DIFFERENTIAL PROTEIN PROFILING OF PATHOGENIC AND NON PATHOGENIC *Escherichia coli*

THESIS

Submitted to

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology, Solan

In the partial fulfillment of the requirement for the degree of

Master of Technology in Biotechnology

By

Monika Choudhary

Roll No. 152551

M. Tech Biotechnology [4th semester]

JUIT Solan (H.P)

Under the Supervision of

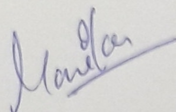
Dr. Jitendraa Vashistt, Assistant Professor (senior grade)

JUIT Solan (H.P)

DECLARATION

I hereby declare that the present work on **"Differential protein profiling of pathogenic and non pathogenic *Escherichia coli*"** is a record of original work done by me under the supervision of **Dr. Jitendraa Vashistt** Assistant **Professor (Senior Grade)** at JUIT Solan (H.P), during July 2016 to May 2017 at **Microbiology lab of Department of Biotechnology and Bioinformatics**, Jaypee University of Information Technology, Solan.

I also declare that no part of this thesis has previously been submitted to any University or any examining body for acquiring any degree.



DATE: 2nd May. 2017

MONIKA CHOUDHARY

PLACE: JUIT Waknaghat

M. Tech Biotechnology

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
(Established by H.P. State Legislature vide Act No. 14 of 2002)
P.O. Waknaghat, Teh. Kandaghat, Distt. Solan - 173234 (H.P.) INDIA
Website : www.juit.ac.in
Phone No. +91-01792-257999 (30 Lines)
Fax : +91-01792-245362

CERTIFICATE

This is to certify that thesis entitled "**Differential protein profiling of pathogenic and non pathogenic *Escherichia coli***", submitted by **Monika Choudhary** in partial fulfillment for the award of degree of **Master of Technology in Biotechnology** to Jaypee University of Information Technology, Waknaghat, Solan has been made under my supervision.

This report has not been submitted partially or fully to any other University or Institute for the award of this or any other degree or diploma.

Jitendraa

Date 2nd May, 2017

Place JUIT, Waknaghat

Dr. Jitendraa Vashistt

Assistant Professor (Senior Grade)

CONTENTS

Page No.

1. Abbreviations i-ii

2. List of Figures iii

3. Abstract 1

4. CHAPTER 1: Introduction 2-4

5. CHAPTER 2: Review of Literature

2.1 Diarrhoea 6

2.2 *Escherichia coli* 6

2.2.1 Origin of *E. coli* pathogenicity 6-7

2.3 Pathogenicity caused by *E. coli*

2.3.1 Enteropathogenic *E. coli* (EPEC) 7

2.3.2 Enterotoxigenic E. coli (ETEC) 7

2.3.3 Enteroaggregative *E. coli* (EAEC) 8

2.3.4 Enterohaemorrhagic *E. coli* (EHEC) 9

2.3.5 Enteroinvasive *E. coli* (EIEC) 9

2.3.6 Diffusely adherent E. coli (DAEC) 9

2.4 Diagnosis 9-10

2.5 Treatment 10-11

2.6 Proteome dynamics 11

2.7 Protein expression studies 11-12

2.8 Development in proteome studies 12-13

AIMS AND OBJECTIVES 14

6. CHAPTER 3: Material and methods

Work plan 16

3.1 Materials 17

3.2 Procedure

3.2.1 Culturing of *E. coli* strains and confirmation with PCR amplification 17-18

3.2.2 Growth curve 18

3.2.3 Protein extraction from all *E. coli* strains 18

3.2.2.1 Secretory proteins from media 19
3.2.2.2 Cytosolic protein extraction 19
3.2.2.3 Protein precipitation using TCA-Acetone 19

3.2.4 Protein Quantification using Bradford Assay 19-20

3.2.5 Protein separation using SDS Polyacrylamide gel electrophoresis 20

3.2.6 Protein separation with 2D PAGE
3.2.6.1 Rehydration of 11cm broad range IEF strips 20
3.2.6.2 Isoelectric focusing of strips 21
3.2.6.3 Equilibration and SDS-PAGE 21
3.2.6.4 Staining and destaining 21

7. CHAPTER 4: Result and Discussion
4.1 Culturing of *E. coli* cells and 23
4.2 Confirmation of *E. coli* strains with PCR 23
4.3 Growth characteristics of pathogenic and non pathogenic *E. coli* 24
4.4 Protein Extraction and Precipitation 25
4.5 Protein Quantification using Bradford Assay 25
4.6 Protein separation using SDS PAGE 26-28
4.7 Secretory protein analysis using 2D gel electrophoresis 28-29

8. CHAPTER 5: Conclusion 30

9. References 31-37

10. Appendix 38-40

ABBEREVERATIONS

A/E	-	attaching and effacing
APS	-	ammonium persulphate
βME	-	β-mercaptoethanol
BFP	-	Bundle forming pili
bp	-	base pair
BSA	-	Bovine serum albumin
ºC	-	degree Celsius
DAEC	-	diffusely adherent *E. coli*
2D	-	2 dimensional
DNA	-	deoxyribonucleic acid
EAEC	-	Enteroaggregative *E. coli*
EDTA	-	Ethidium diamine tetra acetate
EHEC	-	Enterohemorrhagic *E. coli*
EIEC	-	Enteroinvasive *E. coli*
EPEC	-	Enteropathogenic *E. coli*
ETEC	-	Enterotoxigenic *E. coli*
IEF	-	Isoelectric Focusing
IPG	-	Immobilized pH Gradient strips
LB	-	Luria broth
LEE	-	locus of enterocyte effacement
LTs	-	heat labile enterotoxins
μg	-	Microgram
μl	-	Micro litre
min	-	Minutes
ml	-	Milli Litre
MNEC	-	neonatal meningitis *E. coli*
nm	-	Nanometer
ORS	-	oral rehydration solution
PAGE	-	Polyacrylamide gel electrophpresis
PAI	-	pathogenicity island
PCR	-	Polymerase chain reaction
pI	-	Isoelectric point

PTM	-	post-translational modifications
rpm	-	Revolution per minute
rRNA	-	Ribosomal ribonucleic acid
SDS	-	Sodium deodecyl sulphate
STs	-	heat stable enterotoxins
Stx	-	shiga toxin
TCA	-	Tri Carboxylic acid
TEMED	-	Tetramethylethylenediamine
UPEC	-	Uropathogenic *E. coli*
V/V	-	Volume/Volume
vol	-	Volume
W/V	-	Weight/ Volume
WHO	-	world health organization

LIST OF FIGURES

Page no.

Figure 1 Schema of pathogenicity caused by diarrheagenic *E. coli* 8

Figure 2 PCR reaction for the amplification of V3 region of 16s rRNA 18

Figure 3 Pure colonies of *E. coli* strains after streaking on MacConkey agar plate 23

Figure 4 PCR product of V3 region of 16s rRNA using 1.2% agarose gel 23

Figure 5 Comparative analysis of growth characteristics of commensal and pathogenic *E. coli* strains 24

Figure 6 Standard curve for BSA protein estimation using Bradford assay 25

Figure 7 SDS polyacrylamide gel electrophoresis of cytosolic proteins of pathogenic and non pathogenic *E. coli* strains 26

Figure 8 SDS polyacrylamide gel electrophoresis of secretory proteins of pathogenic and non pathogenic *E. coli* strains 27

Figure 9 SDS polyacrylamide gel electrophoresis of secretory proteins of non pathogenic *E. coli* and pathogenic ETEC 28

Figure 10 2D polyacrylamide gel electrophoresis of secretory proteins of *E. coli* 26

ABSTRACT

In India, diarrhea is third most leading cause of deaths in children under five years of age. A variety of pathogens including bacteria, viruses and protozoan are responsible for the cause of diarrhea. Higher morbidity and mortality rates associated with diarrhea may be attributed to the insufficient diagnostic techniques and improper management of diarrhea symptoms. Commensal strains of *E. coli* constitute essential component of human gastrointestinal tract, however pathogenic ones cause debilitating infection in the host. *E. coli* strains causing gastrointestinal infection are categorized as diarrhoeagenic *E. coli* and are classified into six various pathotypes. Each of these pathotype expresses different genomic elements to demonstrate unique mechanism of pathogenicity. The most widely used characterization method is polymerase chain reaction (PCR) targeting unique virulence genes. But the high cost, sophisticated procedure and requirement of pre training make limited use this method in diagnosing diarrhoeagenic *E. coli*.

This study was conducted to obtain differential protein profiles of pathogenic and non pathogenic *E. coli*. Initially growth characteristics of pathogenic and non pathogenic *E. coli* were observed and then secretory and cytosolic proteome were extracted from all strains. Then cytosolic and secretory proteome were analyzed on SDS PAGE gel to differentiate between pathogenic and non pathogenic strains. 2D analysis of secretory proteome of pathogenic and non pathogenic *E. coli* showed differential expression profiles. The present study clearly indicated expression of unique proteins in pathogenic *E. coli* proteome as compared to commensal strains.

CHAPTER-1

INTRODUCTION

Diarrhea is a prime killer among children under five years of age, accounting for 9 per cent of child mortality worldwide (UNICEF data; 2016). Nine percentage deaths transcribe about 1,400 young children each day, or ~0.5 million children per year. More than 90% of diarrhea cases are reported in developing countries. In India diarrhea is the third most common cause of death in under-five children, responsible for 13% deaths, which is about 300,000 children each year (Lakshminarayanan and Jayalakshmy, 2015).

