strain, H10407. *Frontiers in Microbiology*, ***8***, 1525.

Hofmann, A. F. (1999). Bile acids: the good, the bad, and the ugly. *Physiology,* ***14*****(1)**, 24-29.

Holmgren, J., & Svennerholm, A. M. (1985). Vaccine development for the control of cholera and related toxin-induced diarrhoeal diseases. *Ciba Foundation Symposium* (**Vol. 112**, pp. 242-270).

Hosangadi, D., Smith, P. G., Kaslow, D. C., & Giersing, B. K. (2018). WHO consultation on ETEC and Shigella burden of disease, Geneva, 6–7th April 2017: Meeting report. *Vaccine*.

Hung, D. T., & Mekalanos, J. J. (2005). Bile acids induce cholera toxin expression in *Vibrio cholerae* in a ToxT-independent manner. *Proceedings of the National Academy of Sciences,* ***102*****(8)**, 3028-3033.

Isidean, S., Riddle, M., Savarino, S., & Porter, C. (2011). A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine,* ***29*****(37)**, 6167-6178.

Jiang, Z. D., Lowe, B., Verenkar, M. P., Ashley, D., Steffen, R., Tornieporth, N., von Sonnenburg, F., Waiyaki, P. & DuPont, H. L. (2002). Prevalence of enteric pathogens among international travelers with diarrhoea acquired in Kenya (Mombasa), India (Goa), or Jamaica (Montego Bay). *The Journal of Infectious Diseases*, ***185*****(4)**, 497-502.

Jiang, Z. D., & DuPont, H. L. (2017). Etiology of travellers' diarrhoea. *Journal of Travel Medicine*, ***24*(suppl_1)**, S13-S16.

Jobling, M. G., Poole, S. T., Rasulova-Lewis, F., O'Dowd, A., McVeigh, A. L., Balakrishnan, A., Sincock, S.A., Prouty, M.G., Holmes, R.K. & Savarino, S. J. (2020). Biochemical and immunological characterization of an ETEC CFA/I adhesin cholera toxin B subunit chimera. *PloS One*, ***15*(3)**, e0230138.

Joffré, E., & Sjöling, Å. (2016). The LT1 and LT2 variants of the enterotoxigenic *Escherichia coli* (ETEC) heat-labile toxin (LT) are associated with major ETEC lineages. *Gut microbes*, ***7*(1)**, 75-81.

Joffre, E., Nicklasson, M., Álvarez-Carretero, S., Xiao, X., Sun, L., Nookaew, I., Zhu, B. & Sjöling, Å. (2019). The bile salt glycocholate induces global changes in gene and protein expression and activates virulence in enterotoxigenic *Escherichia coli*. *Scientific Reports,* ***9*(1)**, 1-14.

Johnson, T. J., & Nolan, L. K. (2009). Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and Molecular Biology Reviews*, ***73*(4)**, 750-774.

Johnson, W., Lior, H., & Johnson, K. (1978). Heat-stable enterotoxin from *Escherichia coli*:factors involved in growth and toxin production. *Infection and Immunity,* ***20*(2)**, 352- 359.

Kaper, J. B., Nataro, J. P., & Mobley, H. L. (2004). Pathogenic *Escherichia coli*. *Nature Reviews Microbiology*, ***2*(2)**, 123-140.

Karjalainen, T. K., Evans, D. G., Evans, D. J., Graham, D. Y., & Lee, C.-H. (1991). Catabolite repression of the colonization factor antigen I (CFA/I) operon of *Escherichia coli*. *Current Microbiology,* ***23*(6)**, 307-313.

Kaur, P., Chakraborti, A., & Asea, A. (2010). Enteroaggregative *Escherichia coli*: an emerging enteric food borne pathogen. *Interdisciplinary Perspectives on Infectious Diseases*, ***2010***.

Kesty, N. C., Mason, K. M., Reedy, M., Miller, S. E., & Kuehn, M. J. (2004). Enterotoxigenic *Escherichia coli* vesicles target toxin delivery into mammalian cells. *The EMBO Journal,* ***23*(23)**, 4538-4549.

Kharat, V. B., Ahmed, M., Jiang, Z. D., Riddle, M. S., & DuPont, H. L. (2017). Colonization factors in enterotoxigenic *Escherichia coli* strains in travelers to Mexico, Guatemala, and India compared with children in Houston, Texas. *The American journal of Tropical Medicine and Hygiene*, ***96*(1)**, 83.

Klipstein, F. A., Engert, R. F., Houghten, R. A., & Rowe, B. (1984). Enzyme-linked immunosorbent assay for *Escherichia coli* heat-stable enterotoxin. *Journal of Clinical Microbiology*, ***19*(6)**, 798-803.

Kortman, G. A., Boleij, A., Swinkels, D. W., & Tjalsma, H. (2012). Iron availability increases the pathogenic potential of *Salmonella typhimurium* and other enteric pathogens at the intestinal epithelial interface. *PLoS One, **7***(**1**).

Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam,S., Wu, Y., Sow, S.O., Sur, D., Breiman, R.F. and Faruque, A.S. & Levine, M. M.(2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *The Lancet*, ***382***(**9888**), 209-222.

Kumar, P., Luo, Q., Vickers, T. J., Sheikh, A., Lewis, W. G., & Fleckenstein, J. M. (2014).EatA, an immunogenic protective antigen of enterotoxigenic *Escherichia coli*, degradesintestinal mucin. *Infection and Immunity, **82***(**2**), 500-508.

Labrie, V., Beausoleil, H. E., Harel, J., & Dubreuil, J. D. (2001). Binding to sulfatide and enterotoxicity of various *Escherichia coli* STb mutants. *Microbiology*, ***147***(**11**), 3141- 3148.

Lamberti, L. M., Bourgeois, A. L., Fischer Walker, C. L., Black, R. E., & Sack, D. (2014). Estimating diarrhoeal illness and deaths attributable to *Shigellae* and enterotoxigenic *Escherichia coli* among older children, adolescents, and adults in South Asia and Africa. *PLoS Neglected Tropical Diseases*, ***8***(**2**), e2705.

Lanata, C. F., Fischer-Walker, C. L., Olascoaga, A. C., Torres, C. X., Aryee, M. J., Black, R. E., & Child Health Epidemiology Reference Group of the World Health Organization and UNICEF. (2013). Global causes of diarrhoeal disease mortality in children < 5 years of age: a systematic review. *PloS One*, ***8***(**9**), e72788.

Lasaro, M. A., Rodrigues, J. F., Mathias-Santos, C., Guth, B. E. C., Balan, A., Sbrogio- Almeida, M. E., & Ferreira, L. C. S. (2008). Genetic diversity of heat-labile toxin expressed by enterotoxigenic *Escherichia coli* strains isolated from humans. *Journal ofBacteriology*, ***190***(**7**), 2400-2410.

Li, Y., Luo, Q., Shi, X., Lin, Y., Qiu, Y., Lv, D., Jiang, Y., Chen, Q., Jiang, M., Ma, H. & Hu, Q. (2017). Phenotypic and genotypic characterization of clinical Enterotoxigenic *Escherichia coli* Isolates from Shenzhen, China. *Foodborne Pathogens and Disease*, ***14***(**6**), 333-340.

Liévin-Le Moal, V., & Servin, A. L. (2013). Pathogenesis of human enterovirulent bacteria: lessons from cultured, fully differentiated human colon cancer cell lines. *Microbiology and Molecular Biology Reviews, **77***(**3**), 380-439.

Lim, J. Y., Yoon, J. W., & Hovde, C. J. (2010). A brief overview of *Escherichia coli* O157: H7 and its plasmid O157. *Journal of Microbiology and Biotechnology*, ***20***(**1**), 5.

Lindenthal, C., & Elsinghorst, E. A. (1999). Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***67***(**8**), 4084-4091.

Lindenthal, C., & Elsinghorst, E. A. (2001). Enterotoxigenic *Escherichia coli* TibA glycoprotein adheres to human intestine epithelial cells. *Infection and Immunity*, ***69***(**1**), 52-57.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods, 25*(**4**), 402-408.

Lowry, O., Rosebrough, N., Farr, A., & Randall, R. (1951). Protein measurementwith the Folin phenol reagent. *Journal of Biological Chemistry*, ***193***(**1**), 265-275.

Luo, Q., Kumar, P., Vickers, T. J., Sheikh, A., Lewis, W. G., Rasko, D. A., Sistrunk, J. &Fleckenstein, J. M. (2014). Enterotoxigenic *Escherichia coli* secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells. *Infection and Immunity*, ***82***(**2**), 509-521.

Madhavan, T. V., & Sakellaris, H. (2015). Colonization factors of Enterotoxigenic *Escherichia coli*. *Advances in applied microbiology* (**Vol. 90, pp. 155-197**): Elsevier.

Magistro, G., Magistro, C., Stief, C. G., & Schubert, S. (2017). The high-pathogenicity island (HPI) promotes flagellum-mediated motility in extraintestinal pathogenic *Escherichia coli*. *PLoS One, **12***(**10**).

Makvana, S., & Krilov, L. R. (2015). *Escherichia coli* infections. *Pediatrics in Review*, ***36***(**4**),167-171.

Mammarappallil, J. G., & Elsinghorst, E. A. (2000). Epithelial cell adherence mediated by the enterotoxigenic *Escherichia coli* Tia protein. *Infection and Immunity*, ***68***(**12**), 6595-6601.

Maniatis, T. J (1982). *Molecular Cloning*.

Mansour, A., Shaheen, H. I., Amine, M., Hassan, K., Sanders, J. W., Riddle, M. S., Armstrong, A.W., Svennerholm, A.M., Sebeny, P.J., Klena, J.D., Young, S.Y., & Frenck Jr, R. W. (2014). Pathogenicity and phenotypic characterization of enterotoxigenic *Escherichia coli* isolates from a birth cohort of children in ruralEgypt. *Journal of Clinical Microbiology*, ***52***(**2**), 587-591.

Margulieux, K. R., Srijan, A., Ruekit, S., Nobthai, P., Poramathikul, K., Pandey, P., Serichantalergs, O., Shrestha, S.K., Bodhidatta, L. & Swierczewski, B. E. (2018). Extended-spectrum β-lactamase prevalence and virulence factor characterization of enterotoxigenic *Escherichia coli* responsible for acute diarrhoea in Nepal from 2001 to 2016. *Antimicrobial Resistance and Infection Control*, ***7***(**1**), 1-7.

Mathur, R., Reddy, V., Naidu, A. N., & Krishnamachari, K. A. (1985). Nutritional status and diarrhoeal morbidity: a longitudinal study in rural Indian preschool children. *Human Nutrition: Clinical Nutrition*, ***39***(**6**), 447-454.

McConnell, M. M., Thomas, L. V., Willshaw, G. A., Smith, H. R., & Rowe, B. (1988). Genetic control and properties of coli surface antigens of colonization factor antigen IV (PCF8775) of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***56***(**8**), 1974-1980.

Mendez Arancibia, E., Pitart, C., Ruiz, J., Marco, F., Gascon, J., & Vila, J. (2009).

Evolution of antimicrobial resistance in enteroaggregative *Escherichia coli* and enterotoxigenic *Escherichia coli* causing traveller's diarrhoea. *Journal of Antimicrobial Chemotherapy*, ***64*****(2)**, 343-347.

Merson, M. H., Morris, G. K., Sack, D. A., Wells, J. G., Feeley, J. C., Sack, R. B., Creech,W.B., Kapikian, A.Z. & Gangarosa, E. J. (1976). Travelers' diarrhoea in Mexico: a prospective study of physicians and family members attending a congress. *New England Journal of Medicine*, ***294*****(24)**, 1299-1305.

Meyer, T. S., & Lamberts, B. L. (1965). Use of coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide-gel strips. *Biochimica et Biophysica Acta*, ***107*****(1)**, 144-145.

Miller, J., & Lee, K. (1984). *Experiments in molecular genetics*.

Mims, C. A., Nash, A. A., & Stephen, J. (2001). Mims' pathogenesis of infectious disease. *Gulf Professional Publishing*.

Minodier, P. (2011). *Escherichia coli* [esh″ə-rik′e-ə co′lī]. *Euro Surveill,* ***16***, 19815.

Mirhoseini, A., Amani, J., & Nazarian, S. (2018). Review on pathogenicity mechanism of enterotoxigenic *Escherichia coli* and vaccines against it. *Microbial pathogenesis,* ***117***,162-169.

Mondal, D., Haque, R., Sack, R. B., Kirkpatrick, B. D., & Petri Jr, W. A. (2009). Attribution of malnutrition to cause-specific diarrhoeal illness: evidence from a prospective study of preschool children in Mirpur, Dhaka, Bangladesh. *The American Journal of Tropical Medicine and Hygiene*, ***80*****(5)**, 824.

Mondal, D., Minak, J., Alam, M., Liu, Y., Dai, J., Korpe, P., Liu, L., Haque, R. & Petri Jr, W. A. (2012). Contribution of enteric infection, altered intestinal barrier function, and maternal malnutrition to infant malnutrition in Bangladesh. *Clinical Infectious Diseases*, ***54*****(2)**, 185-192.

Mondal, I., Bhakat, D., Chowdhury, G., Manna, A., Samanta, S., Deb, A. K., Mukhopadhyay, A.K. & Chatterjee, N. S. (2022). Distribution of virulence factors and its relatedness towards the antimicrobial response of enterotoxigenic *Escherichia coli* strains isolated from patients in Kolkata, India. *Journal of Applied Microbiology*, ***132*****(1)**, 675-686.

Montero, D., Vidal, M., Pardo, M., Torres, A., Kruger, E., Farfán, M., O'Ryan, M., Luo,Q., Fleckenstein, J., Del Canto, F. & Vidal, R. (2017). Characterization of enterotoxigenic *Escherichia coli* strains isolated from the massive multi-pathogen gastroenteritis outbreak in the Antofagasta region following the Chilean earthquake, 2010. *Infection, Genetics and Evolution*, ***52***, 26-29.

Morris, G. K., Merson, M. H., Sack, D. A., Wells, J. G., Martin, W. T., Dewitt, W. E., Feeley, J.C., Sack, R.B. & Bessudo, D. M. (1976). Laboratory investigation of diarrhoea in travelers to Mexico: evaluation of methods for detecting enterotoxigenic

Echerichia coli. *Journal of Clinical Microbiology*, ***3*(5)**, 486-495.

Moseley, S. L., Huq, I., Alim, A. A., So, M., Samadpour-Motalebi, M., & Falkow, S. (1980). Detection of enterotoxigenic *Escherichia coli* by DNA colony hybridization. *Journal of Infectious Diseases*, ***142*(6)**, 892-898.

Muir, A., & Hopfer, U. (1985). Regional specificity of iron uptake by small intestinal brush-border membranes from normal and iron-deficient mice. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* ***248*(3)**, G376-G379.

Nada, R. A., Armstrong, A., Shaheen, H. I., Nakhla, I., Sanders, J. W., Riddle, M. S., Young, S. & Sebeny, P. (2013). Phenotypic and genotypic characterization of enterotoxigenic *Escherichia coli* isolated from US military personnel participating in Operation Bright Star, Egypt, from 2005 to 2009. *Diagnostic Microbiology and Infectious Disease*, ***76*(3)**, 272-277.

Nair, G. B., & Takeda, Y. (1998). The heat-stable enterotoxins. *Microbial Pathogenesis*, ***24*(2)**, 123-131.

Nataro, J. P., & Kaper, J. B. (1998). Diarrhoeagenic *Escherichia coli*. *Clinical Microbiology Reviews,* ***11*(1)**, 142-201.

Nawar, H. F., King-Lyons, N. D., Hu, J. C., Pasek, R. C., & Connell, T. D. (2010). LT-IIc, a new member of the type II heat-labile enterotoxin family encoded by an *Escherichia coli* strain obtained from a nonmammalian host. *Infection and Immunity*, ***78*(11)**, 4705-4713.

Nazarian, S., Gargari, S. L. M., Rasooli, I., Alerasol, M., Bagheri, S., & Alipoor, S. D. (2014). Prevalent phenotypic and genotypic profile of enterotoxigenic *Escherichia coli* among Iranian children. *Japanese Journal of Infectious Diseases*, ***67*(2)**, 78-85.

Nicklasson, M., Sjöling, Å., Lebens, M., Tobias, J., Janzon, A., Brive, L., & Svennerholm, A. M. (2008). Mutations in the periplasmic chaperone leading to loss of surface expression of the colonization factor CS6 in enterotoxigenic *Escherichia coli* (ETEC) clinical isolates. *Microbial Pathogenesis*, ***44*(3)**, 246-254.

Nicklasson, M., Sjöling, Å., Von Mentzer, A., Qadri, F., & Svennerholm, A.-M. (2012). Expression of colonization factor CS5 of enterotoxigenic *Escherichia coli* (ETEC) is enhanced *in vivo* and by the bile component Na glycocholate hydrate. *PLoS One,* ***7*(4)**,e35827.

Norton, E. B., Lawson, L. B., Freytag, L. C., & Clements, J. D. (2011). Characterization of a mutant *Escherichia coli* heat-labile toxin, LT (R192G/L211A), as a safe and effectiveoral adjuvant. *Clinical and Vaccine Immunology*, ***18*(4)**, 546-551.

Nunes, M., Penna, F., Franco, R., Mendes, E., & Magalhaes, P. (2011). Enterotoxigenic *Escherichia coli* in children with acute diarrhoea and controls in Teresina/PI, Brazil: distribution of enterotoxin and colonization factor genes. *Journal of Applied Microbiology,* ***111*(1)**, 224-232.

Nyholm, O., Halkilahti, J., Wiklund, G., Okeke, U., Paulin, L., Auvinen, P., Haukka, K. & Siitonen, A. (2015). Comparative genomics and characterization of hybrid Shigatoxigenic and enterotoxigenic *Escherichia coli* (STEC/ETEC) strains. *PLoS One*, ***10*(8)**, e0135936.

O'Mahony, M. C., Noah, N. D., Evans, B., Harper, D., Rowe, B., Lowes, J. A., Pearson, A. & Goode, B. (1986). An outbreak of gastroenteritis on a passenger cruise ship. *Epidemiology and Infection*, ***97*(2)**, 229-236.

Olive, D. M. (1989). Detection of enterotoxigenic *Escherichia coli* after polymerase chain reaction amplification with a thermostable DNA polymerase. *Journal of Clinical Microbiology*, ***27*(2)**, 261-265.

Ørskov, F., Ørskov, I., Evans, D. J., Sack, R. B., Sack, D. A., & Wadström, T. (1976). Special *Escherichia coli* serotypes among enterotoxigenic strains from diarrhoea in adults and children. *Medical Microbiology and Immunology*, ***162*(2)**, 73-80.

Oyofo, B. A., Subekti, D. S., Svennerholm, A. M., Machpud, N. N., Tjaniadi, P., Komalarini, T. S., Setiawan, B., Campbell, J.R., Corwin, A.L. & Lesmana, M. (2001). Toxins and colonization factor antigens of enterotoxigenic *Escherichia coli* among residents of Jakarta, Indonesia. *The American Journal of Tropical Medicine and Hygiene*, ***65*(2)**, 120-124.

Pacheco, A. R., & Sperandio, V. (2015). Enteric pathogens exploit the microbiota-generatednutritional environment of the gut. *Metabolism and Bacterial Pathogenesis*, **279-296**.

Pasqua, M., Michelacci, V., Di Martino, M. L., Tozzoli, R., Grossi, M., Colonna, B., Morabito, S. & Prosseda, G. (2017). The intriguing evolutionary journey of enteroinvasive *E. coli* (EIEC) toward pathogenicity. *Frontiers in Microbiology*, ***8***, 2390.

Patel, S. K., Dotson, J., Allen, K. P., & Fleckenstein, J. M. (2004). Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***72*(3)**, 1786-1794.

Perez-Casal, J. O. S. E., Swartley, J. S., & Scott, J. R. (1990). Gene encoding the major subunit of CS1 pili of human enterotoxigenic *Escherichia coli*. *Infection and Immunity*, ***58*(11)**, 3594-3600.

Puiprom, O., Chantaroj, S., Gangnonngiw, W., Okada, K., Honda, T., Taniguchi, T., & Sawanpanyalert, P. (2010). Identification of colonization factors of enterotoxigenic *Escherichia coli* with PCR-based technique. *Epidemiology and Infection,* ***138*(4)**, 519-524.

Qadri, F., Das, S. K., Faruque, A. S. G., Fuchs, G. J., Albert, M. J., Sack, R. B., & Svennerholm, A. M. (2000). Prevalence of toxin types and colonization factors in enterotoxigenic *Escherichia coli* isolated during a 2-year period from diarrhoeal patients in Bangladesh. *Journal of Clinical Microbiology*, ***38*(1)**, 27-31.

Qadri, F., Khan, A. I., Faruque, A. S. G., Begum, Y. A., Chowdhury, F., Nair, G. B., Salam, M.A., Sack, D.A. & Svennerholm, A. M. (2005). Enterotoxigenic *Escherichia coli* and *Vibrio cholerae* diarrhoea, Bangladesh, 2004. *Emerging Infectious Diseases*, ***11*****(7)**, 1104.

Qadri, F., Svennerholm, A. M., Faruque, A. S., & Sack, R. B. (2005). Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clinical Microbiology Reviews,* ***18*****(3)**, 465-483.

Qadri, F., Saha, A., Ahmed, T., Al Tarique, A., Begum, Y. A., & Svennerholm, A. M. (2007). Disease burden due to enterotoxigenic *Escherichia coli* in the first 2 years of life in an urban community in Bangladesh. *Infection and Immunity*, ***75*****(8)**, 3961-3968.

Rao, M. R., Abu-Elyazeed, R., Savarino, S. J., Naficy, A. B., Wierzba, T. F., Abdel-Messih, I., Shaheen, H., Frenck Jr, R.W., Svennerholm, A.M. & Clemens, J. D. (2003). High disease burden of diarrhoea due to enterotoxigenic *Escherichia coli* amongrural Egyptian infants and young children. *Journal of Clinical Microbiology*, ***41*****(10)**, 4862-4864.

Ridley, A. M., Toszeghy, M. J., Cawthraw, S. A., Wassenaar, T. M., & Newell, D. G. (2008). Genetic instability is associated with changes in the colonization potential of Campylobacter jejuni in the avian intestine. *Journal of Applied Microbiology*, ***105*****(1)**, 95-104.

Rivera, F., Ochoa, T., Maves, R., Bernal, M., Medina, A., Meza, R., Barletta, F., Mercado,E., Ecker, L., Gil, A.I. & Hall, E.R. (2010). Genotypic and phenotypic characterization of enterotoxigenic *Escherichia coli* strains isolated from Peruvian children. *Journal of Clinical Microbiology,* ***48*****(9)**, 3198-3203.

Rivera, F. P., Medina, A. M., Aldasoro, E., Sangil, A., Gascon, J., Ochoa, T. J., Vila, J. &Ruiz, J. (2013). Genotypic characterization of enterotoxigenic *Escherichia coli* strainscausing traveler's diarrhoea. *Journal of Clinical Microbiology*, ***51*****(2)**, 633-635.

Rocha, L., Ozaki, C., Horton, D., Menezes, C., Silva, A., Fernandes, I., Magnoli, F.C., Vaz, T.M., Guth, B.E. & Piazza, R. (2013). Different assay conditions for detecting the production and release of heat-labile and heat-stable toxins in enterotoxigenic *Escherichia coli* isolates. *Toxins,* ***5*****(12)**, 2384-2402.

Roy, K., Hamilton, D., Allen, K. P., Randolph, M. P., & Fleckenstein, J. M. (2008). The EtpA exoprotein of enterotoxigenic *Escherichia coli* promotes intestinal colonization and is a protective antigen in an experimental model of murine infection. *Infection and Immunity*, ***76*****(5)**, 2106-2112.

Roy, K., Hilliard, G. M., Hamilton, D. J., Luo, J., Ostmann, M. M., & Fleckenstein, J. M. (2009). Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature*, ***457*****(7229)**, 594-598.

Roy, K., Kansal, R., Bartels, S. R., Hamilton, D. J., Shaaban, S., & Fleckenstein, J. M. (2011). Adhesin degradation accelerates delivery of heat-labile toxin by enterotoxigenic *Escherichia coli*. *Journal of Biological Chemistry*, ***286*****(34)**, 29771-29779.

Rychlik, W. J. S. W., Spencer, W. J., & Rhoads, R. E. (1990). Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Research*, ***18*****(21)**, 6409- 6412.

Sabui, S., Ghosal, A., Dutta, S., Ghosh, A., Ramamurthy, T., Nataro, J. P., Hamabata, T. & Chatterjee, N. S. (2010). Allelic variation in colonization factor CS6 of enterotoxigenic *Escherichia coli* isolated from patients with acute diarrhoea and controls. *Journal of Medical Microbiology*, ***59*****(7)**, 770-779.

Sabui, S., Dutta, S., Debnath, A., Ghosh, A., Hamabata, T., Rajendran, K., Ramamurthy,T., Nataro, J.P., Sur, D., Levine, M.M. & Chatterjee, N.S., (2012). Real-time PCR-based mismatch amplification mutation assay for specific detection of CS6-expressingallelic variants of enterotoxigenic *Escherichia coli* and its application in assessing diarrhoeal cases and asymptomatic controls. *Journal of Clinical Microbiology,* ***50*****(4)**, 1308-1312.

Sabui, S., Debnath, A., Ghosal, A., Wajima, T., Hamabata, T., Ramamurthy, T., Ghosh, A.N., Basak, S. & Chatterjee, N. S. (2016). Characterization of oligomeric assembly of colonization factor CS6 from enterotoxigenic *Escherichia coli*. *Microbiology,* ***162*****(1)**, 72-83.

Sack, D. A., Wells, J., Merson, M., Sack, R. B., & Morris, G. (1975). Diarrhoea associated with heat-stable enterotoxin-producing strains of *Escherichia coli*. *The Lancet*, ***306*****(7928)**, 239-241.

Sack, D. A., Kaminsky, D. C., Sack, R. B., Wamola, I. A., Orskov, F., Orskov, I., Slack, R.C., Arthur, R.R. & Kapikian, A. Z. (1977). Enterotoxigenic *Escherichia coli* diarrhoea of travelers: a prospective study of American Peace Corps volunteers. *The Johns Hopkins Medical Journal*, ***141*****(2)**, 63-70.

Sack, R. B., Gorbach, S. L., Banwell, J. G., Jacobs, B., Chatterjee, B. D., & Mitra, R. C. (1971). Enterotoxigenic *Escherichia coli* isolated from patients with severe cholera-likedisease. *The Journal of Infectious Diseases*, **378-385**.

Sack, R. B., Hirschhorn, N., Brownlee, I., Cash, R. A., Woodward, W. E., & Sack, D. A. (1975). Enterotoxigenic Escherichia-coli-associated diarrhoeal disease in Apache children. *New England Journal of Medicine*, ***292*****(20)**, 1041-1045.

Sack, R. B. (1975). Human diarrhoeal disease caused by enterotoxigenic *Escherichia coli*. *Annual Review of Microbiology*, ***29*****(1)**, 333-354.

Sack, R. B. (1990). Travelers' diarrhoea: microbiologic bases for prevention and treatment. *Reviews of Infectious Diseases,* ***12*****(Supplement_1)**, S59-S63.

Sack, R. B. (2011). The discovery of cholera-like enterotoxins produced by *Escherichia coli* causing secretory diarrhoea in humans. *The Indian Journal of Medical Research,*

***133*(2)**,171.

Sahl, J. W., & Rasko, D. A. (2012). Analysis of global transcriptional profiles of enterotoxigenic *Escherichia coli* isolate E24377A. *Infection and Immunity,* ***80*(3)**, 1232-1242.

Sahl, J. W., Sistrunk, J. R., Fraser, C. M., Hine, E., Baby, N., Begum, Y., Luo, Q., Sheikh,A., Qadri, F., Fleckenstein, J.M. & Rasko, D.A. (2015). Examination of the enterotoxigenic *Escherichia coli* population structure during human infection. *mBio,* ***6*(3)**, e00501-00515.

Sambrook, J., & Russell, D. W. (2006). The Inoue method for preparation and transformation of competent *E. coli*: "ultra-competent" cells. *Cold Spring Harbor Protocols*, ***2006*(1)**, 10-1101.

Sears, C. L., & Kaper, J. B. (1996). Enteric bacterial toxins: mechanisms of action and linkageto intestinal secretion. *Microbiological Reviews*, ***60*(1)**, 167-215.

Sekirov, I., Russell, S. L., Antunes, L. C. M., & Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiological Reviews,* ***90*(3)**, 859-904.

Servin, A. L. (2014). Pathogenesis of human diffusely adhering *Escherichia coli* expressing Afa/Dr adhesins (Afa/Dr DAEC): current insights and future challenges. *Clinical Microbiology Reviews*, ***27*(4)**, 823-869.

Shah, N., DuPont, H. L., & Ramsey, D. J. (2009). Global etiology of travelers' diarrhoea: systematic review from 1973 to the present. *The American Journal of Tropical Medicine and Hygiene*, ***80*(4)**, 609-614.

Shaheen, H. I., Khalil, S. B., Rao, M. R., Elyazeed, R. A., Wierzba, T. F., Peruski, L. F., Putnam, S., Navarro, A., Morsy, B.Z., Cravioto, A. & Clemens, J.D. (2004). Phenotypic profiles of enterotoxigenic *Escherichia coli* associated with early childhood diarrhoea in rural Egypt. *Journal of Clinical Microbiology,* ***42*(12)**, 5588-5595.

Shaheen, H. I., Abdel Messih, I. A., Klena, J. D., Mansour, A., El-Wakkeel, Z., Wierzba, T. F., Sanders, J.W., Khalil, S.B., Rockabrand, D.M., Monteville, M.R. andRozmajzl, P.J., & Frenck, R. W. (2009). Phenotypic and genotypic analysis of enterotoxigenic *Escherichia coli* in samples obtained from Egyptian children presenting to referral hospitals. *Journal of Clinical Microbiology*, ***47*(1)**, 189-197.

Sherlock, O., Vejborg, R. M., & Klemm, P. (2005). The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infection and Immunity*, ***73*(4)**, 1954-1963.

Shetty, V. A., Kumar, S. H., Shetty, A. K., Karunasagar, I., & Karunasagar, I. (2012). Prevalence and characterization of diarrhoeagenic *Escherichia coli* isolated from adults and children in Mangalore, India. *Journal of Laboratory Physicians*, ***4*(01)**, 024-029.

Shevchuk, N. A., Bryksin, A. V., Nusinovich, Y. A., Cabello, F. C., Sutherland, M.,

& Ladisch, S. (2004). Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Research*, ***32*(2)**, e19-e19.

Simuyandi, M., Chilengi, R., Connor, S. B., Voeglein, J. B., Laban, N. M., Mwila-Kazimbaya, K., Chisenga, C.C., Mwaba, J., Sack, D.A. & Chakraborty, S. (2019). Enterotoxigenic *Escherichia coli* toxins and colonization factors among Zambian children presenting with moderate to severe diarrhoea to selected health facilities. *Archives of Microbiology and Immunology*, ***3*(4)**, 173-184.

Sjöling, Å., Wiklund, G., Savarino, S., Cohen, D., & Svennerholm, A.-M. (2007). Comparative analyses of phenotypic and genotypic methods for detection of enterotoxigenic *Escherichia coli* toxins and colonization factors. *Journal of Clinical Microbiology,* ***45*(10)**, 3295-3301.

Smith, J. L., Fratamico, P. M., & Gunther, N. W. (2007). Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathogens and Disease*, ***4*(2)**, 134-163.

Sommerfelt, H., Steinsland, H., Grewal, H. M., Viboud, G. I., Bhandari, N., Gaastra, W., Svennerholm, A.M. & Bhan, M. K. (1996). Colonization factors of enterotoxigenic *Escherichia coli* isolated from children in north India. *Journal of Infectious Diseases*, ***174*(4)**, 768-776.

Sperandio, V., Nguyen, Y. (2012). Enterohemorrhagic *E. coli* (EHEC) pathogenesis. *Frontiers in Cellular and Infection Microbiology*, ***2*,** 90.

Steinsland, H., Valentiner-Branth, P., Perch, M., Dias, F., Fischer, T. K., Aaby, P., Mølbak, K. & Sommerfelt, H. (2002). Enterotoxigenic *Escherichia coli* infections and diarrhoea in a cohort of young children in Guinea-Bissau. *The Journal of Infectious Diseases*, ***186*(12)**, 1740-1747.

Steinsland, H., Valentiner-Branth, P., Gjessing, H. K., Aaby, P., Mølbak, K., & Sommerfelt, H. (2003). Protection from natural infections with enterotoxigenic *Escherichia coli*: longitudinal study. *The Lancet*, ***362*(9380)**, 286-291.

