COMPARATIVE GENOME ANALYSIS OF DIARRHEAGENIC AND NON-PATHOGENIC STRAINS OF *ESCHERICHIA COLI*

A PROJECT

Submitted in partial fulfillment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

BIOTECHNOLOGY

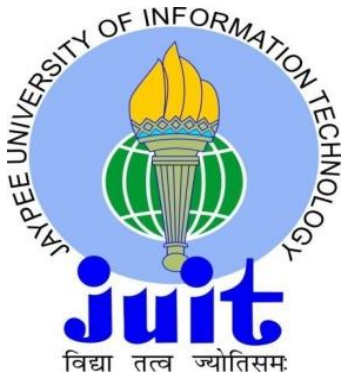
Under the supervision of

Dr. Jitendraa Vashistt

By

Natasha Thakur {133807}

To



JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT, SOLAN – 173234

HIMACHAL PRADESH, INDIA

MAY - 2017

CERTIFICATE

This is to certify that the Project Work entitled **"Comparative genome analysis of diarrheagenic and non-pathogenic strains of *Escherichia coli*"** submitted by **Natasha Thakur** in partial fulfillment for the award of degree of B.Tech in Biotechnology from Jaypee University of Information & Technology, Waknaghat has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of any other degree, diploma or such other titles.

Dr. Jitendraa Vashistt

Assistant Professor (Senior Grade)

Department of Biotechnology & Bioinformatics

Jaypee University of Information & Technology

Waknaghat, District Solan, Himachal Pradesh.

Date:

ACKNOWLEDGEMENT

This project is an outcome of continual work over and intellectual support from numerous sources. Therefore, I would like to express my sincere thanks to the people who have helped me the most throughout my project.

I owe my deepest gratitude to my supervisor Dr. Jitendraa Vashistt, Assistant Professor, Department of Biotechnology & Bioinformatics for suggesting the title of this project for taking keen interest, valuable guidance and constant encouragement as well as for providing necessary information regarding the project and also for their support throughout the project.

I am highly indebted to Ms Nutan, Ph.D. Scholar for making it possible to carry out this work. Without her continuous optimism, enthusiasm, encouragement and support this study would hardly have been accomplished. Furthermore I would also like to acknowledge with much appreciation the crucial role of Mr. Himanshu Singh for his technical assistance.

At last but not the least I m deeply grateful to my colleagues who treasured me for my hard work and my parents for their motivation, continual support and encouragement in all the stages of my life.

Natasha Thakur

133807

Date:

TABLE of CONTENTS

LIST of FIGURES

LIST of TABLES

LIST of ABBREVIATIONS

DEC	Diarrheagenic *Escherichia coli*
EPEC	Enteropathogenic *E. coli*
EHEC/STEC	Enterohemorrhagic (Shiga toxin-producing) *E. coli*
EAEC	Enteroaggregative *E. coli*
ETEC	Enterotoxigenic *E. coli*
EIEC	Enteroinvasive *E. coli*
DAEC	Diffusely adherent *E. coli*
AIEC	Adherent invasive *E. coli*
CD	Crohn's Disease
IBD	Inflammatory Bowel Disease
PCR	Polymerase Chain Reaction
AE	Attaching and Effacing
LT	Heat-labile Enterotoxins
ST	Heat-stable Enterotoxins
LEE	Locus of Enterocyte Effacement
CF	Colonization Factors
CS	Coli Surface
CFA/I	Colonization Factor Antigen I
AMP	Adenosine Monophosphate
ADP	Adenosine diphosphate
HC	Hemorrhagic Colitis
HUS	Hemolytic Uremic Syndrome
pINV	Invasive Plasmid
T3SS	Type Three Secretion System
AAF	Aggregative Adherence Fimbriae
Aap	Antiaggregation protein
AST	Antibiotic Susceptibility Testing
MIC	Minimum Inhibitory Concentration
NCBI	National Center for Biotechnology Information

ATP	Adenosine Triphosphate
ºC	Degree Centigrade
Stx	Shiga Toxin
GM1	Monosialotetrahexosyl Ganglioside
HEp-2 cells	Human epithelial type 2 cells
IS	Insertion Sequence
kDa	Kilodalton
AIDA	1-aminoindan-1,5-dicarboxylic acid
Rpm	Revolutions per minute
TAE	Tris-acetate-EDTA
DNA	Deoxyribonucleic acid
T6SS	Type six secretion system

ABSTRACT

Diarrhea is the most important and common cause of deaths in young children in India. It is the disturbance in the normal mechanism of secretion and absorption of water resulting in excess loss of water in faeces. It is the state of having minimum of three loose or liquid bowel movements each day and most commonly it can last for few days. It is responsible for causing dehydration because of fluid loss. Diarrheagenic *Escherichia coli* (DEC) is one of foremost causative agent of diarrhea. It is considered to be the most common source of epidemic and endemic diarrhea universally. Although the majority of diarrheal illness is resolved itself without any treatment and dehydration can typically be managed by oral rehydration therapy. In severe cases that have a greater risk of causing deaths, it is necessary to prevent diarrhea. In order to achieve this aim, present study has been carried out which focuses on rapid identification of the pathogen. Genomes were compared and two unique proteins clpV1 family T6SS ATPase and type six secretion system protein were discovered to be present only in the pathogenic strains of *E. coli.* With the help of these proteins we can clearly differentiate between the pathogenic and non-pathogenic strains of *E. coli* and can directly identify and detect the pathogen in the stool samples.

CHAPTER 1

INTRODUCTION

1.0 Introduction

Diarrheal illness is a critical health issue in society and a foremost reason of diseased condition and deaths in children and youngsters [1]. *E. coli* strains involved in diarrheal diseases are majorly the chief etiological agents of diarrhea. The most specific combinations depending on the group of virulence determinants acquired determine the presently known *E. coli* pathotypes, which are togetherly known as diarrheagenic *E. coli* (DEC) [2].

The genus *Escherichia* was termed after the German paediatrician "Theodor Escherich". It includes gram-negative, coliform bacilli that are facultative anaerobes and forms the part of the family *Enterobacteriaceae* [3]. *Escherichia coli* naturally colonize the gastrointestinal tract of human infants in few hours after birth and both the organism as well as the host derives mutual benefit for decades [4]. *Escherichia coli* as commensal bacteria rarely cause disease and is broadly distributed inhabiting the large intestine of humans and warm-blooded animals [5]. Some of the non-pathogenic strains of *E. coli* can cause intestinal and extraintestinal infection in immunocompromised individuals or when gastrointestinal barriers are violated. Pathogenic *E. coli* strains can cause urinary tract infection, sepsis/meningitis, and enteric/diarrheal disease [6].

On the basis of virulence mechanisms, host colonization sites, and the subsequent clinical symptoms and consequences, the DEC pathotypes are classified into different categories as Enteroaggregative *E. coli,* Enteropathogenic *E. coli,* Enterotoxigenic *E. coli,* Enteroinvasive *E. coli,* and Enterohemorrhagic (Shiga toxin-producing) *E. coli* [7]. DAEC pathotype called as diffusely-adherent *E. coli* pathotype is an additional, less well-defined pathotype which consists of the strains that adhere to epithelial cells in a diffused distribution [2]. All these pathotypes represents a set of clones that share definite virulence factors [7].

However, it is critical to note that the plasticity of the *E. coli* genome has interrupted the identification of some *E. coli* isolates as a pathotype. This is because certain isolates possess the major virulence characteristics of different pathotypes and are thus considered as more virulent hybrid pathogenic strains [8]. Additionally, few *E. coli* strains classified as adherent invasive *E. coli* (AIEC) pathotype, serve as potential agents for Crohn's disease (CD). It is an inflammatory bowel disease (IBD), which is well known to be caused by a combination of number of factors that includes genetic factors, factors affecting microbiota of intestines, environment affecting factors, and enteric disease causing organisms [9-10]. Diarrheal illness due to DEC infections is a significant health concern among children and adults in developing countries. The reason constitutes their association with morbidity and mortality of children mainly affecting the age group of less than five years [7].

Hence, this study mainly aims at rapid identification of diarrheagenic *E. coli* in one step by comparing the molecular signature sequences of pathogenic and commensal *E. coli* strains. This will bypass the tedious and laborious conventional microbiological protocols from isolation of pure colonies to biochemical analysis and will help in rapid pathogen identification followed by appropriate treatment regimen. Comparison of genomes of commensal *E. coli* strains and pathogenic strains of *E. coli* will result in elucidation of novel genes and proteins that are specific to commensal or pathogenic variants of *E. coli*. The rationale of these screening is to identify the genes/proteins as a molecular marker for rapid differentiation of pathogenic and non-pathogenic *E. coli*.

CHAPTER 2

AIM AND OBJECTIVES

2.0 Aim and objectives

Aim: Comparative genome analysis of diarrheagenic and non-pathogenic strains of *Escherichia coli*.

To achieve the above mentioned aim, following **Objectives** were designed:

- To compare the genomes of commensal and diarrheagenic strains of *E. coli*.
- To explore the unique/novel protein by manual mining of pathogenic and commensal *E. coli* genomes.
- To design PCR primers against the unique gene/protein.
- To design PCR protocol in order to differentiate between commensal and pathogenic strains of diarrheagenic *E. coli*.