A number of pathogens are known as etiologic agents of diarrhea that includes bacteria, protozoan and viruses. The causative bacterial agents include diarrheagenic *Escherichia coli, Shigella, Aeromonas*, *Campylobacter*, *Salmonella*, *Vibrio*, and *Yersinia.* Chief causative viruses which are responsible for outbreak of diarrhea include rotaviruses, noroviruses, astroviruses, and adenoviruses. Intestinal parasites such as *Giardia lamblia* and *Entamoeba histolytica* are major etiologic protozoan are also responsible for diarrhea. (Hashmey *et al* 1997; Finkbeiner *et al;* 2008, Navaneethan *et al*., 2008; Kittigul *et al;* 2009)

Escherichia coli is most abundant gram negative, rod shaped, facultative anaerobe of the human gut biota which belongs to *Enterobacteriaceae* family having mutual relationships. *E. coli* usually remains harmless and restricted to the intestinal lumen. However, in the immunocompromised host, or when gastrointestinal barriers are invaded, normal nonpathogenic" *E. coli* can cause infection. Apart from non pathogenic *E. coli* certain isolates have been involved in causing disease. *E. coli* infections are restricted to mucosal lining of organs which includes two main categories (Kaper *et al*; 2004)

- Diarrheagenic *E. coli* which includes six pathovars: enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC) enteroinvasive *E. coli* (EIEC), diffusely adherent *E. coli* (DAEC), and enteroaggregative *E. coli* (EAEC).
- Extraintestinal *E. coli* which includes uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (MNEC)

DNA microarray has been widely used essential tool for monitoring whole-genome-wide expression profiles at the transcriptome level. Similarly, proteomics can be used to compare changes in the expression levels of many proteins at particular

time under same environmental conditions. Unlike transcriptomics which focuses on gene expression, proteomics examines the levels of proteins and their changes in response to different conditions. The studies on proteomes under well-defined conditions can be a better means for understanding of complex biological mechanisms and may allow inference of unknown functions of protein. Most of all, proteomic approaches provide information about post-translational modifications which cannot be obtained from mRNA expression profiles; these approaches have proven critical to our understanding of proper physiological protein function, translocation, and sub-cellular localization (Liochev *et al;* 1994; Han *et al*., 2006).

Transcriptome expressions are complementary to expression of proteome. But studies at proteome level have advantage over tanscriptome that it determine changes in protein expression, efficiency of translational machinery, protein folding, packaging into vesicles, secretion, transport and localization and also extend of degradation at particular condition and time. These studies help in understanding of complex biological processes and may help to understand the function of protein upto certain extend (Calvo *et al;* 2009). Hence, proteomics approaches offers a way to determine post translational modifications which cannot be determine with mRNA expression studies.

As a vast array of pathogens are potential cause of diarrhea and diagnosis is required before treatment. Although acute diarrhea can be cure with oral rehydration solutions, but in severe cases it is required to have prescribed medicines (The treatment of diarrhea; WHO guidelines). Medication is possible only after characterization of pathogen. The duration of diarrhea is approximately three days, so a diagnostic technique required that should be fast, reliable and cost effective. A number of molecular based diarrhea diagnostic kits are available (Platts-Mills *et al.,* 2012). But the problem in *E. coli* diagnosis is that it has 6 pathotypes and the mechanism of infection is different in each pathotype. The aim of this study was towards differential protein profiles of pathogen versus commensal counterparts, so that unique pathogenic markers can be further utilized for the development of protein based diagnosis kit.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Diarrhea

Diarrhea is a gastrointestinal disease that is an outcome of altered movement of ions and water. Each day about 8-9 liters of fluid along with electrolytes are absorbed by intestinal cells whereas only few amounts about 100-200 ml are excreted in stools. But when enteric pathogens infect intestine, balance between absorption and excretion get disturbed and consequences in diarrheal disease. Pathogens get attached to intestinal cells with different mechanisms and results in altered transport of ions though Na^+/glucose symporter, Na^+/H^+ antiporters and Cl^-/HCO_3^- exchanger, also reduced water absorption through aquaporin (Hodges and Gill., 2010). This altered transport of ions results in access release of essential ions and water from body either though loose stools or vomiting that cause dehydration in the body.

2.2 *Escherichia coli*

E. coli is rod shaped, gram negative, facultative anaerobic bacteria belongs to Enterobacteriaceae family. It is essential part of human gut microbiota and present in infant's gastro intestine within few hours after birth. In 1885 a German bacteriologist firstly discovers *E. coli* and now it's mostly studied model organism. *E. coli* have more than 700 serotype depending on their H (flagellar), O (somatic) and K (capsular) surface antigen profiles (Edwards and Ewing., 1972; Engdaw and Temesgen., 2016).

As *E. coli* is most extensively studied model organism, its complete genome has been sequenced in 1997. *E. coli* K-12 strain contains complete genome of 4,639,221-base pair and 2436 protein coding genes (Blattner *et al*., 1997).

2.2.1 Origin of *E. coli* pathogenicity

Evolution of pathogenic bacteria associated with addition or deletion of mobile genetic elements. Pathogens require new traits for their survival and fitness during infection. These new traits are acquired by the mechanism of horizontal gene transfer (HGT). Comparison between pathogenic and non pathogenic bacteria found presence of large clusters of virulence genes generally known as pathogenicity islands (Dean *et al.,* 2006). These pathogenicity islands can either be found on plasmids or may integrate into the genome. The genome size of pathogenic *E. coli* can be up to 1Mb larger than non pathogenic *E. coli*. This is because of presence of pathogenicity islands (Rasko *et al.,* 2008). HGT help recipient to survive under new environment

conditions. According to Wirth *et al.,* in 2006 bacteria exposed to multiple HGT events are able to survive under new selective conditions and these organisms are virulent cause of epidemic. Hence under new selective pressure recipients of HGT will survive over non pathogenic strains.

The genome size of pathovars and commensal *E. coli* is variable. For example: EPEC contain 400 more genes than *E. coli* K-12 and around 650 fewer genes than EHEC and 770 fewer genes than UPEC, suggesting it required smaller genes to become EPEC than other pathovars (Croxen *et al*., 2012).

2.3 Pathogenicity caused by *E. coli*

2.3.1 Enterotoxigenic *E. coli* (ETEC): Major cause of traveler's diarrhea in developing countries and also fatal for children under 5 years. Colonization factors of ETEC bind to epithelial cells of small intestine. Carbohydrates moieties of glycolproteins act as target for binding of colonization factors (Jansson *et al.,* 2006). Major cause for pathogenicity is secretion of heat stable enterotoxins (STs) and heat labile enterotoxins (LT). Secreted STs results in decreased Na^+ absorption and activation of cystic fibrosis transmembrane conductance regulators (CFTR) that is a consequence of impairment in G-protein coupled receptor signalling. Pathogenicity of LT toxins is similar to cholera toxins (having AB_5 subunits). After secretion, these toxins have lipopolysaccharide coating that helps in attachment on the surface of host cell. The B subunit bind to the cell surface, and toxin get internalized to cytosol. The A subunit ADP-ribosylate Gsα subunit of G protein and results in activation of CFTR that cause secretion of Cl^- ions (Fleckenstein *et al.,* 2010).

2.3.2 Enteropathogenic *E. coli* (EPEC) causes diarrhea by attaching and effacing (A/E) mechanism, where pathogen get attached to epithelial cells of intestine leads to accumulation of polymerized actin. A gene called locus of enterocyte effacement (LEE) which is present on 35kb pathogenicity island (PAI) is responsible for mechanism of attaching and effacing. LEE is responsible for the expression of intimin protein which is a 94kD outer membrane protein that mediates the attachment of this pathotype to host epithelial cells. Apart from attaching intimin also invigorate the response of mucosal T_H1 cells. Expression of LEE gene is highly regulated as it is responsible for type III secretion system which is responsible for secretion of bacterial proteins to the host cytoplasm for example translocated intimin receptor (Tir) which act as receptor for intimin. (Kaper *et al.,* 2004 and Deng *et al.,* 2004)

Bundle forming pili (BFP) is another factor responsible for initial binding of EPEC to small bowel. BFP is encoded by EPEC adherence factor (EAF) plasmid. This BFP are rope like fimbriae that interact with both, other EPEC bacteria and host receptors containing N-acetyl-lactosamine. (Croxen *et al.,* 2010)