Stoll, B. J., Rowe, B., Glass, R. I., Gross, R. J., & Huq, I. (1983). Changes in serotypes of enterotoxigenic *Escherichia coli* in Dhaka over time: usefulness of polyvalent antisera. *Journal of Clinical Microbiology*, ***18*(4)**, 935-937.

Subekti, D. S., Lesmana, M., Tjaniadi, P., Machpud, N., Daniel, J. C., Alexander, W. K.,Campbell, J.R., Corwin, A.L., Beecham III, H.J., Simanjuntak, C. & Oyofo, B. A.(2003). Prevalence of enterotoxigenic *Escherichia coli* (ETEC) in hospitalized acute diarrhoea patients in Denpasar, Bali, Indonesia. *Diagnostic Microbiology and InfectiousDisease*, ***47*(2)**, 399-405.

Svennerholm, A. M., & Glenn, G. (2016). Vaccines against enterotoxigenic *Escherichia coli*. *New Generation Vaccines*, **742-750**.

Svenungsson, B., Lagergren, Å., Ekwall, E., Evengård, B., Hedlund, K. O., Kärnell, A., Löfdahl, S., Svensson, L. & Weintraub, A. (2000). Enteropathogens in adult patients with diarrhoea and healthy control subjects: a 1-year prospective study in a Swedish clinic for infectious diseases. *Clinical Infectious Diseases*, ***30*****(5)**, 770-778.

Szabó, E. 1., Skedsmo, A., Sonnevend, A., Al-Dhaheri, K., Emödy, L., Usmani, A., & Pal, T. (2005). Curli expression of enterotoxigenic *Escherichia coli*. *Folia Microbiologica,* ***50*****(1)**, 40-46.

Tacket, C. O. (2004). Plant-derived vaccines against diarrhoeal diseases. *Expert Opinion on Biological Therapy*, ***4*****(5)**, 719-728.

Tang, H., Li, Y., Zhang, P., Guo, J., Huang, Z., Xu, H., Hou, Q., Wang, C., Zeng, M., Jin, H. & Xu, X. (2015). Surveillance for diarrhoeagenic *Escherichia coli* in Shanghai, 2012-2013. *Zhonghua liu Xing Bing xue za zhi= Zhonghua Liuxingbingxue Zazhi*, ***36*****(11)**, 1263-1268.

Tapader, R., Bose, D., & Pal, A. (2017). YghJ, the secreted metalloprotease of pathogenic *E. coli* induces hemorrhagic fluid accumulation in mouse ileal loop. *Microbial Pathogenesis*, ***105***, 96-99.

Thomas, L. V., McConnell, M. M., Rowe, B., & Field, A. M. (1985). The possession of threenovel coli surface antigens by enterotoxigenic *Escherichia coli* strains positive for the putative colonization factor PCF8775. *Microbiology*, ***131*****(09)**, 2319-2326.

Tickell, K. D., Sharmin, R., Deichsel, E. L., Lamberti, L. M., Walson, J. L., Faruque, A. S. G., Pavlinac, P.B., Kotloff, K.L. & Chisti, M. J. (2020). The effect of acute malnutrition on enteric pathogens, moderate-to-severe diarrhoea, and associated mortality in the Global Enteric Multicenter Study cohort: a post-hoc analysis. *The Lancet Global Health*, ***8*****(2)**, e215-e224.

Tobias, J., Lebens, M., Källgård, S., Nicklasson, M., & Svennerholm, A. M. (2008). Role of different genes in the CS6 operon for surface expression of enterotoxigenic *Escherichia coli* colonization factor CS6. *Vaccine*, ***26*****(42)**, 5373-5380.

Toma, C., Lu, Y., Higa, N., Nakasone, N., Chinen, I., Baschkier, A., Rivas, M. & Iwanaga, M. (2003). Multiplex PCR assay for identification of human diarrhoeagenic *Escherichia coli*. *Journal of Clinical Microbiology*, ***41*****(6)**, 2669-2671.

Torres, A. G., Zhou, X., & Kaper, J. B. (2005). Adherence of diarrhoeagenic *Escherichia coli* strains to epithelial cells. *Infection and Immunity*, ***73*****(1)**, 18-29.

Torres, O. R., González, W., Lemus, O., Pratdesaba, R. A., Matute, J. A., Wiklund, G., Sack, D.A., Bourgeois, A.L. & Svennerholm, A. M. (2015). Toxins and virulence factors of enterotoxigenic *Escherichia coli* associated with strains isolated from indigenous children and international visitors to a rural community in Guatemala. *Epidemiology and Infection*, ***143*****(8)**, 1662-1671.

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O. & Calteau, A (2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS*

Genetics, ***5*(1)**, e1000344.

Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, ***76*(9)**, 4350-4354.

Trachman, J. D., & Maas, W. K. (1998). Temperature regulation of heat-labile enterotoxin (LT) synthesis in *Escherichia coli* is mediated by an interaction of H-NS protein with the LT A-subunit DNA. *Journal of Bacteriology,* ***180*(14)**, 3715-3718.

Trainor, E., Iturriza-Gómara, M., Ngwira, B., & Cunliffe, N. (2016). Detection of enterotoxigenic *E. coli* in hospitalised children with and without diarrhoea in Blantyre, Malawi. *Paediatrics and International Child Health*, ***36*(2)**, 102-105.

Turner, S. M., Chaudhuri, R. R., Jiang, Z. D., DuPont, H., Gyles, C., Penn, C. W., Pallen, M.J. & Henderson, I. R. (2006). Phylogenetic comparisons reveal multiple acquisitions of the toxin genes by enterotoxigenic *Escherichia coli* strains of different evolutionary lineages. *Journal of Clinical Microbiology*, ***44*(12)**, 4528-4536.

Vaandrager, A. B. (2002). Structure and function of the heat-stable enterotoxin receptor/guanylyl cyclase C. *Molecular and Cellular Biochemistry*, ***230*(1)**, 73-83.

Van Asbeck, B., Marcelis, J., Marx, J., Struyvenberg, A., Van Kats, J., & Verhoef, J. (1983). Inhibition of bacterial multiplication by the iron chelator deferoxamine: potentiating effect of ascorbic acid. *European Journal of Clinical Microbiology,* ***2*(5)**, 426-431.

Vidal, R.M., Muhsen, K., Tennant, S.M., Svennerholm, A.M., Sow, S.O., Sur, D., Zaidi, A.K., Faruque, A.S., Saha, D., Adegbola, R. & Hossain, M.J., (2019). Colonizationfactors among enterotoxigenic *Escherichia coli* isolates from children with moderate- to-severe diarrhoea and from matched controls in the Global Enteric Multicenter Study (GEMS). *PLoS Neglected Tropical Diseases*, ***13*(1)**, e0007037.

Vilchez, S., Becker-Dreps, S., Amaya, E., Perez, C., Paniagua, M., Reyes, D., Espinoza, F. & Weintraub, A. (2014). Characterization of enterotoxigenic *Escherichia coli* strains isolated from Nicaraguan children in hospital, primary care and community settings. *Journal of Medical Microbiology*, ***63*(Pt 5)**, 729.

von Mentzer, A., Tobias, J., Wiklund, G., Nordqvist, S., Aslett, M., Dougan, G., Sjöling, Å. & Svennerholm, A.-M. (2017). Identification and characterization of the novel colonization factor CS30 based on whole genome sequencing in enterotoxigenic *Escherichia coli* (ETEC). *Scientific Reports,* ***7*(1)**, 12514.

Wajima, T., Sabui, S., Fukumoto, M., Kano, S., Ramamurthy, T., Chatterjee, N. S., & Hamabata, T. (2011). Enterotoxigenic *Escherichia coli* CS6 gene products and their roles in CS6 structural protein assembly and cellular adherence. *Microbial Pathogenesis,* ***51*(4)**, 243-249.

Walker, C. L.F., Sack, D., & Black, R. E. (2010). Etiology of diarrhoea in older children, adolescents and adults: a systematic review. *PLoS Neglected Tropical Diseases*, ***4(8)***, e768.

Walker, R. I., Steele, D., Aguado, T., & Ad Hoc ETEC Technical Expert Committee. (2007). Analysis of strategies to successfully vaccinate infants in developing countries against enterotoxigenic *E. coli* (ETEC) disease. *Vaccine*, ***25*****(14)**, 2545-2566.

Wang, H., Zhong, Z., Luo, Y., Cox, E., & Devriendt, B. (2019). Heat-stable enterotoxins ofenterotoxigenic *Escherichia coli* and their impact on host immunity. *Toxins*, ***11*****(1)**, 24.

Wennerås, C., & Erling, V. (2004). Prevalence of enterotoxigenic *Escherichia coli*-associateddiarrhoea and carrier state in the developing world. *Journal of Health, Population andNutrition*, **370-382**.

William, S., Feil, H., & Copeland, A. (2012). Bacterial genomic DNA isolation using CTAB. *Sigma*, ***50***(6876).

Wolf, M. K., Andrews, G. P., Tall, B. D., McConnell, M. M., Levine, M. M., & Boedeker, E. C. (1989). Characterization of CS4 and CS6 antigenic components of PCF8775, a putative colonization factor complex from enterotoxigenic *Escherichia coli* E8775. *Infection and Immunity*, ***57*****(1)**, 164-173.

Wolf, M. K., De Haan, L. A., Cassels, F. J., Willshaw, G. A., Warren, R., Boedeker, E. C., & Gaastra, W. (1997). The CS6 colonization factor of human enterotoxigenic *Escherichia coli* contains two heterologous major subunits. *FEMS Microbiology Letters,**148***(**1**), 35-42.

Yano, T., Oliveira, M. S., Fontes, C. F., De Almeida, A. C. P., & de Castro, A. P. (1982). Detection of heat-labile (LT) enterotoxin of enterotoxigenic *Escherichia coli* by the radial immune hemolysis test: a modification for clinical use. *Medical Microbiology and Immunology*, ***171*****(3)**, 171-178.

Youssef, M., Shurman, A., Bougnoux, M. E., Rawashdeh, M., Bretagne, S., & Strockbine, N. (2000). Bacterial, viral and parasitic enteric pathogens associated with acute diarrhoeain hospitalized children from northern Jordan. *FEMS Immunology & Medical Microbiology*, ***28*****(3)**, 257-263.

Zegeye, E. D., Govasli, M. L., Sommerfelt, H., & Puntervoll, P. (2018). Development of anenterotoxigenic *Escherichia coli* vaccine based on the heat-stable toxin. *Human Vaccines & Immunotherapeutics*.

Zhang, W., & Sack, D. A. (2012). Progress and hurdles in the development of vaccines againstenterotoxigenic *Escherichia coli* in humans. *Expert Review of Vaccines*, ***11*****(6)**, 677-694.

Zimmermann, M. B., Chassard, C., Rohner, F., N'Goran, E. K., Nindjin, C., Dostal, A., Utzinger, J., Ghattas, H., Lacroix, C. & Hurrell, R. F. (2010). The effects of iron fortification on the gut microbiota in African children: a randomized controlled trial in Cote d'Ivoire. *The American Journal of Clinical Nutrition,* ***92*****(6)**, 1406-1415.

Chapter 9

APPENDIX

Reagents and Solutions:

1. Luria-Bertani (LB) broth

Reagents	Amount (g L^{-1})
Tryptone	10.00
Yeast extract	5.00
Sodium chloride	10.00
Final pH 7.0 ± 0.2	

2. LB-agar

Reagents	Amount (g L^{-1})
Tryptone	10.00
Yeast extract	5.00
Sodium chloride	10.00
Agar	18 - 20
Final pH 7.0 ± 0.2	

3. MacConkey agar

Ingredients	Amount (g L^{-1})
Peptide digest of animal tissue	17.00
Agar	13.50
Lactose	10.00
Sodium chloride	5.00
Bile salts	1.50
Proteose peptone	3.00
Neutral red	0.03
Crystal violet	0.001
Final pH at 25˚C : 7.1 ± 0.2	
Additional 0.8% agar was added if and when required.	

4. CFA broth

Ingredients	Percentage
Cassamino acid	1%
Yeast extract	0.15%
Magnesium sulphate	0.005%
Manganese chloride	0.0005%
Final pH 7.4 ± 0.2	

5. M9 Minimal Salts

Ingredients	Amount (g L^{-1})
Disodium Phosphate	6.78
Monopotassium Phosphate	3.00
Sodium Chloride	0.5
Ammonium Chloride	1.0
Glucose 20%	20 mL
1M $MgSO_4$	2 mL
1M $CaCl_2$	0.1 mL
Final pH 7.4 ± 0.2	

6. SOB medium

Reagent	Amount (for 100 mL)
Tryptone	2.0 g
Yeast extract	0.5 g
1M NaCl	1 mL
1M KCl	0.25 mL
2M Mg^{2+} stock (filter sterilized)	1 mL

7. SOC medium

SOB + 20 mM glucose (filter sterilized)

8. Phosphate Buffered Saline (PBS)

Reagent	Amount (g L^{-1})
Di-sodium hydrogen phosphate	2.67
Potassium dihydrogen phosphate	1.088
Potassium chloride	0.1998
Sodium chloride	7.946
Volume adjusted to 1L with double distilled water	
Final pH 7.4	

9. Tris Acetate EDTA (TAE) buffer

Reagent	Final Concentration (mmolar L^{-1})
Tris Base	40
Glacial acetic acid	20
EDTA	1
Volume adjusted to 1L with double distilled water	
Final pH 8.0	

10. Tris-EDTA Buffer

Reagent	Final Concentration (mmolar L^{-1})
Tris-HCl (1M); pH 7.4-8	40
Glacial acetic acid	20
EDTA	1
Volume adjusted to 1L with double distilled water	
Final pH 8.0	

11. S-Buffer

CTAB solution: 10% in 0.7M NaCl.

4.1 g NaCl was dissolved in 80 mL MilliQ water. While stirring, 10 g CTAB was added. To dissolve, the solution was heated at 65ºC. The final volume was adjusted to 100 mL with MilliQ water. Final pH 8.0 was made with EDTA.

12. 2M Mg^{2+} stock

Reagent	Amount (grams per 100 mL)
Magnesium chloride	20.33
Magnesium sulfate	24.65

13. 10X DNA loading dye

Reagent	Amount
Glycerol	3.9 mL
10% SDS	0.5 mL
0.5M EDTA	0.2 mL
Bromophenol Blue (BB)	25 mg
Xylene Cyanol (XC)	25 mg
Volume adjusted to 10 mL with double distilled water	
Store at -20°C	

14. Nutrient agar Stab

Reagent	Amount (g L^{-1})
Nutrient Broth	8.00
Agar	8.00
Sodium Chloride	5.00

15. Dulbecco′s Modified Eagle′s Medium (DMEM) Complete Media

Reagent	Amount
DMEM	13.5 grams per Litre
Sodium Bicarbonate	3.7 grams per Litre
Fetal Bovine Serum (FBS)	100 mL
Antibiotics	10 mL
Non-Essential Amino Acids	10 mL
Volume adjusted to 1L with MilliQ water	
Final pH 7.4	

16. Dulbecco′s Modified Eagle′s Medium (DMEM) Incomplete Media

Reagent	Amount
DMEM	13.5 grams per Litre
Sodium Bicarbonate	3.7 grams per Litre
Fetal Bovine Serum (FBS)	5 mL
Non-Essential Amino Acids	10 mL
Volume adjusted to 1L with MilliQ water	
Final pH 7.4	

17. 5X SDS Gel-loading dye

Reagent	Amount
Tris-HCl pH6.8	100 mM
SDS	4% (w/v)
Bromophenol blue	0.2% (w/v)
Glycerol	20% (w/v)
B-mercaptoethanol	200 mM
Volume adjusted to 1L with MilliQ water	
Final pH 7.4	

18. Buffer for polyacrylamide gel electrophoresis

Reagent	Final Concentration
Solution A	
Acrylamide	29.2%
Bis acrylamide	0.8%
The solution was stored in light resistant container.	
Solution B	
Tris HCl (pH 8.8)	1.5 M
SDS	0.4%
Solution C	
Tris HCl (pH 6.8)	0.5 M
SDS	0.4%

19. SDS-PAGE Running buffer

Reagent	Final Concentration
Tris base	0.025 M
Glycine	0.192 M
SDS	0.1%
The pH will be 8.3. No need for pH adjustment.	

20. Staining solution

Reagent	Final Concentration
Methanol	50%
Acetic acid	50%
Coomassie Brilliant blue	0.1%

21. Distaining solution

Reagent	Final Concentration
Methanol	5%
Acetic acid	7.5%
dH_2O	87.5%

22. Transfer buffer (for Western blot)

Reagent	Amount ($g\ L^{-1}$)
Tris	5.8
Glycine	2.9
SDS	0.05

23. X-Gal (2mL)

X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactose, 100 mg) was dissolved in 2 mL of N,N´- dimethyl formamide to make a stock of 50 mg/mL X-Gal. The tube was covered with aluminium foil and stored in -20°C until use.

24. IPTG stock (0.1 M)

24 mg of Isopropyl-β-D-thiogalactopyranoside (IPTG) was dissolved in 1 mL of N, N´-dimethylformamide. The solution was filter sterilized and the tube was covered with aluminium foil. The solution was stored in 4 °C until use.

25. LA plate with antibiotics/IPTG/X-Gal

LA plate with appropriate antibiotic(s) was incubated at 37 °C for 30 minutes and 100 µl of 0.1M IPTG and 20 µl of 50 mg/mL X-Gal were spreaded on the LA plate. The plate was incubated for 30 minutes at 37 °C.

26. Cell lysis buffer

Reagent	Final Concentration
Potassium phosphate Buffer, pH 7.8	50 mM
NaCl	400 mM
KCl	100 mM
Glycerol	10%
Triton X-100	0.5%
Imidazole	10 mM

Prepare the potassium phosphate buffer (pH 7.8) by mixing 0.3 mL KH_2PO_4 and 4.7 mL K_2HPO_4 from their respective 1 M stock.

27. Plasmid DNA Purification Reagents

Reagent	Final Concentration
Cell Resuspension Solution	
Tris-HCl (pH 7.5)	50 mM
EDTA	10 mM
RNase A	100 µg/mL
Cell Lysis Solution	
NaOH	0.2 M
SDS	1%
Neutralization Solution	
Potassium Acetate (pH 4.8)	1.32 M
Column Wash Solution	
Potassium Acetate	80mM
Tris-HCl (pH 7.5)	8.3mM
EDTA	40µM
95% Ethanol	55%

28. Transformation Buffer

Reagent	Final Concentration
PIPES	10 mM
$CaCl_2$•$2H_2O$	15 mM
KCl	250 mM
Adjust pH to 6.7 with 10 M KOH	
Then add, $MnCl_2$•$4H_2O$	1.09 g for 100 mL
Filter sterilize	

29. Antibiotic Stock

Unless otherwise indicated, the antibiotic powder was dissolved in autoclaved MilliQ water and filter sterilized. The stock solutions aliquots were stored at -20ºC.

30. Purification of DNA from Agarose gel

Reagent	Final Concentration
Membrane Wash Solution	
Potassium Acetate, pH 5.0	10 mM
95% Ethanol	80 %
EDTA (pH 8.0)	16.7µM
Membrane Binding Solution	
Guanidine Isothiocyanate	4.5 M
Potassium Acetate, pH 5.0	0.5 M

31. Reagents for ELISA

A. Blocking Buffer

5% skimmed milk in PBS- T (0.05% Tween 20 in PBS, pH 7.4)

B. Wash Buffer

0.05% Tween 20 in PBS, pH 7.4

C. Coating Buffer

50 mM sodium carbonate, pH 9.6 and 20 mM Tris-HCL, pH 8.5

D. Primary/Secondary Antibody Solution

Primary/secondary antibody was diluted in 1x blocking buffer

E. Stopping Solution

2 M H_2SO_4

32. Z buffer

Reagent	Final Concentration
Sodium hydrogen phosphate	0.06 M
Sodium dihydrogen phosphate	0.04 M
Potassium chloride	0.01 M
Magnesium sulfate	0.001 M
B -mercaptoethanol (BME)	0.05 M
Adjust pH to 7.0 and store at 4˚C	

33. Phosphate buffer, 0.1 M

Reagent	Final Concentration
Sodium hydrogen phosphate	0.06 M
Sodium dihydrogen phosphate	0.04 M
Adjust pH to 7.0	
Phosphate buffer is stable at room temperature and does not need to be made fresh each time.	

34. ONPG (Ortho-Nitrophenyl-ß-galactoside)

The ONPG was dissolved to a final concentration of 4mg/mL in 0.1M phosphate buffer pH 7.0.It should be prepared fresh each day prior to the experiment.

Primers Used

Table 9.1: **List of the oligonucleotide sequences of the primers used in multiplex PCR**

Target Gene	Primer Name	Oligonucleotide Sequences (5' → 3')	Product Size (bp)	Gene Bank Accn. No.
CS1	Forward	TTGACCTTCTGCAATCTGA	324	M58550
	Reverse	CATCTGCATGGATTGAAAG		
CS2	Forward	TAACTGCTAGCGTTGATCC	368	Z47800
	Reverse	ATTAGTTTGCTGGGTGCTT		
CS3	Forward	GGTGGGTGTTTTGACTCTT	264	M35657
	Reverse	TGTTCGTTACCTTCAGTGG		
CS4	Forward	TTTTGCAAGCTGATGGTAG	250	X97493
	Reverse	TCTGCAGGTTCAAAAGTCA		
CS5	Forward	CGGATTGGATATACCGTTT	453	X63411
	Reverse	TCAACAGCAAATGTTACCG		
CS6	Forward	ATCCAGCCTTCTTTTGGTA	321	U04844
	Reverse	ACCAACCATAACCTGATCG		
CS14	Forward	ACTGTGACAGCCAGTGTTG	394	X97491
	Reverse	AAACGACGCCTTGATACTT		
CS17	Forward	AACCTATTCTTCGGCTTCA	290	X97495
	Reverse	GCGCAGTTCCTTGTGTG		
CS21	Forward	ACAGTCCGCGTAGCAAT	400	EU107092.1
	Reverse	GTAAAACAGTTGTAGAGG		
CFA/I	Forward	GCTCTGACCACAATGTTTG	364	S73191
	Reverse	TTACACCGGATGCAGAATA		
CFA/III (CS8)	Forward	GCCTTCTGGAAGTCATCAT	438	D37957
	Reverse	CTGCCACATACTCCCAG		
tia	Forward	CTCTGGCTGATGAGAGC	172	U20318.1
	Reverse	TCATAGCCCACTGCAAG		
etpA	Forward	CAGACAGCTACACCAAC	221	AY920525.2
	Reverse	CGATTGAGTCGTCTCAG		
leoA	Forward	AGGCGAATCTGAAAGGC	315	AF170971.1
	Reverse	CATTCTTCTCGACAAAGG		

tibA	Forward	GTGCTTAACAATACTGGG	396	AF109215.1
	Reverse	TTCCAGCAGCATATTGAC		
eatA	Forward	CTGTAAATGGCGCTTATC	465	AY163491.2
	Reverse	TTAATGTTCCCACTCCTG		
elt	Forward	CACACGGAGCTCCTCAG	324	M17874.1
	Reverse	CAAACTAGTTTTCCATACTG		
est	Forward	GCTAAACCAGTAGAGTC	149	M34916.1
	Reverse	CACCCGGTACAAGCAGG		

Table 9.2: **List of Primers used for qRT-PCR**

Target Gene	Primer Name	Oligonucleotide Sequences (5' → 3')	Product Size (bp)	Gene Bank Accn. No.
***csfD* (CS5)**	Forward	ACCCTGAGTACCGCAGTAGA	93	X63411.1
	Reverse	AACTGTTGGACTAGCGCTCC		
***cssB* (CS6)**	Forward	CGGTTGAGATGACGATCCCT	90	U04844.1
	Reverse	GCTTTCCGATCTGCTGTCCA		
eatA	Forward	GGGCCTCTTGCAAACTATCT	103	AY163491.2
	Reverse	CTGATGACAGAACGCCAACTA		
parC	Forward	ATCCGCTGGTTGATGGTCAG	92	EU561348.1
	Reverse	TTTCGACAGACGGGATTCGG		

Table 9.3: **Primers for isogenic non-polar mutants and ETEC complement strains**

Target Gene	Primer Number	Primer Name	Oligonucleotide Sequences (5' → 3')	Product Size (bp)	Gene Bank Accn. No.
CS5 (A)	I	Forward	AGTGAATGCAGCAGGTC	438	AJ224079.2
	II	Reverse	ATCGCTCAAGACGTGTAATTCTCATCCTTCTTTGTCT		
pKD13(B)	III	Forward	TTACACGTCTTGAGCGAT	1362	AY048744.1
	IV	Reverse	AACTATTCCGGCAAAGGAATTCACTGATCAGTGATAAG		

CS5 (C)	V	Forward	TCCTTTGCCGGAATAGTT	372	AJ22407 9.2
	VI	Reverse	ATCATTCAGTGTATCTGGG		
CS5 IP	VII	Forward	GGGCTTTTGTGAACGGATT	402	AJ22407 9.2
	VIII	Reverse	CTGCCTTGGCATTCATATC		
CS5 complement	IX	Forward	ATGAAGAAAAATTTACTGATAACT	612	AJ22407 9.2
	X	Reverse	TTAAATATCAACAGCAAATGTTA		
eatA (A)	I	Forward	TCGGGAGGATAATTAGGG	381	FN6494 18.1
	II	Reverse	ATCGCTCAAGACGTGTAAAGCTATCCAACATAATGCG		
pKD13 (B')	III	Forward	TTACACGTCTTGAGCGAT	1362	AY0487 44.1
	IV	Reverse	GTACAAGGAGCTATCTATTATTCACTGATCAGTGATAAG		
eatA (C)	V	Forward	AATAGATAGCTCCTTGTAC	351	FN6494 18.1
	VI	Reverse	ATATAGACAAACCAAACGC		
eatA IP	VII	Forward	AACGCTATGTCCTGATAG	262	FN6494 18.1
	VIII	Reverse	AAGCGAAACATCGCATCG		
eatA complement	IX	Forward	ATGAATAAAGTGTTCTCTCTT	4092	FN6494 18.1
	X	Reverse	GAAATAATAACGGAAGTTAGC		
pKD46	1	Forward	ATCACCACTCTTCGCCAG	235	AY0487 46.1
	2	Reverse	TACAGGATTCATTGTCCTG		

Table 9.4: **Primers for Promoter Study and deletion of promoter regions**

Target Gene	Primer Number	Primer Name	Oligonucleotide Sequences (5' → 3')	Product Size (bp)	Gene Bank Accn. No.
CS6 promoter	I	Forward	TCCCCCCGGGATAACTAACTGAAAAACAATG	573	NC022333.1
	II	Reverse	GCTCTAGATATAACACCATTACTATTGCT		
CS5 promoter	I	Forward	TCCCCCCGGGACAATCATTATACTATATTGAC	207	NZCP023347.1
	II	Reverse	GCTCTAGATTTATTCTCATCCTTCTTTG		
EatA promoter	I	Forward	TCCCCCCGGGCTGTTTAAATTCGGGAGG	406	FN649418.1
	II	Reverse	GCTCTAGATTTTGATAGCTGCGCAGC		

CS6 promoter	1 prCS6 (-573)	Forward	TCCCCCCGGGTATAACTAACTGAAAAACAATG	597	NC022333.1
	2 prCS6 (-512)	Forward	TCCCCCCGGGACAAGCCGCAGCCGCC	538	
	3 prCS6 (-456)	Forward	TCCCCCCGGGAATGCGCGGTGGAAATG	480	
	4 prCS6 (-350)	Forward	TCCCCCCGGGACTACGATGGTCAGGTTG	374	
	4A prCS6 (-335)	Forward	TCCCCCCGGGTTGCGGAGGCTATGGC	359	
	4B prCS6 (-320)	Forward	TCCCCCCGGGCCCTGGTACGAGCGCT	344	
	4C prCS6 (-305)	Forward	TCCCCCCGGGTGAACAAAATGACGAAAGC	329	
	4D prCS6 (-290)	Forward	TCCCCCCGGGAAGCAGGTATGCCTGAAA	314	
	4E prCS6 (-275)	Forward	TCCCCCCGGGAAAGCGTGCGTATTGCC	299	
	4F prCS6 (-260)	Forward	TCCCCCCGGGCCTGAAAACACAACCCG	284	
	5 prCS6 (-255)	Forward	TCCCCCCGGGAAACACAACCCGCTACG	279	
	6 prCS6 (-78)	Forward	TCCCCCCGGGTCAATACGGTGCTCCG	102	
	7 prCS6 (-23)	Forward	TCCCCCCGGGTATAGCAATAGTAATGGTGTT	47	
	prCS6(+25)	Reverse	GCTCTAGAAATTAAACCAATTGTTTTCTTC		

Table 9.5: Primers used for cloning of *eatA* in pet22b vector

Target Gene	Primer Number	Primer Name	Oligonucleotide Sequences (5' → 3')	Product Size (bp)	Gene Bank Accn. No.
eatA	*reatA* I	Forward	TTTTT*GGATCC*CTACGATAAGAATGGAGT	1483	FN649418.1
	reatA II	Reverse	TTTGG*CTCGAG*CGACCGTACGCCTTTGATT		