CHAPTER 3

REVIEW OF LITERATURE

3.0 Review of Literature

Diarrheagenic *Escherichia coli* have evolved several methods for causing diseased condition in their host such as they have specialized fimbriae that allow them to bind to certain intestinal epithelial cells and produce toxins [11]. Complete understanding of pathogenesis has resulted in the discovery of precise diagnostic tools for classification of *E. coli* strains into many pathotypes [12].

3.1 Enteropathogenic *Escherichia coli* (EPEC)

The term enteropathogenic *E. coli* (EPEC) was used to describe various strains of *E. coli* epidemiologically linked to a continuous chain of outbreaks of infant's chronic diarrhea and persistent diarrhea in the year 1940 and 1950 [13-14]. EPEC when originally recognized by serotyping are defined as those *E. coli* strains that have the capability to cause diarrheal illness, to create a histopathology on the epithelium of intestine called as "the attaching and effacing (AE) lesion", and the incapability to form Shiga toxins and heat-labile (LT) or heat-stable (ST) enterotoxins [2]. Largely, the majority of typical EPEC strains fall into well-known O serotypes. Classical EPEC O serogroups comprises of O55, O86, O111, O114, O119, O127, and O142 [7]. The H6 and H2 antigens are the most common H antigens associated with EPEC [15-18]. H34 is less common EPEC type and certain typical EPEC strains are considered as non-motile (H-) in conventional tests [15-19].

The unique histopathology produced by this category of *E. coli* is known as "attaching and effacing (A/E) lesions". It is due to the close attachment of bacteria to the epithelial cells of intestine and elimination of enterocyte microvilli [20]. The development of micro ulcers and removal of the cells at the place of EPEC attachment was primarily mentioned in pigs that were infected experimentally [21] and later in biopsies from infected children [20].

A protein known as intimin encoded by *eae* gene is positioned on the locus of enterocyte effacement (LEE) and facilitates the attachment of bacteria to the surface of cell membranes [22]. The *eae* gene is the one which is used at present for the molecular diagnosis of this group. Mechanism of pathogenesis of these bacteria has not been fully known up till now and may include factors besides those directly liable for A/E lesions plus more specific intestinal cells [2, 23].

3.2 Enterotoxigenic *Escherichia coli* (ETEC)

ETEC is one of the vital but less familiar causes of diarrheal illness in every age group in regions with poor hygiene and insufficient drinking water [12]. Moreover, 20-60% of the estimated one billion international travellers yearly, people travelling to undeveloped countries will undergo from traveller's diarrhea [24]. In about 30-70% of traveller's diarrhea ETEC are the most commonly detected causative agents [25]. The detection of cholera toxin as a result of the development of rabbit ileal loop assay was also required for pure cultures of *E. coli* that have been isolated from stools of infants and adults showing same symptoms as that of cholera. When living cultures and filtrates of these strains were introduced into the isolated rabbit ileal loops in 1968, they displayed cholera like secretary response which led to the production of the heat labile enterotoxin of *E. coli* and identification of its pathotype [26-27].

ETEC strains stick to the epithelial cells of intestine by means of diverse collection of surface structures that are proteinacious in nature called as "colonization factors" (CFs) which could be fimbrial or non-fimbrial [28]. The current nomenclature termed these factors as "coli surface (CS) antigen", but certain old names still continue e.g. "colonization factor antigen I (CFA/I)". Even though more than 25 CFs have been identified up till now, but on several strains CF is not recognized which can be considered for the techniques used for identifying, true lack of CFs or at present unrevealed ones [29].

ETEC strains cause diarrhea by initially following the adhesion and colonization and then releasing enterotoxins that are heat labile and heat stable and generally encoded by plasmid [30]. The bacteria of this category produces little STs as a preprotoxin of 72-amino acids that can be generated into an active toxin of 18-19 amino acid known as STa and a 42 amino acid toxin named as STb. STa is generally formed by each of animal and human strains, whereas STb is primarily visualized in veterinary strains [31]. Heat-labile enterotoxin similar to cholera toxin forms a part of AB5 family of toxins that are heterohexameric molecules comprising of a single A subunit and five B subunits [32]. The A1 domain consists of activated toxin and is coupled to the A2 domain through a disulfide bond [33]. The A2 section is the spiral part of the molecule and holds the A subunit with the B pentamer which joins permanently to GM1 ganglioside as receptors on outer cell membranes [31]. The toxin subsequently gets incorporated inside and the A subunit known as ADP adds one or more ADP-ribose to the guanine nucleotide-binding protein, resulting in enhancing the amount of intracellular cyclic AMP leading to the diseased condition of diarrhea.

3.3 Enterohaemorrhagic (Shiga toxin-producing) *E. coli* (STEC/EHEC)

EHEC/STEC represents a recognized group of food borne pathogens spread worldwide. The ability to make one or extra Shiga toxin (Stx) family cytotoxins [34] represents the most important virulence characteristic of this pathogroup of *E. coli*. A broad range of infections from mild and almost unapparent diarrhea to highly serious infection such as hemorrhagic colitis (HC) and production of a life-threatening condition called as hemolytic uremic syndrome (HUS) are caused by EHEC/STEC [7]. Infants and children are largely affected, and even though the occurrence of infection varies in different regions, the impact and significance of EHEC/STEC infections in public health is enormous, being the chief reason of acute renal failure in infants in various countries [7].

E. coli O157:H7 serotype was the foremost to correlate to HC and HUS incidents in the beginning 1980s, and from that time it has been responsible for several outbreaks and sporadic cases of severe diseases worldwide, hence considered as the prototype of this disease-causing group of bacteria [35]. A few serogroups involving O26, O45, O103, O111, O121 and O145 can be marked to cause infection in humans [36]. Recently, the appearance of a number of particular clones, for example, the hybrid O104:H4 enteroaggregative *E. coli* having Stx2 genes responsible for causing a severe outbreak of HUS in Germany in 2011 [37], the development of a new O26:H11 clone in Europe [38], and few other hybrid clones [39] recommends that the mobility of genes and definitely the host background are the main features concerned with their pathogenic potential.

Shiga toxin family composed of Stx1, consists of correlated structures and almost identical biological activity. It is basically similar to *Shigella dysenteriae* toxin varying in only one amino acid and Stx2 showing not more than 60% homology of amino acid to Stx1 [32,40]. Almost sixteen potential fimbriae like operons that are not widely considered are identified in strains of this category [28, 41]. Newly, a pilus responsible for attachment and biofilm creation known as "hemorrhagic coli pilus" which is a type IV pilus have been also recognized in this category [42]. 29 different intimin categories with differences in the C-terminal part of the molecule have thus far been known [43-44]. The production of A/E lesions by STEC is enough for causing non bloody diarrhea but Shiga toxin is necessary for the prevalence of bloody diarrhea, hemorrhagic colitis and HUC [2, 45].

The main attributes of disease causing determinants of EHEC strains are chromosomally encoded; however plasmids would possibly play a very major role in the mechanism of pathogenesis of these strains. Plasmid pO157 is usually found in approximately 99-100% of O157:H7 serotype that has been identified from clinical isolates of human. This plasmid has been linked with the haemolytic functioning and attachment to epithelial cells of intestine.

However, the complete interpretation of functioning of plasmids in the mechanism of pathogenesis of this category is held up because of lack of a consistent model of infected humans [30, 46-47].

3.4 Enteroinvasive *E. coli* (EIEC)

Bacillary dysentery in contrast to amoebic dysentery was reported in 1887 and the agent responsible for causing the disease was *Bacillus dysenteriae* in 1898 by Shiga at the time of epidemic of 89,400 incidents [48]. The significance of *Shigella* strains on medical basis resulted in their severance from *E. coli* and the recently formed genus along with its four species can be separated from *E. coli* on the premise of physiological and biochemical attributes. On the other hand, the detection of strains that might cause dysentery and acted as intermediate between *Shigella* and *E. coli* in biochemical attribute in 1944 caused the division of two genera to be cross examined [49]. It has been revealed that EIEC strains and *Shigella* species are generally very closely associated in terms of biochemically, genetically, and pathogenetically such that it has been considered to classify them together as one species in genus *Escherichia* [50-51].

Attainment of the invasive plasmid (pINV) determines the capability to attack the tissues of the host [52-54] is most likely to be the only one main event that has possibly given rise to the discovery of both *Shigella* and EIEC from commensal *E. coli.* Almost one third of this huge sole copy plasmid encodes IS elements and consists of a 30 kb region allowing the bacteria to attack enteric epithelial cells [55]. Various constituents of type three secretion system (T3SS) like translocators, transcriptional activators, few effectors and chaperones are coded by this section with the functioning of the Inv encoded genes that are controlled internationally by VirB and MxiE [30]. Also several chromosomal genes that are not particularly for *Shigella* species but are present on the chromosome are necessary for disease causing mechanism [55].