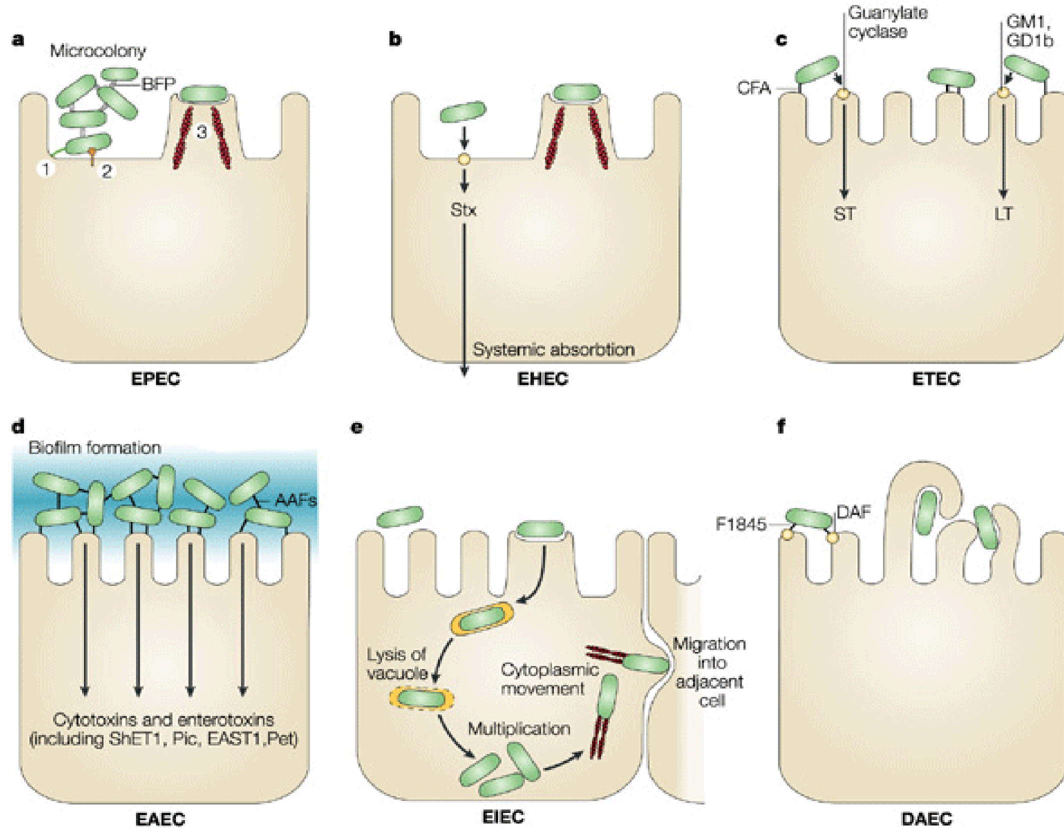


Figure 1- Schema of pathogenicity caused by diarrheagenic *E. coli*. Each of six pathogenic *E. coli* have unique features of attachment to intestinal cells. **a**) EPEC interact to small bowel by A/E mechanism. **b**) EHEC also induce A/E lesions but also produce stx toxin. **c**) ETEC induces secretion of STs and LTs. **d**) EAEC induces formation of thick biofilms. **e**) EIEC induces the lyses of phagosome. **f**) DAEC having diffusely adherence binding patterns are generated with the help of fimbrial (Dr) and afimbrial (Afa) adhesions (Kaper *et al*., 2004)

2.3.3 Enteroaggregative *E. coli* (EAEC) second most common cause for traveler's diarrhea. EAEC colonization may occur in mucosa of either small intestine or in large intestine (Kaper *et al.,* 2004). The aggregative adhesion includes stacked brick pattern with epithelial cells that involves pathogenic factors encoded by pAA virulence plasmid. Aggregative adherence fimbriae which are encoded by pAA plasmid are

responsible for adherence of EAEC (Boisen *et al.,* 2008). Pathogenicity also involves secretion of toxins though type IV secretion system. Two toxins, *Shigella* enterotoxin 1 (ShET1) and enteroaggregative *E. coli* ST (EAST1) are encoded by same chromosomal locus from opposite strands (Sheikh *et al.,* 2001). Still complete mechanism of pathogenicity through these toxins is not completely understood.

2.3.4 Enterohaemorrhagic *E. coli* (EHEC) is responsible for major outbreaks in North America, Japan and some parts of Europe (WHO: Outbreaks of *E. coli* O104:H4 infection). This pathovar infect host via attaching and effacing mechanism. Major virulence factors for pathogenicity are encoded by virulence plasmid called pO157. Another important factor for pathogenicity is Shiga toxin (Stx) but EHEC lacks secretory mechanism to secretion of Stx. The release of Stx is only a response of lambdoid phage mediated lysis which may be a result of antibiotic therapy (Wagner *et al.,* 2002).

2.3.5 Enteroinvasive *E. coli* (EIEC): it causes pathogenicity similar to *Shigella* like bloody diarrhea. It infects epithelial cells and reaches submucosa where it disrupts tight junctions and cause inflammation. This strain also contains virulence vector that encodes type three secretion system responsible for its invasion, survival and apoptosis of macrophages (Ogawa *et al.,* 2008).

2.3.6 Diffusely adherent *E. coli* (DAEC): associated with diarrhea in children and recurring urinary tract infections in adults. Diffusely adherence binding patterns are generated with the help of fimbrial (*Dr*) and afimbrial (*Afa*) adhesions. These adhesion molecules gel attached to decay-accelerating factors associated on the surface of intestinal and urinary epithelial cells. As a result brush border microvilli effaced from intestine (Servin *et al*., 2005).

2.4 Diagnosis

Most common conventional methods for detection includes culture based techniques and microscopic examination. Culture based techniques are time and labor intensive also have low detection yield. Hence, are not suitable before antibiotic prescriptions. Similarly microscopic detection requires time and labor, also needs substantial equipment and training (Guerrant *et al.,* 2001). Antigen based detection has been used since 1970s and important tool for diarrhea diagnosis for determination of rotavirus, norovirus, astrovirus, Entamoeba, Cryptosporidium Giardia, Clostridium and

Campylobacter (Kirby *et al*., 2010). Again the problem of antigen based detection is high cost and only limited number of pathogens can be detected. Molecular detection methods such as PCR based amplification of RNA or DNA have promising role in the diagnosis of infectious diseases. These methods offer excellent sensitivity over traditional diagnostics (Sjöling *et al.,* 2014).

EPEC infection can be detected via amplifying bundle-forming pilus (*bfp* ie., 326bp product) gene through PCR. (Gunzburg *et al*.,1995). Presence of ETEC can be detected via PCR amplification of genes for ST toxins, LT toxins or any of 22 different colonization factors. Gene *eltB* (273) encode for LT, *st1* (166) STp and *estA* (64) encode STh (Sjoling *et al* 2007). The detection of EAEC presence is associated with amplification of three aggregative adherence genes of plasmid. These genes include *aaP* (310bp), *aggR* (457bp) and *AAprobe* (629bp). EHEC confirmation can be attain with the PCR amplification of either stx_1 (180bp), stx_2 (255bp), *eaeA* (384bp) or *EHEC hylA* (534bp) (Paton and Paton., 1998). As 220 kb plasmid is major virulence determinant of EIEC and *Shigella*. This plasmid contains a gene called ipaH (150bp) that can be used as diagnostic tool for EIEC detection (Thiem *et al*., 2004).

As *E. coli* is major cause of diarrhea worldwide and major challenge in detection is 6 pathogenic strains. Among these strains, antibiotic treatment against EHEC should be avoidable whereas; use of antibiotics is a treatment against other strains. The use of suitable treatment therapy must be decided in short time after infection. Hence, the method of detection must be rapid, inexpensive and sensitive. In 2008 Guion *et al.,* have developed a real time multiplex PCR detection method where PCR products are identified based on melting-point curve analysis. This method is capable of diagnosing all pathotypes of *E. coli* at once. Limitation of this diagnosing method is that it is not efficient to diagnose pathogen directly from stool sample. Hence cannot be used in laboratories.

2.5 Treatment

Dehydration is major symptom and severe problem during diarrhea. Loss of fluids within 24 hours during diarrhea may vary from 5ml/kg (normal) to 200ml/kg or more (severe). Sodium deficiency in case of severe dehydration is about 70-110mM/litre (Ribeiro *et al*., 1994). To prevent dehydration primary treatment is to drink oral rehydration solution (ORS) contains salt concentration upto 3g/litre and avoid commercial fruit juice or carbonated beverages.

Zinc deficiency is also associated with high risk of diarrhea. Supplementation with zinc ion can reduce the risk of diarrhea. According to WHO guidelines, if children under 5 years are supplemented with 10-20mg Zn per day upto 14 days, risk of diarrhea will reduced for 2 or 3 months.

Antibiotic treatment is not prescribed until pathogen diagnosed. And it is not advisable to treat children below 5 years with antimicrobial agents because these agents neither prevent dehydration nor improve nutritional status. Also the use of antibiotics may results in antibiotic associated diarrhea, which is a consequence of disruption of commensal gut microflora. (Barbut and Meynard, 2002)

2.6 Proteome dynamics

Proteins concentrations in biological samples vary from organ to organ, at a time one protein may have high concentration in an organ, at same time it may have very low concentration in other organ. The expressions of proteins are highly dependent on the need of an organism, but it is not obvious that their levels are always regulated at mRNA levels. The roles of protein are highly regulated by post-translational modifications (PTMs) which increase the complexity of proteome. Another challenge of proteome studies is the detection of low abundant proteins as their relative concentration is very less; it is difficult to detect them. (Hortin *et al.,* 2006 and GE Healthcare Life sciences).

2.7 Protein expression studies

Gene expressions studies at transcriptomic level are conjectural as gene regulation takes place at particular locus on DNA. There are variations in mRNA copy number and expressed protein at a particular time in a cell (Taniguchi *et al*., 2011). As proteins are functional units performing almost all catalytic functions in a cell, complete information about the amino acid sequence and tertiary (3D) structure of proteins must required for determining its role in biological system. Studies at proteome level have advantage over tanscriptome studies that it determine changes in protein expression and also extend of expression at particular condition and time. These studies help in understanding of complex biological processes and may help to understand the function of protein upto certain extend. Proteomic approaches offers a way to determine post translational modifications which cannot be determine with mRNA expression studies (Li *et al*., 1998).

Separation of proteins using this technique is still one of the most widely used because of its optimal cost, robustness, adequate resolution, and ability to separate entire and intact proteins. Identified protein spots can be further sequenced with the help of mass spectrometry (Hao *et al.,* 2015). In 1994 Marc Wilkins defined the concept of proteome which creates new possibilities in elucidation of biopatho-mechanisms and the discovery of novel biomolecular markers (Wilkins, 2009 and Hao *et al.,* 2015).

Proteins has widely used as clinical tool since 1847 simultaneously Quantitative and qualitative determination of hundreds of proteins in biological fluids in clinical laboratories have developed over succeeding years. The use current methods have ability to separate different components of complex biological fluids. The highest resolution techniques for analysis of proteins can separate a maximum of a few thousand components at particular time (Hortin *et al*., 2006). Proteomic technology allows identification of large of proteins in single analysis, greatly accelerated identification of new biomarker for detection of diseases.