Vectors used:

pKD13

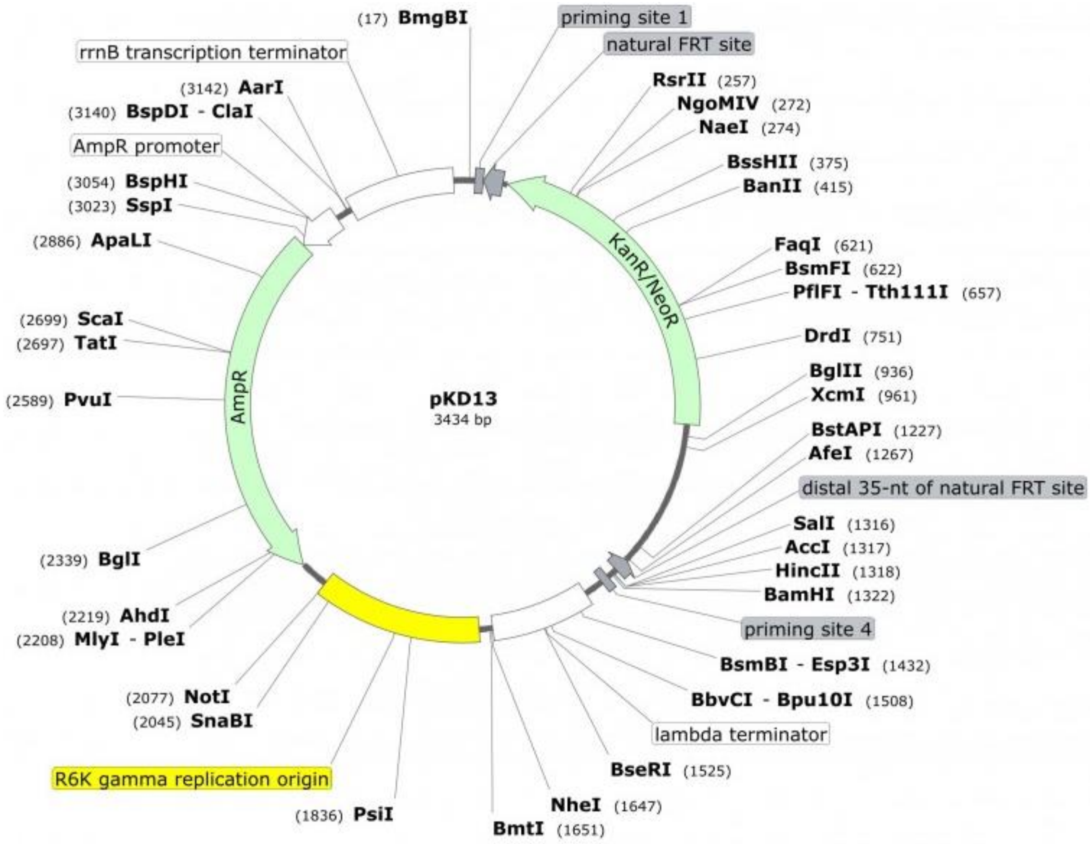

Figure 9.1: Template plasmid for gene disruption. The resistance gene is flanked by FRT sites. 3434 base pairs. (Reference: Datsenko *et al.,* 2000)

pKD46

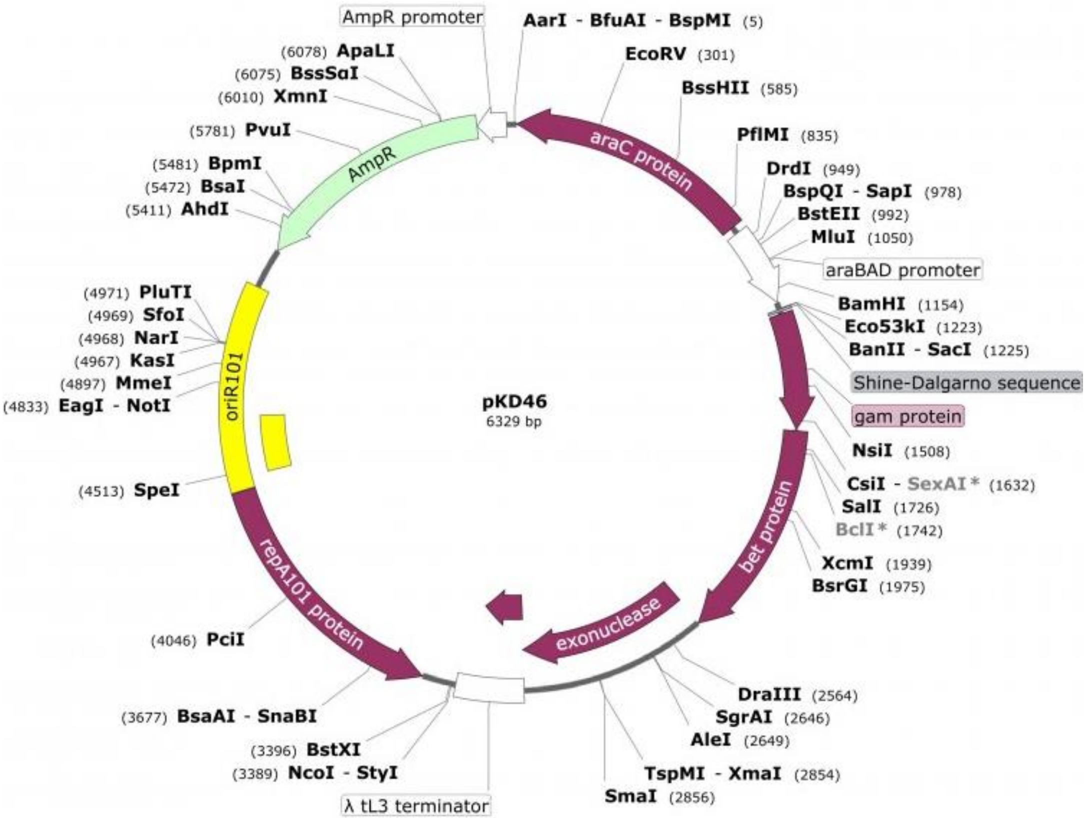

Figure 9.2: A plasmid for RED-recombineering – site-targeted modification of any locus in E. coli. The plasmid shows temperature-sensitive replication (not greater than 30°C). (Reference: Datsenko *et al.,* 2000)

pCP20

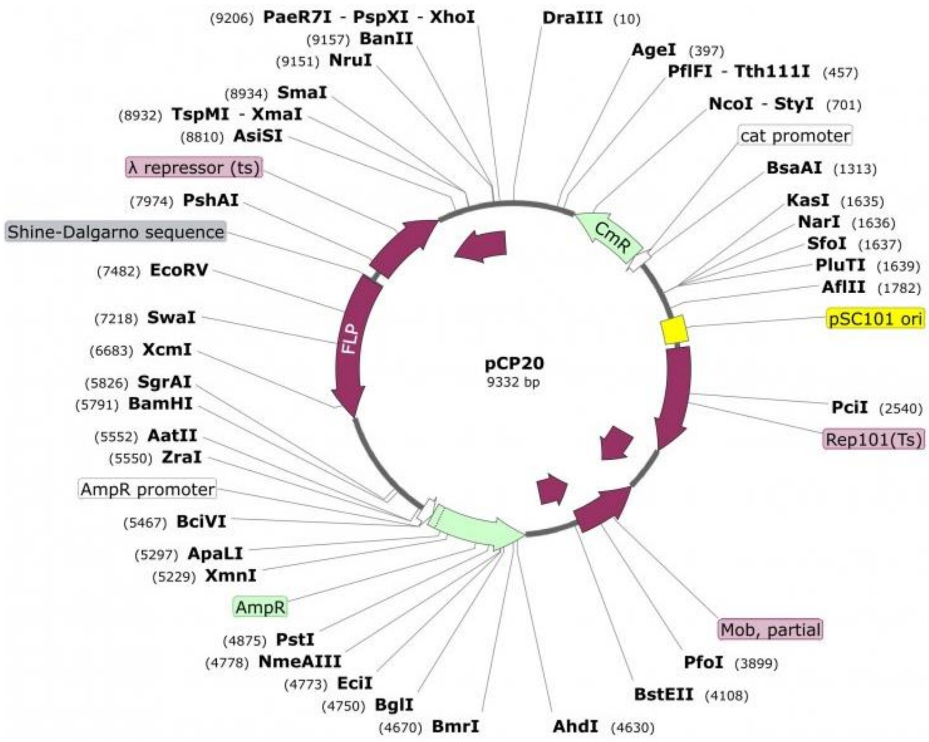

Figure 9.3: A plasmid for introducing deletions in E. coli genome. The plasmid carries gene coding for temperature-sensitive version of Rep101 protein needed for replication with the pSC101 origin.

pGEM®-T Easy Vector

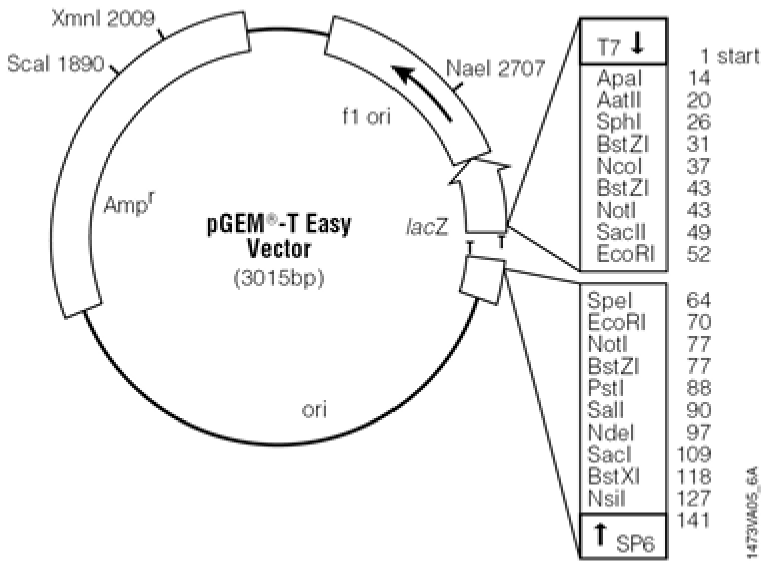

Figure 9.4: pGEM®-T Easy Vector (Promega) Map. The top strand shown corresponds to the RNA synthesized by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesized by SP6 RNA polymerase.

pBAD-TOPO TA expression vector

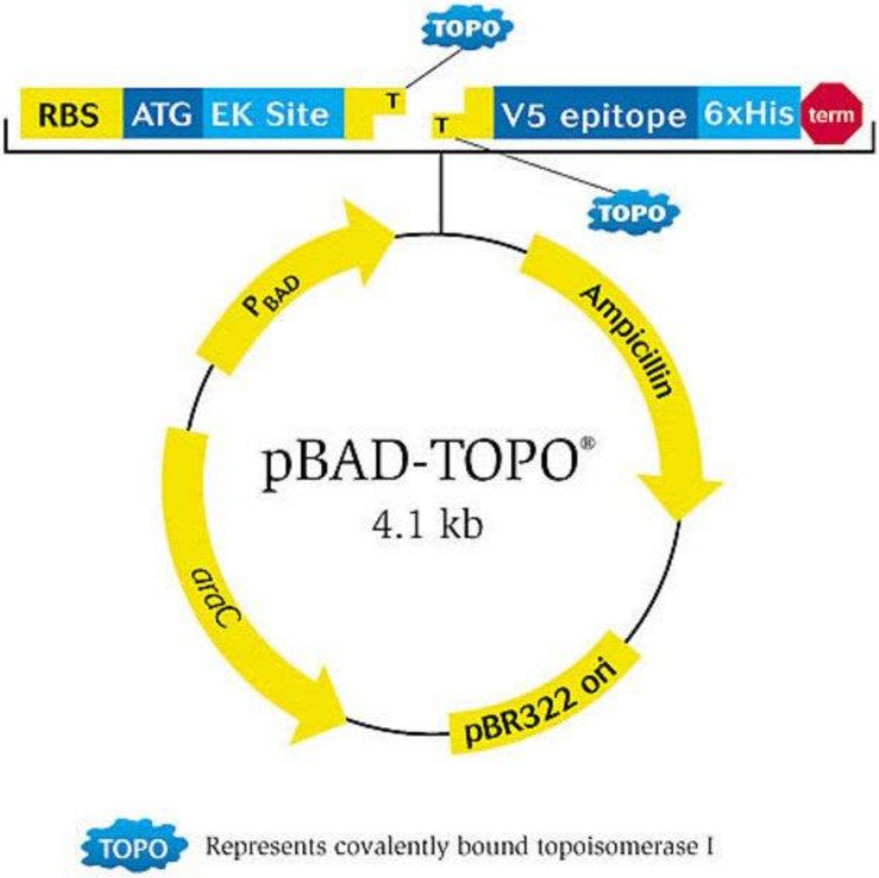

Figure 9.5: Vector map of pBAD-TOPO. Expression in E. coli is driven by the araBAD promoter (PBAD). The AraC gene product encoded on the pBAD-TOPO® plasmid positively regulates this promoter.

pTL61T Vector

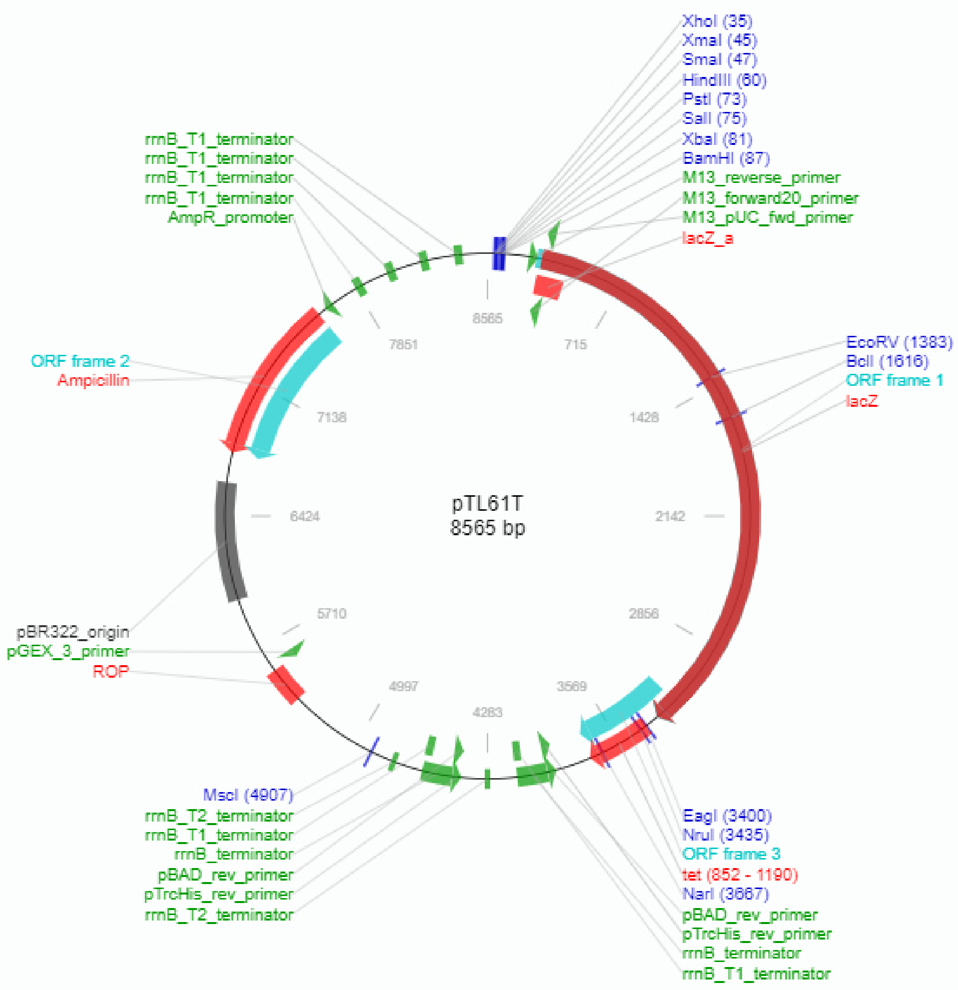

Figure 9.6: A promoterless *lacZ* fusion plasmid used for promoter region identification.

pET-22b(+)

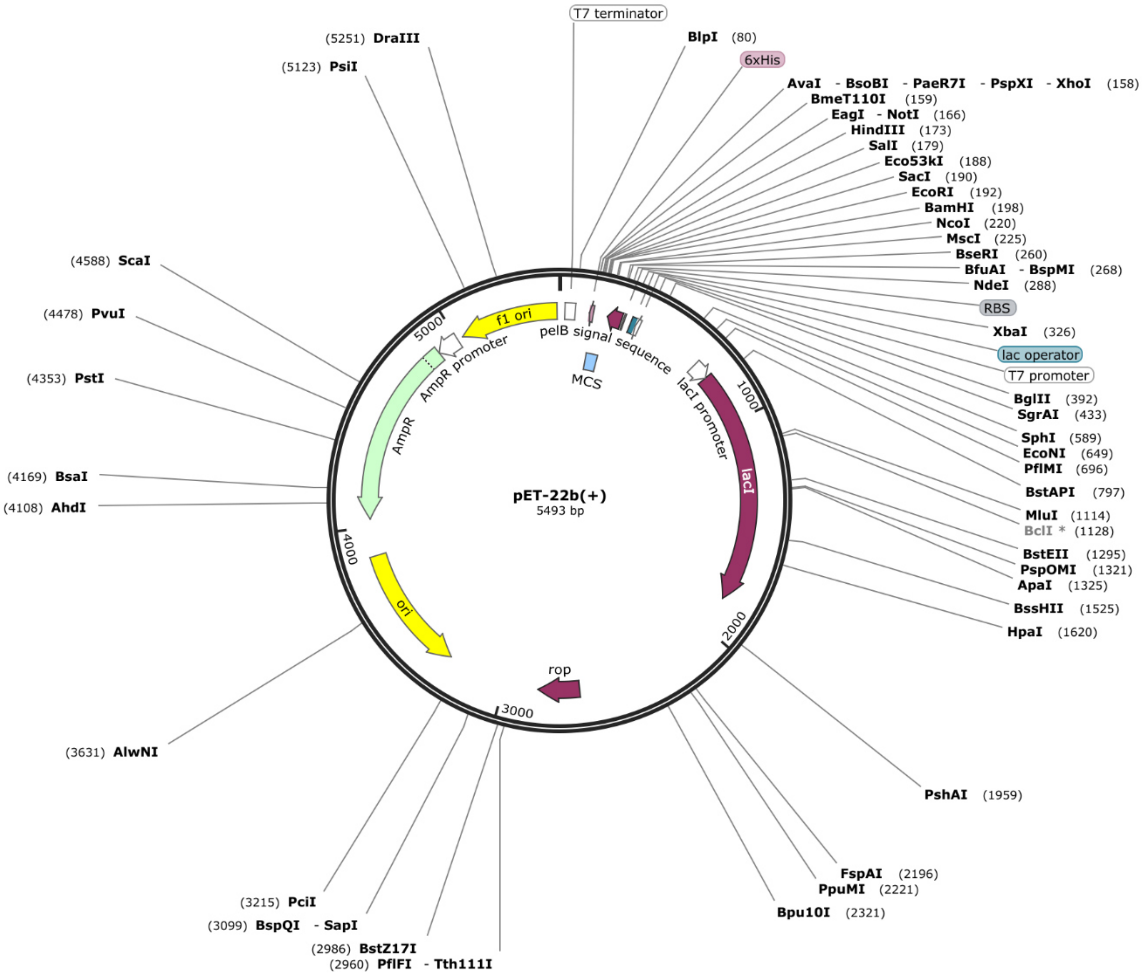

Figure 9.7: Bacterial vector that encodes a signal sequence for inducible expression of proteins in the periplasm.

Chapter 10

PUBLICATIONS

&

CONFERENCES

JOURNAL ARTICLES PUBLISHED

- **Bhakat, D.**, Debnath, A., Naik, R., Chowdhury, G., Deb, A.K., Mukhopadhyay, A.K. and Chatterjee, N.S., (2019). Identification of common virulence factors present in enterotoxigenic *Escherichia coli* isolated from diarrhoeal patients in Kolkata, India. *Journal of Applied Microbiology*, *126*(1), pp.255-265.
- **Bhakat, D**., Mondal, I. and Chatterjee, N.S., (2020). EatA, a non-classical virulence factor, of Enterotoxigenic *Escherichia coli* (ETEC) is modulated by the host factors during pathogenesis. *International Journal of Infectious Diseases, 101* (supplement 1), pp.3-4.
- Mondal, I., **Bhakat, D.**, Chowdhury, G., Manna, A., Samanta, S., Deb, A. K., Mukhopadhyay, A. K., & Chatterjee, N. S. (2022). Distribution of virulence factors and its relatedness towards the antimicrobial response of enterotoxigenic *Escherichia coli* strains isolated from patients in Kolkata, India. *Journal of Applied Microbiology*, *132*(1), 675-686.
- **Bhakat, D.**, Mondal, I., Mukhopadhyay, A. K., & Chatterjee, N. S. (2021). Iron influences the expression of colonization factor CS6 of enterotoxigenic *Escherichia coli. Microbiology*, *167*(9), 001089.

CONFERENCES ATTENDED

- Poster presented in 26th ISCB International Conference (ISCBC-2020) held in Nirma Institute of Pharmacy International Conference (NIPiCON) at Nirma University, Ahmedabad, India. **Title:** Understanding the Prevalence and Expression of Virulence Factors harbored by Enterotoxigenic *Escherichia coli*. **Bhakat, D**., Mondal, I. and Chatterjee, N.S., 2020.
- Poster presented in 15th Asian Conference on Diarrhoeal Disease and Nutrition (ASCODD 2020) held in Pan Pacific Sonargaon Hotel, Dhaka, Bangladesh. **Title:** Understanding modulation of Colonization factor CS6 of enterotoxigenic *E. coli* expression by host factors. **Bhakat, D**., Mondal, I. and Chatterjee, N.S., 2020.
- Poster presented in 31st ECCMID, the European Congress of Clinical Microbiology and Infectious Diseases, (Online). **Title**: Revealing the region-specific prevalence and expression of virulence factors present in enterotoxigenic *Escherichia coli*. Mondal, I., **Bhakat, D**., Mukhopadhyay, A. K. and Chatterjee, N.S., 2021

- Oral presentation at International e-Conference - 2021 on "Covid-19: Challenges and Opportunities in Pharmaceutical Research." **Title:** Interpreting the Expression profile of prevalent colonisation factor CS6 of Enterotoxigenic *E. coli* in Kolkata region. . Mondal, I., **Bhakat, D**., Mukhopadhyay, A. K. and Chatterjee, N.S., 2021

Journal of
Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Identification of common virulence factors present in enterotoxigenic *Escherichia coli* isolated from diarrhoeal patients in Kolkata, India

D. Bhakat[1], A. Debnath[1], R. Naik[1], G. Chowdhury[2], A.K. Deb[3], A.K. Mukhopadhyay[2] and N.S. Chatterjee[1]

1 Division of Biochemistry, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India
2 Division of Bacteriology, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India
3 Division of Epidemiology, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India

Keywords
colonization factor, diarrhoea, enterotoxigenic *Escherichia coli*, ETEC, nonclassical virulence factor, vaccine.

Correspondence
Nabendu Sekhar Chatterjee, Division of Biochemistry, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India.
E-mail: nschatterjee@rediffmail.com

2017/2315: received 24 November 2017, revised 17 July 2018 and accepted 23 August 2018

doi:10.1111/jam.14090

Abstract

Aims: Development of an effective vaccine against enterotoxigenic *Escherichia coli* (ETEC) is largely dependent on the conscientious understanding of different virulence associated factors from diverse geographical areas. So, the objective of this study is to elucidate the distribution of enterotoxins, CF and NCVF in clinical ETEC strains isolated between 2008 and 2014 from two hospitals in Kolkata, India.
Methods and Results: Multiplex PCR method was used for detection of two enterotoxin genes, 11 common CFs and five common NCVFs. Among the 350 tested ETEC strains, 61% strains possessed *est*+*elt* genes, 25% *est* and 14% *elt*. Among 56% CF positive ETEC strains, CS21 was the prevalent one (37%) followed by CS6 (36%). NCVF genes were present in 59% of the ETEC strains; *eatA* was the most prevalent (65%) followed by *etpA* (51%). There were 29% strains negative for any CFs or NCVFs.
Conclusions: We conclude that a pattern exists between CS6, *eatA* and toxins. We observed *est* with or without *elt*, CS6 with or without CS5 and with or without *eatA* were present in 24% of clinical ETEC strains (59/250) analysed. CS21 has emerged as another predominant CF but it had diverse CFs and NCVFs.
Significance and Impact of the Study: Prevalence of ETEC virulence factors would help in tracking ETEC globally and suggests the need of a multivalent ETEC vaccine.

Introduction

Enterotoxigenic *Escherichia coli* (ETEC) is one of the important bacterial pathogens contributing to diarrhoeal diseases worldwide (Kotloff *et al.* 2013). A Global enteric multi-centric prospective case–control study done in developing countries with four sites in Africa and three in Asia (Kolkata, India; Mirzapur, Bangladesh and Karachi, Pakistan) underlined the importance of this issue by highlighting that ETEC featured in every site as one of the four causative organism of diarrhoea among children younger than 5 years of age. In Bangladesh, ETEC affects infant, children <5 years of age and also adults living in or travelling to the country (Qadri *et al.* 2005). Annually, ETEC is estimated to cause 200 million diarrhoeal episodes during epidemic or sporadic outbreaks and approximately 380 000 deaths globally (WHO 2006).

In order to be able to colonize the surface of the intestinal epithelium, the bacteria must attach first where bacterial adhesins come into play and act on the cognate receptors present on host epithelium. This helps the bacteria to withstand intestinal peristalsis. In ETEC, adhesion to the small intestine is mediated by colonization factors (CFs), a major virulence determinant for initiating pathogenesis. Intestinal colonization is followed by the release of well-characterized, heat-labile (LT) and/or heat-stable (ST) enterotoxins (Gaastra and Svennerholm 1996) which trigger secretory diarrhoea.

With the inclusion of newly discovered CS30 (Von Mentzer *et al.* 2017), there are now 23 well-defined CFs (Gaastra and Svennerholm 1996; Qadri *et al.* 2005). CFs along with aforesaid toxins, are referred to as the classical virulence factors. Additional virulence-related factors have been identified (Del Canto *et al.* 2011) and these are termed 'nonclassical' virulence factors such as *tibA*, *tia*, *etpA*, *eatA*, *leo.* TibA is a glycosylated adhesion factor that promotes invasion, bacterial aggregation and biofilm formation (Elsinghorst and Kopecko 1992). Tia acts as an invasin as well as an adhesin (Elsinghorst and Weitz 1994). EtpA involved in adhesion (Fleckenstein *et al.* 2006). EatA is a member of the serine protease autotransporter that facilitates faster delivery of heat-labile toxin (Patel *et al.* 2004). LeoA is a cytoplasmic protein with GTPase activity, required for maximal LT secretion (Fleckenstein *et al.* 2000).

Only 6–8 of the 23 known CFs are most frequently detected in ETEC, of which the CFA/I and CS1 to CS6 are predominant (Qadri *et al.* 2005). CS6 was frequently detected (37%) among ETEC clinical isolates in our region i.e. the eastern India (Ghosal *et al.* 2007; Sabui *et al.* 2012; Dutta *et al.* 2013). Also in Thailand and Peruvian children, CS6 was mostly detected (Puiprom *et al.* 2010; Rivera *et al.* 2010). The CS14 and CFA/IV were the most common CFs in Teresina, Brazil (Nunes *et al.* 2011). In Bangladesh, the predominant CFs were CS6, CS5, CFA/I, CS7, CS17, CS1, CS3 and CS14 (Begum *et al.* 2014). In Iranian children, CFA/I, CS3, CS2 and CS5 were the most common CFs (Nazarian *et al.* 2014). ETEC isolates from a birth cohort in rural Egypt showed that most strains had CFA/I followed by CS6 as their major CF gene (Mansour *et al.* 2014). CS23 (Del Canto *et al.* 2012) and CS30 were recently discovered in 2017 (Von Mentzer *et al.* 2017).

Among nonclassical virulence factors (NCVFs), *eatA*, *etpA*, *etpB*, *irp2*, *fyuA* are the predominant (Del Canto *et al.* 2011; Guerra *et al.* 2014). In strains isolated from Bangladesh between 1998 and 2011, *eatA* and *etpA* were present in more than half of the ETEC isolates (Luo *et al.* 2015). In ETEC strains isolated from Bolivian children aged <5 years, *clyA* was the most common (92·6%), followed by *eatA* (34·8%), *east* (31·4%), *tibC* (20·4%), *tia* (12%) and *leoA* (6·7%) (Gonzales *et al.* 2013).

As ETEC is the major bacterial causative organism for diarrhoea in developing countries (Kosek *et al.* 2003; Black *et al.* 2010), it is important to characterize the inter-relationship between the virulence factors of them and the transmission potential. This would help to track the spread of ETEC strains round the globe. In this study, we analyzed the distribution of the genes encoding enterotoxins, commonly occurring CFs and NCVFs in clinical ETEC isolates collected through an ongoing hospital based diarrhoeal surveillance during 2008–2014. We have used previously established molecular methods to identify virulence factors of the commonly circulating strains in this region.

Materials and methods

Bacterial strains and culture conditions

Under the hospital based surveillance project of the National Institute of Cholera and Enteric Diseases (NICED), stool samples were collected from diarrhoea patients admitted at the Infectious Diseases and Beliaghata General Hospital (ID & BG Hospital, Kolkata) and Dr. B.C. Roy Post Graduate Institute of Pediatric Sciences during the period of 2008–2014 (limited to first 1000 cases). The diarrhoeal types of the patients were given in Table S1. This study was approved by the duly constituted Institutional Ethics Committee of ICMR-NICED. In this tenure, total 350 ETEC strains as sole and mixed pathogen were obtained from the archive of NICED. Thirty-five strains isolated from nondiarrhoeic stool collected from community served as comparators. For ETEC detection, stool samples were plated on MacConkey agar plates and incubated at 37°C for 18 h. The colonies fermenting the lactose and morphologically resembling *E. coli* were isolated and cryo-preserved in Luria Bertani broth containing 15% glycerol and stored at −80°C for further use. These strains were grown in LB media at 37°C for overnight to isolate genomic DNA.

Isolation of genomic DNA

Total genomic DNA was isolated as described previously (Ghosal *et al.* 2007). In brief, overnight grown ETEC culture was pelleted down, resuspended in water, and then boiled for 10 min. After centrifugation, the supernatant was used as template for PCR. The amplified products were detected by 1% agarose gel electrophoresis.

Multiplex PCR for gene detection

The primer sets for multiplex PCR were designed by keeping similar T_m values. Primers were tested for degeneracy and cross-reactivity using BLAST program. (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch). Eleven sets of primers were designed for the detection of CFs, two primer set for toxin gene and five pair of primers for the NCVF (Table 1). STh variant was considered for the detection of the ST toxin. The PCR conditions used for amplification were as follows: initial denaturation at 95°C for 3 min, followed by 30 cycles of denaturation at 95°C for 50 s, annealing at 52°C for 40 s and elongation at 72°C for 40 s and finally a 5 min extension at 72°C. The amplified PCR products were resolved on 1% agarose gel. The specificity of the primer pairs were confirmed by using known strains. ETEC H10407 was used for confirmation of the toxins, CFA/I and all the NCVFs amplicons. Strains used in our earlier studies (Ghosal *et al.* 2007; Sabui *et al.* 2010) were used for confirmation of all other classical CFs except CS21. All the amplicons were confirmed by sequencing using an automated DNA sequencer (ABI 3730 DNA Analyzer; Applied Biosystems, Foster City, CA) during confirmation of the primer sets. CS21 primer pair was confirmed by sequencing the first three amplicons from the strains used in this study.

Statistical analysis

For comparison of two variables, a chi-square test was used; wherever appropriate, Fisher's exact test was also conducted. A $P < 0{\cdot}05$ was considered statistically significant.

Results

ETEC isolates of Kolkata

In this study, we used 350 archived samples of ETEC out of which 171 were isolated as sole pathogen and 179 strains were isolated as mixed pathogens (Table S2) from

Table 1 List of primers used in this study

Virulence factor	Target gene	Primer name	Oligonucleotide sequences (5′ → 3′)	Product size (bp)	Gene bank accession no.
CS1	*cooA*	Forward	TTGACCTTCTGCAATCTGA	324	M58550
		Reverse	CATCTGCATGGATTGAAAG		
CS2	*cotA*	Forward	TAACTGCTAGCGTTGATCC	368	Z47800
		Reverse	ATTAGTTTGCTGGGTGCTT		
CS3	*cstH*	Forward	GGTGGGTGTTTTGACTCTT	264	M35657
		Reverse	TGTTCGTTACCTTCAGTGG		
CS4	*csaB*	Forward	TTTTGCAAGCTGATGGTAG	250	X97493
		Reverse	TCTGCAGGTTCAAAAGTCA		
CS5	*csfA*	Forward	CGGATTGGATATACCGTTT	453	X63411
		Reverse	TCAACAGCAAATGTTACCG		
CS6	*cssA*	Forward	ATCCAGCCTTCTTTTGGTA	321	U04844
		Reverse	ACCAACCATAACCTGATCG		
CS14	*csuA1*	Forward	ACTGTGACAGCCAGTGTTG	394	X97491
		Reverse	AAACGACGCCTTGATACTT		
CS17	*csbA*	Forward	AACCTATTCTTCGGCTTCA	290	X97495
		Reverse	GCGCAGTTCCTTGTGTG		
CS21	*lngA*	Forward	ACAGTCCGCGTAGCAAT	400	EU107092.1
		Reverse	GTAAAACAGTTGTAGAGG		
CFA/I	*cfaB*	Forward	GCTCTGACCACAATGTTTG	364	S73191
		Reverse	TTACACCGGATGCAGAATA		
CS8	*cofA*	Forward	GCCTTCTGGAAGTCATCAT	438	D37957
		Reverse	CTGCCACATACTCCCAG		
Tia	*tia*	Forward	CTCTGGCTGATGAGAGC	172	U20318.1
		Reverse	TCATAGCCCACTGCAAG		
EtpA	*etpA*	Forward	CAGACAGCTACACCAAC	221	AY920525.2
		Reverse	CGATTGAGTCGTCTCAG		
LeoA	*leoA*	Forward	AGGCGAATCTGAAAGGC	315	AF170971.1
		Reverse	CATTCTTCTCGACAAAGG		
TibA	*tibA*	Forward	GTGCTTAACAATACTGGG	396	AF109215.1
		Reverse	TTCCAGCAGCATATTGAC		
EatA	*eatA*	Forward	CTGTAAATGGCGCTTATC	465	AY163491.2
		Reverse	TTAATGTTCCCACTCCTG		
LT	*eltB*	Forward	CACACGGAGCTCCTCAG	324	M17874.1
		Reverse	CAAACTAGTTTTCCATACTG		
STh	*estA*	Forward	GCTAAACCAGTAGAGTC	149	M34916.1
		Reverse	CACCCGGTACAAGCAGG		

hospital admitted patients during 2008–2014. All the strains were positive for toxin genes, LT and STh; STp was not included in this study. ETEC harbouring *est*+*elt* genes together was most common (61%), followed by ETEC with *est* gene (25%) (Fig. 1) and *elt* gene (14%). We also analysed 35 asymptomatic ETEC comparator strains isolated from nondiarrhoeic community members for the presence of toxin genes. Among these strains, we found 60% strains with *elt* followed by 26% ETEC with *est* and 14% ETEC isolates with *est*+*elt genes*. The difference in the proposition of clinical strains with *est*+*elt* gene were significantly higher compared to such dual presence in comparator strains ($P < 0{\cdot}01$).

Table 2 Distribution of virulence factors in ETEC age-wise

ETEC strains*	Age group ≤5 year (*n* = 174)	>5 year (*n* = 176)	*P* Value†
elt gene	20	29	0·210
est gene	41	48	
est+*elt* gene	113	99	
At least one VF gene present‡	138	112	0·001
No VF gene present	36	64	
Classical CF§ present	110	86	0·007
Classical CF absent	64	90	
Nonclassical VF¶ present	115	90	0·005
Nonclassical VF absent	59	86	

*ETEC isolated from diarrhoeal patients.
†Chi square or Fisher's exact was conducted as appropriate.
‡Classical colonization factor (CF) genes or nonclassical virulence factor (VF) genes.
§Classical CF genes studied here by PCR. Please see Materials and methods for details.
¶Nonclassical VF genes studied here by PCR. Please see Materials and methods for details.