The site of infection of *Shigella* and EIEC is colonic mucosa in which entry of M cells, macrophages and epithelial cells takes place causing diarrhea with high water content also called watery diarrhea, which in critical incidents could also be followed by the starting of very little dysenteric stools having blood and mucus [56]. EIEC strains, additionally, can turn out to form a 63 kDa toxin referred as Sen which encouraged the enterotoxic activity identified in the strains carrying the gene [57].

3.5 Enteroaggregative *E. coli* (EAEC)

In both the developed and developing countries, EAEC is the second main cause of traveller's diarrhea after ETEC. This pathotype has latest been identified as diarrheagenic *E. coli.* EAEC are generally accepted as the root cause of endemic and epidemic diarrhea globally and in recent times has been known to cause acute diarrhea in infants and youngsters in developed countries [12]. The organisms in this category have also been responsible for causing persistent diarrhea. EAEC most commonly causes watery diarrhea; however it can often be followed by mucus or blood [28, 58-60]. Examining a group of *E. coli* strains causing diarrhea and that were not the part of EPEC serogroups displayed numerous strains have also sticked to HEp-2 cells and the observable characteristics were totally dissimilar from that of EPEC [61-62]. This type of adherence termed as "diffuse", was then subcategorized into aggregative and true diffuse adherence [63].

Lack of appropriate animal models and the variability of virulence factors caused the dearth of details concerning the spread of EAEC, pathogenicity and epidemiology. Nevertheless, there are many features that are considered for EAEC pathogenesis such as colonization of intestinal mucosa, mucoid biofilm development and explanation of various enterotoxins, cytotoxins and mucosal swelling or redness [28, 58-60, 64-65]. Colonization by EAEC of intestinal mucosa generally occurs by means of aggregative adherence fimbriae (AAF) which is encoded by a 55-65 MDa plasmid called pAA.

Aggregative adherence fimbriae I (AAF/I) was the first one to be cloned and distinguished from EAEC strain 17-2 [60, 66]. A probe acquired from this adhesin did not identify O42 which is the second EAEC prototype and consequently new fimbriae were identified in this strain known as AAF/II [67]. There are two other adherence factors "AAF/III and AAF/IV" along with non-fimbrial adhesin that are reported but few strains do not carry any of these known fimbriae. The attachment of EAEC to enteric tissue is through antigenically different adhesins [68, 69]. A transcriptional activator recognized as "AggR," encoded by pAAs, controls the biogenesis of AAFs [70]. It is the most important EAEC virulence regulator that regulates different virulence genes encoded by pAAs and chromosomes [58, 60].

Attachment of EAEC strains to the mucous membrane is determined by the production of a broad, aggregating mucus layer within which they survive and this biofilm formation has been credited to the activity of "*fis"* and "*yafK"* genes [71, 72]. A protein secretory in nature and of 10 kDa encoded by pAA and termed as "antiaggregation protein (Aap) or dispersin" helps the bacteria to move over the surface of the cells for successive aggregation and attachment [30, 73]. Dispersin is extremely immunogenic and is transported by means of an ATP binding cassette (ABC) called as the "Aat apparatus" [74]. Both the genes are used for recognition and categorization of EAEC isolates; however it has been reported that dispersin gene can be discovered in DAEC as well as in commensal *E. coli* [75].

3.6 Diffusely adherent *E. coli* (DAEC)

DAEC is a diversified category that creates a diffuse adherence pattern on HEp-2 and HeLa cells. It is linked with watery diarrhea that can turn out to be in young children and infants for longer period of time in both developing and urbanized countries [28, 76]. It has been revealed that the threat of diarrhea coupled with DAEC usually rises in children from 18 months to 5 years of age. It has been noted that the enteric carriage of diffusely adherent strains are prevalent in youngsters and adults [76].

Two different kinds of adhesins showing the DA pattern have been reported until now resulting in classification of DAEC strains into "AIDA-I-dependent group" and those where adhesins is encoded by various correlated operons comprising of fimbrial as well as afimbrial adhesins. The collection of these proteins is togetherly termed as "Afa-Dr adhesins" [28, 77]. The former afimbrial adhesin (*afa*) operon included in this collection of proteins was fully considered and sequenced in 1984 [78], and afterwards one more operon in this collection with the adhesins receptor were reported [79, 80]. AIDA-I is a 100 kDa surface protein coupled with diffusely adherent characteristics and was reported by Benz [81]. He also showed that this kind of attachment was not regularly seen in DEAC isolates [2, 82].

The mechanism of pathogenesis of DAEC mainly appears to be carried out through interactions of Afa/Dr adhesin with host cells. An extra released "autotransporter toxin (Sat)" has been seen to be involved in pathogenesis. On the basis of mannose-resistant diffuse adhesion of DEAC strains to cultured epithelial HEp-2 or HeLa cells, phenotypic detection of these strains has been carried out [2, 62-63]. Although, the adhesion assay is not particularly for detection of Afa/Dr DAEC because many other pathogenic *E. coli* such as EPEC strains may perhaps illustrate this adhesion pattern [83, 77].

Several phenotypic assays have been introduced, but not any of them proved suitable and globally to be used for detection and classification of the entire Afa/Dr DAEC isolates [77]. Colony hybridization by means of a variety of probes has also been used in epidemiological studies [84-86]. But this kind of technique is quite labour intensive and time consuming and is not appropriate for its use on individual strains. Designing of PCR methods mediated the characterization of every well-known Afa/Dr adhesins [87].

3.7 Novel proteins:

The protein ClpV1 family T6SS ATPase and type VI secretion system protein are the unique proteins present only in pathogenic strains of *E. coli* according to this study.

The Type VI secretion system (T6SS) consists of numerous proteins and form a “multi protein complex” broadly present in Gram negative bacteria. Its main function is to carry and release the effector toxins at targeted location [88-91]. Type VI secretion system offers competitive benefit to almost every bacteria in an environment containing various species. This is because it releases anti-bacterial toxins which consist of peptidoglycan hydrolase, phospholipase and shows DNase activity, into the bacterial cells that have been targeted [92-95]. As a result it maintains the populations of bacteria and environment colonization [96]. All these properties determine it as a chief player in pathogenesis of bacteria as well as in inter-bacterial competition. ClpV1 related with T6SS is a major component of pathogenicity of bacteria and this secretion system is basically characterised by the existence of an AAA+ Clp-like ATPase.

CHAPTER 4

MATERIALS AND METHODS

4.0 Materials and Methods

4.1 Materials

4.1.1 Bacterial Strains

A total of 10 diarrheagenic *E. coli* strains were taken. Out of which six strains were of EPEC pathotype, one was of EAEC pathotype and the remaining three were commensal *E. coli*. These previously characterized *E. coli* strains in Table 1 were screened for AST and MIC.

Table 1: Diarrheagenic *E. coli* strains

S. no.	***E. coli* strains**
1.	52s (EPEC)
2.	40s
3.	56s (EAEC)
4.	311
5.	377 (EPEC)
6.	431 (EPEC)
7.	329
8.	373 (EPEC)
9.	502 LF (EPEC)
10.	478 LF (EPEC)

4.1.2 Comparative genome analysis of pathogenic and commensal strains of *E. coli*

Diarrheagenic and commensal strains of *E. coli* were taken from NCBI along with their accession number and the genomes of these strains were downloaded.

Comparison was made between the three diarrheagenic and three commensal *E. coli* strains listed in Table 2 and Table 3 respectively.

Table 2: Diarrheagenic *E. coli* Strains

S.no.	***E. coli* Strain**	**NCBI Accession No.**	**Disease**
1.	*Escherichia coli 042*	NC_017626	Persistent diarrhea
2.	*Escherichia coli ETEC H10407*	NC_017633	Diarrhea
3.	*Escherichia coli UMNK88*	NC_017641	Diarrhea

Table 3: Commensal *E. coli* Strains

S.no.	***E. coli* Strain**	**NCBI Accession No.**
1.	***Escherichia coli str. K-12***	NC_000913
2.	*Escherichia coli DH1*	NC_017625
3.	*Escherichia coli BL21*	NC_012892

4.1.3 Primer Designing by NCBI BLAST Tool

Table 4: Primers for *clpV1* gene

Target gene	**Protein**	**Amplicon size (bp)**	**Sequence (5'-3')**	**Tm**	**GC%**
clpV1 F	*clpV1 family T6SS ATPase*	246	GTGACGCGGCTCATGATTTC	59.97	55.00
clpV1 R			GATTGTCGAACACGCCGATG	59.97	55.00

Table 5: Primers for gene of *type VI secretion system*

Protein	**Amplicon size (bp)**	**Sequence (5'-3')**	**Tm**	**GC%**
Type VI Secretion System Protein	280	CACAGACTGTACGCACTGGT	59.97	55.00
		TCACGAAACCGGTGGACAAA	60.11	50.00

4.2 Methodology

4.2.1 Protocol for AST and MIC of diarrheagenic *E. coli* strains listed in Table 1.

Single bacterial colony was inoculated into 10 mL Mueller Hinton Broth and incubated for 4 to 5 hours.

The O.D. of bacterial cultures was adjusted to obtain 0.5 McFarland turbidity.