2.8 Development in proteome studies

First proteomic analysis conducted since 40 years ago, continuous advancements are added day by day in the techniques for better expression studies. There are three types of approaches to study proteome, they includes: Gel based, non gel based and predictive proteomics. Gel based approaches are being used to separate proteins in gel matrix (Klose, 1975). Most widely used gel based technique to study *E. coli* proteome is 2D gel electrophoresis. In 1975 the technique of two-dimensional gel electrophoresis (2-DE) was developed by O'Farrell and Klose. In 1992 VanBogelen *et al.,* have developed 2D protein maps for *E. coli* using carrier ampholytes containing IEF in first dimension. The limitation of this method is variability of results among different laboratories. Then the use of an immobilized pH gradient gel came into the picture and now a day number of different pH gradient strips are available commercially. University hospital of Geneva and Department of Medical Biochemistry of Geneva University has established SWISS-2DPAGE database in 1993 (Appel *et al.,* 1993). SWISS-2DPAGE database for *E. coli* has been established in year 1996 by collaboration of number of institutes (Pasquali *et al.,* 1996). These 2D maps were established by using IPG strips for first dimension from pH range 4-5, 4.5-5.5, 5-6, 5.5-6.7, 6-7, 6-11, 6-9 that contain 206 spots out of which 180 protein

entries have been identified (http://world-2dpage.expasy.org/swiss-2dpage/). The use of these wide range gels allowed the visualization of 70% of the *E. coli* proteome (Tonella *et al.,* 2001).

The separation potential of 2D gels have been increased with sample pre fractionation which help to improve protein loading capability with potential to discriminate number of migrating proteins. The advantage of sample pre fractionation is enhanced sensitivity (Molloy *et al.,* 2000).

Use of chromatography for sample fractionation can achieve hundreds of fractions for single 2D gel analysis that is essential to visualize proteins with low expression. Hence allow better quantitative and qualitative analysis of 2D gels (Han *et al.,* 2006). Maximum protein expression can be achieved with combination of 2D, LC and MS/MS, which is helpful in all stages of drug discovery (Lee, 2002).

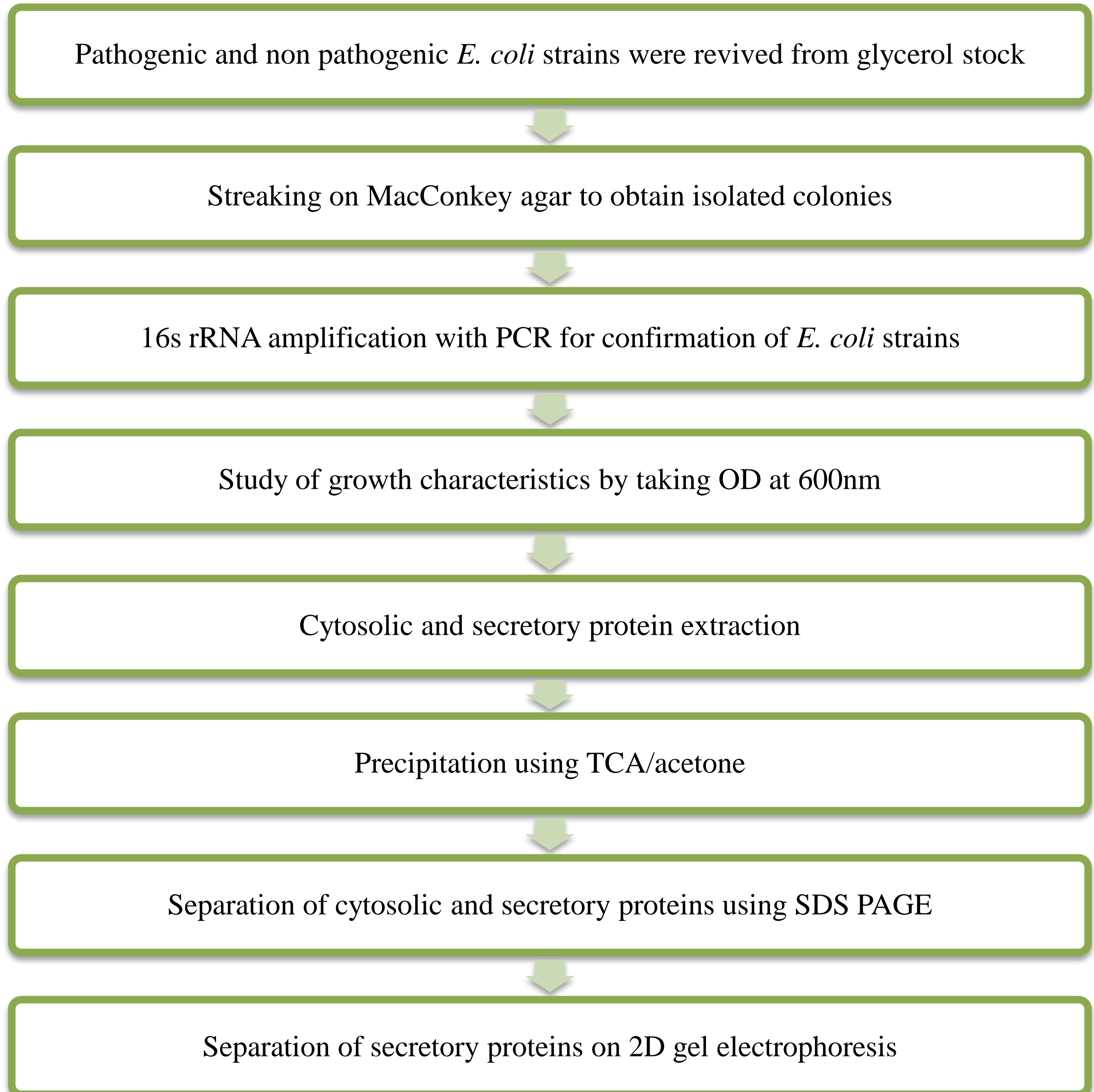
AIM AND OBJECTIVES

- To study growth characteristics of pathogenic and commensal *E. coli.*
- To extract cytosolic and secretory proteome from three pathotypes and commensal *E. coli*.
- To separate and compare secretory and cytosolic proteome of commensal *E. coli* with pathogenic *E. coli* strains using SDS –PAGE technique.
- To compare secretory proteome of commensal *E. coli* with pathogenic strains *E. coli* using 2 dimension gel electrophoresis.

CHAPTER 3

MATERIALS AND METHODS

WORK PLAN



3.1 Materials

Luria broth, MacConkey agar, BSA stock (20mg/ml), bradford reagent, ammonium persulphate (APS), triton X-100, idoacetamide were purchased from Hi-media Mumbai, India. Sodium dodecyl sulphate (SDS), glycerol, trichloroacetic acid, acetone, glacial acetic acid, ethanol, methanol, urea were purchased from Loba chemie Mumbai, India. Coomassie Brilliant blue G250, TEMED, β mercaptoethanol were purchased from Sisco research laboratory Mumbai, India. Acrylamide and bis-acrylamide, tris base, EDTA, glycine, broad range protein marker, 11cm broad range IEF strips, DTT, biolytes, CHAPS were bought from Bio-rad laboratories, USA. PCR master mix was purchased from takara, Japan and primers were used from Eurofins scientific USA.

3.2 Procedure

3.2.1 Culturing of *E. coli* strains and confirmation with PCR amplification

1. 20μl from stocks of each *E. coli* strains was inoculated into 10ml of LB.
2. Cultures were incubated at 120 rpm for overnight at 37°c.
3. Each strain was streaked on MacConkey agar plates and incubated at 37°c for overnight, to make ensure cultures are pure.
4. Isolated colonies were picked and dissolve in 10μl nuclease free water.
5. Colonies were boiled at 95°c for 10 min.
6. Centrifuged at 3000 rpm for 3 minutes.
7. Supernatant contains DNA and used for the PCR reaction
8. PCR reaction mixture was prepared using

Nucleas free water	-	4.5μl
DNA template	-	2μl
PCR master mix (takara)	-	7.5μ
Forward primer	-	0.5μl
Reverse primer	-	0.5μl

9. The following PCR reaction was run in Applied Biosystem thermocycler.

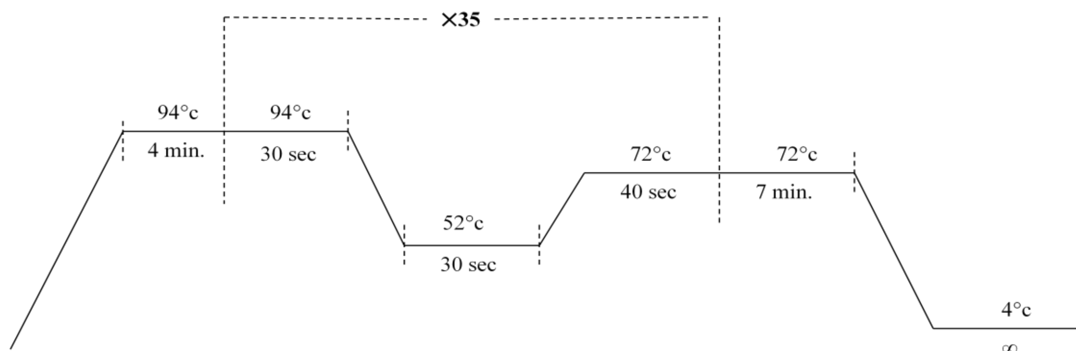


Figure 2 - PCR reaction for the amplification of V3 region of 16s rRNA. The reaction was allowed to run 35 cycles.

3.2.2 Growth curve

1. Single isolated colony was picked and inoculated in 10ml of LB.
2. Tubes were allowed to incubate at 120rpm for overnight at 37°c.
3. OD was taken at 600nm with ThermoScientific spectrophotometer.
4. Equal number of cells (*i.e.*,~1.8×10^7 cells/ml) from each strain was inoculated in 100ml LB.
5. Again cultures were allowed to incubate at 120rpm for overnight at 37°c.
6. Growth characteristics of each strain were determined by measuring OD at 600nm after interval of 1 hour using ThermoScientific spectrophotometer using LB as blank.