Toxin genes in ETEC isolates

Distribution of 350 ETEC strains in different age groups was examined. We found that of 350 ETEC isolates, 174 (49·71%) strains were from children ≤5 years of age and the rest 176 (50·28%) strains were from patients who were above 5 years of age (Table 2 and Fig. 2). The *est*+*elt* ETEC strains were identified in more abundance than isolates with only *elt* or *est* in both the age groups. The *est*+*elt* containing ETEC strains were 65 and 56% in age groups ≤5 and >5, respectively. The *elt* harbouring strains were less in age group ≤5 (11%) compared to the age group >5 (17%). The proportion of *est* containing strains were although in higher abundance than *elt,* but similar in both lower and higher age groups (24 and 27% respectively).

Toxin genes identified over time in ETEC isolates

ETEC strains isolated pertaining the period (2008–2014) are listed in Table 3. It can be noted that in 2011, only 10 ETEC strains were isolated from patients. In the year 2008, the strains positive for both the toxin genes (*est* and *elt*) were nearly double than strains positive for either the *est* or the *elt* gene. During the years 2009 and 2011, the number of est positive was higher than the strains positive for the *est*+*elt* or *elt*. In the years 2012 and 2013, the number of *elt* and *est* positive strains were negligible and *elt*+*est* strains dominate. In the year 2014, the strains positive for only *est* and only *elt* toxin again emerged. However, from 2012 onwards, *est*+*elt* harbouring strains were identified in more abundances than the other toxin genes. In 2012, the

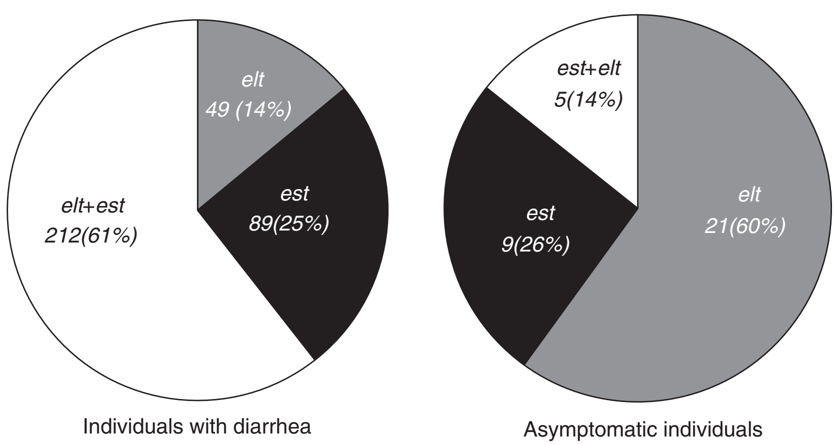

Figure 1 Distribution of ETEC toxins during the study period (2008–2014). ETEC clinical strains were treated as case when isolated from individuals with diarrhoea (*n* = 350) and samples from asymptomatic individuals were used as control (*n* = 35). *elt represents* strains containing only the heat labile toxin; *est* represents strains with heat stable toxin only; and *elt*+*est* represents the strains that contain both the toxins. (□) *elt*+*est*; (■) *elt*; (■) *est*.

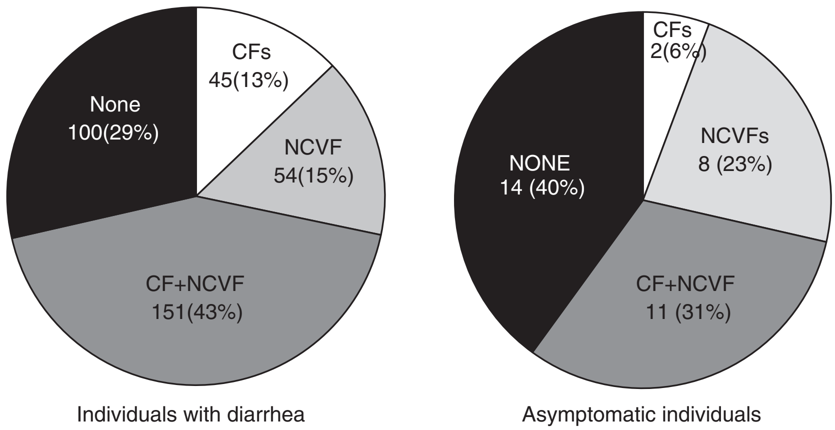

Figure 2 Distribution of ETEC virulence factors during the study period (2008–2014). Clinical isolates from individuals with diarrhoea were taken as case (n = 350) and isolates from asymptomatic individuals were used as control (n = 35). Here, CFs – classical colonization factors; NCVFs – non classical virulence factors. CFs denotes the isolates positive only for the classical colonization factors; NCVFs represents strains positive only for the nonclassical virulence factors; CF+NCVF represents strains positive for both the virulence factors; and NONE represents strains negative for both the virulence factors. We tested common virulence factors and the results shown are based on the screened factors. (□) CFs; (■) NCVF; (■) CF+NCVF; (■) None.

Table 3 Distribution of virulence factors in ETEC year-wise

ETEC strains*	2008 (n = 34)	2009 (n = 60)	2010 (n = 58)	2011 (n = 10)	2012 (n = 58)	2013 (n = 45)	2014 (n = 85)	P value†
elt gene	9	12	12	2	0	1	13	<0·05
est gene	9	26	25	5	6	0	18	
est+*elt* gene	16	22	21	3	52	44	54	
At least one VF gene present‡	24	37	49	9	40	32	59	0·130
No VF gene present	10	23	9	1	18	13	26	
Classical CF present§	17	31	38	7	34	24	45	0·618
Classical CF absent	17	29	20	3	24	21	40	
Nonclassical VF¶ present	19	31	40	7	31	25	52	0·492
Nonclassical VF absent	15	29	18	3	27	20	33	

*ETEC isolated from diarrhoeal patients.
†Chi-squared or Fisher's exact was conducted as appropriate.
‡Classical colonization factor (CF) genes or nonclassical virulence factor (VF) genes.
§Classical CF genes studied here by PCR; please see Materials and methods for details.
¶Nonclassical VF genes studied here by PCR; please see Materials and methods for details.

presence of dual toxin genes in the ETEC strains was 89%, in 2013 about 100% and in 2014 it was 64% of the total isolates. We did not find any *elt*-containing strains during 2012 and only one *elt* positive strain in 2013. The *est*-containing strains were in low abundance in 2012 and no *est* containing strain in 2013.

CF genes in ETEC isolates

Among the 11 CFs analysed in this study for detection, we found that CS21 and CS6 (Fig. 3) were the most common CF harboured. The other common CF genes were CS5 and CFA/I. Seventy two ETEC strains (37%) had CS21, out of which only 20 (10%) strains harboured this gene alone. CS21 was mainly associated with CFA/I, CS17, CS1 and CS3. There were 71 (36%) strains harbouring CS6, of which 23 (12%) strains had CS6 alone. CS5 was mainly present in CS6-positive ETEC strains (92%), showing higher co-occurrence of two genes in clinical ETEC isolates (P < 0·01). However, we found only CS5 genes in four ETEC isolates. In addition, only eight strains had other CF genes along with CS6+CS5. CFA/I, a CF prevalent in other parts of the world was not as frequent as CS6 and CS21 in our settings 36 (18%). CFs detected in very low proportions were CS4 (3%) and CFA/III (2%).

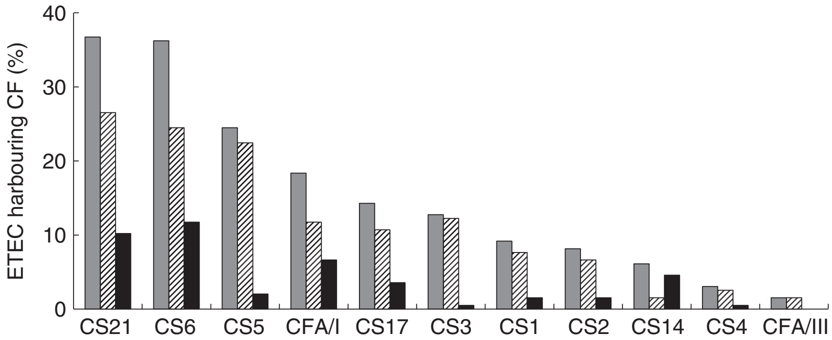

Figure 3 Distribution of colonization factors (CFs) among the ETEC isolates during the study period (2008–2014). There were 196 strains that showed presence of one or more CFs. CS21 and CS6 were among the most common CFs detected in our settings. Detection was done by using multiplex PCR amplification of genomic DNA isolated from ETEC strains as described in materials and methods. Here, 'total' mean the total number of the strains where that gene was present; with 'others' represents that isolates that have that particular gene and other genes along with it; and 'alone' represents the strains having only that particular gene. (▒) total (%); (▨) with others (%); (■) alone (%).

CF genes in ETEC isolates over age

Of 196 ETEC isolates positive for CFs, 110 (56%) strains containing CFs were from children below 5 years of age and 86 (44%) strains were from patients above 5 years of age. The CS21 (38%) and CS6 (37%) harbouring strains were identified more than other CF genes in the age group ≤5 years. CS6 and CS21 (40%) were also commonly identified in age group >5 years. CS5 was the third abundant CF; 25% in lower and 23% in higher age group. CS4, CS14 and CFA/III were detected in low abundance in both the age groups. CFA/I (13%) was more in age group ≤5 than age group >5 (7%). Age and CF were found to be dependent variables (P = 0·001).

CF genes detected in ETEC isolates over time

CS6 was detected in about 41% isolates in 2008; 35% isolates in 2009; 32% isolates in 2010; 57% isolates in 2011; 24% isolates in 2012; 29% isolates in 2013; 49% isolates in 2014. CS21 was detected more in the years 2010, 2012 and 2013 (34, 43 and 50% strains respectively). Among rarely encountered CF gene CFA/III was detected one in 2008 and two isolates in 2012. Similarly, four CS4-containing strains were detected in 2012 and one in 2009 and 2014. CS6 was found less from the year 2012 onwards, where CS21 was more. On the other hand, we identified more CS6-containing strains than that of CS21-containing strains in other years of study.

NCVF genes in ETEC isolates in the present study

We detected nonclassical VFs in 205 (59%) ETEC strains isolated from patients with diarrhoea. *eatA* (65%) and *etpA* (51%) were the common NCVF harboured by the isolates. Following these, the other common VF genes were *tibA* (18%) and *tia* (16%). The *leoA* gene was the least detected NCVF gene in our setting (3%) (Fig. 4). ETEC strains with *eatA* was mostly detected in isolates that had at least one classical CF (P < 0·01) compared to

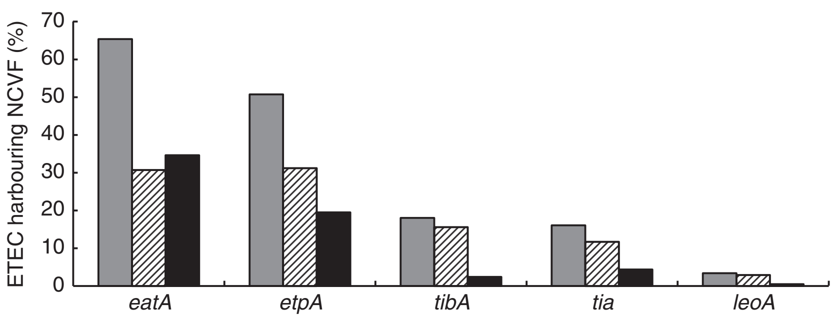

Figure 4 Distribution of nonclassical virulence factors (NCVFs) among the ETEC isolates during the study period (2008–2014). There were 205 strains that harboured NCVF genes. Among these five NCVFs, *eatA* was the prevalent one followed by *etpA*. Detection of NCVFs was performed by multiplex PCR of genomic DNA isolated from ETEC clinical isolates. Here, 'total' mean the total number of the strains where that gene was present; with 'others' represents that isolates that have that particular gene and other genes along with it; and 'alone' represents the strains having only that particular gene. (▒) total (%); (▨) with others (%); (■) alone (%).

other NCVFs. However no such preponderance was observed in case of *etpA* ($P = 0{\cdot}084$). *etpA* was found in 27% of CF negative ETEC strains. There were 19 comparator strains collected from individuals that had the NCVF gene. We did not observe any preponderant distribution between NCVF and clinical strains ($P = 0{\cdot}624$). The detection of NCVFs was observed more in isolates containing toxin genes among the clinical isolates ($P = 0{\cdot}036$).

NCVF genes in ETEC isolates in the current study

One hundred and fifteen (56%) ETEC strains were isolated from children ≤5 years of age and 90 (44%) ETEC isolates were from patients of age group >5 years. *eatA* was the prevalent NCVF in all ages without any age group prediction ($P = 0{\cdot}76$), as was the case for other NCVF ($P = 0{\cdot}361$). *tibA* was identified more in strains isolated from ≤5 age group. Distribution of *tia* was nearly the same in both the age groups. All 7 *leo*-containing ETEC were isolated from patients belonging to <5 years age group.

NCVF genes in ETEC isolates over time

Nonclassical VF *eatA* was more in ETEC isolates in the years 2008 to 2014. *etpA* was the next common one. In 2013 and 2014 both *eatA* and *etpA* were found to be nearly in same number. Nonclassical VF was found to be an independent variable with time (P=0·492).

Correlation of virulence factor genes in ETEC isolates

CF genes and nonclassical VF genes were analyzed for distribution in addition to the toxin genes in 350 ETEC clinical isolates. Of these strains, 250 (72%) ETEC harboured at least one or more classical CF genes and/or NCVF genes. There were 100 (29%) ETEC isolates that did not contain either classical CF or nonclassical VF genes analysed in this study. Of250 strains, 196 (56%) ETEC strains harboured classical CF genes and 205 (59%) ETEC strains had NCVF genes. Altogether, 151 strains (43%) had both CF as well as NCVF genes, 45 (13%) ETEC showed only presence of classical CF, whereas 54 (15%) had the NCVF genes.

In the 35 asymptomatic control strains, 21 (60%) ETEC harboured at least one classical CF genes, nonclassical VF genes or classical and nonclassical factors together. However, 14 (40%) strains did not contain any of these genes. Eleven strains (31%) had both CF as well as nonclassical VF genes, 2 (6%) ETEC had classical CF present, whereas 8 (23%) had the nonclassical VF genes tested in this study. We found an association between the clinical strains and the CFs present ($P = 0{\cdot}033$), whereas the NCVFs did not have any association with the clinical strains ($P = 0{\cdot}624$).

Among the 196 ETEC strains containing the classical CF genes, 23 (12%) strains had *elt* gene, 45 (23%) strains had *est* gene, whereas 128 (65%) were positive for *elt*+*est* genes. Of the 35 asymptomatic control ETEC strains, 13 contained the classical CF genes; four strains (31%) had *elt* gene; four strains (31%) had *elt+est* gene; and 5 (38%) strains had *est* gene. We found an association between the clinical strains and the toxin genes ($P = 0{\cdot}031$) as well as the CF genes ($P = 0{\cdot}031$). However, no association was observed between toxin and CF genes in the clinical isolates ($P = 0{\cdot}114$).

Of the 205 nonclassical VF-containing ETEC strains, 21 (10%) strains had *elt* gene, 58 (28%) strains had *est* gene, whereas 126 (62%) had *elt+est* toxin genes. Of the 100 strains without the classical CF or nonclassical VF genes, 18 (18%), 19 (19%) and 63 (63%) strains had *elt*, *est* and *elt+est* toxin genes, respectively. We observed an association between the clinical strains and the toxin genes ($P = 0{\cdot}036$), although, no association was found with the nonclassical VF genes ($P = 0{\cdot}624$). However, there was also an association between toxin and the nonclassical VF genes in clinical isolates ($P = 0{\cdot}036$). Furthermore, an association was also found between CFs and nonclassical VFs in the clinical isolates ($P < 0{\cdot}01$).

Circulating strains of ETEC in Kolkata during the period of study

The top 5 circulating strains had *est*+*elt*+CS6+CS5+*eatA* (11%); *est*+*etpA* (5%); *elt*+CS6 (3%); *elt*+*est*+*etpA* (3%); *est*+*elt*+CS21+CS3+*etpA* (2%). Although ETEC strains containing CS21 were also common, these strain had wide variety of combinations.

Discussion

Vaccine development against ETEC has been proposed to be an important preventive strategy. Since there are multiple important virulence factors in ETEC, a design of multivalent vaccine may provide wide protection against range of ETEC strains. However, due to presence of various combinations of these virulence factors among ETEC strains rather makes the prevention much more difficult. Construction of multivalent chimeric proteins may also be a strategy to develop vaccines. In both the cases, it is important to know the prevalent virulence factors so that the important factors are targeted for a wide range of protection. Therefore, to gain knowledge about the common circulating strain and to find out the prevalent virulence factors in a geographic region, knowledge of distribution of these factors are necessary.

The distribution of the virulence genes vary in ETEC isolated from one geographic region to another. In this study, we screened for toxigenic genes, CF genes and nonclassical virulence genes. As STh variant is most common than STp, we checked for the STh variant in the present study. We observed the presence of *elt*+*est* ETEC strains as the commonly occurring phenomenon followed by *est*-positive strains and *elt*-containing strains are least in number in the clinical isolates. This information is in support of earlier studies (Dutta *et al.* 2013; Kharat *et al.* 2017). This trend has also been seen by researchers in Bangladesh, indicating a shift from their earlier reports (Begum *et al.* 2014). Our results are in contrast with global prevalence of toxin genes where the presence of LT is much higher than ST either alone or in combination with LT (ST+LT) (Isidean *et al.* 2011). Contrasting in Bangladesh (1996–1998) and Shanghai (2012–2013), ST was the major toxin type and least being ST+LT (Qadri *et al.* 2000; Tang *et al.* 2015). ST is higher than LT in Shenzhen, China (Li *et al.* 2017). In our community based comparator strains, we found that the pattern is reverse i.e., ST+LT was the least prevalent toxin and LT was the predominant one. This enterotoxin distribution pattern of asymptomatic control strains is similar to the results obtained from a cohort study in Egypt (Mansour *et al.* 2014). While in Northern Colombia (Guerra *et al.* 2014) and Guatemala (Torres *et al.* 2015) the trend was LT > ST > ST+LT. Global enteric multicenter study (GEMS) in several developing countries reported that ETEC strains with either ST or together with LT were associated with moderate to severe diarrhoea (Kotloff *et al.* 2013). It should be noted that comparing the small number of control strains with the archived strains is a limitation of this study and to obtain a correct picture, a study with larger number of paired strains should be analysed to understand the geographic distribution of the toxin genes with respect to vaccine development.

Kolkata strains analyzed in the present study had at least one common CF (CS21, CS6, CS5, CFA/I) in 56% of the isolates. In Bangladesh, 56% isolates were positive for the screened CFs (Qadri *et al.* 2000). In rural Egypt, ETEC isolates from a birth cohort of children showed that only 33% isolates were associated with any screened CFAs (Mansour *et al.* 2014). ETEC in children with acute diarrhoea and controls in Teresina/PI, Brazil overall, CF genes were identified in 53·3% (32/60) of the ETEC strains isolated from 19 (67·6%) children, 15 (78·9%) with diarrhoea and four (21·1%) controls (Nunes *et al.* 2011). In Peruvian children effected with ETEC, CFs were identified in 64% of diarrhoeal samples and 37% of control samples (Rivera *et al.* 2010). ETEC isolates from Shenzhen, China had one or multiple CFs in 54% (91/168) samples (Li *et al.* 2017).

In our settings, CS21 and CS6 distributions in ETEC were similar and were the predominant CFs. Our previous findings highlighted the presence of CS6 as the major CF (Ghosal *et al.* 2007; Sabui *et al.* 2012; Dutta *et al.* 2013) as was seen in Guatemala, Egypt and Shenzhen, China (Mansour *et al.* 2014; Torres *et al.* 2015; Li *et al.* 2017). It has also been identified in some study as one of the predominant CF responsible for traveller's diarrhoea (Rivera *et al.* 2013). CS21 was the prevalent CF in regions like North Colombia and traveller's coming back to Spain (Rivera *et al.* 2013; Guerra *et al.* 2014). A recent study describes that CS21 is expressed in high percentage in isolates from India and Latin America (Kharat *et al.* 2017). CS21 was also detected as the prevalent CF in ETEC isolated from Chilean children in 2010 (Del Canto *et al.* 2017). There is striking co-occurrence of CS6 and CS5 in ETEC isolates. We previously reported frequent occurrence of CS6 and CS5 together (Ghosal *et al.* 2007; Sabui *et al.* 2012; Dutta *et al.* 2013). CS4, CS5 and CS6 together forms the CF antigen group IV (CFA/IV). CS6 is known to be expressed along or together with either CS4 or CS5 (Gaastra and Svennerholm 1996; Wolf *et al.* 1997). However, we found only one ETEC strain with CS6+CS4 in the current study. We did not find any significant co-occurrence of CS6 with CS21 or any other CF other than CS5. On the contrary, reports have indicated a strong association between CS21 and CFA/I (Guerra *et al.* 2014).

We screened only the common nonclassical VF genes. Analysis showed the presence of at least one of these factors in 59% of the strains. Among these, the most prevalent was being *eatA* in clinical strains. The *etpA* was in second in rank. It was observed that there was no significant association of *etpA* with classical CF, this may be due to the reason that EtpA interacts with highly conserved regions of flagellin and this interaction promotes intestinal adherence and colonization (Roy *et al.* 2009). However, EatA, being an extracellular protease, was co-identified with the CFs, suggesting that EatA might play an important role in adhesion of the pathogen. ETEC strains isolated from clinical samples and from asymptomatic individuals did not reveal any association of these factors. Similar result was obtained in a study in Bolivia (Gonzales *et al.* 2013). Similar to Bolivian study, we found co-occurrence of the toxin gene profiles in clinical ETEC isolates. The *leoA* gene was detected in low frequencies and in most of the cases where no classical genes were detected. These observations coincide with results from Chile, Colombia and Guatemala (Del Canto *et al.* 2011; Guerra *et al.* 2014; Torres *et al.* 2015).

From these findings it can be concluded that in this region the ETEC strains have same CFs and their association is different than other regions.

Over a third of the samples did not show the presence of either classical or nonclassical virulence genes. Globally there were reports of ETEC isolates negative for CFs and NCVFs. In Bangladesh only 56% of the isolates were positive for the screened CFs (Qadri *et al.* 2000) and 51% strains were negative during 2007–2012 (Begum *et al.* 2014). In Chile, 16% of the ETEC strains were negative for classical and nonclassical adhesins (Del Canto *et al.* 2011). Forty-one percent of the strains were negative for any CFs from strains isolated from Thailand (Puiprom *et al.* 2010). For isolates from North Colombia, 25% were negative for any CF (Guerra *et al.* 2014). Among Iranian children, 33·3% of the ETEC isolates did not express CFs (Nazarian *et al.* 2014). Similarly, 46% samples were CFs negative for isolates from Shenzhen, China (Li *et al.* 2017). In ETEC isolated from travellers 31% of all strains did not possess any of the tested CFs (Rivera *et al.* 2013). In ETEC isolates from indigenous children and international visitors to a rural community in Guatemala, a large percentage analysed for CFs lacked a detectable CF (Torres *et al.* 2015). These results indicated the possibility of the existence of other CFs that are yet to be identified. Studies with ETEC strains that were negative for the known CFs had led to the identifications of new adhesion factors like CS23 (Del canto *et al.* 2012) and recently CS30 (Von Mentzer *et al.* 2017). Genomic analysis of ETEC strains with no known CFs had led to the identification of structures related to CU-pili of γ_2 and κ families and β and π-CU pili (Del Canto *et al.* 2017). These CU-pili are the new addition in the long list of ETEC adhesions. However, these CFs, along the NCVFs, that were included in this study does not fully check for all the known CFs, which is a limitation of the study. Inclusion of these CFs might reduce the number of CF negative strains. These findings suggest that there could be more unknown CFs that are yet to be identified and CF negative strains of this region should be explored more to identify new CFs which will aid in ETEC vaccine development.

As reported previously ETEC is mainly detected in children below 5 years and 5–14 years (Dutta *et al.* 2013). GEMS study was conducted in children <59 months of age (Kotloff *et al.* 2013). However, our results showed that the age groups above 5 years were also affected in this region. This is similar to the results of a surveillance study reported from Shanghai (Tang *et al.* 2015) and Shenzhen, China (Li *et al.* 2017). Though ETEC was mainly associated with children less than 5 years of age but it also affects the adults which are too considered for studies related to ETEC. The age is significantly related to the type of diarrhoea; severe diarrhoea is related to age group >5 years whereas no diarrhoea is related to age group ≤5 years (6·8% *vs* 77%; $P < 0{\cdot}00$). The target for different medication for ETEC infection has to be children but adult population is to be included.

The distribution of the toxin genes were similar during 2008–2011; after this period there was a significant increase in *est*+*elt* containing strains in 2012–2014. On the contrary, we found variations in occurrence of CS6 and CS21 year wise, not following any pattern. Some variation was observed in case of CFA/I as well. However, variation was not observed in the distribution between *eatA* and *etpA* in the yearly ETEC isolates. Therefore, it may be suggested to monitor these variations for vaccine development.

In conclusion, it can be said that the CS6-containing ETEC strains continue to be the common strain. A recognizable pattern can be observed between CS6+CS5 and *eatA* along with *est* gene alone or in combination with *elt*. CS21 has emerged as another predominant CF. However; the CS21-containing strains had diverse CFs and nonclassical virulence genes. To the best of our knowledge there had been no such study based on nonclassical virulence genes conducted in this region.

From our findings and studies in different regions it can be suggested that development of ETEC vaccine has to be region specific. ETEC strains are antigenically diverse displaying different toxins, CFs and NCVFs. The combination of these factors has to be considered for an effective vaccine development.

Our study was based on detection of common CFs and nonclassical virulence genes in clinical isolates archived from hospital based surveillance, which could be seen as a limitation. The availability of fewer number of control community strains may be a drawback of this study. We only checked for the common CFs in this region but advance study with all the known CFs has to be analyzed for a much bigger aspect. However, the pattern of distribution of ETEC virulence genes identified in the eastern state of India could provide useful information which can be used for effective vaccine development in future. Recognizing the predominant pattern and analyzing the co-occurrence of classical and NCVFs may help to track the movement of ETEC around the globe and prioritize the potential vaccine candidate.

Acknowledgements

The authors are grateful to the Indian Council of Medical Research, internal to our organization for the financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Debjyoti Bhakat has been supported by the funds from Indian Council of Medical Research (ICMR file no. – 3/1/3/JRF-2015/HRD-LS/96/40713/57). The authors thank Dr. Samiran Panda for his

contribution in study design, data analysis and manuscript preparation.

Conflict of Interest

The authors have no conflict of interest to declare.

References

Begum, Y.A., Baby, N.I., Faruque, A.S., Jahan, N., Cravioto, A., Svennerholm, A.M. and Qadri, F. (2014) Shift in phenotypic characteristics of enterotoxigenic *Escherichia coli* (ETEC) isolated from diarrheal patients in Bangladesh. *PLoS Negl Trop Dis* **8**, e3031.

Black, R.E., Cousens, S., Johnson, H.L., Lawn, J.E., Rudan, I., Bassani, D.G., Jha, P., Campbell, H. *et al.* and Child Health Epidemiology Reference Group of WHO and UNICEF (2010) Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* **375**, 1969–1987.

Del Canto, F., Valenzuela, P., Cantero, L., Bronstein, J., Blanco, J.E., Blanco, J., Prado, V., Levine, M. *et al.* (2011) Distribution of classical and non-classical virulence genes in enterotoxigenic *Escherichia coli* isolates from Chilean children and tRNA gene screening for putative insertion sites for genomic islands. *J Clin Microbiol* **49**, 3198–3203.

Del Canto, F., Botkin, D.J., Valenzuela, P., Popov, V., Ruiz-Perez, F., Nataro, J.P., Levine, M.M., Stine, O.C. *et al.* (2012) Identification of coli surface antigen 23, a novel adhesin of enterotoxigenic *Escherichia coli*. *Infect Immun* **80**, 2791–2801.

Del Canto, F., O'Ryan, M., Pardol, P., Torres, A., Gutiérrez, D., Cádiz, L., Valdés, R., Mansilla, A. *et al.* (2017) Chaperone-Usher pili loci of colonization factor-negative human enterotoxigenic *Escherichia coli*. *Front Cell Infect Microbiol* **6**, 200.

Dutta, S., Guin, S., Ghosh, S., Pazhani, G.P., Rajendran, K., Bhattacharya, M.K., Takeda, Y., Nair, G.B. *et al.* (2013) Trends in the prevalence of diarrheagenic *Escherichia coli* among hospitalized diarrheal patients in Kolkata, India. *PLoS ONE* **8**, e56068.

Elsinghorst, E.A. and Kopecko, D.J. (1992) Molecular cloning of epithelial cell invasion determinants from enterotoxigenic *Escherichia coli*. *Infect Immun* **60**, 2409–2417.

Elsinghorst, E.A. and Weitz, J.A. (1994) Epithelial cell invasion and adherence directed by the enterotoxigenic *Escherichia coli tib* locus is associated with a 104-kilodalton outer membrane protein. *Infect Immun* **62**, 3463–3471.

Fleckenstein, J.M., Lindler, L.E., Elsinghorst, E.A. and Dale, J.B. (2000) Identification of a gene within a pathogenicity island of enterotoxigenic *Escherichia coli* H10407 required for maximal secretion of the heat-labile enterotoxin. *Infect Immun* **68**, 2766–2774.

Fleckenstein, J.M., Roy, K., Fisher, J.F. and Burkit, M. (2006) Identification of a two-partner secretion locus of enterotoxigenic *Escherichia coli*. *Infect Immun* **74**, 2245–2258.

Gaastra, W. and Svennerholm, A.M. (1996) Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol* **4**, 444–452.

Ghosal, A., Bhowmick, R., Nandy, R.K., Ramamurthy, T. and Chatterjee, N.S. (2007) PCR-based identification of common colonization factor antigens of enterotoxigenic *Escherichia coli*. *J Clin Microbiol* **45**, 3068–3071.

Gonzales, L., Sanchez, S., Zambrana, S., Iñiguez, V., Wiklund, G., Svennerholm, A.M. and Sjöling, Å. (2013) Characterization of enterotoxigenic *Escherichia coli* isolates recovered from children with diarrhea during a 4-year period (2007 to 2010) in Bolivia. *J Clin Microbiol* **51**, 1219–1225.

Guerra, J.A., Romero-Herazo, Y.C., Arzuza, O. and Gómez-Duarte, O.G. (2014) Phenotypic and genotypic characterization of enterotoxigenic *Escherichia coli* clinical isolates from northern Colombia, South America. *Biomed Res Int* **2014**, 236260.

Isidean, S.D., Riddle, M.S., Savarino, S.J. and Porter, C.K. (2011) A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* **29**, 6167–6178.

Kharat, V., Ahmed, M., Jiang, Z.D. and Dupont, H. (2017) Colonization factors in enterotoxigenic *Escherichia coli* (ETEC) isolates from travelers to Mexico, Guatemala and India compared with children in Houston, Texas. *Am J Trop Med Hyg* **96**, 83–87.

Kosek, M., Bern, C. and Guerrant, R.L. (2003) The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ* **81**, 197–204.

Kotloff, K.L., Nataro, J.P., Blackwelder, W.C., Nasrin, D., Farag, T.H., Panchalingam, S., Wu, Y., Sow, S.O. *et al.* (2013) Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* **382**, 209–222.

Li, Y., Luo, Q., Shi, X., Lin, Y., Qiu, Y., Lv, D., Jiang, Y., Chen, Q. *et al.* (2017) Phenotypic and genotypic characterization of clinical enterotoxigenic *Escherichia col* isolates from Shenzhen, China. *Foodborne Pathog Dis* **14**, 333–340.

Luo, Q., Qadri, F., Kansal, R., David, A.R., Sheikh, A. and Fleckenstein, J.M. (2015) Conservation and Immunogenicity of novel antigens in diverse isolates of enterotoxigenic *Escherichia coli*. *PLoS Negl Trop Dis* **9**, e0003446.

Mansour, A., Shaheen, H.I., Amine, M., Hassan, K., Sanders, J.W., Riddle, M.S., Armstrong, A.W., Svennerholm, A.M. *et al.* (2014) Pathogenicity and phenotypic characterization of enterotoxigenic *Escherichia coli* isolates from a birth cohort of children in rural Egypt. *J Clin Microbiol* **52**, 587–591.