A sterile cotton swab was dipped into the bacterial culture and pressed against wall of test tube in order to remove excess culture.

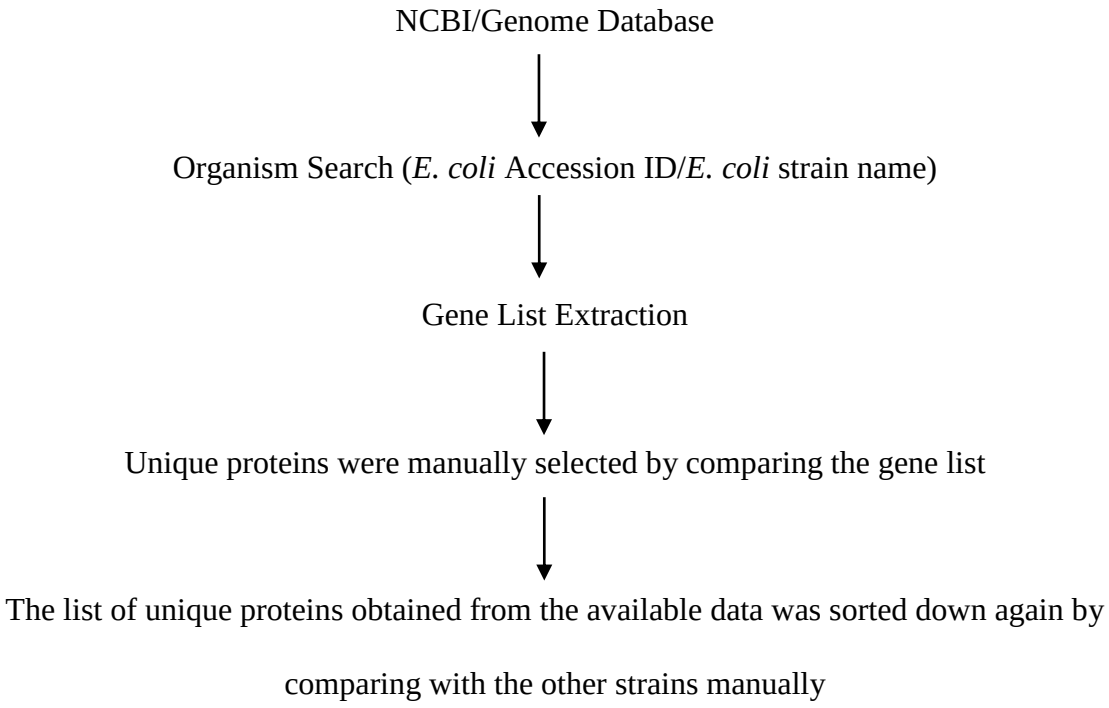
Bacterial culture were spread on the Mueller Hinton agar plates and dried for 5 minutes.

Antibiotic disks & E-strips of respective antibiotics were placed with the help of sterile forceps on Mueller Hinton agar plate and incubated for 16 to 18 hours at 37ºC.

Plates were examined and the results were interpreted according to the ICMR SOPs.

4.2.2 Methodology for Comparative Genome Analysis

E. coli strains selected for genome comparison are listed in Table 2 and Table 3. Gene list of each strain was extracted through Scripts and analysis was done manually.



4.2.3 Culturing and Isolation of Bacteria

Luria Broth (LB): It is a nutritionally rich medium used for culturing and maintenance of *E. coli* strains. It helps the *E. coli* strains to grow more rapidly as all the nutrients and essential growth nutrients are readily provided to them.

MacConkey Agar: To selectively isolate diarrheagenic *E. coli*, the culture was streaked on MacConkey agar and was incubated at 37ºC for 18-24 hours.

4.2.4 Protocol for genomic DNA isolation by using Heat Boiling Method

DNA isolation of *E. coli* strains was done by using heat boiling method. In this protocol, firstly, 2-3 colonies of overnight grown *E. coli* culture were taken and were dispensed into 20µl of nuclease free water. These dispensed colonies were heated in PCR machine for 95ºC for 10 minutes. It was then centrifuged at 3000 rpm for 5 minutes. Supernatant contains DNA and hence can act as DNA template in PCR reaction.

4.2.5 PCR Assay

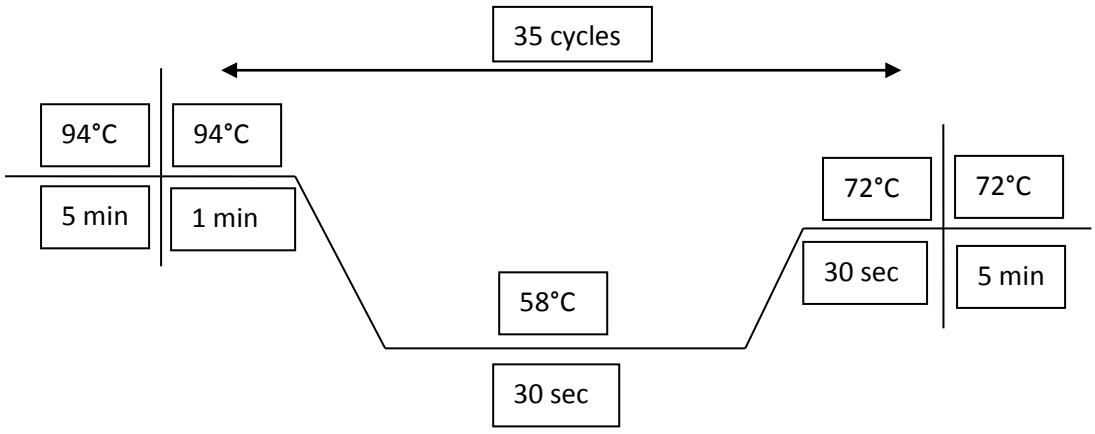
For validation of the results of comparative genome analysis PCR was set up. It is a common laboratory technique which is used to make million copies of a particular gene of DNA invitro from very minute quantity of DNA and is most commonly used in therapeutic research labs. After the completion of PCR cycle, the PCR products were made to run on 1.5% agarose gel in 1X TAE. The gel was observed in Gel Doc in order to check the amplification. Annealing temperature of gene of T6SS was: 56°C and of clpV1 gene was: 58°C.

PCR components:

Total Reaction volume: 15µl

1. Forward primer 0.5 µl
2. Reverse primer 0.5 µl
3. Nuclease free water 4.5 µl
4. DNA Template 2.0 µl
5. PCR Mastermix (Emerald) 7.5 µl

PCR conditions for *clpV1 gene*:



CHAPTER 5

RESULTS

5.0 Results

5.1 Antimicrobial Susceptibility Test Results for Diarrheagenic *E. coli* strains

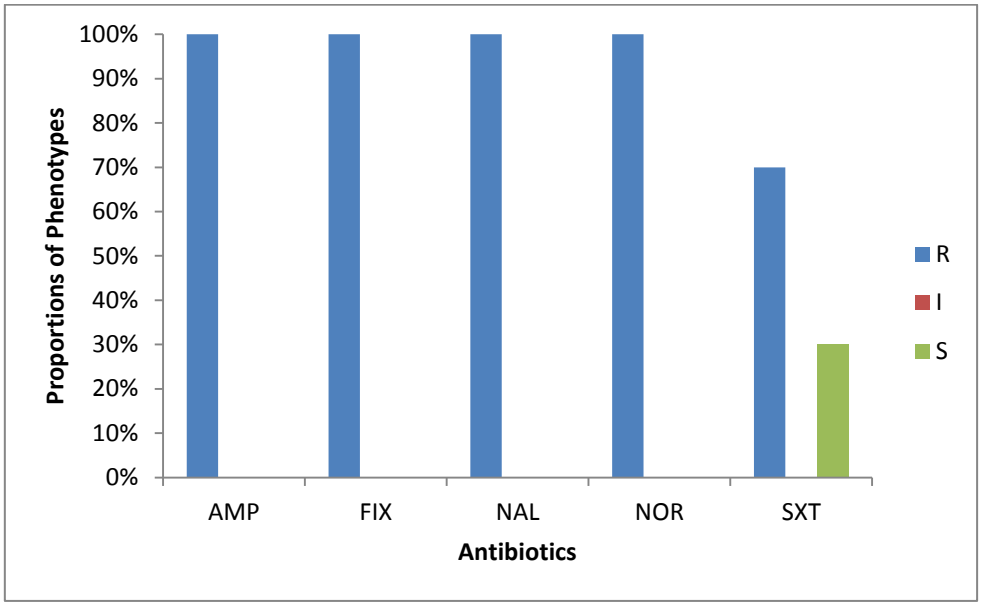


Figure1: Represents antibiogram patterns exhibited by diarrheagenic *E. coli* against commonly used antibiotics. X-axis represents the antibiotics used in antibiotic susceptibility test; AMP (Ampicillin;10µg), FIX (Cefixime;5µg), NAL (Nalidixic acid;30µg), NOR (Norfloxacin;10µg), SXT (Trimethoprimsulfamethoxazole;23.75µg). Y-axis represents proportion of Resistant (R), Intermediate (I) and Sensitive (S) phenotypes.