3.2.3 Protein extraction from all *E. coli* strains (Marcus *et al.* 2010)

1. All strains are inoculated with approximately same number of cells *i.e.,* ~1.8×10^7 cells/ml which corresponds to OD of 5.5
2. Cultures were allowed to incubate at 120rpm for overnight at 37°c.
3. After 10 hours OD reaches about 1.5 which means cells had completed log phage and now they are entering in stationary phase.
4. Cultures transferred to 50ml falcon tubes, and pellets were harvested by centrifugation at 8,000rpm for 7 minutes at 4°c.
5. Supernatant collected in sterile falcons as it contain excretory proteins and cells in pellet washed with Tris-Cl at 8,000rpm for 7 mints at 4°c.

3.2.3.1 Secretory proteins from media

1. Supernatant from each culture along with LB as control was transferred in round bottom flasks (RBF) and freezed at -80°c meanwhile lyophilizer turned on and temperature maintained at -80°c.
2. After 3hours freezed samples in RBF were loaded onto lyophilizer and vacuum pump turned on and kept there for ̴40hours.
3. RBF removed after releasing eliminating vacuum conditions.
4. Each sample solublized in 10ml of distilled water and then precipitated using TCA acetone.

3.2.3.2 Cytosolic protein extraction:

1. 7.5ml of chilled lysis buffer added in each cell fraction.
2. Samples were sonicated at interval of 60:45:30 seconds and simultaneously thaw on ice for 10 seconds after each interval.
3. Centrifuged at 7,000rpm for 8 min at 4°c.
4. Pellet containing cell debris discarded whereas supernatant was transferred in fresh tubes as it contains membrane proteins and precipitated.

3.2.3.3 Protein precipitation using TCA-Acetone: (TCA/Acetone Protein Precipitation Protocol; Cornell University**)**

1. In each sample 10% (W/V) TCA and 40% (V/V) chilled acetone added.
2. Incubated at -20°c for overnight.
3. Samples were thawed on ice for about 3hours.
4. Centrifuged at 5,000rpm for 10min at 4°c.
5. Supernatant was discarded and pellet was washed with 1ml chilled acetone.

3.2.4 Protein Quantification using Bradford Assay (Bradford, 1976)

1. BSA stock solution was prepared (1mg/ml)
2. From BSA stock volume in increasing order added in different wells (2µl, 4µl, 8µl, 16µl, 32µl and 40µl) and in blank well no BSA added.
3. From precipitated protein samples 2µl sample was added.

4. In each well volume made up to 40µl with distilled water (38µl, 36µl, 32µl, 24µl, 8µl and in last well no water added) and in blank well 40µl distilled water added.
5. 400µl of Bradford reagent added in each well including blank.
6. 96 well plate was incubated for 5 min at room temperature.
7. OD at 595 nm wavelength was taken

3.2.5 Protein separation using SDS Polyacrylamide gel electrophoresis using Bio-rad min gel assembly (Laemmli, 1970)

1. Resolving gel mounted in the electrophoresis apparatus. After polymerization (about 30 min), then stacking gel poured and comb was placed.
2. After 20 minutes, the comb was removed carefully and wells were rinsed immediately with water or with the SDS-PAGE running buffer to remove any unpolymerized acrylamide.
3. **Sample preparation**- known concentration of sample was mixed with 2X sample buffer and then boiled for 5 min at 95°c.
4. Then the gel placed in running chamber, and filled with 1X running buffer.
5. Protein samples were loaded carefully in the each well and allowed to run at 100V.
6. After complete run, the gel was removed from running chamber and placed into staining chamber.
7. Staining chamber was placed onto rocker for at least one hour or overnight.
8. Staining dye was removed, and destaining solution was added in the same chamber again the chamber was kept onto rocker.
9. Destaining solution was removed several times after each hour.
10. Protein bands were viewed on white background.

3.2.6 Protein separation with 2D PAGE (Biorad instruction manual)

3.2.6.1 Rehydration of 11cm broad range IPG strips

1. Sample (200µg protein) was prepared in 200µl of rehydration buffer.
2. Sample was loaded on rehydration tray and IPG strips were gently placed (gel side down) onto sample after removing plastic covering.

(During IPG strip loading, must ensure to remove air bubbles)

3. Strips were overlay with 2ml mineral oil and allowed to rehydrate overnight after covering tray properly.

3.2.6.2 Isoelectric focusing of IPG strips

1. Mineral oil was removed from strips rehydration tray.
2. Paper wicks were loaded on each end of focusing tray and 50µl of deionized water was added to each strip.
3. Strips were placed (gel side down) carefully over IEF tray with forceps and overlay with mineral oil.
   (Care is required to ensure + end of strip must paced towards + marked end of IEF tray)
4. IEF tray covered properly and 11cm IEF strips were run for 20-35,000 volt hours, with 50µA/IPG at 20°C.
   (250V-30min, 500V-30min, 2000V-1hr, 5000V-2hr, 8000V-2hr)

3.2.6.3 Equilibration and SDS-PAGE

1. 2ml equilibration buffer I was added onto each IPG strip and kept on shaking for 10 min.
2. After decanting equilibration buffer I, equilibration buffer II was added and again IPG strips were kept on shaking for 10 min.
3. 12.5% resolving gel was prepared and strips were loaded (gel side towards face).
   (Care required ensuring no air bubble should trap beneath IPG strips)
4. 2% low melting agarose overlay to fix IPG strip.
5. The gel was placed in running chamber, and filled with 1X running buffer.
6. Gel was allowed to run at 80V for about two and half hours.

3.2.6.4 Staining and destaining

1. After complete run, the gel was removed from running chamber and placed into staining chamber.
2. Staining chamber was placed onto rocker for at least one hour or overnight.
3. Staining dye was removed, and destaining solution was added in the same chamber again the chamber was kept onto rocker.
4. Destaining solution was removed several times after each hour.
5. Protein bands were visualized on white background.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Culturing of *E. coli* cells

Total eleven strains of *E. coli* have taken which includes nine pathotypes and two non pathogenic. All strains were grown in Luria broth. MacConkey agar plates were used to ensure pure culture, as it is differentiate between lactose fermenting and non lactose fermenting microbes.

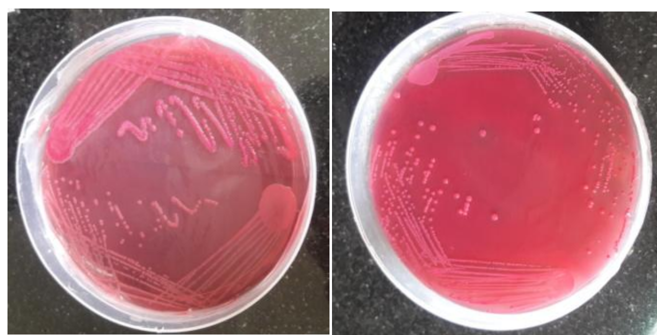


Figure 3- Pure colonies of *E. coli* strains after streaking on MacConkey agar plate. Pink colonies on MacConkey agar plate ensure the growing bacteria is gram negative.

4.2 Confirmation of *E. coli* strains with PCR

PCR reaction for the amplification of 16rRNA of all *E. coli* strains was used as gold standard for *E. coli* confirmation. PCR product was run on 1.2% agarose gel along with 100bp ladder. Amplification of V3 region of 16s rRNA (542 bp product) confirm that the used culture is of *E. coli*; hence we can proceed with further analysis.

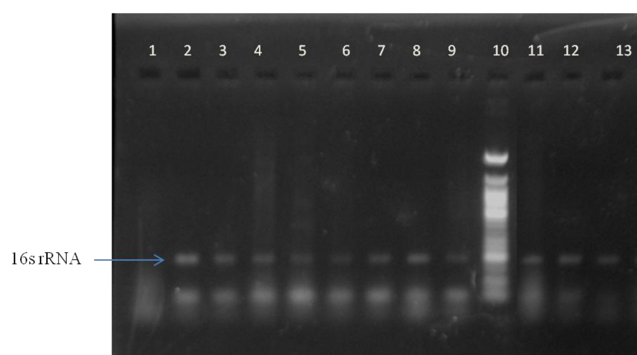


Figure 4 – PCR product of V3 region of 16s rRNA using 1.2% agarose gel. Lane 1 represent negative control, lane 2 *E. coli* DH5α, lane 3 *E. coli* ATCC 25922, lane 4, 5, 6 pathogenic ETEC, lane 7, 8, 9 pathogenic EPEC, lane 10 100bps ladder form new emerald biology and lane 11, 12, 13 pathogenic EAEC.

4.3 Growth characteristics of pathogenic and non pathogenic *E. coli*

E. coli cells give maximum absorption at 600nm, which can be used as standard for determining growth of cells. As cells divide number of cells increase respectively, hence OD will also increase. According to Sezonov *et al*., 2007, OD of 0.3 corresponds to 5×10^7 cells/ml, hence it is possible to determine cell concentration by measuring OD. Here, pathogenic and non pathogenic strains were cultured using approximately equal number of cells (~1.8×10^7 cells/ml), and growing characteristics have determined.

Optical density at 600nm was taken for all cultured strain and used to plot graph between time (on x-axis) and OD (on y-axis). This graph expresses different phases of bacterial growth *i.e.,* Lag, Log and Stationary phase.

According to the obtained growth curve we can say that commensal *E. coli* have more doubling time as compare to other pathogenic strains. As non pathogenic *E. coli* enters in lag phase after 4 hours, whereas pathogenic strains takes only 2 hours to enter log phase, it may be possible because non pathogenic *E. coli* take more time to get established in new environment.