Nazarian, S., Gargari, S.L., Rasooli, I., Alerasol, M., Bagheri, S. and Alipoor, S.D. (2014) Prevalent phenotypic and genotypic profile of enterotoxigenic *Escherichia coli* among Iranian children. *Jpn J Infect Dis* **67**, 78–85.

Nunes, M.R., Penna, F.J., Franco, R.T., Mendes, E.N. and Magalhães, P.P. (2011) Enterotoxigenic *Escherichia coli* in children with acute diarrhoea and controls in Teresina/PI, Brazil: distribution of enterotoxin and colonization factor genes. *J Appl Microbiol* **111**, 224–232.

Patel, S.K., Dotson, J., Allen, K.P. and Fleckenstein, J.M. (2004) Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infect Immun* **72**, 1786–1794.

Puiprom, O., Chantaroj, S., Gangnonngiw, W., Okada, K., Honda, T., Taniguchi, T. and Sawanpanyalert, P. (2010) Identification of colonization factors of enterotoxigenic *Escherichia coli* with PCR-based technique. *Epidemiol Infect* **138**, 519–524.

Qadri, F., Das, S.K., Faruque, A.S., Fuchs, G.J., Albert, M.J., Sack, R.B. and Svennerholm, A.M. (2000) Prevalence of toxin types and colonization factors in enterotoxigenic *Escherichia coli* isolated during a 2-year period from diarrheal patients in Bangladesh. *J Clin Microbiol* **38**, 27–31.

Qadri, F., Svennerholm, A.M., Faruque, A.S.G. and Sack, R.B. (2005) Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin Microbiol Rev* **18**, 465–483.

Rivera, F.P., Ochoa, T.J., Maves, R.C., Bernal, M., Medina, A.M., Meza, R., Barletta, F., Mercado, E. *et al.* (2010) Genotypic and phenotypic characterization of enterotoxigenic *Escherichia coli* strains isolated from Peruvian children. *J Clin Microbiol* **48**, 3198–3203.

Rivera, F.P., Medina, A.M., Aldasoro, E., Sangil, A., Gascon, J., Ochoa, T.J., Vila, J. and Ruiz, J. (2013) Genotypic characterization of enterotoxigenic *Escherichia coli* strains causing traveler's diarrhea. *J Clin Microbiol* **51**, 633–635.

Roy, K., Hilliard, G.M., Hamilton, D.J., Luo, J., Ostmann, M.M. and Fleckenstein, J.M. (2009) Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature* **457**, 594–598.

Sabui, S., Ghosal, A., Dutta, S., Ghosh, A., Ramamurthy, T., Nataro, J.P., Hamabata, T. and Chatterjee, N.S. (2010) Allelic variation in colonization factor CS6 of enterotoxigenic *Escherichia coli* isolated from patients with acute diarrhoea and controls. *J Med Microbiol* **59**, 770–779.

Sabui, S., Dutta, S., Debnath, A., Ghosh, A., Hamabata, T., Rajendran, K., Ramamurthy, T., Nataro, J.P. *et al.* (2012) Real-time PCR-based mismatch amplification mutation assay for specific detection of CS6-expressing allelic variants of enterotoxigenic *Escherichia coli* and its application in assessing diarrheal cases and asymptomatic controls. *J Clin Microbiol* **50**, 1308–1312.

Tang, H., Li, Y., Zhang, P., Guo, J., Huang, Z., Xu, H., Hou, Q., Wang, C. *et al.* (2015) Surveillance for diarrheagenic *Escherichia coli* in Shanghai, 2012-2013. *Zhonghua Liu Xing Bing Xue Za Zhi* **36**, 1263–1268.

Torres, O.R., González, W., Lemus, O., Pratdesaba, R.A., Matute, J.A., Wiklund, G., Sack, D.A., Bourgeois, A.L. *et al.* (2015) Toxins and virulence factors of enterotoxigenic *Escherichia coli* associated with strains isolated from indigenous children and international visitors to a rural community in Guatemala. *Epidemiol Infect* **143**, 1662–1671.

Von Mentzer, A., Tobias, J., Wiklund, G., Nordqvist, S., Aslett, M., Dougan, G., Sjöling, Å. and Svennerholm, A.M. (2017) Identification and characterization of the novel colonization factor CS30 based on whole genome sequencing in enterotoxigenic *Escherichia coli* (ETEC). *Sci Rep* **7**, 12514.

WHO. (2006) Future directions for research on enterotoxigenic *Escherichia coli* vaccines for developing countries. *Wkly Epidemiol Rec* **81**, 97–104.

Wolf, M.K., De Haan, L.A., Cassels, F.J., Willshaw, G.A., Warren, R., Boedeker, E.C. and Gaastra, W. (1997) The CS6 colonization factor of human enterotoxigenic Escherichia coli contains two heterologous major subunits. *FEMS Microbiol Lett* **148**, 35–42.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Table S1. Diarrhoea type in relation with the age of patients.

Table S2. Distribution of different enteric pathogens as mixed infections with ETEC in our study.

Conclusion: HPAI H5 viruses bearing the N6 subtype neuraminidase and mammalian adaptation markers in the polymerase genes are capable of causing lethal infection in ferrets as well as transmitting to uninfected animals in a direct contact setting. The heterogeneity in mammalian virulence among H5N6 viruses underscores the necessity of in vivo evaluation of H5Nx viruses as they continue to evolve and pose a threat to human health.

https://doi.org/10.1016/j.ijid.2020.09.045

0004

Correlation of host immune response to internalized and extracellular bacterial strains of streptococcus pneumoniae of various serotypes

N. Palanisamy [1,*], F. Mohd Nor [1], E. Abdul Rahim [2], Z. Mohd-Zain [1], J. Hussaini [1], E.H. Wong [3], A. Adnan [1]

[1] *Universiti Teknologi MARA, Faculty of Medicine, Sg. Buloh, Selangor, Malaysia*
[2] *Universiti Teknologi MARA, Institute of Medical Molecular and Biotechnology, Selangor, Malaysia*
[3] *Taylor's University, School of Medicine, Bandar sunway, Malyasia*

Background: Innate and adaptive immune response against pneumococcal infection is important for bacterial clearance. Pneumococcal cell wall plays an important role in host pathogen interaction. On the host, Toll-like receptors (TLRs) identifies the molecular components on pathogens which activates the host immune responses. It also triggers other pathogen specific responses and signalling molecules to inhibit internalization and colonization of this organism. The aim of this study was to investigate the regulation of genes encoding adherence and internalization to host immune response.

Methods and materials: *S. pneumoniae* strains of six serotypes (1, 3, 5, 19F, 23F, 14) were used. Their extracted cell walls were challenged against A549 human lung epithelial cell line. Expressions of 84 genes associated to host immune response were performed by RT2 Profiler PCR Array. Fold differences in gene expression were determined using the $2^{-\Delta\Delta Ct}$ method. Bacterial adherence was also measured simultaneously upon infection and expression of neuraminidases (NanA and NanB) were quantified.

Results: Bacterial adhesion assay showed strain of serotype 14 had the highest capability of adherence while serotype 23F has the weakest ability to adhere. In comparison to the host immune response of these strains, it was observed that strains of serotype 23F had significant upregulation of pro-inflammatory cytokines; CCL2,IL-8,IL-1B,IL-6, and CLEC4E. Other downstream TLR signalling molecules, IRAK1,IRAK2 were also up-regulated in this strain. Generally, regulation of the TLRs, pro-inflammatory cytokines (IL1A/B IL6, IL8), other downstream signalling molecules and adaptors showed serotypes 1, 3, 5 and 23F to be upregulated whilst serotype 14 showed downregulated expression.

Conclusion: The adherence assay reflects the bacteria found on the cell surface and not the internalized bacteria. It suggest that cell wall of *serotype* 23F induces greater pro-inflammatory host response as compared to cell wall of serotype 14. This finding suggests that serotype 23F internalize better compared to serotype 14, thus evoke a stronger signalling cascade as compared to serotype 14. This data suggest serotype 23F to be more invasive compared to serotype 14. This is supported with the increased expression of neuraminidase NanA in serotype 23F.

https://doi.org/10.1016/j.ijid.2020.09.046

0005

High virulence of ST 238 *Leptospira* interrogans isolated from small mammal captured in human leptospirosis suspected area in Selangor, Malaysia

N. Philip [*], N.N. Azhari, Z. Sekawi, V.K. Neela

Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, SELANGOR/MY

Background: Leptospirosis is a zoonotic and re-emerging disease caused by the *Leptospira* spp. Human leptospirosis presents with a wide range of symptoms and often being misdiagnosed with other diseases. It is well-known that the broad clinical presentations and severity of leptospirosis depend on the individual's age and immune status. However, the *Leptospira* species also contribute to the vast variation in clinical presentations. Hence, this study aimed to evaluate the virulence of three pathogenic (ST 238 *L. interrogans,* ST 242 *L. weillii,* ST 143 *L. borgepetersenii*) and six novel *Leptospira* spp. (*L. semungkisensis, L. fletcheri, L. congkakensis, L. jelokensis, L. perdikensis, L. langatensis*) isolated from small mammals and environments respectively in Malaysia in the hamster model.

Methods and materials: A four to six weeks hamsters were intraperitoneally injected with leptospires and observed up to 14 days for environment isolates and up to 21 days for small mammal isolates. Any moribund animals and the surviving animals were humanely euthanized and the kidney, liver and lungs were harvested for detection of leptospires and histopathology examination.

Results: ST 238 *L. interrogans* showed high virulence as four animals infected with this isolate died within the seven days of infection and gross finding showed haemorrhage in the lung. ST 242 *L. weillii* induced a chronic infection in the kidney as all animals survived up to 21st days though all animals infected with this isolate showed inactivity and weight loss after 6th days of infection. Animals infected with ST 143 *L. borgepetersenii* and the six environmental isolates showed no clinical symptoms.

Conclusion: Altogether, these data showed that local isolate of *L. interrogans* is highly pathogenic and as this strain was isolated from small mammal in human leptospirosis suspected area, a precautionary measure should be considered when coming in contact with this rodent and when visiting the place. The three pathogenic *Leptospira* spp. induced different level of virulence in the hamster model and future study through comparative genomic is needed to infer the virulence determinants of each isolate.

https://doi.org/10.1016/j.ijid.2020.09.047

0006

EatA,a non-classical virulence factor,of Enterotoxigenic *Escherichia coli* (ETEC) is modulated by the host factors during pathogenesis

D. Bhakat [1,*], I. Mondal [2], N.S. Chatterjee [2]

[1] *ICMR-National Institute of Cholera and Enteric Diseases, Biochemistry, Kolkata, West Bengal, India*
[2] *ICMR-National Institute of Cholera and Enteric Diseases, Biochemistry, Kolkata, India*

Background: Enterotoxigenic *Escherichia coli* (ETEC) is one of the leading causes of watery diarrhoea in developing countries, particularly among children less than five years and in travellers. Transmission is caused by contaminated food and water and is established in the small intestine via colonization factors (CFs).Recently it is shown that the pathogenicity of ETEC also depends on the presence of additional virulence factors termed as non-classical virulence factors (NCVF). One of the most common NCVF isolated world-wide is EatA. EatA, an auto transporter, is important for ETEC to adhere to epithelial cells and takes part in intestinal colonization by digesting EtpA, secreted by ETEC. During infection pathogenic bacteria use specific host factors to modulate virulence and stress responses. In the continuing fight against diarrhoeal diseases, studies on ecological niches and their influence on survival and virulence of ETEC are important to generate new knowledge to inhibit or control the disease.

Methods and materials: Well-characterized ETEC strain harboring *est* and *elt* genes along with non-classical virulence, EatA was primarily used for this study. For the experiments, ETEC strain was grown in minimal media (M9) supplemented with 0.2% glucose; relevant additives were supplemented when required. RNA was isolated by using trizol and qRT-PCR was done to check the expression.

Results: Our experiment showed that individual bile component sodium chenodeoxycholate, sodium cholatehydrate, sodium taurodeoxycholate modulates EatA expression and was upregulated by 19-fold, 10-fold and 7-fold respectively. EatA expression was most at 0.0125% mucin concentration. EatA expression was positively upregulated by 40-fold when supplemented with 0.2 mM iron salt. In presence of NaCl, expression of EatA is upregulatedby 3-fold. EatA expression is downregulated in presence of glucose. EatA expression is most at pH 6. Maximal expression of EatA was seen at 37 °C. In rabbit model,the expression of EatA was most at the highest cell density ($1x10^8$ cells/ml). In in-vivo condition lower cell density did not produce increased fluid.

Conclusion: EatA expression depends on the host's intestinal factors. Understanding the pathogenesis of ETEC in different intestinal conditions generate important information that can be exploited towards developing methods of controlling infection and a strategy for vaccine development.

https://doi.org/10.1016/j.ijid.2020.09.048

0007

Cytopathological changes associated with the experimental inoculation of Brucella Abortus in laboratory animal (*cavia porcellus*)

W. Ahmad [1,*], K. Rafique [2], T. Jamil [3], Q. Abbas [3], I. Khan [3], M. Younus [1]

[1] *University of Veterinary and Animal Sciences, Narowal Campus, Department of Pathobiology, Narowal/PK*
[2] *University of Veterinary and Animal Sciences, Jhang campus, Department of Pathobiology, Jhang/PK*
[3] *University of Veterinary and Animal Sciences, Jhang campus, Department of Clinical Sciences, Jhang/PK*

Background: Brucellosis is a zoonotic disease caused by Brucella Abortus which is a gram negative facultative intracellular pathogen, affecting both humans and animals. Transmission of bacteria has usually been through occupational hazards as well as through oral route where it lodges in the mucosa, beneath the sub mucosa these bacteria are ingested by phagocytes. It has been observed that the disease induces hepatomegaly and measurable splenomegaly related to accelerated lymphohistiocytic cells inside the spleen with predominant increases in the share of splenic macrophages, but it may also produce other degenerative and necrotic changes in other vital organs and lymph nodes. The intracellular life of Brucella limits touch to the host innate and adaptive immune responses, protects the organism from the effects of some antibiotics, and drives the only features of pathology in infected hosts which is normally separated into three wonderful levels; the incubation, acute and the extension phase.

Methods and materials: The study comprised of 3 male, 3 female and 3 control animals. Among these, the male and female animals were inoculated with doses of inoculum 1×10^8 CFU, 3×10^7 CFU and 1×10^7 CFU respectively. The reason of different doses was meant to check out extent of pathological lesions in the target laboratory animal. The bacterial strain isolate was obtained from the spleens of inoculated mice cultured in broth C (BBLTM, USA).

Results: Grossly, lymph nodes and mesenteric network were severely hyperemic and congested while, heart, liver, spleen, kidneys, lungs and heart were also congested, enlarged with pale mucous fluid present all over the dorsal surface. Intestines were swollen while strauss reaction and orchitis and anemic uterus were observed in male and female animals respectively. Microscopically, sloughing of germinal epithelium of seminiferous tubules, and infiltration of inflammatory cells were observed in epidydimus. Emphysematous alveoli and thickened alveolar walls in lungs, increased bowman's space in kidneys, fragmentation in heart muscle fibers, lymphohistiocytic cells in spleen and vacuolation changes were observed in these cells.

Conclusion: The bacterium also induces specific gross and microscopic changes in all body organs apart from liver and spleen which were considered as the main predilection site of the bacteria.

https://doi.org/10.1016/j.ijid.2020.09.049

0008

Effect of Haskap (*Lonicera caerulea*) on *streptococcus pneumoniae* infected aged-mouse

M. Minami [1,*], M. Nakamura [2]

[1] *Nagoya City University, Bacteriology, Nagoya, Japan*
[2] *Nakamura pharmacy, Sapporo, Japan*

Background: Pneumonia is a high fatal disease in the world today. As *streptococcus pneumoniae* is most popular in bacterial pneumonia, improvement of pneumococcal pneumonia is considered to be one of the important factors in the longevity of the elderly. Haskap (*Lonicera caerulea*) is a fruit of a plant which has been eaten by the Ainu people (indigenous people in Hokkaido: northern part of Japan) as a fruit of longevity. However, there have been no scientific evidences of the Haskap as longevity so far. In this study, we examined whether Haskap fruit extract (HKP) is effective in aged-mouse infected with *S. pneumoniae*.

Methods and materials: 72 weeks old BALB / c female mice were orally administered HKP. On the second day of oral administration, mice were infected with *S. pneumoniae* ATCC 49619 strain intranasally. After that murine general condition such as survival rate were evaluated until 5 days. Lung tissue was aseptically isolated from mice at the time of death or 5 days after oral administration, and the number of colonies was counted after culturing on a blood agar medium at 37 °C under 5% CO_2 for 24 hours.Further, alveolar macrophage were collected from murine lung tissue treated with the HKP for 4 days. As a phagocytic assay, both alveolar macrophage and *S. pneumoniae* were mixed for 1 hour. Thereafter, the mixed samples were cultured on a blood agar medium, and then the number of colonies was evaluated.

Results: Although all untreated mice died up to 5 days, all HKP treated mice survived. Moreover, in the lung, the number of bacterial colonies from HKP treated mouse was decreased compared to untreated mouse. In the phagocytic assay, the number of bacteria

Received: 28 December 2020 | Revised: 11 June 2021 | Accepted: 29 June 2021

DOI: 10.1111/jam.15206

Journal of Applied Microbiology sfam

ORIGINAL ARTICLE

Distribution of virulence factors and its relatedness towards the antimicrobial response of enterotoxigenic *Escherichia coli* strains isolated from patients in Kolkata, India

Indranil Mondal[1] | **Debjyoti Bhakat**[1] | **Goutam Chowdhury**[2] | **Asis Manna**[3] | **Sandip Samanta**[4] | **Alok Kumar Deb**[5] | **Asish Kumar Mukhopadhyay**[2] | **Nabendu Sekhar Chatterjee**[1]

[1]Division of Biochemistry, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India

[2]Division of Bacteriology, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India

[3]Infectious Diseases and Beliaghata General Hospital, Kolkata, India

[4]Dr. B.C.Roy Post Graduate Institute of Pediatric Sciences, Kolkata, India

[5]Division of Epidemiology, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India

Correspondence
Nabendu Sekhar Chatterjee, Division of Biochemistry, National Institute of Cholera & Enteric Diseases, P33 C.I.T. Road, Scheme XM, Beliaghata, Kolkata, 700 010, India.
Email: nschatterjee@rediffmail.com

Abstract

Aim: Enterotoxigenic *Escherichia coli* (ETEC) is one of the most widely recognized diarrhoeal pathogens in developing countries. The advancement of ETEC vaccine development depends on the antigenic determinants of the ETEC isolates from a particular geographical region. So, the aim here was to comprehend the distribution of virulence determinants of the clinical ETEC strains of this region. Additionally, an attempt was made to find any correlation with the antimicrobial response pattern.

Methods and results: Multiplex PCR was employed to identify virulence determinants followed by confirmatory singleplex PCR. For observation of antibiotic response, the Kirby-Bauer method was used. Out of 379 strains, 46% of strains harboured both the enterotoxins ST and LT, whereas 15% were LT only. Among the major colonization factors (CFs), CS6 (41%) was the most prevalent followed by CFA/I (35%) and CFA/III was the lowest (3%). Among the minor CFs, CS21 (25%) was most prevalent, while CS15 showed the lowest (3%) presence. Among the non-classical virulence factors, EatA (69%) was predominant. ETEC strains harbouring CS6 showed resistance towards the commonly used drug Ciprofloxacin (70%).

Conclusion: CS6 and *elt*+*est* toxin genes co-occurred covering 51% of the isolates. CS21 was found in most strains with *est* genes (43%). EatA was found to occur frequently when ST was present alone or with LT. CS6-harbouring strains showed an independent correlation to antimicrobial resistance.

Significance and Impact of the Study: This study would aid in identifying the commonly circulating ETEC isolates of Kolkata, India, and their prevalent virulence determinants. Knowledge of antibiotic resistance patterns would also help in the appropriate use of antibiotics. Furthermore, the study would aid in identifying the multivalent antigens suitable for region-specific ETEC vaccines with maximum coverage.

KEYWORDS

antimicrobial agents, colonization factor, Enterotoxigenic *Escherichia coli*, prevalence, virulence factors

INTRODUCTION

Enterotoxigenic *Escherichia coli* (ETEC) is an enteric pathogen estimated to cause about 400 million cases of diarrhoea and 157,000 deaths in children <5 years of age. Additionally, the bacteria also cause 44 million cases and 89,000 deaths in older children & adults. ETEC causes 8.5 million Disability Adjusted Life Years (DALYs) or 1 million Years Lived with Disability (YLDs) (Bourgeois, Wierzba & Walker, 2016). ETEC is the leading cause of diarrhoea in developing countries and among travellers from developed countries to these endemic regions (Arduino & Dupont, 1993). In developed nations like the United States, ETEC is acknowledged as a major cause of foodborne disease.

ETEC disease burden estimation of all ages has increased significantly in the number of deaths from 59,200 in 2013 to 74,100 in 2015 (Hosangadi, Smith, Kaslow, et al., 2019). Globally ETEC is detected frequently in Southeast Asian, African, and Middle East regions of the world. In Iraq, ETEC was detected in almost 75% of food items consumed by humans followed by diarrhoea during 2016–2017 (Taha & Yassin, 2019).

Global Enteric Multicentric Study (GEMS) indicated that ETEC was one of the four important pathogens related to diarrhoeal illnesses among children <5 years old in Africa and South Asia (Anderson IV et al., 2019). Although, it was previously thought that ETEC burden was lower, GEMS and MAL-ED molecular re-analysis data indicated the opposite spectrum. ETEC was among the top four pathogens causing diarrhoea-associated deaths among children <5 years old as estimated by CHERG (Hosangadi, Smith, & Giersing, 2019).

In the central region of Kenya, about 6%–7% of all diarrhoeal children were detected positive for ETEC (Mbuthia et al., 2018). In India, according to the global burden of diseases (GBD) 2015, around 6% of all diarrhoeal deaths in children aged less than 5 years were due to ETEC. In this eastern region of India, almost 5% of diarrheagenic *Escherichia coli* (DEC) was ETEC (Dutta et al., 2013). Out of the detected DEC, ETEC was the third most prevalent cause of pathogenesis (about 13.6%) between 2008 and 2012 in Karnataka, India (Singh et al., 2019). Surveillance study suggested that about 4% of clinical cases were ETEC among all diarrhoea infected patients hospitalized in Kolkata (Nair et al., 2010).

Pathogenesis of ETEC causes the release of electrolytes and water during watery diarrhoea. This was due to the release of plasmid-encoded two enterotoxins—heat-labile (LT) and heat-stable (ST). LT, encoded by *elt* gene, is an 84-kDa protein consists of an AB5 multimeric structure, where a pentamer of B chains has a membrane-binding function and the A chain is responsible for toxin activity. Structurally and functionally, LT is the mirror of cholera toxin. ST, encoded by *est* gene, is a non-antigenic low-molecular-weight peptide of 18–19 amino acids. It is of two variants, STp and STh, named after their discovery from porcine and humans respectively (Rao, 1985). ETEC strains produce either ST or LT or both; however, the ratio of different types of toxin-producing ETEC varies in different geographic regions (Qadri et al., 2005).

To release the enterotoxins, the bacteria first attach to the epithelium of the small intestine through different colonization factors (CF), also known as antigenic fimbriae called colonization factor antigens (CFA), a major virulence determinant for initiating pathogenesis (Gaastra & Svennerholm, 1996). CFAs encoded by ETEC are mainly divided into four major groups—CFA/I, CFA/II, CFA/III, and CFA/IV. These CFA groups are further subdivided according to their antigenic properties as coli surface antigens (CS). CFA/I is a single member of the first group. CFA/II group encodes CS3 alone or in combination with CS1 and CS2. CFA/III or CS8 is also a single member of this group. CFA/IV group encodes CS6 alone or in combination with CS4 and CS5. Rigid fimbriae CS1, CS2, CS4, CS5 and CFA/I are ~6 nm in diameter, whereas flexible fibrialle CS3 is ~3 nm in diameter (Vidal et al., 2019). Afimbrial CS6 is an assembly of two subunits, CssA and CssB associated tightly in an equal (1:1) stoichiometry (Sabui et al., 2016). Beside these main CF groups, there also exists minor CF which also helps in adherence of ETEC to the intestinal epithelium. Minor CFs include CS7, CS12, CS13, CS14, CS15, CS17, CS18, CS19, CS20, CS21, CS22, CS23, CS26, CS27, CS28, CS30 and PCF071 (Del Canto et al., 2012; Gaastra & Svennerholm, 1996; Mentzer et al., 2017). With the identification of CS30, there are 23 well-characterized CFs to date (Bhakat et al., 2018). Two more CFs are known in addition to these strains making the total number of CFs 25. The CFs together with the toxins are considered classical virulence factors. Out of the known CFs, only a few are prevalent such as CS6, CS5, CFA/I. The CFs like CS19, CS13, CS18 and CS15 are generally less common globally.

Besides the CFs, studies have indicated that other putative factors playing a role in ETEC pathogenesis, and these are known as 'non-classical virulence factors (NCVFs)'. EatA, EtpA, LeoA, TibA and Tia are the most commonly found NCVF in ETEC strains worldwide. EatA is a serine protease autotransporter that alleviates the delivery of LT toxin (Patel et al., 2004). EtpA is a secreted 170-kDa glycoprotein that plays a role in adherence of the bacteria to the intestinal mucosa (Fleckenstein et al., 2006). LeoA, a cytoplasmic protein, acts to maximize LT secretion and invasion of the epithelial cells by acting as a molecular bridge (Fleckenstein et al., 2000). TibA and Tia are the invasins which are autotransporters. They are predicted to promote adherence and promote the formation of biofilm in the host (Fleckenstein et al., 2010). More than 50% of the strains had EatA and EtpA as their NCVFs isolated from ETEC isolates in Bangladesh between 1998 and 2011 (Luo et al., 2015). In Bolivian children positive for ETEC, ClyA was the most commonly detected NCVF followed by EatA (Gonzales et al., 2013).

Until now, there is no effective licensed efficacious ETEC vaccine available and it allows indiscriminate use of antibiotics to curb the disease. As ETEC is prevalent throughout the globe and its variants are geo-region specific, so it is mandatory to know the disparity in the antibiotic response profile (Medina et al., 2015). About 20%–50% of all ETEC cases in developing countries emerge due to the wide range of antimicrobial resistance among the strains (Ericsson, 2003). Along with this, the emergence of multidrug-resistant ETEC is making the use of antibiotics unavailing against these pathogens (Diemert, 2006). Hence, unwrapping the antibiotic response of clinical ETEC isolates becomes a matter of immense importance.

According to WHO, the development of a practical and effective vaccine against ETEC is of high priority, especially for low- to middle-income countries which can be achieved by including prevalent toxins, CFs and NCVFs harboured by clinical strains. The strategy of developing a vaccine includes attenuated ETEC expressing CFs, multi-epitope fusion antigens, inactivated fimbriated ETEC, ST toxoids. As well as triggering intestinal secretory IgA antibodies to counter CFs and prevent ETEC from colonizing the digestive system is critical for effective immunization (Vidal et al., 2019).

The prevalence of these CFs and NCVFs is geographically based and changes with time. So, information about the prevalence of ETEC virulence determinants over a longer period of time is important for designing an effective ETEC vaccine with broader coverage. Therefore, monitoring the prevalence of various CFs and NCVFs changes with time turns out to be fundamental for defining the multivalent ETEC vaccine that would work over a wide scope of the topographical areas and effective for the long term.

In this study, the aim was to identify the commonly circulating virulence determinants of ETEC isolates from Kolkata, India. The antimicrobial response pattern among the strains having the most predominant virulence determinant was also studied. This study is expected to improve the knowledge about the strains from this region of India and their antimicrobial response pattern that would help to prescribe antibiotics for diarrhoeagenic patients with ETEC infection.

MATERIALS AND METHODS

Isolation and Identification of ETEC

In the hospital-based epidemiological surveillance study at the National Institute of Cholera and Enteric Diseases (NICED), stool specimens were collected from diarrhoeal patients admitted at the Infectious Diseases and Beliaghata General Hospital (ID & BG Hospital, Kolkata) and Dr. B.C. Roy Post Graduate Institute of Pediatric Sciences during 2015–2019. During this period, a total of 379 strains were positive for ETEC using the PCR-based analysis and collected for further investigation from the archive of NICED.

ETEC strains were plated onto MacConkey agar (SRL) plates and incubated for 16–18 h at 37°C. Colonies fermenting lactose on MacConkey agar plate having round and pink morphology were reconfirmed as ETEC and later used for experiments. These strains were further plated on Luria Bertini (LB) agar (Sigma Aldrich) plates. The culture from this nonselective medium was grown in LB and finally stored in 15% glycerol at −80°C for further use.

Extraction of DNA templates

DNA templates were prepared by the boil lysis method. Few colonies of ETEC from agar plate were taken in Tris-EDTA buffer (pH-6.8) and boiled for 10 min followed by immediate cooling in ice. Then the mix was centrifuged and an aliquot of the supernatant was used as a DNA template for PCR.

Detection of virulence determinants by multiplex PCR

For molecular detection of virulence determinants, a total of three primer sets for toxin detection, 24 primer sets for CFs and five primer sets for Non-classical virulence determinants were used (The details of primers are given in Table S1). The PCR reaction was carried in a total volume of 20 μl under the following condition-denaturation at 95°C for 10 min, denaturation at 94°C for 1 min, annealing at 52°C for 30 s, extension at 72°C for 60 s/kb of amplicon and a final extension at 72°C for 10 min. The virulence determinants tested positive in multiplex PCR was confirmed by simplex PCR under the same condition using respective primers for each of the genes.

Antimicrobial susceptibility test

The Kirby–Bauer method was used for bacterial susceptibility towards antimicrobial agents. For the antimicrobial susceptibility test, representative isolates were selected on the basis of having predominant CF. The antibiotic discs (Becton, Dickinson and Company) used are as follows: Imipenem (Ipm)-10 μg; Chloramphenicol (C)-30 μg; Cefapime (Fep)-30 μg; Doxycycline (D)-30 μg; Ceftriaxone (Cro)-30 μg; Tetracycline (Te)-30 μg; Norfloxacin (Nor)-10 μg; Sulfamethoxazole W/ Trim (Sxt); Ciprofloxacin (Cip)-5 μg; Ampicillin (Am)-10 μg; Azithromycin (Azm)-15 μg; Nalidixic Acid (Na)-30 μg; Erythromycin (E)-15 μg; Streptomycin (S)-10 μg.

The test was performed on the Mueller–Hinton (MH) agar plates following Clinical and Laboratory Standards Institute (CLSI) guidelines (Hsueh et al., 2010). The isolates were inoculated in MH Broth and incubated at 37°C up to 0.5 McFarland standard turbidity. The culture was streaked thoroughly onto MH Agar plates using cotton swabs. After air-drying the plates, antibiotic discs were placed and after overnight incubation at 37°C, the zone of inhibition was measured as per the manufacturer's protocols.

Statistical analysis

For comparison of two variables, a chi-square test was used. A $p < 0.05$ was considered statistically significant.

RESULTS

Toxin types of ETEC isolates

Archived isolates of ETEC were analysed in this study. Of 379 ETEC isolates, 38% ($n = 142$) strains were yielded along with other pathogens and 62% ($n = 237$) strains were isolated where ETEC was the sole pathogen. All the 379 strains screened were positive for toxin genes. Among the ETEC isolates, 175 strains harboured both *elt* + esth toxin genes (46%) followed by the strains harbouring *esth* gene alone (37%) and strains harbouring the only *elt* as toxin gene (15%). Only 1% of strains harboured *elt* + estp toxin gene and the frequency of *estp* only strains was also 1%. Only one strain was detected positive for the presence of *elt* + esth + *estp* (Table 1).

Distribution of toxin genes among ETEC isolates identified over time

ETEC strains isolated over 2015–2019 showed that the cases of ETEC infection were highest in 2015 (Table 1). The cases gradually decreased over the next 4 years. In all these years, *elt*+esth was the most frequently detected toxin gene combination among the ETEC isolates followed by *esth* only strains, making the *elt*+esth toxin gene combination the most prevalent toxin type in this region of India. The *elt* gene toxin was detected almost three times more in 2016 ($n = 34$) than in 2015 ($n = 10$). Among the isolates during these 5 years, the presence of *estp*+elt was observed in alternate years, that is, in 2015, 2017 and 2019. The *estp* harbouring strains were also detected in these above-mentioned years only (Table 1).

Age-wise distribution of toxin genes among ETEC isolates

Among the archived isolates, 49% of ETEC strains were isolated from patients younger than 5 years of age and 45% were above the age of 18 years (Table 2). The rest 6% were in between the age of 5–18 years.