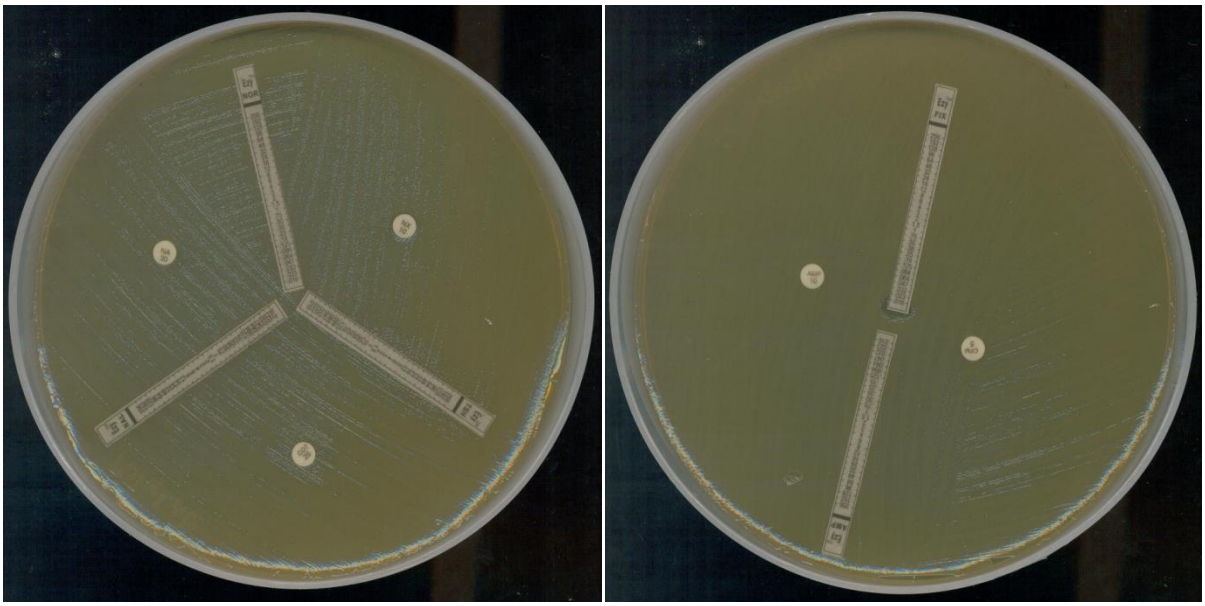


Figure 2: AST and MIC Results of 502LF (EPEC) *E. coli*

Table 6: Antimicrobial Susceptibility Test results for Diarrheagenic *E. coli* strains

(Interpretation is according to Standard Operative Procedures Bacteriology 2015, ICMR)

Strains	**Antimicrobial agent**	**Zone diameter (mm)**	**Interpretation**	**MIC (µg/ml)**	**Interpretation**
52s	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	22	Sensitive	0.50	Sensitive
40s	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	22	Sensitive	0.25	Sensitive
56s	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant
311	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	9	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant
377	Ampicillin (AMP)	Zero	Resistant	48	Resistant

	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	9	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant
431	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	9	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant
329	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	9	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	20	Sensitive	0.125	Sensitive
373	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant
502 LF	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant

478 LF	Ampicillin (AMP)	Zero	Resistant	>256	Resistant
	Cefixime (FIX)	Zero	Resistant	>256	Resistant
	Nalidixic acid (NAL)	Zero	Resistant	>256	Resistant
	Norfloxacin (NOR)	Zero	Resistant	>256	Resistant
	Sulfamethoxazole (SXT)	Zero	Resistant	>256	Resistant

5.2 Results of Comparative Genomics

1	NC_017626 (Pathogenic)	NC_017633 (Pathogenic)	NC_017641 (Pathogenic)
2	thr operon leader peptide	thr operon leader peptide	thr operon leader peptide
3	bifunctional aspartate kinase/homoserine	bifunctional aspartate kinase/homoserine	bifunctional aspartate kinase/homoserine
4	homoserine kinase	homoserine kinase	homoserine kinase
5	threonine synthase	threonine synthase	threonine synthase
6	hypothetical protein	hypothetical protein	hypothetical protein
7	hypothetical protein	hypothetical protein	hypothetical protein
8	transposase	peroxide stress protein YaaA	peroxide stress protein YaaA
9	DNA-binding protein	sodium:alanine symporter family protein	sodium:alanine symporter family protein
10	peroxide stress protein YaaA	transaldolase	transaldolase 1
11	sodium:alanine symporter family protein	molybdopterin adenylyltransferase	molybdopterin adenylyltransferase
12	transaldolase 1	succinate-acetate/proton symporter SatP	succinate-acetate/proton symporter SatP
13	molybdopterin adenylyltransferase	hypothetical protein	hypothetical protein
14	succinate-acetate/proton symporter SatP	hypothetical protein	hypothetical protein
15	hypothetical protein	molecular chaperone DnaK	molecular chaperone DnaK
16	hypothetical protein	molecular chaperone DnaJ	molecular chaperone DnaJ
17	molecular chaperone DnaK	IS4 family transposase	IS4 family transposase
18	molecular chaperone DnaJ	protein hokC	protein hokC
19	protein hokC	Na+/H+ antiporter NhaA	Na+/H+ antiporter NhaA
20	Na+/H+ antiporter NhaA	transcriptional activator NhaR	transcriptional activator NhaR
21	transcriptional activator NhaR	hypothetical protein	hypothetical protein
22	hypothetical protein	hypothetical protein	fimbrial usher protein

Figure 3: Gene Extraction List of three strains of diarrheagenic *E. coli*

1	NC_000913 (Commensal)	NC_017625 (Commensal)	NC_012892 (Commensal)
2	thr operon leader peptide	thr operon leader peptide	thr operon leader peptide
3	Bifunctional aspartokinase/homoserine	bifunctional aspartokinase I/homoserine	bifunctional aspartate kinase/homoserine
4	homoserine kinase	homoserine kinase	homoserine kinase
5	L-threonine synthase	threonine synthase	threonine synthase
6	DUF2502 family putative periplasmic protein	hypothetical protein	hypothetical protein
7	peroxide resistance protein, lowers	hypothetical protein	hypothetical protein
8	putative transporter	inner membrane transport protein	peroxide stress protein YaaA
9	transaldolase B	transaldolase B	sodium:alanine symporter family protein
10	molybdochelatase incorporating molybdenum into	molybdenum cofactor biosynthesis protein MogA	transaldolase
11	succinate-acetate transporter	hypothetical protein	molybdopterin adenylyltransferase
12	UPF0174 family protein	hypothetical protein	succinate-acetate/proton symporter SatP
13	UPF0412 family protein	hypothetical protein	hypothetical protein
14	chaperone Hsp70, with co-chaperone DnaJ	molecular chaperone DnaK	hypothetical protein
15	chaperone Hsp40, DnaK co-chaperone	molecular chaperone DnaJ	molecular chaperone DnaK
16	IS186 transposase	Gef protein	molecular chaperone DnaJ
17	regulatory protein for HokC, overlaps CDS of	pH-dependent sodium/proton antiporter	IS4 family transposase
18	toxic membrane protein, small	transcriptional activator NhaR	protein hokC
19	sRNA antisense regulator blocking mokC, and	hypothetical protein	Na+/H+ antiporter NhaA
20	sodium-proton antiporter	hypothetical protein	transcriptional activator NhaR
21	transcriptional activator of nhaA	outer membrane usher protein	hypothetical protein
22	IS1 transposase B	fimbrial chaperone	fimbrial usher protein
23	IS1 repressor TnpA	fimbrial protein	hypothetical protein
24	30S ribosomal subunit protein S20	hypothetical protein	transposase

Figure 4: Gene Extraction List of three strains of commensal *E. coli*

Table 7: Total number of Resulted Proteins from the Gene Extraction List

Strains of *E. coli* (diarrheagenic/commensal)	**Total no. of proteins**
Escherichia coli 042 (NC_017626)	5216
Escherichia coli ETEC H10407 (NC_017633)	5197
Escherichia coli UMNK88 (NC_017641)	5273
Escherichia coli str. K-12 **(**NC_000913**)**	4701
Escherichia coli DH1(NC_017625)	5360
Escherichia coli BL21(NC_012892)	4526

Table 8: According to the gene extraction list, proteins uniquely present in the pathogenic strains of *E. coli*

type VI secretion system protein	alpha/beta hydrolase
ClpV1 family T6SS ATPase	fimbria adhesin EcpD
type VI secretion system protein ImpI	usher protein EcpC
group II intron reverse transcriptase/maturase	Replication protein 14
copper resistance protein B	protein ninB
flagellar M-ring protein FliF	protein ninX
MbtH family protein	DNA circularization protein
hydrolase YafV	replication endonuclease
phage baseplate protein	phage N-6-adenine-methyltransferase
type III effector	endoribonuclease YbeY
protein AbrB	cyclic pyranopterin phosphate synthase MoaA
toxin RelE	molybdopterin synthase catalytic subunit

- This list was again sorted down by comparing with the other remaining commensal strains to be specific.
- Three unique proteins were found mainly playing a key role in pathogenesis of *E. coli* that are type III effector protein, clpV1 family T6SS ATPase and type VI secretion system protein.