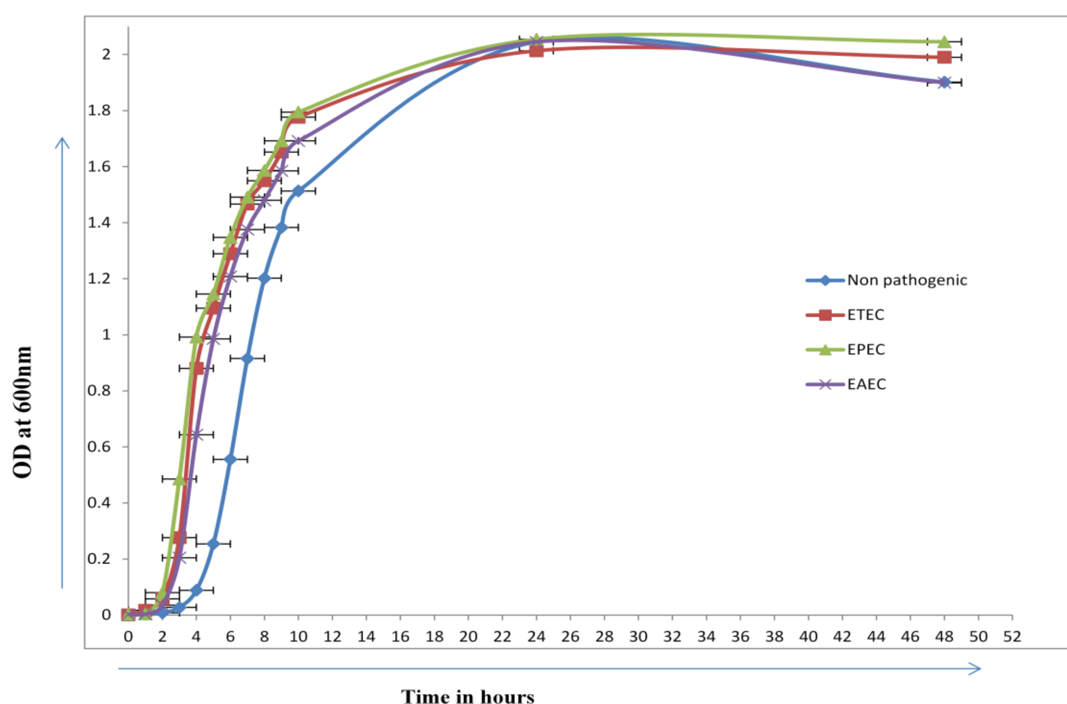


Figure 5- Comparative analysis of growth characteristics of commensal and pathogenic *E. coli* strains. x-axis represents OD at 600nm and y-axis represents time in hours. Blue line represents the growth pattern of non pathogenic *E. coli,* whereas red purple and green represents the growth pattern of pathogenic *E. coli*.

4.4 Protein Extraction and Precipitation

After protein extraction we cannot ensure whether the protein has extracted or not, until unless it is quantified. Since white pellets were observed after precipitation with TCA Acetone, hence we assume that our protein has extracted successfully. Precipitated protein pellets were dissolved in 10mM tris-Cl pH 7.4 for quantification with Bradford assay.

4.5 Protein Quantification using Bradford Assay

Standard plot of BSA protein was plotted using 10mM tris as blank to determine concentration of *E. coli* proteins from all samples.

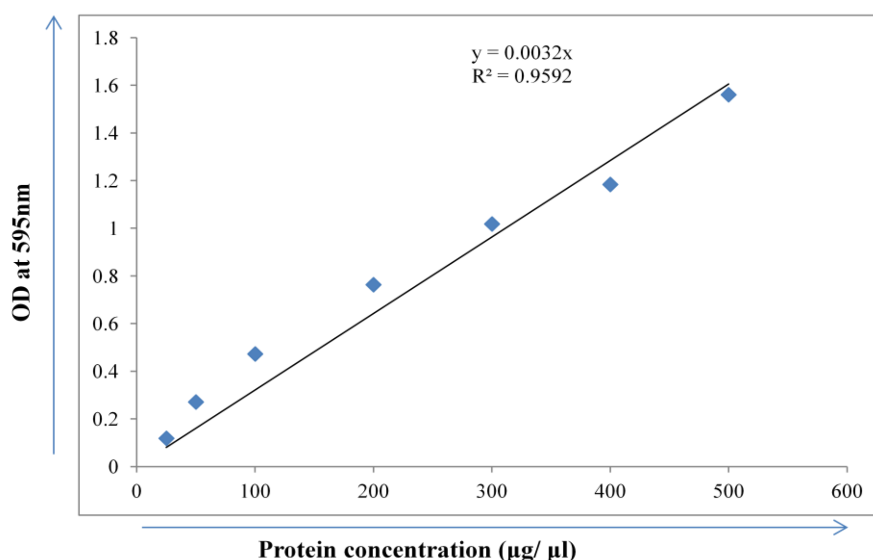


Figure 6- Standard curve for BSA protein estimation using Bradford assay. x-axis represents OD at 595nm and y-axis represents protein concentrations in µg/µl.

The standard plot from bradford assay was obtained that have R^2 value of 0.9592 and

$$y=0.0032x$$

$$\text{Hence, } x=y/0.0032$$

The above equation can be used to determine the concentrations in cytosolic as well as secretory protein samples. (because here, x is protein concentration and y is OD at 595nm)

Proteins concentrations (mg/ml) were determined using above equation

i.e., $x= y/0.0032\times\text{dilution factor}$

4.6 Protein separation using SDS PAGE

Sample loading was most critical step in SDS PAGE, too high concentration results in protein aggregates which are visible as smear or large spot. On other hand if loaded concentration is not optimal, only most abundant proteins will be visible or no results will be obtained.

Here, equal final concentration *i.e.,* 1mg/ml from each proteins sample was loaded in each well which help us to understand the expression of proteins at a particular time.

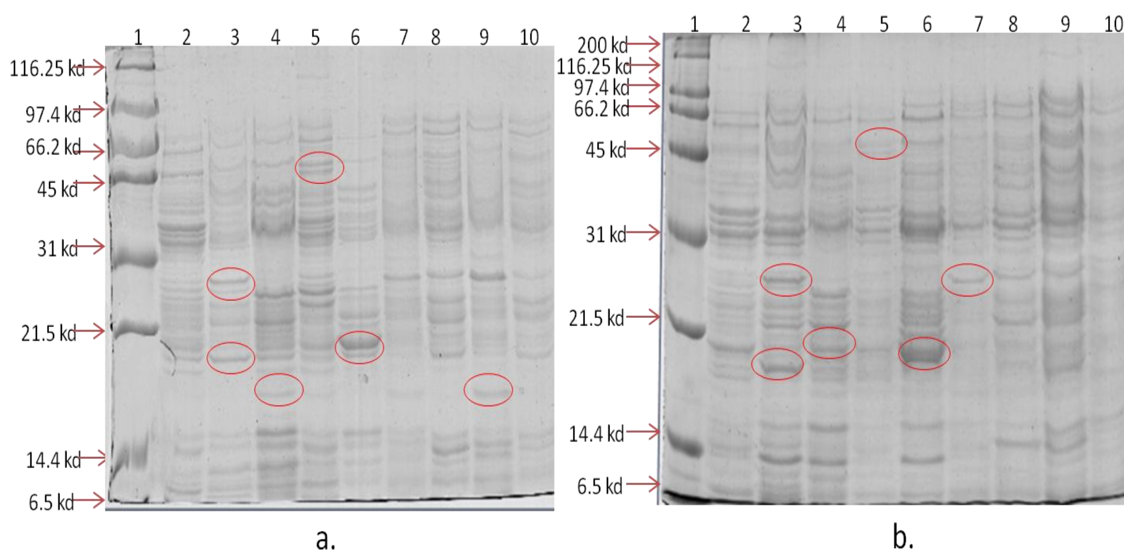


Figure 7– SDS polyacrylamide gel electrophoresis of cytosolic proteins of pathogenic and non pathogenic *E. coli* strains. In both images a) and b), lane 1 broad range protein marker (Bio-rad), lane 2 *E. coli* DH5α, lane 3 *E. coli* ATCC 25922, lane 4 EPEC, lane 5 EAEC, lane 6 ETEC, lane 7 EPEC, lane 8 EAEC, and lane 9 ETEC and lane 10 EAEC. Each well was loaded with 20µg of protein sample (image captured with Bio-rad GS-800 scanner)

From above SDS PAGE gel, difference in protein expression of pathogenic and non pathogenic *E. coli* can be visualized easily. Although results from SDS-PAGE cannot be used to interpret any final conclusion as a single band may contain number of proteins. But after visualizing SDS PAGE image we determine comparison between number of expressed proteins between pathogenic and non pathogenic and can proceed with further analysis.

In the case of Secretory proteins again equal concentration (1mg/ml) of proteins has loaded. Here, LB has used as a control to ensure whether proteins are coming from *E. coli* strain or from media itself.