When the distribution of toxin genes was studied in these three different age groups, that is, <5 years, 5–18 years, and >18 years of age, it was found that in all age groups, *elt*+est harbouring strains were more than *elt* only or *est* only ETEC isolates (Table 2). In age groups <5 years *elt*+*est* harbouring strains were 43%, in the age group 5–18 years 48% and of isolates from adults it was 51% whereas 18%, 9% and 11% of strains harboured *elt* gene respectively. Four ETEC isolates from children of ≤5 year's age had both *elt*+ *estp* as toxin gene and this combination has not been found in patients of more than 5 years of age. Two strains in the age group 5–18 years were found harbouring e*stp* gene only. Only 1 ETEC isolate harbouring *elt*+ *esth*+ *estp was* found in the patient of 5–18 years of age group. However, when analysed it was observed that distribution of toxin genes is not statistically significantly associated with age groups ($p = 0.357$).

Level of dehydration in comparison with toxin types

When recorded data of ETEC patients were collected on the level of dehydration they had during infection, it was observed

TABLE 1 ETEC distribution over the years with toxin genes

Year	Toxin genes						Total
	elt	*est*	*elt+esth*	*elt+estp*	*estp*	*elt+esth+estp*	
2015	10	36	56	2	2	0	106
2016	34	33	34	0	0	0	101
2017	2	25	38	1	1	1	68
2018	5	22	27	0	0	0	54
2019	5	23	20	1	1	0	50
Total	56	139	175	4	4	1	379

Toxin genes were detected using PCR method. Please see Material and Methods for details.

elt—heat labile toxin gene; *esth*—heat stable toxin gene; human variant; *estp*—heat stable toxin gene; porcine variant.

TABLE 2 Occurrence of ETEC in different age group with toxin genes

Age	Toxin genes			
	elt (%)	*est* (%)	*elt*+*est* (%)	*p* value
0–5 ($n = 185$)	33(17.8)	72(38.4)	80(43.2)	0.357
>5–18 ($n = 23$)	2(8.0)	10(43.4)	11(47.8)	
18+ ($n = 171$)	21(12.2)	61(35.6)	89(52.0)	

Toxin genes were detected using PCR method. Please see Material and Methods for details.

elt—heat labile toxin gene; *est*,—heat stable toxin gene; includes both human (*esth*) and porcine (*estp*) variant.

TABLE 3 Dehydration level associated with toxin genes

Dehydration level	Toxin genes				
	elt	*est*	*elt*+*est*	Total	*p* value
None	14	31	31	76	0.687
Some	37	101	134	272	
Severe	4	11	16	31	

elt—heat labile toxin gene; est—heat stable toxin gene; includes both human (*esth*) and porcine (*estp*) variant.

that 72% of patients with some dehydration were maximum in number (Table 3). Severe dehydration was observed in the least number of patients (8%). The rest of the 20% of patients had not experienced any dehydration during the infection.

When this data of level of dehydration was compared to the type of toxin produced by ETEC strains, it was revealed that among *elt*+est harbouring ETEC strains, 74% were from patients having 'some' level of dehydration followed by 17% patients with no dehydration and only 9% experienced 'severe' level of dehydration. In *elt*, *est* and *elt*+*est* harbouring strains similar patterns of dehydration levels were observed with a maximum number of isolates with some level of dehydration and the least with a 'severe' level of dehydration. Statistical analysis revealed that there is no significant association between the level of dehydration and toxin genes ($p > 0.05$). The absence of association may be due to the small sample size and unavailability of data about if the patients had taken any over the counter drug before admission to the hospital.

Genotypic distribution of ETEC virulence factors

Among the ETEC isolates, both the classical CF and NCVF were present in 69% of the ETEC strains. Only CF genes were detected in 22% of the strains. Six percent of strains showed the presence of NCVF genes alone as their virulence determinants. No virulence determinants could be detected in 3% of the ETEC isolates in this study (Figure 1a).

The distribution of virulence factors over the years followed a similar pattern with isolates harbouring both CF+NCVF as their virulence determinants and least being the isolates without any detectable virulence factors (Figure 1b).

When the distribution of virulence factors in ETEC isolates was compared to the presence of toxin genes in the isolates, it was revealed that among 46% of strains harbouring CFs had both the *elt*+est toxin genes, followed by 40% of strains having *est* toxin gene. Distribution of the toxin genes among NCVF harbouring ETEC isolates revealed that maximum strains (47%) had both the *elt*+est toxin genes (Table 4).

Distribution of major CFs in ETEC isolates

The application of PCR enabled us to detect the presence of different CFs in this study as well as the prevalence of those CFs among ETEC isolates. Eight major CFs were targeted for detection grouped in CFA/I, CFA/II, CFA/III and CFA/IV (Table 5). The CFA/IV group was detected in 43% of the ETEC isolates followed by the CFA/I, and CFA/II group. The CFA/III was detected in the least number of isolates.

Major CF distribution in relation to toxin genes

Within the CFA/IV group, comprising of CFs CS4, CS5 and CS6, the CS6 was maximally detected (Table 6). CS6 harbouring ETEC isolates were detected in 41% of LT+ST strains. Among both LT-only and ST-only strains, CS6 detected almost in equal frequency, 34% and 36% respectively. Alongside CS6, in ST-only isolates, CFA/I was detected in almost the same number (37%) but in LT+ST strains it was detected in only 28% strains compared to 41% strains positive for CS6 (Table 6). Following CS6 and CFA/I, CS5 is another major CF found commonly in 21% CF positive ETEC isolates. The most prevalent combination of Major CF was found to be CS6 with CS5 followed by CFA/I with minor CF CS21. Among the prevalent major CF combination, 61% of CS6+CS5 ETEC isolates were found to harbour *elt*+*est* as their toxin gene followed by ST-only CS6+CS5 isolates (27%) and only 10% of CS6+CS5 isolate harboured *elt* as their toxin

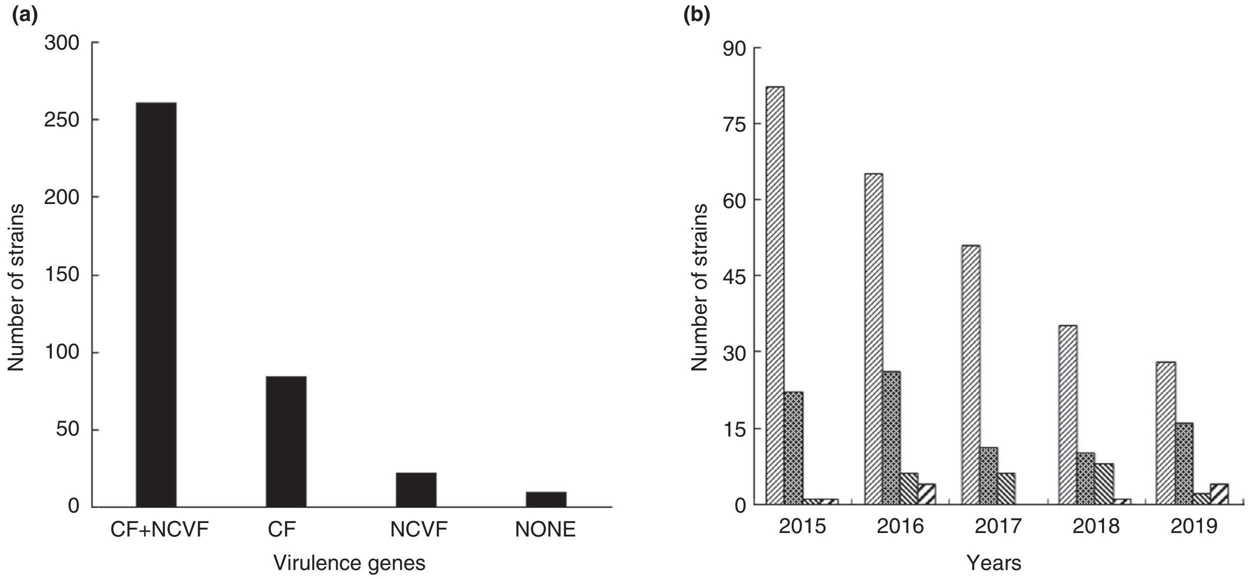

FIGURE 1 Distribution of ETEC virulence factors during the study period (2015–2019). (a) Shows overall distribution of virulence determinants. (b) (▩) CFs; (▧) NCVF; (▨) CF+NCVF; and (◪) None; represents year wise distribution of virulence determinants. A total of 379 clinical samples were screened for CFs—classical colonization factors; and NCVFs— NCVFs. CFs denote the isolates positive only for the classical colonization factors; NCVFs represents strains positive only for the NCVFs; CF+NCVF represents strains positive for both the virulence factors; and NONE represents strains negative for both the virulence factors. We tested a total of 29 virulence factors (24 CFs and 5 NCVFs) and the results shown are based on the screened factors

TABLE 4 ETEC Virulence factors distribution with toxin genes

Virulence factors	Toxin genes			
	elt	*est*	*elt+est*	Total
CF	48	143	163	354
NCVF	31	119	134	284
None	2	2	6	10

est includes both human (*esth*) and porcine (*estp*) variant.

Classical colonization factor (CF) genes studied here by PCR; please see Materials and Methods for details.

Non-classical virulence factor (VF) genes studied here by PCR; please see Materials and Methods for details.

No tested virulence factor detected.

TABLE 5 ETEC CF distribution with toxin genes

Major CF groups	Toxin genes			
	elt	*est*	*elt+est*	Total
I	19	52	51	122
II	6	22	55	83
III	4	3	3	10
IV	23	59	81	163

est includes both human (*esth*) and porcine (*estp*) variant.

Classical colonization factor (CF) genes, Classical CF genes studied here by PCR; please see Materials and Methods for details.

TABLE 6 Prevalent CF distribution with toxin genes

Prevalent major CFs	Toxin genes			
	elt	*Est*	*elt+est*	Total
CS6	19	51	73	143
CFA/I	19	52	51	122
CS5	9	21	45	75

est includes both human (*esth*) and porcine (*estp*) variant.

Classical colonization factor (CF) genes, Classical CF genes studied here by PCR; please see Materials and Methods for details.

TABLE 7 Occurrence of the most prevalent combination of CF with toxin genes

Classical genes	Toxin			
	elt	*est*	*elt+est*	Total
CS6+CS5	7	18	40	65
CFA/I+CS21	3	17	21	41
CS6+CFA/I	3	13	20	36
CS6+CS21	0	8	6	14

est heat stable toxin gene; includes both *esth* and *estp*.

gene. Whereas 55% of CS6+CFA/I isolate harboured *elt+est* as their toxin gene. In CFA/I+CS21 isolates, 51% of strains harboured *elt+est* as prevalent toxin gene (Table 7).

Distribution of major CFs over the years

During 2015–2019, CS6 was detected as the predominant CF over other CFs in every year, for example, 35% isolates in 2015, 44% isolates in 2016, 43% isolates in 2017, 24%

TABLE 8 Major CF distribution over the years

Major CF groups	CF	Years 2015	2016	2017	2018	2019	Total	Percent
CFA/I	CFA/I	32	42	16	17	15	122	34
CFA/II	CS1	5	17	6	3	6	37	10
	CS2	8	5	4	3	4	24	7
	CS3	14	24	12	5	7	62	18
CFA/III	CFA/III	1	3	4	1	1	10	3
CFA/IV	CS4	3	3	2	4	1	13	4
	CS5	22	25	16	7	5	75	21
	CS6	37	44	29	16	17	143	40

Classical colonization factor (CF) genes, Classical CF genes studied here by PCR; please see Materials and methods for details.

isolates in 2018, and 34% isolates in 2019 (Table 8). The detection of CFA/I in the ETEC isolates in these years were after CS6. In 2016, CS6 and CFA/I were detected almost in similar abundance in the ETEC isolates. In 2016 and 2017, CS6 was detected in a higher percentage (44% and 43% respectively) than that of 2015, 2018, and 2019. Among the rarely encountered CFs, CFA/III was detected in 3% of isolates only.

Minor CF among ETEC isolates

Seventy-five percent of the total ETEC strains were positive for harbouring at least one minor CF. Among these, 11%strains harboured *elt* toxin gene only, 42% harboured est toxin gene only and 47% harboured *elt*+est toxin gene. Among the minor CFs, CS21 was accounted for the highest presence (31%). In these strains, *elt*+est positive CS21 was maximum, whereas only three strains had the *elt* gene.

During the 2015–2019 period, CS21 was detected maximally over other minor CFs in every year, for example, 21% isolates in 2015, 19% isolates in 2016, 30% isolates in 2017, 27% isolates in 2018, and 18% isolates in 2019 (Table 9), This was followed by detection of CS23 in these years except in 2019, when no strains were found having CS23. In minor CFs, CS15 was found in the least number of strains found only. Four minor CFs, CS7, CS19, CS13, CS26 and CS30 were not detected in any of the ETEC isolates in our study.

Genotypic distribution of NCVF genes in ETEC isolates in the present study

The non-classical VFs were screened and found to be present in 75% of the ETEC strains during the study period 2015–2019. Out of these positive NCVFs strains, EatA (69%) was most commonly detected followed by EtpA (42%), tibA (18%), and tia (12%) (Table 10). LeoA gene (7%), was least detected NCVF in our study. Most of the NCVF harbouring ETEC isolates were in co-presence with at least one or more classical CFs. In this study, only 8% of the ETEC isolates had NCVF genes alone. EatA was mostly detected (42%) along with prevalent classical CF CS6. Six percent of EatA gene was detected in ETEC isolates having no detectable CFs. Out of all the ETEC isolates, 25% isolates were NCVF negative.

Distribution of NCVF genes in ETEC isolates in relation to toxin genes detected

When analysed in comparison with toxin genes, it was observed that *elt*+est-ETEC isolates harboured maximum NCVF in 47% strains (47%). In *elt* only strains, least number of NCVF (11%) was found. EatA, the predominant NCVF, found maximally in *elt*+est- strains (49%), and *elt* only strains, EatA was minimum (9%). Following EatA, EtpA also showed the same pattern of distribution accounting for half of the EtpA strains (50%) produced LT+ST toxins and almost equal percentage produced ST toxins (Table 10).

Year-wise distribution of NCVF genes among ETEC isolates

During the span of study from 2015 to 2019, EatA was the prevalent non-classical VF detected in our region followed by EtpA as the next common one. EatA was identified in the highest number in 2015, whereas EtpA was most detected in 2016. In 2016, the least common NCVF, LeoA was identified only in one strain. NCVF genes were found to be an independent variable over time (Table 11).

TABLE 9 Minor CF distribution over the years

Minor CFs	Years 2015	2016	2017	2018	2019	Total	Percent
CS21	23	20	21	15	9	88	25
CS23	20	8	6	3	0	37	10
CS20	18	0	10	0	3	31	9
CS28	17	1	3	2	3	26	8
CS27	12	6	0	1	1	20	6
PCF071	4	5	7	3	1	20	6
CS22	13	0	0	1	0	14	4
CS12	8	2	1	0	0	11	3
CS18	2	2	7	0	0	11	3
CS14	7	2	1	0	2	12	3
CS17	6	5	2	0	2	15	4
CS15	4	0	0	3	5	12	3
CS19	0	0	0	0	0	0	0
CS13	0	0	0	0	0	0	0
CS26	0	0	0	0	0	0	0
CS30	0	0	0	0	0	0	0

Classical colonization factor (CF) genes, Classical CF genes studied here by PCR; please see Materials and methods for details.

TABLE 10 Occurrence of ETEC NCVFs with toxin genes

NCVF	Toxin genes *elt*	*est*	*elt+est*	Total
EatA	17	82	97	196
EtpA	10	50	59	119
Tib	5	21	24	50
Tia	10	12	13	35
LeoA	3	8	9	20

est includes both human (*esth)* and porcine *(estp)* variant.

Non-classical virulence factor (VF) genes. Non-classical VF genes studied here by PCR; please see Materials and methods for details.

Antibiotic susceptibility of ETEC isolates

From this study, CS6 was identified as the prevalent CF during 2015–2019. The antibiotic susceptibility in the CS6-harbouring ETEC isolates were tested for antibiotic response against 14 antimicrobial agents (Table 12). All CS6-harbouring isolates were resistant to one or more antimicrobial agents with the most frequent resistance found against 10 μg streptomycin (99.5%) and 15 μg erythromycin (97.7%). More than 80% of strains were resistant to 30 μg Nalidixic Acid and 15 μg Azithromycin. All the ETEC isolates were sensitive to 10 μg Imipenem. The majority of the strains were sensitive to 30 μg Chloramphenicol. When combinations of different antibiotic resistance were analysed compared to the toxin types in CS6-harbouring strains, a very few strains showed a similar pattern of resistance to different antibiotic combinations (Table S3).

To compare the results, we also tested 10 strains, in which no known virulence factors were detected. Among these isolates, almost a similar pattern of response to antimicrobial agents was found (Table 12).

DISCUSSION

Examining the presence of toxin genes in clinical ETEC strains during this study period, it was observed that the *elt+est* harbouring ETEC strains were most common, and *elt* only strains were found in the lowest number. This result is comparable with the findings from studies conducted in Bangladesh (Begum et al., 2014). Findings from a previous study in this region of Kolkata also indicated the same pattern of the presence of toxin genes in ETEC isolates (Bhakat et al., 2018). Global enteric multicenter study (GEMS) on Asia and Africa reported 68% were either ST only or LT/ST ETEC strains (Vidal et al., 2019). In contrast, the global prevalence of *elt* toxin gene is much higher than *est*, either alone or together with *elt* (*est+elt*). A study revealed that LT toxin genes are present in approximately 60% of the field ETEC strains associated with diarrheal incidents in humans, either LT alone (27%) or in combination with ST (33%) (Isidean et al., 2011).

Results presented here trends for the presence of at least one CF are present in 91% of ETEC isolates. This can be

Journal of Applied Microbiology sfam

TABLE 11 NCVF distribution with time

	Years						
NCVF	2015	2016	2017	2018	2019	Total	Percent
EatA	64	46	41	29	16	196	69
EtpA	23	41	24	15	16	119	42
TibA	19	11	13	5	2	50	18
Tia	10	7	7	6	5	35	12
LeoA	12	1	4	2	1	20	7

Non-classical virulence factor (VF) genes. Non-classical VF genes studied here by PCR; please see Materials and methods for details.

TABLE 12 Response of CS6 harbouring ETEC isolates and ETEC isolates having no detectable virulence factors to antimicrobial agents

Antibiotics	Number of CS6 resistant strains (%)	Number of none resistant strains (%)
Imipenem (Ipm)—10 μg	0 (0)	0 (0)
Chloramphenicol (C)—30 μg	8 (6)	2(20)
Cefapime (Fep)—30 μg	34 (24)	3 (30)
Doxycycline (D)—30 μg	39 (27)	2 (20)
Ceftriaxone (Cro)—30 μg	46 (32)	2 (20)
Tetracycline (Te)—30μg	47 (33)	2 (20)
Norfloxacin (Nor)—10μg	51 (36)	5 (50)
Sulfamethoxalone W/Trim (Sxt)	58 (41)	3(30)
Ciprofloxacin (Cip)—5 μg	100 (70)	6 (60)
Ampicillin (Am)—10 μg	107 (75)	8 (80)
Azithromycin (Azm)—15 μg	120 (84)	7 (70)
Nalidixic Acid (Na)—30 μg	124 (87)	10 (100)
Erythromycine (E)—15 μg	140 (98)	10(100)
Streptomycine (S)—10 μg	142 (99)	10 (100)

None refers to those ETEC isolates in which no virulence factors were detected by our tested virulence factors by PCR method.
Antibiotic disks were used (BD, USA) for this study using Kirby–Bauer Disc diffusion method.

compared to the frequencies of 23–94% isolates positive for CF reported in a systematic review done by Isidean et al., 2011. In Bangladesh, 49% of the ETEC isolates were positive for at least one CF (Begum et al., 2014). A study on Nicaraguan children reported that at least 50% of ETEC strains were positive for one CF (Vilchez et al., 2014). In Shenzhen, China, 54% of ETEC isolates had one or more CF present in ETEC strains (Li et al., 2017). The high presence of CF in ETEC isolates in our study (Table S2) was most probably due to the inclusion of 24 CFs in the detection protocol.

During an earlier analysis, findings from Kolkata strains between 2008 and 2014 showed CS21 and CS6 were the predominant CFs (Bhakat et al., 2018). In this study, the distribution of CS6 is predominant followed by CFA/I, rather than CS21. A similar trend in the presence of CS6 was also observed in China and Guatemala (Li et al., 2017; Torres et al., 2014). In the Kolkata region of India, CS6 was mostly detected in the clinical ETEC isolates (Dutta et al., 2013; Ghosal et al., 2007). The CS6 was the predominant CF in ETEC isolates from Zambia (Simuyandi et al., 2019), Guatemala (Torres et al., 2014). A study in Bali, Indonesia indicated CFA/I as the predominant CF (Subekti et al., 2003). Our revelations are also in accordance with GEMS discovery from studies conducted on Asian and African countries where CFA/I as well as CS1-CS6 are some major CF antigens. In rural Egypt, most of the ETEC strains had CFA/I, followed by CS6 as the predominant CF (Shaheen et al., 2004). CFA/I and CS17 are the most common virulence determinant detected in Bolivia (Rodas et al., 2011). In accordance with our findings, a study on the ETEC population in Nepal revealed that CS21 (62.6%) and CS6 (30.2%) were the most prevalent CFs (Margulieux et al., 2018). Another study in our laboratory earlier detected CS21 as the prevalent CF followed by CS6 among clinical ETEC isolates from Kolkata (Bhakat et al., 2018). In Bangladesh, frequently detected CFs were

CS5, CS6 and CS1 (Begum et al., 2014). The co-occurrence of CS5 and CS6 in ETEC strains was found in our study. This trend was also similar to the previous study (Bhakat et al., 2018). This is in accordance with the study that stated CS6 is expressed alone or in combination with CS5 or CS4 (Gaastra & Svennerholm, 1996; Wolf, 1997). This study found only three ETEC strains showing co-occurrence of CS6 and CS4.

In our study, we detected most of the ETEC strains having 1–5 CFs except a few strains having more than the average number of CFs (Table S2). So, there might be a possibility of having mixed strains in archived stock. This might suggest that during stock preparation of archived strains there could be some technical limitations which might give rise to mixed strains.

Analysis with respect to the presence of non-classical virulence determinants showed that three-fourth of the ETEC strains contained at least one or more of these factors. Among these, EatA was the most predominant one followed by EtpA. A study on ETEC clinical isolates of Northern Colombia, South America also indicated EatA as one of the most prevalent NCVF detected (Guerra et al., 2014). EatA was generally identified in strains having the presence of CFs, recommending a significant part to advance intestinal colonization. A similar trend of the presence of EatA and CFs were also found in Chilean ETEC isolates (Del Canto et al., 2011). However, any noteworthy connection of conjunction of EtpA and CFs was not found. This might be because of the cooperation of EtpA with highly conserved flagellin and accordingly elevating adherence to the intestinal wall (Roy et al., 2009). The other three NCVFs Tia, TibA and LeoA were found in low recurrence. Among these, LeoA was found in the most minimal recurrence and two strains with LeoA were negative for any classical CF genes. This outcome repudiates with the examination from Chile, Colombia and Guatemala where the greater part of the LeoA strains were negative for any classical genes (Del Canto et al., 2011; Guerra et al., 2014; Torres et al., 2014).

Among the remaining one-fourth of strains negative for NCVFs, only 10 strains were negative for known CFs. There were only 4% of CF negative strains that were positive for NCVF, that is: Tia, TibA and EtpA. So, despite the inclusion of most of the discovered CFs and NCVFs, we were unable to detect any factor responsible for colonization in 10 strains suggesting that there may be additional colonization determinants yet to be perceived.

Globally, reports of ETEC isolates negative for any virulence factors other than toxins are common. In Bangladesh, 51% of ETEC strains were negative for any CFs (Begum et al., 2014). In Thailand, 41% of ETEC isolates were negative for CFs (Puiprom et al., 2010). Similarly, in 46% of strains, no CFs were detected in Shenzen, China, and 33% in Iranian children (Li et al., 2017; Nazarian et al., 2014). This study identified just 3% of ETEC isolates negative for any virulence determinants. This was possible with the inclusion of more CF in the study design. This is an indication that for maximum coverage on the epidemiological survey we should follow the broad consideration of epidemiologic data about ETEC Classical and Non-Classical virulence determinants. This should be an important aspect in tracking the diversity of ETEC isolates in this region of India and prioritize the choice of the potential vaccine candidate, explicitly for this geographic realm of India.

This study on the antimicrobial response pattern of CS6-harbouring ETEC isolates towards common antimicrobial agents used for the treatment of ETEC-related diarrhoea, demonstrated that almost all these strains were resistant to early-generation drug erythromycin and streptomycin. More than 70% of CS6 harbouring ETEC isolates were ampicillin-resistant which is comparable to maximum resistance to ampicillin in Bolivian children (Rodas et al., 2011). A similar observation was reported in a study on Peruvian children that showed 64% of ETEC strains were resistant to ampicillin (Medina et al., 2015). The results obtained here are in accordance with a previous study in India where ETEC resistance to ciprofloxacin was reported (Chakraborty et al., 2001). In this investigation, it was observed that more than 60% of CS6-harbouring ETEC isolates of this region showed resistance towards nalidixic acid, azithromycin and tetracycline which were also observed in ETEC strains of Bangladesh showing similar resistance to the above antibiotics (Rahman et al., 2020). A study in Eastern Odisha, India revealed that more than 50% of ETEC strains were resistant to ceftriaxone and 2.4% of those isolates were resistant to imipenem (Moharana et al., 2019), whereas our investigation showed only 6% strains were resistant to ceftriaxone and no strains were resistant to imipenem. In this study, the absence of major similarity in resistance towards different combinations of antibiotics shows that there is a huge genetic diversity and acquiring pattern of resistance is also very diverse among the CS6-positive isolates.

WHO recommended using the prevalent CF around the world which includes CFA/I, CS1-CS6 along with LT-B of ETEC for developing vaccine candidates (Bourgeois et al., 2016). From our examination, it may be presumed that strains holding both the enterotoxins LT and ST, alongside classical CF CS6, CS5, CFA/I and CS21 along with NCVF EatA are the probable combination of virulence determinates coursing in this region.

ACKNOWLEDGEMENTS

The authors thank the Indian Council of Medical Research, internal to our organization for the financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Mr. Indranil Mondal has been supported by the Department of Biotechnology, Govt. of India, Grant No. BT/PR21476/MED/29/1135/2016.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

IM, DB and NSC conceptualized the study. IM, DB and GC performed experiments. IM, DB, AKM and NSC prepared the original draft of the manuscript. AM, SS, AKD and AKM collaborated in this study. All authors did data analysis, draft review, editing and approval.

ORCID

Nabendu Sekhar Chatterjee https://orcid.org/0000-0001-7043-0725

REFERENCES

Anderson, J.D., Bagamian, K.H., Muhib, F., Amaya, M.P., Laytner, L.A., Wierzba, T. et al. (2019) Burden of enterotoxigenic *Escherichia coli* and shigella non-fatal diarrhoeal infections in 79 low-income and lower middle-income countries: a modelling analysis. *The Lancet Global Health*, 7(3), e321–e330.

Arduino, R.C. & Dupont, H.L. (1993) Travellers' diarrhoea. *Baillière's Clinical Gastroenterology*, 7(2), 365–385.

Begum, Y.A., Baby, N.I., Faruque, A.S.G., Jahan, N., Cravioto, A., Svennerholm, A.-M. et al. (2014) Shift in phenotypic characteristics of enterotoxigenic *Escherichia coli* (ETEC) isolated from diarrheal patients in Bangladesh. *PLoS Neglected Tropical Diseases*, 8(7), e3031.

Bhakat, D., Debnath, A., Naik, R., Chowdhury, G., Deb, A.K., Mukhopadhyay, A.K. et al. (2018) Identification of common virulence factors present in enterotoxigenic *Escherichia coli* isolated from diarrhoeal patients in Kolkata, India. *Journal of Applied Microbiology*, 126(1), 255–265.

Bourgeois, A.L., Wierzba, T.F. & Walker, R.I. (2016) Status of vaccine research and development for enterotoxigenic *Escherichia coli*. *Vaccine*, 34(26), 2880–2886.

Chakraborty, S., Deokule, J.S., Garg, P., Bhattacharya, S.K., Nandy, R.K., Nair, G.B. et al. (2001) Concomitant infection of enterotoxigenic *Escherichia coli* in an outbreak of cholera caused by *Vibrio cholerae* O1 and O139 in Ahmedabad, India. *Journal of Clinical Microbiology*, 39(9), 3241–3246.

Del Canto, F., Botkin, D.J., Valenzuela, P., Popov, V., Ruiz-Perez, F., Nataro, J.P. et al. (2012) Identification of coli surface antigen 23, a novel adhesin of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, 80(8), 2791–2801.

Del Canto, F., Valenzuela, P., Cantero, L., Bronstein, J., Blanco, J.E., Blanco, J. et al. (2011) Distribution of classical and nonclassical virulence genes in enterotoxigenic *Escherichia coli* Isolates from Chilean Children and tRNA gene screening for putative insertion sites for genomic Islands. *Journal of Clinical Microbiology*, 49(9), 3198–3203.

Diemert, D.J. (2006) Prevention and self-treatment of traveler's diarrhea. *Clinical Microbiology Reviews*, 19(3), 583–594.

Dutta, S., Guin, S., Ghosh, S., Pazhani, G.P., Rajendran, K., Bhattacharya, M.K. et al. (2013) Trends in the prevalence of diarrheagenic *Escherichia coli* among hospitalized diarrheal patients in Kolkata, India. *PLoS One*, 8(2), e56068.

Ericsson, C.D. (2003) Travellers' diarrhoea. *International Journal of Antimicrobial Agents*, 21(2), 116–124.

Fleckenstein, J.M., Hardwidge, P.R., Munson, G.P., Rasko, D.A., Sommerfelt, H. & Steinsland, H. (2010) Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. *Microbes and Infection*, 12(2), 89–98.

Fleckenstein, J.M., Lindler, L.E., Elsinghorst, E.A. & Dale, J.B. (2000) Identification of a gene within a pathogenicity island of enterotoxigenic *Escherichia coli* H10407 required for maximal secretion of the heat-labile enterotoxin. *Infection and Immunity*, 68(5), 2766–2774.

Fleckenstein, J.M., Roy, K., Fischer, J.F. & Burkitt, M. (2006) Identification of a two-partner secretion locus of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, 74(4), 2245–2258.

Gaastra, W. & Svennerholm, A.M. (1996) Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends in Microbiology*, 4(11), 444–452.

Ghosal, A., Bhowmick, R., Nandy, R.K., Ramamurthy, T. & Chatterjee, N.S. (2007) PCR-based identification of common colonization factor antigens of enterotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology*, 45(9), 3068–3071.

Gonzales, L., Sanchez, S., Zambrana, S., Iniguez, V., Wiklund, G., Svennerholm, A.-M. et al. (2013) Molecular characterization of enterotoxigenic *Escherichia coli* isolates recovered from children with diarrhea during a 4-year period (2007 to 2010) in Bolivia. *Journal of Clinical Microbiology*, 51(4), 1219–1225.

Guerra, J.A., Romero-Herazo, Y.C., Arzuza, O. & Gómez-Duarte, O.G. (2014) Phenotypic and genotypic characterization of enterotoxigenic *Escherichia coli* clinical isolates from Northern Colombia, South America. *BioMed Research International*, 2014, 1–11.

Hosangadi, D., Smith, P.G. & Giersing, B.K. (2019) Considerations for using ETEC and Shigella disease burden estimates to guide vaccine development strategy. *Vaccine*, 37(50), 7372–7380.

Hosangadi, D., Smith, P.G., Kaslow, D.C. & Giersing, B.K. (2019) WHO consultation on ETEC and Shigella burden of disease, Geneva, 6–7th April 2017: meeting report. *Vaccine*, 37(50), 7381–7390.

Hsueh, P.R., Ko, W.C., Wu, J.-J., Lu, J.-J., Wang, F.-D., Wu, H.-Y. et al. (2010) Consensus statement on the adherence to Clinical and Laboratory Standards Institute (CLSI) Antimicrobial Susceptibility Testing Guidelines (CLSI-2010 and CLSI-2010-update) for Enterobacteriaceae in Clinical Microbiology Laboratories in Taiwan. *Journal of Microbiology, Immunology, and Infection*, 43(5), 452–455.

Isidean, S.D., Riddle, M.S., Savarino, S.J. & Porter, C.K. (2011) A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine*, 29(37), 6167–6178.