5.2 PCR Assay Results *for gene of type VI secretion system*

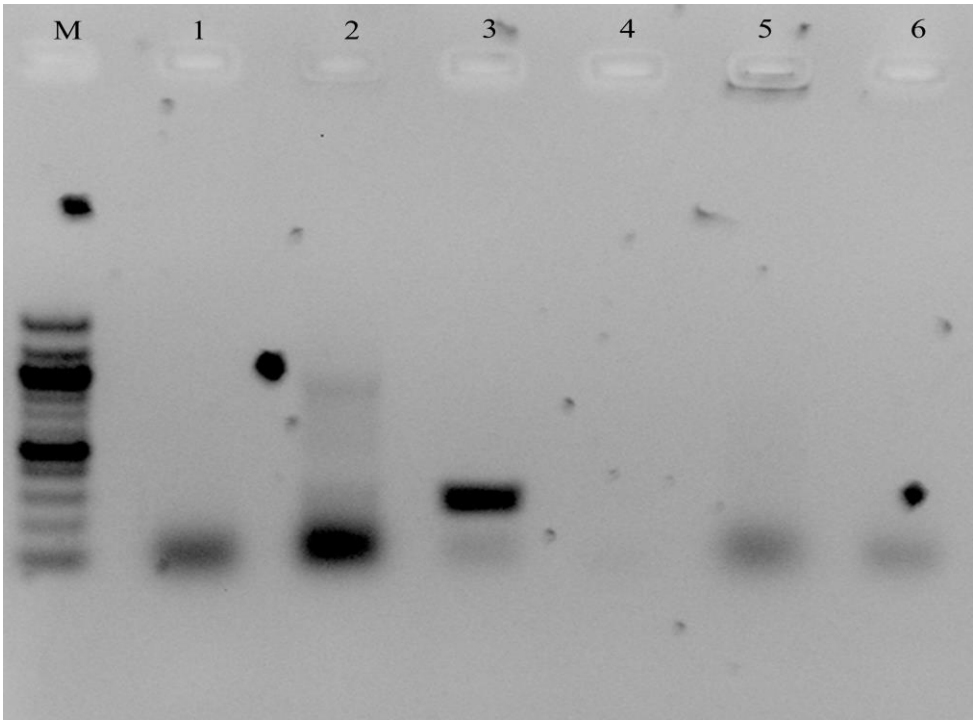


Figure 5: Gel electrophoresis image of 1.5% agarose gel depicting the detection of gene of type VI secretion system in pathogenic *E. coli* and non-pathogenic *E. coli*. Lane M represents ladder of 100 bp (NEB). Lane 1: *E. coli 25922;* Lane 2: *E. coli DH5α;* Lane 3: EPEC pathotype (28s); Lane 4: EAEC pathotype (29s); Lane 5: *E. coli* strain 478 and Lane 6: *E. coli 12579*. The amplicon size of the gene is 280 bp. Only lane 3 shows amplification at approximately 280 bp depicting the presence of the gene in pathogenic *E. coli.*

5.3 PCR Assay Results *for clpV1* gene

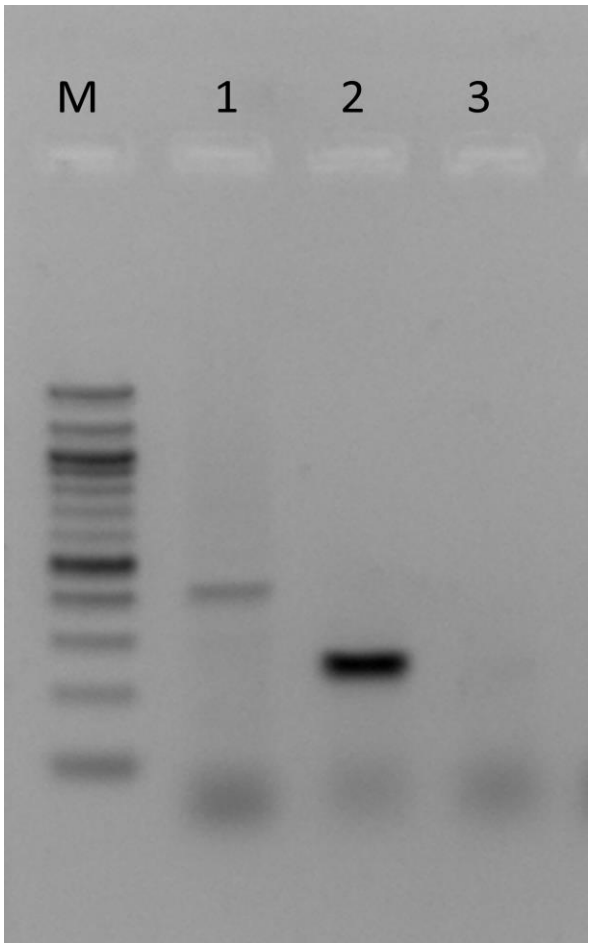


Figure 6: Gel electrophoresis image of 1.5% agarose gel depicting the detection of *clpV1* gene in pathogenic *E. coli*. Lane M represents ladder of 100 bp (NEB), Lane 1: *E. coli DH5α*, Lane 2: EPEC pathotype (28s) and Lane 3 represents negative control. The amplicon size of the gene is 246 bp. In lane 1, *E. coli DH5α* (commensal) shows non-specific amplification at approximately 400 bp and in lane 2, EPEC shows amplification for the gene *clpV1* at approximately 246 bp.

CHAPTER 6

DISCUSSION AND CONCLUSION

6.0 Discussion and Conclusion

In this project, while sorting the listed proteins we were successful in finding two unique genes/proteins that are only present in pathogenic *E. coli* and absent in commensal *E. coli.* The unique proteins were ClpV1 family T6SS ATPase and Type VI Secretion System Protein. *clpV1 gene* is associated with type six secretory system and plays a major role in bacterial pathogenesis. T6SS is responsible for delivering the toxins at target cells.

From the comparative analysis data, it was observed that there is a difference in the number of proteins due to the difference in the size of genomes of selected strains of *E. coli*. According to the gene extraction list, *Escherichia coli DH1*, a commensal strain, have the largest number of proteins among all the selected strains. Difference in the number of proteins justifies that there must be unique proteins that are present either in pathogenic strain or in the commensal strains. Hence, by manual searching, ClpV1 family T6SS ATPase and Type VI Secretion System Protein were discovered.

As a result, it helped in rapid identification of the pathogen directly from the samples overcoming all the tedious and labor intensive processes. This has been validated through PCR. *clpV1 gene* showed amplification in pathogenic *E. coli* whereas commensal strain showed non-specific amplification and it can be modified by standardizing the conditions of the PCR cycle. Gene for type VI secretion system protein showed amplification only in pathogenic *E. coli.* Hence, proves that our protein is a unique protein present only in pathogenic strains and differentiates the pathogenic and non-pathogenic *E .coli* easily. PCR is the best technique for rapid detection of the pathogen.

Therefore, by achieving the aim of rapid and specific detection of the pathogen through this study, we can initiate the proper and adequate treatment of the patient as soon as possible which will further help in curing the disease and reduces the chances of complications and deaths in infants and children.

CHAPTER 7

REFERENCES

7.0 References

1. World Health Organization. *World Health Statistics*. Geneva, Switzerland: WHO Press; 2012.
2. Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. *ClinMicrobiol Rev*. 1998; 11(1): 142–201.
3. Ewing WH. *Edwards and Ewing's Identification of Enterobacteriaceae*. 4th ed. New York: Elsevier; 1986.
4. Drasar, B. S., and M. J. Hill, Human intestinal flora, Academic Press, Ltd., London, United Kingdom. 1974; 36-43.
5. Conway PL. Microbial ecology of the human large intestine. In: Gibson GR, Macfarlane GT, eds. *Human Colonic Bacteria: Role in Nutrition, Physiology and Pathology*. Boca Raton, FL, USA: CRC Press; 1995:1–24.
6. Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. *Nat Rev Microbiol*. 2004; 2(2): 123–140.
7. Gomes TA, Elias WP, Scaletsky IC, Guth BE, Rodrigues JF, Piazza RM, Ferreira LC, Martinez MB. Diarrheagenic Escherichia coli. brazilian journal of microbiology. 2016 Dec 31; 47: 3-0.
8. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M,Finlay BB. Recent advances in understanding entericpathogenic *Escherichia coli*. *Clin Microbiol Rev*.2013; 26(4): 822–880.
9. Rolhion N, Darfeuille-Michaud A. Adherent-invasive *Escherichia coli* in inflammatory bowel disease. *Inflamm BowelDis*. 2007; 13(10): 1277–1283.
10. Cieza RJ, Cao AT, Cong Y, Torres AG. Immunomodulation for gastrointestinal infections. *Expert Rev Anti Infect Ther*.2012; 10(3): 391–400.
11. Smith. A, Microbiology and Pathology. ninth edition: 723.
12. Jafari A, Aslani MM, Bouzari S. Escherichia coli: a brief review of diarrheagenic pathotypes and their role in diarrheal diseases in Iran. Iranian journal of microbiology. 2012 Sep; 4(3): 102-7.
13. Bray J. Isolation of antigenically homogeneous strains of *Bacterium coli neopolitanum* from summer diarrhoea of infants. *J Pathol Bacteriol*. 1945; 57(2): 239–247.