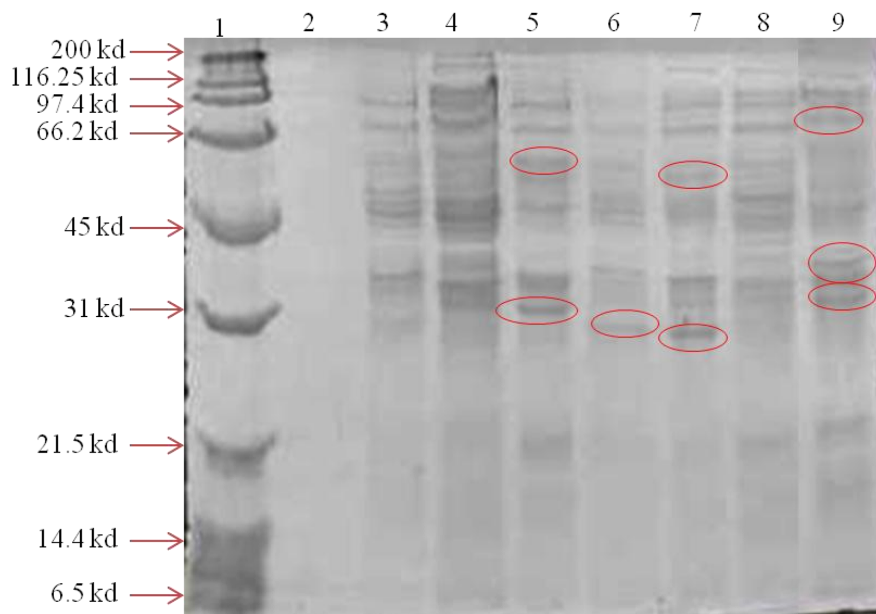


Figure 8– SDS polyacrylamide gel electrophoresis of secretory proteins of pathogenic and non pathogenic *E. coli*. Lane 1 broad range protein marker, lane 2 media control, lane 3 *E. coli* DH5α, lane 4 *E. coli* ATCC 25922, lane 5 EPEC, lane 6 EAEC, lane 7 ETEC lane 8 EPEC and lane 9 ETEC. Each well was loaded with 20µg of protein sample (image captured with Bio-rad GS-800 scanner).

As no protein band has detected in precipitated media control it is sure whatever bands or smear coming, it is from particular *E. coli* strain. From above image the variations in expressed proteins can be visualize variations among pathogenic and non pathogenic secretory proteome.

From figure 8, maximum variations are obtained in ETEC, selected for further analysis. But before proceeding with analysis, SDS PAGE was performed with all strains of ETEC to visualize differences in expressed proteins.

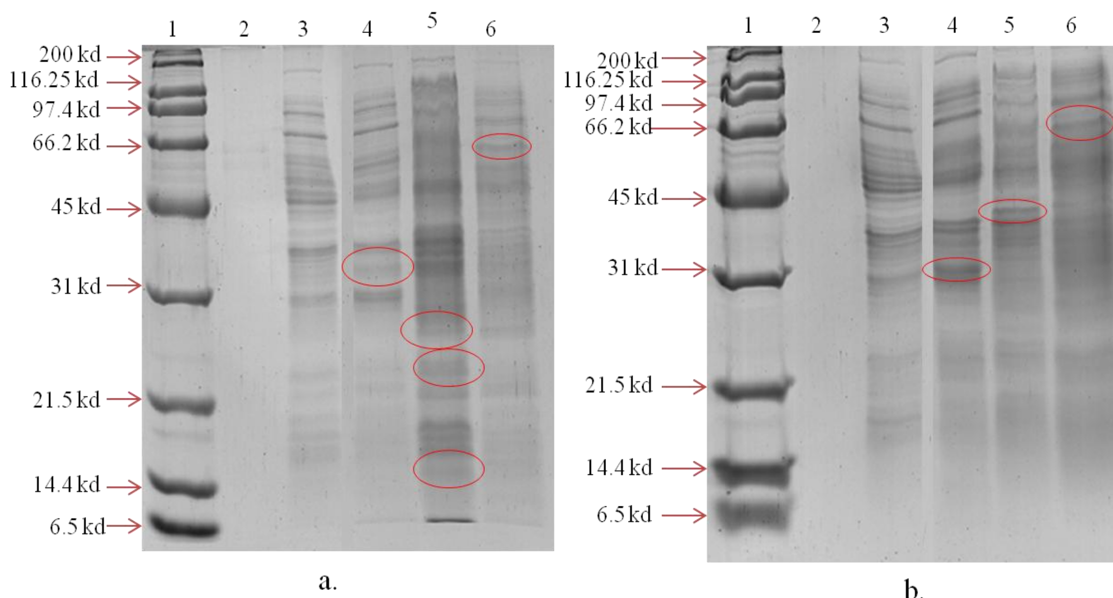


Figure 9– SDS polyacrylamide gel electrophoresis of secretory proteins of non pathogenic *E. coli* and pathogenic ETEC. In both images a) and b), lane 1 broad range protein marker, lane 2 Media control, lane 3 *E. coli* DH5α, lane 4, lane 5 and lane 6 contains pathogenic ETEC. Each well was loaded with 20μg of protein sample (image captured with Bio-rad GS-800 scanner) Highlighted red stops represent variation in expressed proteins in pathogenic and non pathogenic strains.

Again from above image of pathogenic and non pathogenic *E. coli,* visible differences in the expression of secretory proteins were obtained. These differences can represent either the presence or absence of a protein or extent of expression at particular time and conditions.

After obtaining differences in secretory proteome of pathogenic ETEC and non pathogenic *E. coli*, 2D analysis was performed to determine further differences.

4.7 Secretory protein analysis using 2D gel electrophoresis

Separation using 2D gel for the comparison of secretory proteome of pathogenic and non pathogenic *E. coli* results in better prediction of expressed proteins.

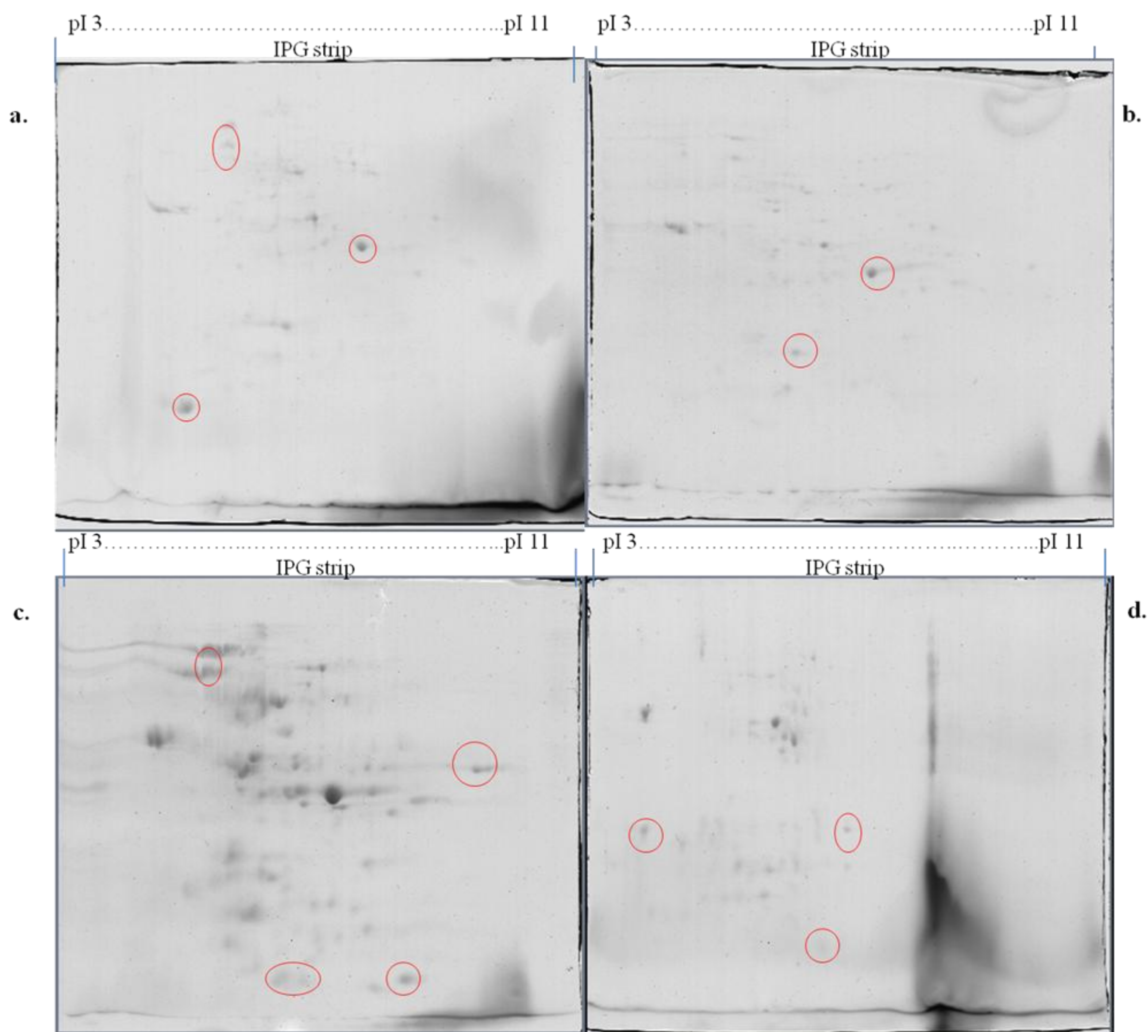


Figure 10 – 2D polyacrylamide gel electrophoresis of secretory proteins of *E. coli*. a) and b) secretory proteome of non pathogenic *E. coli*, c) and d) are secretory proteome of pathogenic ETEC on 3-11 IPG 11 cm IPG strips loaded with 100µg protein from each sample (image captured with Bio-rad GS-800 scanner). Highlighted red stops represent variation in expressed proteins in pathogenic and non pathogenic strains.

From above gels we can predict that there are variations in secretory proteins of pathogenic and non pathogenic *E. coli*. The use of silver staining may reproduce better better prediction from above image. There are some spots present in pathogenic ETEC but absent in non pathogenic *E. coli*, these proteins may be responsible pathogenicity.

CONCLUSION

The study was conducted to determine differences in growth characteristics and proteome of pathogenic and non pathogenic *E. coli*. The study shows differences in the growth of pathogenic and non pathogenic *E. coli*. As non pathogenic *E. coli* require more time to enter log phase as compare to pathogenic strains, this may be a reason for the cause of diarrhea because it may be possible that faster growing pathogenic strains compete the gut microflora.