Li, Y., Luo, Q., Shi, X., Lin, Y., Qiu, Y., Lv, D. et al. (2017) Phenotypic and genotypic characterization of clinical Enterotoxigenic *Escherichia coli* isolates from Shenzhen, China. *Foodborne Pathogens and Disease*, 14(6), 333–340.

Luo, Q., Qadri, F., Kansal, R., Rasko, D.A., Sheikh, A. & Fleckenstein, J.M. (2015) Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic *Escherichia coli*. *PLoS Neglected Tropical Diseases*, 9(1), e0003446.

Marguliex, K.R., Srijan, A., Ruekit, S., Nobthai, P., Poramathikul, K., Pandey, P. et al. (2018) Extended-spectrum β-lactamase prevalence and virulence factor characterization of enterotoxigenic *Escherichia coli* responsible for acute diarrhea in Nepal from 2001 to 2016. *Antimicrobial Resistance & Infection Control*, 7(1), 1–7.

Mbuthia, O.W., Mathenge, S.G., Oyaro, M.O. & Ng'ayo, M.O. (2018) Etiology and pathogenicity of bacterial isolates: a cross sectional study among diarrheal children below five years in central regions of Kenya. *Pan African Medical Journal*, 31(1), 88.

Medina, A.M., Rivera, F.P., Pons, M.J., Riveros, M., Gomes, C., Bernal, M. et al. (2015) Comparative analysis of antimicrobial resistance in enterotoxigenic *Escherichia coli* isolates from two paediatric cohort studies in Lima, Peru. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 109(8), 493–502.

Mentzer, A., Tobias, J., Wiklund, G., Nordqvist, S., Aslett, M., Dougan, G. et al. (2017) Identification and characterization of the novel colonization factor CS30 based on whole genome sequencing in enterotoxigenic *Escherichia coli* (ETEC). *Scientific Reports*, 7(1), 1–11.

Moharana, S.S., Panda, R.K., Dash, M., Chayani, N., Bokade, P., Pati, S. et al. (2019) Etiology of childhood diarrhoea among under five children and molecular analysis of antibiotic resistance in isolated enteric bacterial pathogens from a tertiary care hospital, Eastern Odisha, India. *BMC Infectious Diseases*, 19(1), 1–9.

Nair, G., Ramamurthy, T., Bhattacharya, M., Krishnan, T., Ganguly, S., Saha, D. et al. (2010) Emerging trends in the etiology of enteric pathogens as evidenced from an active surveillance of hospitalized diarrhoeal patients in Kolkata, India. *Gut Pathogens*, 2(1), 4.

Nazarian, S., Gargari, S.L.M., Rasooli, I., Alerasol, M., Bagheri, S. & Alipoor, S.D. (2014) Prevalent phenotypic and genotypic profile of enterotoxigenic *Escherichia coli* among Iranian children. *Japanese Journal of Infectious Diseases*, 67(2), 78–85.

Patel, S.K., Dotson, J., Allen, K.P. & Fleckenstein, J.M. (2004) Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, 72(3), 1786–1794.

Puiporm, O., Chantaroj, S., Gangnonngiw, W., Okada, K., Honda, T., Taniguchi, T. et al. (2010) Identification of colonization factors of enterotoxigenic *Escherichia coli* with PCR-based technique. *Epidemiology and Infection*, 138(4), 519–524.

Qadri, F., Svennerholm, A.-M., Faruque, A.S.G. & Sack, R.B. (2005) Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clinical Microbiology Reviews*, 18(3), 465–483.

Rahman, M.M., Ahmed, P., Kar, A., Sakib, N., Shibly, A.Z., Zohora, F.T. et al. (2020) Prevalence, antimicrobial resistance, and pathogenic potential of enterotoxigenic and enteropathogenic *Escherichia coli* associated with acute diarrheal patients in Tangail, Bangladesh. *Foodborne Pathogens and Disease*, 17(7), 434–439.

Rao, M.C. (1985). Toxins which activate guanylate cyclase: heat-stable enterotoxins. Ciba Foundation Symposium 112 – Microbial Toxins and Diarrhoeal Disease, 112, 74–93.

Rodas, C., Mamani, R., Blanco, J., Blanco, J.E., Wiklund, G., Svennerholm, A.-M. et al. (2011) Enterotoxins, colonization factors, serotypes and antimicrobial resistance of enterotoxigenic *Escherichia coli* (ETEC) strains isolated from hospitalized children with diarrhea in Bolivia. *Brazilian Journal of Infectious Diseases*, 15(2), 132–137.

Roy, K., Hamilton, D., Ostmann, M.M. & Fleckenstein, J.M. (2009) Vaccination with EtpA glycoprotein or flagellin protects against colonization with enterotoxigenic *Escherichia coli* in a murine model. *Vaccine*, 27(34), 4601–4608.

Sabui, S., Debnath, A., Ghosal, A., Wajima, T., Hamabata, T., Ramamurthy, T. et al. (2016) Characterization of oligomeric assembly of colonization factor CS6 from enterotoxigenic *Escherichia coli*. *Microbiology*, 162(1), 72–83.

Shaheen, H.I., Khalil, S.B., Rao, M.R., Abu Elyazeed, R., Wierzba, T.F., Peruski, L.F. et al. (2004) Phenotypic profiles of enterotoxigenic *Escherichia coli* associated with early childhood diarrhea in rural Egypt. *Journal of Clinical Microbiology*, 42(12), 5588–5595.

Simuyandi, M., Chilengi, R., B Connor, S., B. Voeglein, J., M. Laban, N., Mwila-Kazimbaya, K. et al. (2019) Enterotoxigenic *Escherichia Coli* toxins and colonization factors among zambian children presenting with moderate to severe diarrhea to selected health facilities. *Archives of Microbiology & Immunology*, 3(4), 173–184.

Singh, P., Metgud, S.C., Roy, S. & Purwar, S. (2019) Evolution of diarrheagenic *Escherichia coli* pathotypes in India. *Journal of Laboratory Physicians*, 11(4), 346–351.

Subekti, D.S., Lesmana, M., Tjaniadi, P., Machpud, N., Sriwati, Sukarma et al. (2003) Prevalence of enterotoxigenic *Escherichia coli* (ETEC) in hospitalized acute diarrhea patients in Denpasar, Bali, Indonesia. *Diagnostic Microbiology and Infectious Disease*, 47(2), 399–405.

Taha, Z.M. & Yassin, N.A. (2019) Prevalence of diarrheagenic *Escherichia coli* in animal products in Duhok province, Iraq. *Iranian Journal of Veterinary Research*, 20(4), 255–262.

Torres, O.R., González, W., Lemus, O., Pratdesaba, R.A., Matute, J.A., Wiklund, G. et al. (2014) Toxins and virulence factors of enterotoxigenic *Escherichia coli* associated with strains isolated from indigenous children and international visitors to a rural community in Guatemala. *Epidemiology and Infection*, 143(8), 1662–1671.

Vidal, R.M., Muhsen, K., Tennant, S.M., Svennerholm, A.-M., Sow, S.O., Sur, D. et al. (2019) Colonization factors among enterotoxigenic *Escherichia coli* isolates from children with moderate-to-severe diarrhea and from matched controls in the Global Enteric Multicenter Study (GEMS). *PLoS Neglected Tropical Diseases*, 13(1), e0007037.

Vilchez, S., Becker-Dreps, S., Amaya, E., Perez, C., Paniagua, M., Reyes, D. et al. (2014) Characterization of enterotoxigenic *Escherichia coli* strains isolated from Nicaraguan children in hospital, primary care and community settings. *Journal of Medical Microbiology*, 63(Pt_5), 729–734.

Wolf, M.K. (1997) Occurrence, distribution, and associations of O and H serogroups, colonization factor antigens, and toxins of enterotoxigenic *Escherichia coli*. *Clinical Microbiology Reviews*, 10(4), 569–584.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Mondal, I., Bhakat, D., Chowdhury, G., Manna, A., Samanta, S., Deb, A.K., et al. (2021) Distribution of virulence factors and its relatedness towards the antimicrobial response of enterotoxigenic *Escherichia coli* strains isolated from patients in Kolkata, India. *Journal of Applied Microbiology*, 00, 1–12. https://doi.org/10.1111/jam.15206

MICROBIOLOGY

RESEARCH ARTICLE
Bhakat *et al*., *Microbiology* 2021;167:001089
DOI 10.1099/mic.0.001089

MICROBIOLOGY SOCIETY

Iron influences the expression of colonization factor CS6 of enterotoxigenic *Escherichia coli*

Debjyoti Bhakat[1], Indranil Mondal[1], Asish Kumar Mukhopadhyay[2] and Nabendu Sekhar Chatterjee[1,*]

Abstract

Enterotoxigenic *Escherichia coli* (ETEC) is a major pathogen of acute watery diarrhoea. The pathogenicity of ETEC is linked to adherence to the small intestine by colonization factors (CFs) and secretion of heat-labile enterotoxin (LT) and/or heat-stable enterotoxin (ST). CS6 is one of the most common CFs in our region and worldwide. Iron availability functions as an environmental cue for enteropathogenic bacteria, signalling arrival within the human host. Therefore, iron could modify the expression of CS6 in the intestine. The objective of this study was to determine the effect of iron availability on CS6 expression in ETEC. This would help in understanding the importance of iron during ETEC pathogenesis. ETEC strain harbouring CS6 was cultured under increasing concentrations of iron salt to assess the effect on CS6 RNA expression by quantitative RT-PCR, protein expression by ELISA, promoter activity by β-galactosidase activity, and epithelial adhesion on HT-29 cells. RNA expression of CS6 was maximum in presence of 0.2 mM iron (II) salt. The expression increased by 50-fold, which also reduced under iron-chelation conditions and an increased iron concentration of 0.4 mM or more. The surface expression of CS6 also increased by 60-fold in presence of 0.2 mM iron. The upregulation of CS6 promoter activity by 25-fold under this experimental condition was in accordance with the induction of CS6 RNA and protein. This increased CS6 expression was independent of ETEC strains. Bacterial adhesion to HT-29 epithelial cells was also enhanced by five-fold in the presence of 0.2 mM iron salt. These findings suggest that CS6 expression is dependent on iron concentration. However, with further increases in iron concentration beyond 0.2 mM CS6 expression is decreased, suggesting that there might be a strong regulatory mechanism for CS6 expression under different iron concentrations.

INTRODUCTION

In developing countries, acute diarrhoea is among the top major pathogenic agents responsible for deaths in children under the age of 5 years and also among all age groups according to the 2015 Global Burden of Disease Study [1]. Enterotoxigenic *E. coli* (ETEC) is the major etiological agent for acute watery diarrhoea in these countries and is estimated to cause 400 million episodes of diarrhoea and approximately 380000 deaths annually among children <5 years [2]. ETEC is also the causative agent for traveller's diarrhoea to people travelling to these endemic areas and is frequently exported to the developed countries [3].

Diarrhoea by ETEC is caused due to the release of the heat-labile (LT) and/or heat-stable (ST) enterotoxin. Both the toxins stimulate chloride secretion causing water and electrolytes to release into the intestinal lumen and subsequently, watery diarrhoea occurs [2]. For the toxins to be released, ETEC must adhere to the small intestine to withstand intestinal peristalsis. Different colonization factors, the major virulence determinant for initiating pathogenesis, mediates the adhesion or colonization of ETEC onto the surface of the small intestine [4]. With the recent discovery of CS30 [5], there are more than 27 different colonization factors, 24 are well-characterized [2, 4]. Among the colonization factors, CS6 is prevalent globally and it is the predominant colonization factor in this region [6–8]. In contrast to the other ETEC CFs that are fimbrial, CS6 is afimbrial. CS6 consists of two tightly-associated subunits CssA and CssB, organised in an equal (1:1) ratio, to form an external array of higher-order oligomers [9].

Received 18 March 2021; Accepted 03 August 2021; Published 22 September 2021
Author affiliations: [1]Division of Biochemistry, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India; [2]Division of Bacteriology, ICMR-National Institute of Cholera and Enteric Diseases, Kolkata, India.
***Correspondence:** Nabendu Sekhar Chatterjee, nschatterjee@rediffmail.com
Keywords: CS6; enterotoxigenic *Escherichia coli*; expression; ETEC; iron.
Abbreviations: CFs, colonization factors; ETEC, enterotoxigenic *Escherichia coli*; LT, heat labile enterotoxin; RT-PCR, real time - polymerase chain reaction; ST, heat stable enterotoxin.
Three supplementary figures are available with the online version of this article.

Table 1. Strains used in the study

ETEC strain	Culture on MacConkey agar	Other microorganism isolated along with ETEC	CF and toxin profile	Year of isolation
ETEC 4266*	*E. coli*	No	CS6; STh +LT	2008
ETEC 0001	*E. coli*	No	CS6; STh +LT	2008
ETEC 0002	*E. coli*	No	CS6; STh +LT	2009
ETEC 0003	*E. coli*	No	CS6; STh +LT	2010
ETEC 0004	*E. coli*	No	CS6; STh +LT	2012
ETEC 0005	*E. coli*	No	CS6; STh +LT	2014
ETEC 0006	*E. coli*	Yes	CS6; STh +LT	2008
ETEC 0007	*E. coli*	Yes	CS6; STh +LT	2009
ETEC 0008	*E. coli*	Yes	CS6; STh +LT	2010
ETEC 0009	*E. coli*	Yes	CS6; STh +LT	2012
ETEC 0010	*E. coli*	Yes	CS6; STh +LT	2014

*Reference strain.

When ETEC infects the host, it experiences an assortment of host environmental factors which may impact virulence gene expression by sensing the proper microenvironment for the appropriate expression of the virulence factors. Iron is one such essential element in the intestinal environment. In pathogenic bacteria the expression of virulence-associated properties is frequently involved with iron availability [10]. It has been demonstrated that the freely available iron increases the bacterial virulence in at least 18 different bacterial species [11].

Worldwide iron deficiency is one of the most prevalent nutritional disorders affecting infants, young children and women in developing countries and causes significant health consequences for example infection [12]. However, studies have also shown that iron supplements had adverse effects causing an exaggeration in enteric infections and fatality in young children [13, 14]. Numerous pathogens in animal models demonstrated increased virulence due to the administration of excess iron [15].

For the survival of a pathogen in the host, its ability to regulate gene expression throughout an infection is vital. Iron securing is vital for bacterial survival, and furthermore assumes a significant function in bacterial pathogenesis. The knowledge about the regulation of colonization factor CS6 in the human gut is limited and in this study, we attempted to determine the impact of iron accessibility on CS6 expression in ETEC. We found that CS6 expression increased up to 0.2 mM in iron concentration and beyond this concentration, CS6 expression decreased. Similar observations were made for ETEC adhesion and growth suggesting that there might be a strong regulatory mechanism for CS6 expression under different iron concentrations. Results obtained here would help in understanding the importance of iron during ETEC pathogenesis.

METHODS

Bacterial strains used

All the ETEC isolates (Table 1) used in this study were isolated from patients with acute watery diarrhoea admitted to the Infectious Diseases Hospital, Kolkata, and Dr. B C Roy Post Graduate Institute of Paediatric Sciences as routine surveillance system. A well-characterized clinical ETEC strain ETEC4266 harbouring ST-LT gene and colonization factors CS6 was primarily used as a wild-type strain in this study [16]. Other ETEC strains were used from the strain archive of Bacteriology as and when required. The strains used in this study were CS6-only expressing strains. The strains were verified for L5 lineage but the strains were CS5 negative. The strains were a STh producing variant. All the bacterial strains were maintained at −80 °C in Luria-Bertani broth (BD Difco, USA) containing 15% (v/v) glycerol as cryoprotectant.

Bacterial growth and additives

For the routine maintenance and growth, strains were grown on MacConkey Agar plates. For the experiments, Minimal Media (M9) (BD Difco, USA), Luria Bertani Broth (BD Difco, USA), and Colonization Factors expressing (CFA) media were used; minimal media (M9) was supplemented with glucose (Merck, Germany) (0.2% w/v) as carbon source. Relevant iron salts, ferrous sulphate heptahydrate (Sigma-Aldrich, USA), ferrous chloride (Sigma-Aldrich, USA), ferric chloride (Sigma-Aldrich, USA), and ferric nitrate nonahydrate (Sigma-Aldrich, USA) as additives were supplemented for the experiments when required. The iron-chelating agent desferrioxamine mesilate salt or deferoxamine (Sigma-Aldrich, USA) was added to the medium when required.

Table 2. Primers used for PCR and Real-time Reverse Transcriptase PCR (rt RT-PCR)

Gene	Direction	Primers (5′→3′)	Product size (in bp)	GenBank accession no.
cssB (CS6)	Forward	CGGTTGAGATGACGATCCCT	90	U04844
	Reverse	GCTTTCCGATCTGCTGTCCA		
parC	Forward	ATCCGCTGGTTGATGGTCAG	92	EU561348
	Reverse	TTTCGACAGACGGGATTCGG		
cssB (CS6)	Forward	ATCCAGCCTTCTTTTGGTA	321	U04844
	Reverse	ACCAACCATAACCTGATCG		
elt (LT)	Forward	CACACGGAGCTCCTCAG	324	M17874
	Reverse	CAAACTAGTTTTCCATACTG		
esth (STh)	Forward	GCTAAACCAGTAGAGTC	149	M34916
	Reverse	CACCCGGTACAAGCAGG		
CS6 promoter	Forward	TCCCCCCGGGATAACTAACTGAAAAACAATG	573	NC022333
	Reverse	GCTCTAGATATAACACCATTACTATTGCT		
fur	Forward	GATACCAGCGTCGTCAAACT	104	NC017633
	Reverse	ACGTCAGTGCGGAAGATTTAT		

RNA extraction

Cells were collected at the exponential phase by centrifugation and each culture processed individually. RNA was isolated using trizol (Invitrogen, Thermo Fisher Scientific) extraction method as per manufacturer's protocol. Briefly, pelleted cells were suspended in trizol reagent followed by the addition of chloroform, and the mixture was centrifuged at 4 °C to allow for phase separation. The aqueous phase containing the RNA was transferred to a new tube. Next isopropanol was added to the aqueous phase and incubated followed by centrifugation. The pellet was resuspended in 75% (v/v) ethanol and then centrifuged and air-dried after that. Then it was resuspended in RNase free water. The RNA was then subjected to Turbo DNase (Ambion) treatments, followed by inactivation of the reagent and precipitation. RNA concentrations were measured using the Optizen UV-Vis spectrophotometer. Agarose gel electrophoresis was done to check RNA quality and absence of contaminating DNA before use in subsequent experiments. All RNA samples were stored at −80 °C.

cDNA synthesis and quantitative reverse transcription PCR

cDNA was synthesized by Verso cDNA Synthesis Kit (Thermo Scientific) as per user manual. Controls for DNA contamination without reverse transcriptase (-RT) were prepared simultaneously with the synthesized cDNA from the same amount of total RNA. cDNA and controls were stored at −20 °C. The real-time RT-PCR assays were performed on ABI 7500 (Applied Biosystems, Foster City, CA) using SYBR Green I (Applied Biosystems) as the detector. Primers for the CS6, structural subunit gene *cssB* and the *E. coli* housekeeping gene p*arC* were designed by NCBI Primer designing software (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Table 2). The relative expression of the target transcripts was calculated according to Livak method [17] using *parC* as an internal control. *cssA* and *cssB* are expressed equally and in equal stoichiometry. Therefore, measuring any one would be representative of CS6 expression. It is *cssB* transcripts that are being measured here in this study. However, it can be noted in case of any defective assembly process, c*ssA* itself is unstable in *E. coli* and that ETEC may possess a unique system to degrade *cssA* when expressed alone. According to Wajima *et al.*, 2011, [18] stable expression of *cssA* is dependent on *cssB* but not vice versa, and *cssA* is more unstable in the ETEC genetic background than in laboratory strains. Therefore, focusing on *cssA* may produce confusing results in different experiments as our strains are clinical isolates. So, we focused on *cssB* for measuring the changes for CS6.

Enzyme-linked immunosorbent assay (ELISA)

The phenotypic expression of CS6 on the bacterial surface was quantified by ELISA assays as per Elder *et al.*, 1982 [19]. To quantitate the surface expression, 10^7 c.f.u. ml^{-1} of ETEC was used for coating and incubated overnight at 4 °C. Unbound bacteria were decanted and wells were washed with PBS followed by blocking in 5% (w/v) non-fat skimmed milk in PBS. After washing, the bound CS6 was determined by anti-CssB polyclonal antibody as the primary antibody followed by HRP-conjugated secondary antibody. Using a standard curve obtained with purified CS6 the amount of protein (µg) was determined.

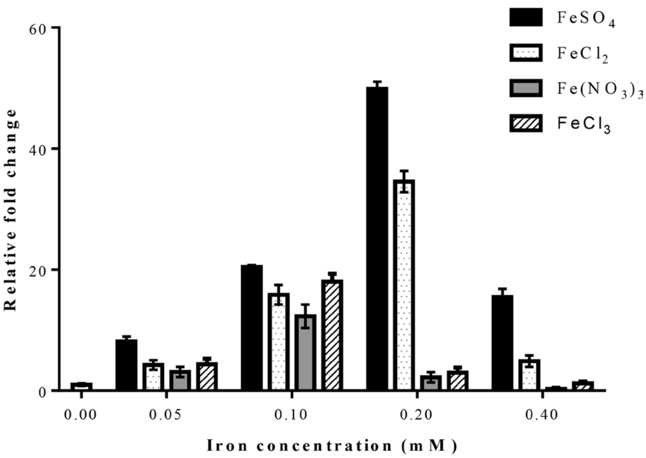

Fig. 1. Effect of iron salt on CS6 RNA expression in ETEC. CS6 RNA expression was quantified by RT-PCR using *parC* as the internal control. The strain was cultured in M9 minimal media with 0.2% glucose as carbon source and supplemented with varying concentrations of different iron salts [$FeSO_4$; $FeCl_2$; $Fe(NO_3)_3$; $FeCl_3$]. The bars show means and the standard error of the means (SEM) of three separate experiments. The results were significant when comparing the expression in presence of iron salts supplementation condition with the media only condition with no supplementation of iron (P<0.05).

LT was quantified by a GM1-based ELISA assay, using an anti-LTB Ab and GM1 as coating antigen [20]. STh was detected by indirect ELISA and protein was coated, which was detected by Anti-ST MAb [21].

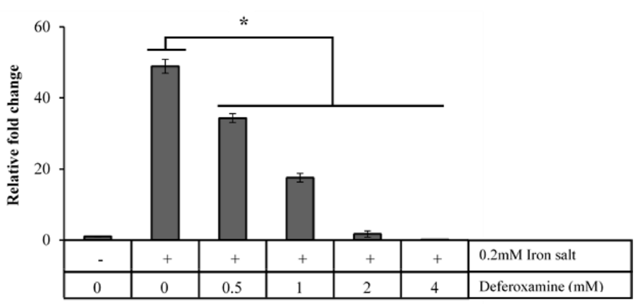

Fig. 2. Effect on CS6 expression in ETEC presence of iron chelation compound. The effect of 0.2 mM iron salt on the expression of the CS6 gene was studied in the presence of deferoxamine (Def) in dose-dependent concentration. The ETEC strain was grown in M9 minimal media with 0.2% glucose as carbon source with/without 0.2 mM $FeSO_4$ in the presence and absence of deferoxamine separately treated as indicated. Then the expression of CS6 was analysed by RT-PCR. The internal control used for the experiment was *parC*. Bars show means and standard error of the means of three separate experiments. *, Indicates significance (P=<0.05), when the expression in presence of iron salts and deferoxamine was compared to the expression only in presence of 0.2 mM iron.

Recombinant DNA techniques and promoter assay

All DNA manipulations were performed using standard molecular biology-based methods. For the promoter activity assay, we used the β-galactosidase expressing promoter-less plasmid pTL61T. We cloned CS6 promoter genes in this plasmid. In short, PCR amplified genomic DNA by using the promoter primers (Table 2) were cloned into plasmid pTL61T to generate plasmid-borne fusions by using the restriction enzymes Xho1 and Xba1 (New England Biolabs). The fusion plasmid was then transformed into the strains of ETEC (Table 1). The LacZ activity of the transcriptional reporter strains was measured by Miller assays [22]. In short, ETEC strains were incubated in LB overnight at 37 °C. Next, bacteria were transferred to M9 media containing desired concentrations of iron and were grown to mid-log phase and then placed on ice for 30 min. Then 200 µl of the culture was resuspended in 'Z' buffer. To this, 30 µl 0.1% (w/v) SDS and 60 µl of chloroform were added to each sample and vortexed. Next, 200 µl of ONPG (o-nitrophenyl-β-galactoside) solution, was added and incubated at 28 °C and the reaction was stopped after getting the desired yellow colour by adding 500 µl of 1M Na_2CO_3. The β-galactosidase activity was expressed as Miller units.

HT-29 cell culture and adhesion assay

The HT-29 epithelial cell lines were cultured in Corning 6-wells plate using DMEM (Dulbecco's Modified Eagle Medium) at 37 °C with 5% CO_2 using a Heracell 150i CO_2 Incubator (Thermo Scientific, USA) to form a monolayer. Sodium bicarbonate, 3.7 g was added to per litre of the media. Foetal bovine serum (FBS) (PAN-Biotech, Germany) 10% (v/v), 1% (v/v) non-essential amino acids (100X) (MP Biomedical, USA) and 1% (v/v) antibiotics (Penicillin 5000 IU per ml and Streptomycin 5 mg per ml; MP Biomedical, USA) were supplemented. The pH of the media was adjusted to ~7.2 and sterilized by filtration under aseptic conditions. The used media was replaced in alternate days and when the monolayer reached ~90% confluency it was used. The cells were treated with trypsin-EDTA to detach the cells and then used for sub-culturing. The stock of the cultured epithelial cells was prepared by using 95% complete growth media and 5% (v/v) DMSO (Sigma-Aldrich, USA) and kept in liquid nitrogen.

Before the adhesion experiments, the epithelial cells were serum-starved for 18 h in an appropriate medium with 0.5% (v/v) FBS. ETEC strains were pre-inoculated in M9 media supplemented with varying concentrations of iron salts and grown up to mid-log phase at 37 °C. Then 10^7 c.f.u. ml^{-1} of bacterial suspension was pelleted down and dissolved in the incomplete culture medium. This suspension was added to epithelial cells and after 3 h of incubation at 37 °C in 5% CO_2, cells were washed with PBS and detached by 0.1% Triton X-100 or by Trypsin-EDTA. The adhered bacteria were counted after serial dilution by plating on MacConkey agar plates.

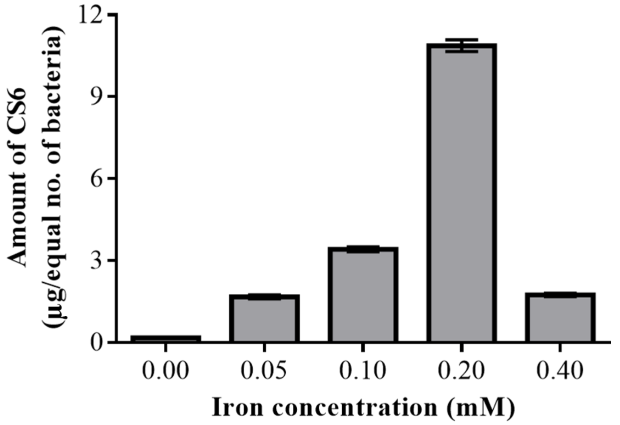

Fig. 3. Phenotypic expression of CS6 on ETEC surface in presence of iron. The level of CS6 on the bacterial surface and total CS6 protein in each strain was determined using 10^7 c.f.u. ml^{-1} of bacteria. CS6 was detected by anti-*cssB* antibody followed by incubation with HRP-conjugated secondary antibody. The amount of CS6 (µg) was determined from the standard curve. The bars show means and the standard error of the means (SEM) of three separate experiments. The results were significant when comparing the expression in presence of iron salts supplementation condition with the media only condition with no supplementation of iron ($P<0.05$).

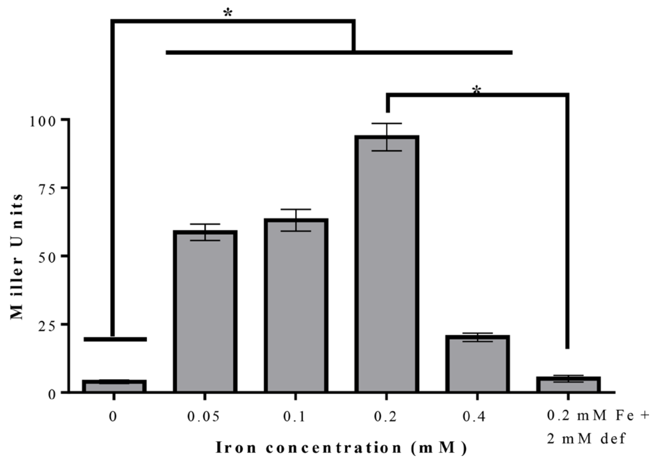

Fig. 4. CS6 promoter activity in response to 0.2 mM iron concentration. The effect of iron on the expression of CS6 promoter was measured by *β*-galactosidase activity assay as per Miller and Lee protocol, 1984 [22]. Promoter constructs were made in no-promoter pTL61T plasmid and transformed into strains used for this study. The transformed strain was grown in M9 minimal media with 0.2% glucose as carbon source and $FeSO_4$ was added at the indicated concentrations with or without 2 mM deferoxamine (Def). Assays were performed in triplicate, and bars show the mean and standard error of the means. *, Indicates significance (P=<0.05), when the promoter activity in the presence of iron salts was compared to the expression under no iron supplementation and also the decrease in the promoter activity when deferoxamine had been added in the presence of iron.

Rabbit ileal loop assay

In vivo adhesion of bacterial strains were also evaluated in the rabbit intestinal lumen. For this rabbit ileal loop assay was performed in young New Zealand white rabbits as described previously [23]. ETEC strains grown in M9 broth with or without iron salts were inoculated in rabbit ileum at different concentrations. The PBS-inoculated loop was used as a negative control. Intestinal loop sections recovered 18 h after the rabbit ileal loop experiment was washed in PBS three times, homogenized, and serially diluted in PBS. The adherent bacterial count was determined by plating these bacterial cultures on MacConkey agar plates.

Ethical approval

All the ETEC strains were isolated from patients following the approved institutional (ICMR-NICED) ethical guidelines (No. A-1/2009-IEC and No. A-1/2015-IEC). All animal experiments were conducted following the standard operating procedure as outlined by committee for the purpose of supervision and control experiments on animals (CPCSEA), Government of India and protocols were approved by the Institutional (ICMR-NICED) Animal Ethics Committee (NICED/CPCSEA/68/GO/(25/294)/2016-IAEC/NSC-1).

Statistical analysis

All the experiments were repeated at least three times. Data were expressed as mean±standard error of means (SEM). One-way ANOVA was used to analyse the data wherever applicable. A P value of <0.05 was considered statistically significant.