14. Robins-Browne RM. Traditional Enteropathogenic *Escherichia coli* of infantile diarrhea. *Rev Infect Dis*.1987; 9(1): 28–53.
15. Trabulsi LR, Keller R, Gomes TAT. Typical and a typical enteropathogenic *Escherichia coli*. *Emerg Infect Dis*.2002; 8(5): 508–513.
16. Ørskov F, Whittam TS, Cravioto A, Ørskov I. Clonal relationships among classic Enteropathogenic *Escherichia coli* (EPEC) belong to different O groups. *J Infect Dis*. 1990; 162(1): 76–81.
17. Whittam TS, McGraw EA. Clonal analysis of EPEC serogroups. *Rev Microbiol*. 1996; 27: 7–16.
18. Gomes TAT, González-Pedrajo B. Enteropathogenic *Escherichia coli* (EPEC). In: Torres AG, ed. *Pathogenic Escherichia coli in Latin America*. Sharjah, United ArabEmirates: Betham Science Publishers Ltd.; 2010: 66–126.
19. Gomes TAT, Griffin PM, Ivey C, Trabulsi LR, Ramos SRTS.EPEC infections in São Paulo. International Symposium onEnteropathogenic *Escherichia coli* (EPEC), São Paulo, SP. *RevMicrobiol Soc Bras Microbiol*. 1996; 27: 25–33.
20. Chen HD, Frankel G. Enteropathogenic *Escherichia coli*: unraveling pathogenesis. *FEMS Microbiol Rev* 2005; 29: 83-98.
21. Staley TE, Jones EW, Corley LD. Attachment and penetration of *Escherichia coli* into intestinal epithelium of the ileum in newborn pigs. *Am J Pathol* 1969; 56: 371-392.
22. McDaniel TK, Jarvis KG, Donnenberg MS, Kaper JB. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *PNAS* 1995; 92: 1664-1668.
23. Cravioto A, Gross RJ, Scotland SM, Rowe B. An adhesive factor found in strains of *Escherichia coli* belonging to the traditional infantile enteropathogenic serotypes. *Curr Microbiol* 1979; 3: 95-99.
24. Hill DR, Beeching NJ. Travelers' diarrhea. *Curr Op Infect Dis* 2010; 23: 481-487.
25. Wiedermann U and Kollaritsch H. Vaccines against traveler's diarrhea and rotavirus disease – a review. Wein Klin Wochenscher 2006; 118: 2-8.
26. De SN, Bahatachary 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.
27. Sack SB. The discovery of cholera-like enterotoxins produced by *Escherichia coli* causing secretory diarrhea in humans. *Indian J Med Res* 2011; 133: 171-178.
28. Croxen MA, Finlay BB. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* 2010; 8: 26-38.

29. Isidean SD, Riddle MS, Savarino SJ, Porter CK. A systemic review of ETEC epidemiology focusing on colonization antigen factor and toxin expression. *Vaccine* 2011; 29: 6167-6178.
30. Johnson TJ, Nolan LK. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol Mol Biol Rev* 2009; 73: 750-774.
31. Nagy B, Fekete PZS. Enterotoxigenic *Escherichia coli* in veterinary medicine. *Int J Med Microbiol* 2005; 295: 443-454.
32. Beddoe T, Paton AW, Le Nours J, Rossjohn J, Paton JC. Structure, biological functions of the AB5 toxins. *Cell* 2010; 35: 411-418.
33. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA Sommerfelt H Steinsland H. Molecular mechanism of enterotoxigenic *Escherichia coli* infection. *Microbes Infect* 2010; 12: 89-98.
34. Melton-Celsa AR. Shiga toxin (Stx) classification, structure, and function. *Microbiol Spectr*. 2014;2(3). EHEC-0024-2013.
35. Kaper JB, O'Brien AD. Overview and historical perspectives. *Microbiol Spectr*. 2014;2(2). EHEC-0028-2014.
36. Gould LH, Mody RK, Ong KL, et al. Increased recognition of non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States during 2000–2010: epidemiologic features and comparison with *E. coli* O157 infections. *Food Pathog Dis*. 2013; 10:453–460.
37. Muniesa M, Hammerl JA, Stefan Hertwig S, Appel B,Brüssow H. Shiga toxin-producing *Escherichia coli* O104:H4: a new challenge for microbiology. *Appl Env Microbiol*.2012; 78:4065–4073.
38. Bletz S, Bielaszewska M, Leopold SR, et al. Evolution of enterohemorrhagic *Escherichia coli* O26 based on single-nucleotide polymorphisms. *Genome Biol Evol*.2013; 5:1807–1816.
39. Rivas M, Chinen I, Guth BEC. Enterohemorrhagic (Shigatoxin-producing) *Escherichia coli*. In: Torres AG, ed. *Escherichia coli in the Americas*. Springer International Publishing; 2016:97–123.
40. Caprioli A, Morabito S, Brugère H, Oswald E. Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet Res* 2005; 36: 289-311.
41. Low AS, Holden N, Rosser T, Roe AJ, Constantinidou C, Hobman JL, *et al*. Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. *Environ Microbiol* 2006; 8: 1033-1047.

42. Xicohtencatl-Cortes J, Monteiro-Neto V, Saldaña Z, Ledesma MA, Puente JL, Girón JA. The type 4 Pili of enterohemorrhagic *Escherichia coli* O157:H7 are multipurpose structures with pathogenic attributes. *J Bacteriol* 2009; 191: 411-421.
43. Mora A, Blanco M, Yamamoto D, Dahbi G, Blanco JE, López C, *et al*. HeLa-cell adherence patterns and actin aggregation of enteropathogenic *Escherichia coli* (EPEC) and Shiga-toxin-producing *E. coli* (STEC) strains carrying different *eae* and *tir* alleles. *Int Microbiol* 2009; 12: 243-251.
44. Gyles CL. Shiga-producing *Escherichia coli*: an overview. *J Anim Sci* 2007; 85(E. Suppl.): E45-E62.
45. Savarino SJ, McVeigh A, Watson J, Cravioto A, Molina J, Echeverria P, *et al*. Enteroaggregative *Escherichia coli* heat-stable enterotoxin is not restricted to enteroaggregative *E. coli*. *J Infect Dis* 1996; 173: 1019-1022.
46. Lim JY, Yoon JW, Hovde CJ. A brief overview of Escherichia coli O157:H7 and its plasmid O157. *J Microbiol Biotechnol* 2010a; 20: 1-10.
47. Lim JY, Hong JB, Sheng H, Shringi S, Kaul R, Besser TE, *et al*. Phenotypic diversity of *Escherichia coli* O157:H7 strains associated with the plasmid O157. *J Microbiol* 2010; 48: 347-357.
48. Peng J, Yang J, Jin Q. The molecular evolutionary history of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infect Genet Evol* 2009; 9: 147-152.
49. Van den Beld MJC, Reubsaet FAG. Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*. *Eur J Clin Microbiol Infect Dis* 2012; 31: 899-904.
50. Benner DJ, Fanning GR, Steigerwalt AG, Orskov I, Orskov F. Polynucleotide sequence relatedness among three groups of pathogenic *Escherichia coli* strains. *J Bacteriol* 1972; 109: 953-965.
51. Lan R, Alles MC, Donohoe K, Martinez MB, Reeves PR. Molecular evolutionary relationship of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 2004; 72: 5080-5088.
52. Silva RM, Toledo MR, Trabulsi LR. Biochemical and cultural characteristics of invasive *Escherichia coli*. *J Clin Microbiol* 1980; 11: 441-444.
53. Hale TL, Sansonetti PJ, Schad PA, Austin S, Formal SB. Characterization of virulence plasmids and plasmid associated outer membrane proteins of *Shigella flexneri, Shigella sonnei* and *Escherichia coli*. *Infect Immun* 1983; 40: 340-350.