RESULTS

RNA expression of CS6 is influenced by iron concentration

The expression of CS6 in ETEC was investigated *in vitro* in the presence of different iron salts by quantitative real-time PCR. It was found that CS6 RNA expression increased significantly in a dose-dependent manner in the presence of iron salts (Fig. 1). The tested concentrations were from 0 to 0.4 mM iron. CS6 RNA expression increased by ~8-fold at 0.05 mM, 20-fold at 0.01 mM, 50-fold at 0.2 mM and 20-fold at 0.4 mM of $FeSO_4$ when compared to RNA expression in untreated conditions. It was observed that the increase in CS6 RNA was maximal in the presence of $FeSO_4$. In the presence of 0.2 mM $FeCl_2$, a maximum increase of 35-fold was observed in respect to untreated conditions. All the data with the supplement of iron salts were significantly modulated ($P<0.05$) when compared to the control or untreated condition with no supplement of iron. A significant increase in the CS6 expression can be seen in 0.1 mM for all iron salts and most significant increase in presence of 0.2 mM iron (II) salts when compared to the control condition with no supplementation of iron salts. There had been a significant decrease in the expression of CS6 in presence of 0.4 mM iron (II) salts when compared to the expression of CS6 at 0.2 mM iron (II) salts. Whereas there had been a significant decrease in the expression of CS6 when iron (III) salts were supplemented at a concentration of 0.2 and

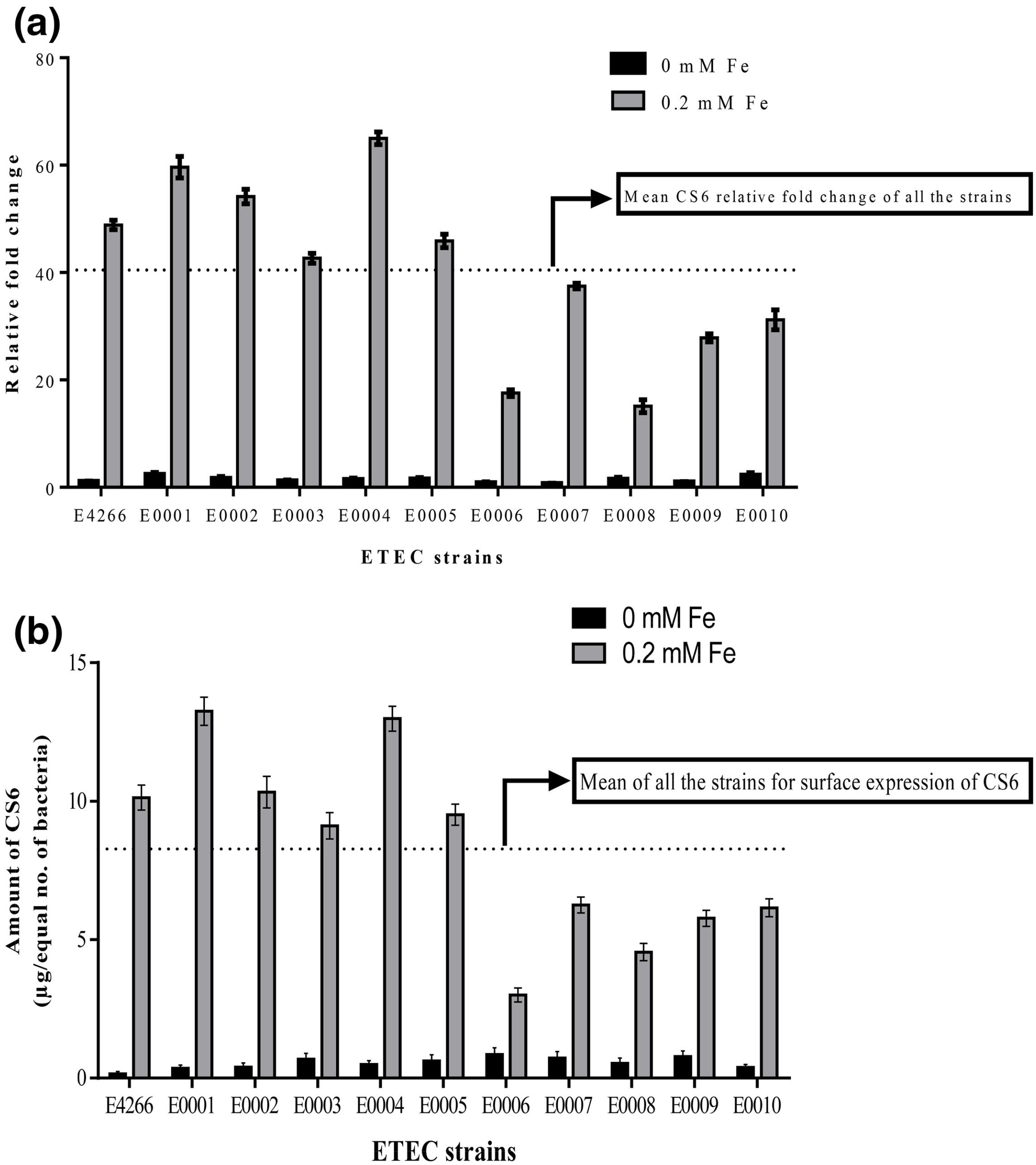

Fig. 5. CS6 expression in different ETEC strains in presence of iron salts. (a) Expression of CS6 in M9 minimal media (supplemented with 0.2% glucose as carbon source) supplemented with iron salts standardized to the level of transcription in cultured in M9 minimal media with 0.2% glucose as carbon source. RNA expression was measured by RT-PCR with *parC* as the internal control. (b) The level of CS6 on the bacterial surface was measured by anti-*cssB* antibody from the standard curve. In both the figures, bars show means and standard error of the means of three separate experiments. Dotted line represents the mean of all the strains.

0.4 mM when compared to the expression of CS6 at a 0.1 mM iron (III) concentration. The increase in presence of Fe (III) was not as much as Fe (II)-mediated induction and was maximal at 0.1 mM salt (12-fold and 18-fold for Fe $(NO)_3$ and $FeCl_3$, respectively) when compared to untreated RNA expression.

In *E. coli* it is already reported that *fur* expression is auto-regulated in response to iron. So for providing a control gene that is expressed under iron limitation we have checked the expression of *fur* in presence of different concentrations of iron ($FeSO_4$) by quantitative real-time PCR. It was found that *fur* RNA expression decreased significantly (P<0.05) in a dose-dependent manner in the presence of iron salts except for an increase of 0.2-fold of *fur* RNA in presence of 0.05 mM $FeSO_4$ (Supplementary file 1, Fig. S1, available in the online version of this article).

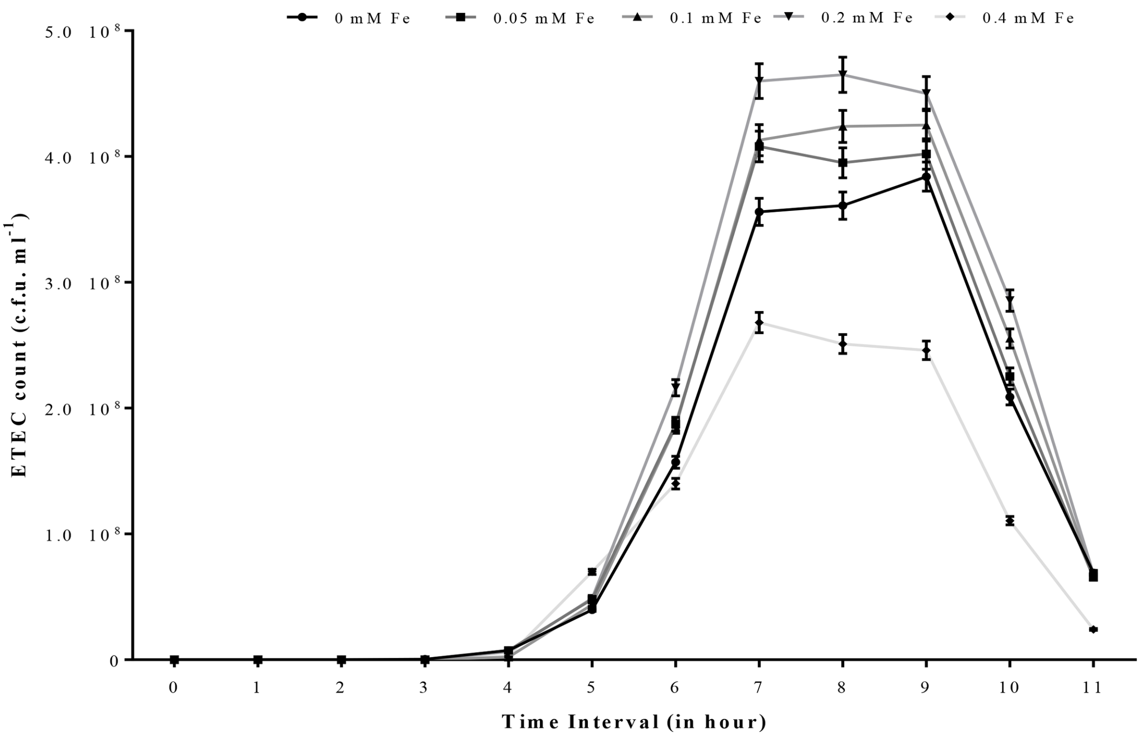

Fig. 6. Effect of iron salt on ETEC growth. ETEC growth was measured by colony counting method. The strain was cultured in M9 minimal media with 0.2% glucose as carbon source and supplemented with varying concentrations of iron salts ($FeSO_4$). The bars show mean and the standard error of the mean (SEM) of three separate experiments.

Expression of CS6 is reduced under iron starvation

To investigate the effect of iron depletion on CS6 expression M9 media, we examined the iron chelator deferoxamine (Def) in a concentration gradient manner along with the supplementation of iron salt at the concentration of 0.2 mM. Real-time PCR confirmed that sequestering of free iron by the addition of deferoxamine (Def) decreased the levels of CS6 expression. On addition of 0.5 and 1 mM deferoxamine, the RNA expression of CS6 was decreased by 1.5-fold and 3-fold respectively, when compared to no-deferoxamine treated condition and in presence of 0.2 mM iron. At 2 mM deferoxamine concentration, the RNA expression of CS6 is nearly the same as without iron salts and at 4 mM deferoxamine concentration the RNA expression of CS6 is downregulated by five-fold with respect to CS6 level in absence of iron. (Fig. 2). RNA fold change of CS6 expression gradually decreased in the presence of increasing concentrations of deferoxamine when compared to no deferoxamine added condition ($P<0.05$). These results indicated that iron availability increases CS6 expression as iron chelation reduces the expression of CS6.

Phenotypic expression of CS6 is induced in the presence of iron

The surface expression of CS6 was studied in the presence of 0.2 mM $FeSO_4$ salt by ELISA. (Fig. 3). The surface expression of CS6 increased gradually by 10-fold, 20-fold and 60-fold in presence of 0.05, 0.1 and 0.2 mM iron salt, respectively, when compared to CS6 grown without iron. However, the surface expression of CS6 decreased by six-fold when the concentration was further increased to 0.4 mM as compared to CS6 of ETEC grown in presence of 0.2 mM iron. The increase in phenotypic expression of CS6 with the addition of iron (II) salts was significant ($P<0.05$) when compared with the untreated condition. The result suggests that the optimum concentration for maximum CS6 expression is 0.2 mM iron (II) concentration.

CS6 promoter activity is upregulated in presence of iron

Since the expression of CS6 was regulated by iron availability, the modulation of the CS6 promoter was evaluated as a measure of β-galactosidase activity in the presence and absence of iron salts. The promoter activity of CS6 was increased in the presence of 0.05 and 0.1 mM iron concentration by ~15-fold and ~16-fold when compared to the promoter activity under untreated conditions. The promoter activity was elevated maximally by 23-fold at 0.2 mM iron concentration. The increase in the promoter activity on addition of iron was significant ($P<0.05$) when the promoter activity was compared with the untreated condition. However, it decreases with further increases in iron salt. Results showed that the decrease was five-fold in presence of 0.4 mM iron salt concentration when compared to the promoter activity under 0.2 mM condition (Fig. 4). The CS6 promoter activity

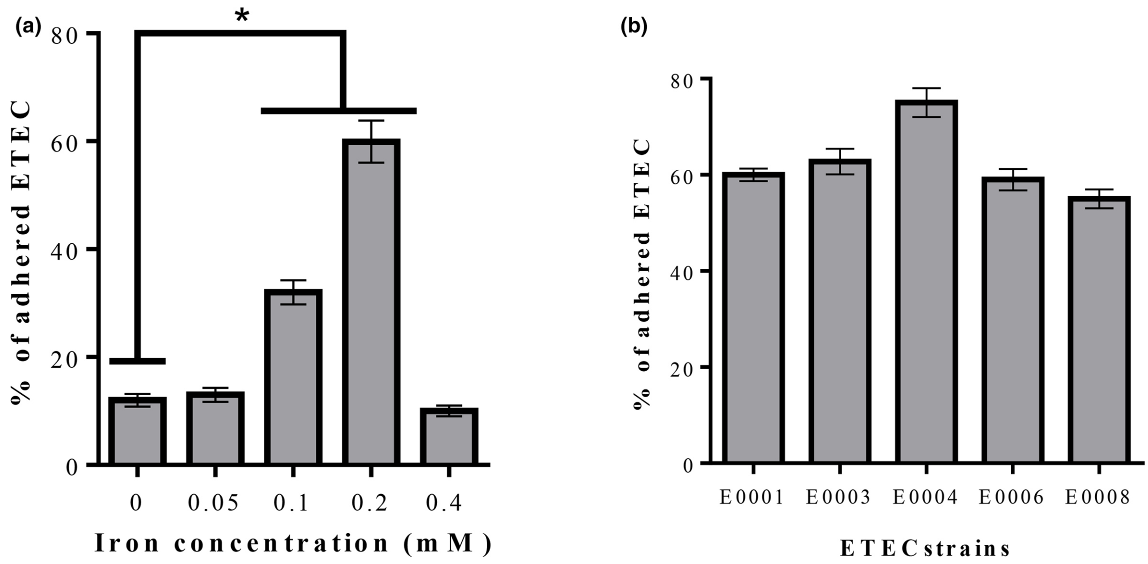

Fig. 7. Effect of iron on ETEC adherence to the epithelial cells. Adherence assays were performed for 3 hours in McCoy's 5A medium with 10% of FBS with or without iron salts. Percent of bacteria adhered to HT-29 cells for (a) reference strain in absence of Fe and the presence of a varying concentration of Fe(II), and (b) different CS6 harbouring ETEC strains in the presence of 0.2 mM iron salt. Assays were performed in triplicate, and bars show the mean and standard error of the means. *, Indicates significance (P=<0.05), when the adherence levels in presence of iron salt were compared with the no iron supplementation condition. The percentage of adhered bacteria was calculated as described in Methods.

was repressed in the presence of 0.2 mM iron salt and deferoxamine by 18-fold when compared to the promoter activity under 0.2 mM condition. This decrease in the promoter activity on the addition of deferoxamine under the iron supplementation condition is significant (P<0.05).

Expression of CS6 RNA in different ETEC strains harbouring CS6 is upregulated in presence of iron

To investigate whether the induction of CS6 RNA was strain-specific, the effect of iron on different ETEC strains harbouring CS6 was tested under the influence of 0.2 mM of iron salt. The experiment revealed that the RNA expression of CS6 had increased when supplemented with iron in the growth media with respect to untreated conditions in all the 11 strains tested (Fig. 5a). The mean increase in RNA expression of all the strains was 40-fold, being 15-fold and 65-fold as the minimum and maximum increase in expression, respectively. CS6 surface expression was also in accordance with the RNA expression data (Fig. 5b). CS6 surface expression was also increased significantly in different strains of ETEC at 0.2 mM iron (II) concentration.

Growth of ETEC is iron salt concentration-dependent

Next, the effect of iron salt on the growth of ETEC strain was examined. The experiment revealed an effect of iron salts on the growth of ETEC in a concentration dependent manner (Fig. 6). It was observed that in the presence of 0.05 mM iron concentration, the ETEC strain grew more by one-fold with respect to the untreated condition. At 0.1 mM iron concentration, the growth of ETEC increased by 1.5-fold. Maximum growth was observed in the presence of 0.2 mM iron which is two-fold higher as compared with the untreated condition. The growth was retarded by two-fold in the presence of 0.4 mM concentration of iron salt. Regression analysis showed that these differences in growth are not statistically significantly different. We had also checked the growth of ETEC in the presence of other iron salts. The results were similar as there was an increase in growth at iron concentrations from 0.05 mM to 0.2 mM for ferrous chloride and the growth was inhibited at higher iron concentrations (Fig. S2a). For ferric salts the growth was enhanced at iron concentrations of 0.05 and 0.1 mM and on addition of higher concentrations of iron salts the growth was retarded (Fig. S2b, c).

Adhesion of ETEC to epithelial cells is increased in presence of iron

Adhesion or colonization to the epithelial layer of the small intestine is a prerequisite for ETEC pathogenesis. To evaluate bacterial adhesion due to effect of iron salts, HT-29 cells were incubated with ETEC strain in the presence and absence of iron salts. It was observed that at 0.05 mM iron concentration, the ETEC adhesion to HT-29 cells was similar to the untreated condition. At 0.1 mM iron concentration, the adhesion of ETEC to HT-cells increased by three-fold. At 0.2 mM iron concentration, the adhesion increased maximally by five-fold compared to the untreated condition. Iron essentially improved the ability of ETEC cells to adhere to HT-29 cells contrasted with

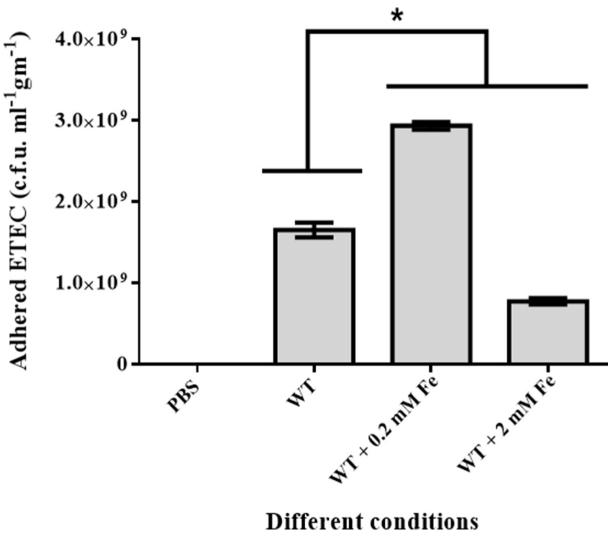

Fig. 8. Effect of iron on ETEC adherence to the rabbit ileal epithelium. Intestinal adherence or colonization was done in rabbit ileal loop. The ileal was divided into loops, each between 8–10 cm in length and interloop was present between loops. Loops were inoculated with 10^7 c.f.u. wild-type ETEC isolate expressing CS6 with or without iron. At 18 h of infection, the rabbit was sacrificed, and 2–3 small sections each of 2 cm from each loop were homogenized in 1 ml of PBS and ETEC adherence was measured in c.f.u. gm^{-1}. *, Indicates significance (*P*=<0.05), when the adherence levels in presence of iron salt were compared with the no iron supplementation condition.

untreated ETEC. However, at 0.4 mM iron concentration, the adhesion decreased by six-fold compared to the 0.2 mM iron condition. (Fig. 7a). When the adherence level in presence of iron salts were compared to the adherence level without iron salts it was found that the results were significant (*P*<0.05) at 0.1 mM and 0.2 mM iron supplementation condition. Therefore, it was observed that the adherence of ETEC increased in a dose-dependent manner with increasing concentrations of iron up to 0.2 mM, after which it started decreasing. The adhesion was further studied with other clinical ETEC strains harbouring CS6 in presence of 0.2 mM iron concentration. The adhesion of these five ETEC strains to HT-29 cells were about five-fold more with respect to untreated conditions (Fig. 7b). This also showed that the adhesion of ETEC was not strain-specific.

ETEC harbouring CS6 isolates were tested here for its adherence to rabbit ileal epithelial cells when supplemented with iron salts. The result suggested that when ETEC was grown in the presence of 0.2 mM iron salts, adherence increased by two-fold, whereas in the presence of 2 mM iron, the adherence was decreased by half in comparison to the wild-type ETEC grown without iron supplementation (Fig. 8). The adherence level increased significantly on addition of 0.2 mM iron salts whereas it decreased significantly when 2 mM iron salts had been supplemented (*P*<0.05).

DISCUSSION

In this study, the regulation of CS6 by iron was investigated and the analysis showed that iron acts as a modulator of CS6 expression. The expression of CS6 was enhanced in the presence of iron salts up to 0.2 mM iron and decreased thereafter with increased iron concentrations of 0.4 mM or more, suggesting that the optimum enhancement could be achieved under 0.2 mM. Though iron is critical for almost all living organisms, in high concentrations it is extremely toxic [24].

Uptake of both forms of iron, Fe(II) and Fe(III) by Feo system and by siderophores respectively are the different iron acquisition systems used by pathogens to obtain iron from the host [25]. Iron fundamentally happens in both of two oxidation states in biological systems as Fe(II) or ferrous, supported by low pH and low oxygen and Fe(III) or ferric state, which is the predominant form of iron under aerobic conditions [10]. Fe(III) is reduced to Fe(II) for subsequent incorporation into haem- or iron-containing proteins [26]. In this study, it was also observed that both forms of iron, ferrous and ferric, induced the expression of CS6 RNA. The modulation of CS6 in the presence of iron was not strain specific; maximum enhanced expression of CS6 was observed under 0.2 mM iron condition in other CS6-positive clinical ETEC strains, followed by decreased RNA expression with further increases in iron concentration. In comparison, under iron starvation conditions, the yield of the CFA/I fimbriae was enhanced in ETEC H10407 [27].

The effect CS6 RNA modulation by iron was reflected on the phenotypic expression of CS6. The expression of virulence factors had shown to be regulated by iron availability in an earlier study [28]. A rise in the expression of flagellin production with additional iron was also seen for extraintestinal pathogenic *Escherichia coli* [29]. Iron increases the synthesis and surface placement of adhesins in trichomonas [30]. On the contrary, a previous study had shown that in cases of colonization factors CS1, CS3 and CFA/I, iron starvation induced their expression [27].

The modulation CS6 expression could be attributed to the upregulated activity of the CS6 promoter in presence of iron up to 0.2 mM. The decrease in CS6 expression might also be due to downregulation of the promoter activity beyond the optimum iron concentration. It is possible that intracellular levels of free iron had reached toxic levels. Iron had been reported to control the promoters involved in *fur* regulation and upon interaction with iron *fur* becomes repressed [10] and our promoter analysis by bioinformatics using 'softberry bacterial promoter prediction tool' (Fig. S3) revealed the presence of *fur* region and RNA analysis showed downregulation of the *fur* gene in presence of 0.2 mM iron salts. However, in case of CFA/I, it was observed that its promoter activity was induced under iron starvation condition [27, 31].

The ability to adhere to the intestinal mucosal layer is a primary requirement for ETEC to initiate pathogenesis. In this current study, it was observed that with the increased CS6 expression in the presence of increased iron, adherence of ETEC to intestinal HT-29 cells and rabbit ileal epithelium also increased. With an increase of iron concentration beyond 0.2 mM, there was decreased adherence due to decreased CS6 levels. These results suggested that iron may contribute to modulate ETEC colonization of the intestinal epithelial layer. A study demonstrated that iron enhanced the adhesion of enteric pathogen *S. typhimurium,* but in contrast, iron showed no effect or even reduced the adhesion of *E. faecalis* and *L. plantarum* [32]. Lactoferrin, an iron chelator, had appeared to disable the capacity of surface-expressed virulence factors and thereby in this way restraining the adherence of ETEC to epithelial cells [33].

Results presented here showed that addition of iron up to the optimum concentration of 0.2 mM resulted in a boosting growth of ETEC. Iron concentration beyond the optimum iron concentration affects ETEC growth. These results are in accord with the finding that when a lot of unabsorbed dietary iron enters the colon enteric microbes can possibly grow out. In infants, there was a significant increase in the sum of pathogenic *E. coli* through iron fortification [34]. It was observed that when the culture medium was supplemented with an excess of iron, the *Shigella* shiga toxin was produced in lesser amounts but the growth of bacteria was enhanced [28]. Free iron may lead to the rapid bacterial growth by abolishing the natural defensive bactericidal mechanisms in tissue fluids [11].

It has been shown that iron sequestering by lactoferrin inhibits the growth of enteric pathogens [33]. In pathogens like *Staphylococcus aureus, Staphylococcus epidermidis, Escherichia coli* and *Enterobacter,* deferoxamine had a bacteriostatic effect as it depleted available iron and created an iron-scarce environment, as the bacteria could not use iron for their multiplication [35]. Chelating of iron by deferoxamine reduced the expression of CS6, both at RNA and protein levels. Iron chelation therapy has worked successfully to control infection against *S. aureus*, *Plasmodium falciparum, Yersinia enterocolitica, Vibrio vulnificus, Klebsiella pneumonia,* and *Aeromonas hydrophila* [36]. It was previously demonstrated that mice were less prone to infection with *Salmonella typhimurium* when serum iron levels were reduced. However, when normal serum iron levels was restored by injecting sufficient iron, infection increased [28]. These perceptions on the impacts of iron deprivation in biological systems propose the significance of iron to the microbes.

The LT and ST toxins of ETEC showed a different expression pattern compared to CS6 expression in the presence of iron. We found that iron richness had inhibited the secretion of the toxins (Fig. S4). Previous work with Shiga-like toxin I of enterohaemorrhagic *E. coli* showed toxin induction at iron starvation. Toxins such as shiga toxin, *Pseudomonas aeruginosa* A toxin, hemolysin and SLT of *E. coli* and diphtheria toxin were reduced in the presence of an increased iron concentration [28]. In contrast, it was found that LT secretion was inhibited in ETEC H10407 strain producing CFA/I fimbriae under iron starvation [27]. As a part of innate defence, the host limits iron availability by employing iron-binding proteins such as transferrin and lactoferrin so that invading pathogens cannot survive in low iron conditions and this low iron level acts as a signal to the pathogens for induction of virulence [37].

CONCLUDING REMARKS

In summary, our data showed that the free iron availability acts as a modulator for the expression of ETEC colonization factor CS6 which supports the hypothesis that iron can increase the virulence and growth of enteric pathogens. Our study suggests that ETEC is capable of expressing CS6 in consonance with the concentration of iron salts.

Funding information

Our research was supported in part by the Department of Biotechnology, Govt. of India, Grant No. BT/PR21476/MED/29/1135/2016. Mr Debjyoti Bhakat was supported by a fellowship from Indian Council of Medical Research [ICMR File No.- 3/1/3/JRF-2015/HRD-LS/96/40713/57] Govt. of India. Mr Indranil Mondal was supported by the above mentioned Department of Biotechnology grant. The funders had no role in study design, data obtaining, analysis, decision to publish, or preparation of the manuscript.

Author contributions

D.B., I.M., N.S.C., conceptualized the study. D.B., I.M., performed experiments. All authors prepared the original draft of the manuscript and did data analysis, draft review, editing and approval.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Hosangadi D, Smith PG, Kaslow DC, Giersing BK. WHO consultation on ETEC and Shigella burden of disease, Geneva, 6–7th April 2017: Meeting report. *Vaccine* 2019;37:7381–7390.
2. Qadri F, Svennerholm A-M, Faruque ASG, Sack RB. Enterotoxigenic *Escherichia coli* in developing countries: Epidemiology, microbiology, clinical features, treatment, and prevention. *Clin Microbiol Rev* 2005;18:465–483.
3. Black RE. Epidemiology of travelers' diarrhea and relative importance of various pathogens. *Rev Infect Dis* 1990;12 Suppl 1:S73-9.
4. Gaastra W, Svennerholm A-M. Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol* 1996;4:444–452.
5. von Mentzer A, Tobias J, Wiklund G, Nordqvist S, Aslett M, *et al*. Identification and characterization of the novel colonization factor CS30 based on whole genome sequencing in enterotoxigenic *Escherichia coli* (ETEC). *Sci Rep* 2017;7:12514.
6. Sabui S, Dutta S, Debnath A, Ghosh A, Hamabata T, *et al*. Real-time PCR-based mismatch amplification mutation assay for specific detection of CS6-expressing allelic variants of enterotoxigenic *Escherichia coli* and its application in assessing diarrheal cases and asymptomatic controls. *J Clin Microbiol* 2012;50:1308–1312.
7. Begum YA, Baby NI, Faruque ASG, Jahan N, Cravioto A, *et al*. Shift in phenotypic characteristics of enterotoxigenic *Escherichia coli* (ETEC) isolated from diarrheal patients in Bangladesh. *PLoS Negl Trop Dis* 2014;8:e3031.
8. Bhakat D, Debnath A, Naik R, Chowdhury G, Deb AK, *et al*. Identification of common virulence factors present in enterotoxigenic *Escherichia coli* isolated from diarrhoeal patients in Kolkata, India. *J Appl Microbiol* 2019;126:255–265.
9. Sabui S, Debnath A, Ghosal A, Wajima T, Hamabata T, *et al*. Characterization of oligomeric assembly of colonization factor CS6 from enterotoxigenic *Escherichia coli*. *Microbiology (Reading)* 2016;162:72–83.
10. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. *FEMS Microbiol Rev* 2003;27:215–237.

11. Bullen JJ, Rogers HJ, Spalding PB, Ward CG. Iron and infection: The heart of the matter. *FEMS Immunol Med Microbiol* 2005;43:325–330.
12. WHO U. Focusing on anaemia: Towards an integrated approach for effective anaemia control: *Joint statement by the World Health Organization and the United Nations Children's Fund*. 2004.
13. Sazawal S, Black RE, Ramsan M, Chwaya HM, Stoltzfus RJ, *et al*. Effects of routine prophylactic supplementation with iron and folic acid on admission to hospital and mortality in preschool children in a high malaria transmission setting: Community-based, randomised, placebo-controlled trial. *The Lancet* 2006;367:133–143.
14. Oppenheimer SJ. Iron and its relation to immunity and infectious disease. *J Nutr* 2001;131:616S-635S.
15. Skaar EP. The battle for iron between bacterial pathogens and their vertebrate hosts. *PLoS Pathog* 2010;6:e1000949.
16. Ghosal A, Bhowmick R, Nandy RK, Ramamurthy T, Chatterjee NS. Pcr-based identification of common colonization factor antigens of enterotoxigenic *Escherichia coli*. *J Clin Microbiol* 2007;45:3068–3071.
17. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25:402–408.
18. Wajima T, Sabui S, Fukumoto M, Kano S, Ramamurthy T, *et al*. Enterotoxigenic Escherichia coli cs6 gene products and their roles in cs6 structural protein assembly and cellular adherence. *Microb Pathog* 2011;51.
19. Elder BL, Boraker DK, Fives-Taylor PM. Whole-bacterial cell enzyme-linked immunosorbent assay for streptococcus sanguis fimbrial antigens. *J Clin Microbiol* 1982;16.
20. Sjöling A, Wiklund G, Savarino SJ, Cohen DI, Svennerholm A-M. Comparative analyses of phenotypic and genotypic methods for detection of enterotoxigenic *Escherichia coli* toxins and colonization factors. *J Clin Microbiol* 2007;45:3295–3301.
21. Rocha LB, Ozaki CY, Horton DSPQ, Menezes CA, Silva A, *et al*. Different assay conditions for detecting the production and release of heat-labile and heat-stable toxins in enterotoxigenic *Escherichia coli* isolates. *Toxins (Basel)* 2013;5:2384–2402.
22. Miller J, Lee K. *Experiments in Molecular Genetics*. Yi Hsien Pub. Co, 1984.
23. De SN, Bhattacharya K, Sarkar JK. A study of the pathogenicity of strains of *Bacterium coli* from acute and chronic enteritis. *J Pathol Bacteriol* 1956;71:201–209.
24. Sousa Gerós A, Simmons A, Drakesmith H, Aulicino A, Frost JN. The battle for iron in enteric infections. *Immunology* 2020;161:186–199.
25. Andrews S, Norton I, Salunkhe AS, Goodluck H, Aly WS, *et al*. Control of iron metabolism in bacteria. In: *Metallomics and the cell*. Springer, 2013. pp. 203–239.
26. Schaible UE, Kaufmann SHE. Iron and microbial infection. *Nat Rev Microbiol* 2004;2:946–953.
27. Haines S, Arnaud-Barbe N, Poncet D, Reverchon S, Wawrzyniak J, *et al*. ISCR regulates synthesis of colonization factor antigen I fimbriae in response to iron starvation in enterotoxigenic *Escherichia coli*. *J Bacteriol* 2015;197:2896–2907.
28. Payne SM, Neilands I. Iron and virulence in the family Enterobacteriaceae. *Crit Rev Microbiol* 1988;16:81–111.
29. Magistro G, Magistro C, Stief CG, Schubert S. The high-pathogenicity island (HPI) promotes flagellum-mediated motility in extraintestinal pathogenic *Escherichia coli*. *PloS one* 2017;12:e0183950.
30. Alderete JF, Nguyen J, Mundodi V, Lehker MW. Heme-iron increases levels of AP65-mediated adherence by *Trichomonas vaginalis*. *Microb Pathog* 2004;36:263–271.
31. Karjalainen TK, Evans DG, Evans DJ Jr, Graham DY, Lee C-H. Iron represses the expression of CFA/I fimbriae of enterotoxigenic *E. coli*. *Microbial Pathogenesis* 1991;11:317–323.
32. Kortman GAM, Boleij A, Swinkels DW, Tjalsma H. Iron availability increases the pathogenic potential of *Salmonella* typhimurium and other enteric pathogens at the intestinal epithelial interface. *PloS one* 2012;7:e29968.
33. Ochoa TJ, Cleary TG. Effect of lactoferrin on enteric pathogens. *Biochimie* 2009;91:30–34.
34. Jaeggi T, Kortman GAM, Moretti D, Chassard C, Holding P, *et al*. Iron fortification adversely affects the gut microbiome, increases pathogen abundance and induces intestinal inflammation in kenyan infants. *Gut* 2015;64:731–742.
35. van Asbeck BS, Marcelis JH, Marx JJ, Struyvenberg A, van Kats JH, *et al*. Inhibition of bacterial multiplication by the iron chelator deferoxamine: Potentiating effect of ascorbic acid. *Eur J Clin Microbiol* 1983;2:426–431.
36. Carver PL. The battle for iron between humans and microbes. *CMC* 2018;25:85–96.
37. Sritharan M. Iron and bacterial virulence. *Indian J Med Microbiol* 2006;24:163–164.

Edited by: D. Grainger and M. Van Der Woude