54. Harris JR, Watchsmoth IK, Davis BR, Cohen ML. High molecular-weight plasmid correlates with *Escherichia coli* enteroinvasiveness. *Infect Immun* 1982; 73: 1295-1298.
55. Sansonetti, PJ. *Escherichia coli*, *Shigella*, antibiotic associated diarrhea, and prevention and treatment of gastroenteritis. *Curr Opin Infect Dis* 1992; 5: 66-73.
56. Nataro, JP, Seriwatana J, Fasano A, Maneval D, Guers LD, Noriega F, Dubovsky F, Levine MM, Morris Jr. JG. Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and *Shigella* strains. *Infect Immun* 1995; 63: 4721-4728.
57. Venkatesan MM, Buysse JM, Kopecko DJ. Use of *Shigella flexneri ipa*C and *ipa*H gene sequences for the general identification of *Shigella* spp and enteroinvasive *Escherichia coli*. *J Clin Microbiol* 1989; 27: 2687-2691.
58. Weintraub A. Enteroaggregative *Escherichia coli*: epidemiology, virulence and detection. *J Med Microbiol* 2007; 56: 4-8.
59. Harrington SM, Dudley EG, Nataro JP. Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiol Lett* 2006; 254: 12-18.
60. Scaletsky IC, Silva ML, Trabulsi LR. Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. *Infect Immun* 1984; 45: 534-536.
61. Nataro JP, Baldini MM, Kaper JB, Black RE, Bravo N, Levine MM. Detection of an adherence factor of enteropathogen*ic Escherichia coli* with a DNA probe. *J Infect Dis* 1985; 152: 560-565.
62. Nataro JP, Kaper JB, Robins-Browne R, Prado V, Vial P, Levine MM. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr Infect Dis J* 1987; 6: 829-831.
63. Bouzari S, Jafari A, Farhoudi-Moghaddam AA, Shokouhi F, Parsi M. Adherence of non-enteropathogenic *Escherichia coli* to HeLa cells. *J Med Microbiol* 1994; 40: 95-7.
64. Boisen N, Ruiz-Perez F, Scheutz F, Krogfelt KA, Nataro JP. High prevalence of serine protease autotransporter cytotoxins among strains of Enteroaggregative *Escherichia coli*. *Am J Trop Med Hyg* 2009; 80: 294-301.
65. Huang DB, Mohanty A, Dupont HL, Okhuysen PC, Chiang T. A review of an emerging enteric pathogen: enteroaggregative *Escherichia coli*. *J Med Microbiol* 2006; 55: 1303-1311.

66. Czeczulin JR, Balepur S, Hicks S, Hicks S, Henderson IR, Le Bouguénec C. Aggregative adherence fimbriae II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect Immun* 1997; 65: 4135-4145.
67. Bernier C, Gounon P, Le Bouguénec C. Identification of an aggregative adhesion fimbriae (AAF) type IIIencoding operon in enteroaggregative *Escherichia coli* as a sensitive probe for detecting the AAF-encoding operon family. *Infect Immun* 2002; 70: 4302-4311.
68. Bouzari S, Jafari A, Zarepour M. Distribution of virulence related genes among enteroaggregative *Escherichia coli* isolates: using multiplex PCR and hybridization. *Infect Genet Evol*. 2005; 5: 79-83.
69. Nataro JP, Yikang D, Yingkang D, Walker K. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J Bacteriol* 1994; 176: 4691-4699.
70. Sheikh J, Hicks S, Dall'Agnol M, Phillips AD, Nataro JP. Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol Microbiol* 2001; 41: 983-997.
71. Nataro JP, Deng Y, Maneval DR, German AL, Martin WC, Levine MM. Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infect Immun* 1992; 60: 2297-2304.
72. Sheikh J, Czeczulin JR, Harrington S, Hicks S, Henderson IR, Le Bouguénec C, *et al*. A novel dispersin protein in enteroaggregative *Escherichia coli*. *J Clin Invest* 2002; 110: 1329-1337.
73. Nishi J, Sheikh J, Mizuguchi K, Luisi B, Burland V, Boutin A, et al. The Export of Coat Protein from Enteroaggregative *Escherichia coli* by a Specific ATP binding Cassette Transporter System. *J Biol Chem* 2003; 278: 45680-45689.
74. Monteiro BT, Campos LC, Sircili MP, Franzolin MR, Bevilacqua LF, Nataro JP, *et al*. The dispersin-encoding gene (aap) is not restricted to enteroaggregative *Escherichia coli*. *Diagn Microbiol Infect Dis* 2009; 65: 81-84.
75. Wakimoto N, Nishi J, Sheikh J, Nataro JP, Sarantuya J, Iwashita M, *et al*. Quantitative biofilm assay using a microtiter plate to screen for enteroaggregative *Escherichia coli*. *Am J Trop Med Hyg* 2004; 71: 687-690.
76. Servin A. Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clin Microbiol Rev*. 2005; 18: 264-292.

77. Labigne-Roussel AF, Lark D, Scoolink G, Falkow S. Cloning and expression of an afimbrial adhesin (Afa-I) responsible for P blood group-independent mannoseresistant hemagglutination from a pyelonephritic *Escherichia coli*. *Infect Immun*1984; 46: 251-259.
78. Nowicki B, Barrish JP, Korhonen T, Hull RA, Hull SI. Molecular cloning of the *Escherichia coli* O75X adhesin. *Infect Immun* 1987; 55: 3168-3173.
79. Nowicki B, Labigne, A, Moseley S, Hull R, Hull S, Moulds J. The Dr hemagglutinin, afimbrial adhesins Afa-I and Afa-III, and F1845 fimbriae of uropathogenic and diarrhea-associated *Escherichia coli* belong to a family of hemagglutinins with Dr receptor recognition. *Infect Immun* 1990; 58: 279-281.
80. Benz I, Schmidt MA. Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect Immun* 1989; 57: 1506-1511.
81. Benz I, Schmidt MA. AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesized via a precursor molecule. *Mol Microbiol* 1992; 6: 1539-1546.
82. Bilge SS, Clausen CR, Lau W, Moseley SL. Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *J Bacteriol* 1989; 171: 4281-4289.
83. Bouzari S, Jafari MN, Shokouhi F, Parsi M, Jafari A. Virulence-related sequences and adherence patterns in strains of enteropathogenic *Escherichia coli*. *FEMS Microbiol Letts* 2000; 185: 89-93.
84. Stapleton A, Moseley S, Stamm WE. Urovirulence determinants in *Escherichia coli* isolate causing firstepisode and recurrent cystitis in women. *J Infect Dis* 1991; 163: 773-779.
85. Foxman B, Zhang L, Palin K, Tallman P, Marrs CF. Bacterial virulence characteristics of *Escherichia coli* isolates from first-time urinary tract infection. *J Infect Dis* 1995; 171: 1514-1521.
86. Le Bouguénec C, Archambaud M, Labigne A. Specific detection of the pap, afa, and sfa adhesin-encoding operons in uropathogenic *Escherichia coli* strains by polymerase chain reaction. *J Clin Microbiol* 1992; 30: 189-1193.
87. Le Bouguénec C, Lalioui L, du Merle, L, Jouve M, Courcoux P, Bouzari S**,** et al. Characterization of AfaE adhesins produced by extraintestinal and intestinal human

Escherichia coli isolates: PCR assays for detection of Afa adhesins that do or do not recognize Dr blood group antigens. *J Clin Microbiol* 2001; 39: 1738-1745.

88. Bingle, L. E., Bailey, C. M. & Pallen, M. J. Type VI secretion: a beginner's guide. *Curr Opin Microbiol* 2008; 11: 3–8.
89. Cascales, E. The type VI secretion toolkit. *EMBO Rep* 2008; 9: 735–741.
90. Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y. & Attree, I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* 2009; 10: 104.
91. Coyne, M. J., Roelofs, K. G. & Comstock, L. E. Type VI secretion systems of human gut Bacteroidales segregate into three genetic architectures, two of which are contained on mobile genetic elements. *BMC Genomics* 2016; 17: 58.
92. Hood, R. D. *et al.* A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. *Cell Host Microbe* 2010; 7: 25–37.
93. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* 2011; 475: 343–347.
94. Durand, E., Cambillau, C., Cascales, E. & Journet, L. VgrG, Tae, Tle, and beyond: the versatile arsenal of Type VI secretion effectors. *Trends Microbiol* 2014; 22: 498–507.
95. Alcoforado Diniz, J., Liu, Y. C. & Coulthurst, S. J. Molecular weaponry: diverse effectors delivered by the Type VI secretion system. *Cell Microbiol* 2015; 17: 1742–1751.
96. Sana, T. G. *et al.* Salmonella Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. *Proc Natl Acad Sci USA* 2016; 113; 5044–5051.

CHAPTER 8

APPENDIX

8.0 Appendix

8.1 Bacteriological Media

Media were prepared in distilled water and autoclaved at 15 pounds per square inch for 15 minutes unless indicated.

8.1.1 Luria Broth (HIMEDIA Laboratories)

Procedure: 20.0 grams of LB was suspended in 1000 ml of distilled water.

8.1.2 MacConkey Agar (HIMEDIA Laboratories)

Procedure: 55.07 grams were suspended in 1000 ml distilled water. After autoclaving, it was left for cooling down to 45-50 °C and then poured into sterile petriplates.

8.1.3 Mueller Hinton Agar (HIMEDIA Laboratories)

Procedure: 38.0 grams were suspended in 1000 ml distilled water. It was heated to boiling so that the medium get dissolved completely. After autoclaving, it was left for cooling down to 45-50°C and then poured into sterile petriplates.

8.2 Buffer for Electrophoresis

TAE Buffer (1X)

10 ml of 50X TAE buffer was added in 490 ml of distilled water making up the total volume of 500 ml.