Cytosolic and secretory proteome was extracted efficiently and show visible differences in pathogenic and non pathogenic *E. coli*. The differences in SDS PAGE images for secretory and cytosolic proteome of pathogenic and non pathogenic *E. coli* have obtained which may allow further analysis of differential band to obtain complete information about these proteins. The significant differences in 2D PAGE gel images of pathogenic ETEC and non pathogenic *E. coli* have obtained. The study may helpful for further comparison between secretory proteome of pathogenic and non pathogenic *E. coli* to obtain expression of unique pathogenic protein that can be utilized for development protein based diagnostic tools.

REFERENCES

1. Appel RD, Sanchez JC, Bairoch A, Golaz O, Miu M, Vargas JR, Hochstrasser DF. SWISS-2DPAGE: A database of two-dimensional gel electrophoresis images. Electrophoresis. 1993 Jan 1;14(1):1232-8.

2. Barbut F, Meynard JL. Managing antibiotic associated diarrhoea: probiotics may help in prevention. British Medical Journal. 2002 Jun 8;324(7350):1345-7.

3. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J. The complete genome sequence of *Escherichia coli* K-12. science. 1997 Sep 5;277(5331):1453-62.

4. Boisen N, Struve C, Scheutz F, Krogfelt KA, Nataro JP. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. Infection and immunity. 2008 Jul 1;76(7):3281-92.

5. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry. 1976 May 7;72(1-2):248-54.

6. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings of the National Academy of Sciences. 2009 May 5;106(18):7507-12.

7. Croxen MA, Finlay BB. Molecular mechanisms of *Escherichia coli* pathogenicity. Nature Reviews Microbiology. 2010 Jan 1;8(1):26-38.

8. Dean P, Maresca M, Schüller S, Phillips AD, Kenny B. Potent diarrheagenic mechanism mediated by the cooperative action of three enteropathogenic *Escherichia coli*-injected effector proteins. Proceedings of the National Academy of Sciences of the United States of America. 2006 Feb 7;103(6):1876-81.

9. Deng W, Puente JL, Gruenheid S, Li Y, Vallance BA, Vázquez A, Barba J, Ibarra JA, O'Donnell P, Metalnikov P, Ashman K. Dissecting virulence: systematic and functional analyses of a pathogenicity island. Proceedings of the national academy of sciences of the United States of America. 2004 Mar 9;101(10):3597-602.

10. Edwards PR, Ewing WH. Identification of enterobacteriaceae. Identification of Enterobacteriaceae.. 1972(Third edition).

11. Engdaw TA and Temesgen W 2016 O157:H7 Serotype of *Escherichia coli* as an Important Emerging Zoonosis. International Journal of Microbiological Research 7 (1): 09-17

12. Experimental biosciences introductory laboratory. Preparing Protein Samples for Electrophoresis.(http://www.ruf.rice.edu/~bioslabs/resources_home.html)

13. Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, Wang D. Metagenomic analysis of human diarrhea: viral detection and discovery. PLoS Pathog. 2008 Feb 29;4(2):e1000011.

14. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, Steinsland H. Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. Microbes and Infection. 2010 Feb 28;12(2):89-98.

15. Guerrant RL, Van Gilder T, Steiner TS, Thielman NM, Slutsker L, Tauxe RV, Hennessy T, Griffin PM, DuPont H, Sack RB, Tarr P. Practice guidelines for the management of infectious diarrhea. Clinical infectious diseases. 2001 Feb 1;32(3):331-51.

16. Gunzburg ST, Tornieporth NG, Riley LW. Identification of enteropathogenic *Escherichia coli* by PCR-based detection of the bundle-forming pilus gene. Journal of Clinical Microbiology. 1995 May 1;33(5):1375-7.

17. Han MJ, Lee SY. The *Escherichia coli* proteome: past, present, and future prospects. Microbiology and molecular biology reviews. 2006 Jun 1;70(2):362-439.

18. Han MJ, Lee SY. The *Escherichia coli* proteome: past, present, and future prospects. Microbiology and molecular biology reviews. 2006 Jun 1;70(2):362-439.

19. Hao R, Adoligbe C, Jiang B, Zhao X, Gui L, Qu K, Wu S, Zan L. An Optimized trichloroacetic acid/acetone precipitation method for two-dimensional gel electrophoresis analysis of Qinchuan cattle Longissimus dorsi muscle containing high proportion of marbling. PloS one. 2015 Apr 20;10(4):e0124723.

20. Hashmey R, Genta RM, White AC. Parasites and diarrhea. I: Protozoans and diarrhea. Journal of travel medicine. 1997 Mar 1;4(1):17-31.

21. Hodges K, Gill R. Infectious diarrhea: cellular and molecular mechanisms. Gut Microbes. 2010 Jan 1;1(1):4-21.

22. Hortin GL, Jortani SA, Ritchie JC, Valdes R, Chan DW. Proteomics: a new diagnostic frontier. Clinical chemistry. 2006 Jul 1;52(7):1218-22.

23. http://www.euro.who.int/en/health-topics/disease-prevention/food-safety/outbreaks-of-e.-coli-o104h4-infection

24. Jansson L, Tobias J, Lebens M, Svennerholm AM, Teneberg S. The major subunit, CfaB, of colonization factor antigen i from enterotoxigenic *Escherichia coli* is a glycosphingolipid binding protein. Infection and immunity. 2006 Jun 1;74(6):3488-97.

25. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. Nature Reviews Microbiology. 2004 Feb 1;2(2):123-40.

26. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. Nature Reviews Microbiology. 2004 Feb 1;2(2):123-40.

27. Kirby A, Gurgel RQ, Dove W, Vieira SC, Cunliffe NA, Cuevas LE. An evaluation of the RIDASCREEN and IDEIA enzyme immunoassays and the RIDAQUICK immunochromatographic test for the detection of norovirus in faecal specimens. Journal of Clinical Virology. 2010 Dec 31;49(4):254-7.

28. Kittigul L, Pombubpa K, Taweekate Y, Yeephoo T, Khamrin P, Ushijima H. Molecular characterization of rotaviruses, noroviruses, sapovirus, and adenoviruses in patients with acute gastroenteritis in Thailand. Journal of medical virology. 2009 Feb 1;81(2):345-53.

29. Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. Human Genetics. 1975 Sep 1;26(3):231-43.

30. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. nature. 1970 Aug;227:680-5.

31. Lakshminarayanan S, Jayalakshmy R. Diarrheal diseases among children in India: current scenario and future perspectives. Journal of natural science, biology, and medicine. 2015 Jan;6(1):24.

32. Lee MS. LC/MS Applications. LC/MS Applications in Drug Development. 2002:65-182.

33. Li L, Wu J, Li B. Emergence and significance of proteomics. Chinese Bulletin of Life Sciences. 1998 Dec;11(2):49-50.

34. Liochev SI, Hausladen A, Beyer WF, Fridovich I. NADPH: ferredoxin oxidoreductase acts as a paraquat diaphorase and is a member of the soxRS regulon. Proceedings of the National Academy of Sciences. 1994 Feb 15;91(4):1328-31.

35. Liu J, Gratz J, Maro A, Kumburu H, Kibiki G, Taniuchi M, Howlader AM, Sobuz SU, Haque R, Talukder KA, Qureshi S. Simultaneous detection of six diarrhea-causing bacterial pathogens with an in-house PCR-luminex assay. Journal of clinical microbiology. 2012 Jan 1;50(1):98-103.

36. Luria SE, Burrous JW. Hybridization between *Escherichia coli* and Shigella. Journal of bacteriology. 1957 Oct;74(4):461.

37. Molloy MP, Herbert BR, Slade MB, Rabilloud T, Nouwens AS, Williams KL, Gooley AA. Proteomic analysis of the *Escherichia coli* outer membrane. European journal of biochemistry. 2000 May 1;267(10):2871-81.

38. Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. Clinical microbiology reviews. 1998 Jan 1;11(1):142-201.

39. Navaneethan U, Giannella RA. Mechanisms of infectious diarrhea. Nature clinical practice Gastroenterology & hepatology. 2008 Nov 1;5(11):637-47.

40. Ogawa M, Handa Y, Ashida H, Suzuki M, Sasakawa C. The versatility of Shigella effectors. Nature Reviews Microbiology. 2008 Jan 1;6(1):11-6.

41. Pasquali C, Frutiger S, Wilkins MR, Hughes GJ, Appel RD, Bairoch A, Schaller D, Sanchez JC, Hochstrasser DF. Two-dimensional gel electrophoresis of *Escherichia coli* homogenates: The *Escherichia coli* SWISS-2DPAGE database. Electrophoresis. 1996 Jan 1;17(3):547-55.

42. Paton AW, Paton JC. Detection and Characterization of Shiga Toxigenic*Escherichia coli* by Using Multiplex PCR Assays forstx 1, stx 2, eaeA, Enterohemorrhagic *E. coli* hlyA, rfb O111, andrfb O157. Journal of Clinical Microbiology. 1998 Feb 1;36(2):598-602.

43. Platts-Mills JA, Operario DJ, Houpt ER. Molecular diagnosis of diarrhea: current status and future potential. Current infectious disease reports. 2012 Feb 1;14(1):41-6.

44. Protein sample preparation handbook (2016) GE Healthcare Life sciences Chapter 1:15-28.

45. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. Journal of bacteriology. 2008 Oct 15;190(20):6881-93.

46. Ribeiro Junior H, Ribeiro T, Mattos A, Palmeira C, Fernandez D, Sant'Ana I, Rodrigues I, Bendicho MT, Fontaine O. Treatment of acute diarrhea with oral rehydration solutions containing glutamine. Journal of the American College of Nutrition. 1994 Jun 1;13(3):251-5.

47. Servin AL. Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. Clinical microbiology reviews. 2005 Apr 1;18(2):264-92.

48. Sezonov G, Joseleau-Petit D, D'Ari R. *Escherichia coli* physiology in Luria-Bertani broth. Journal of bacteriology. 2007 Dec 1;189(23):8746-9.

49. Sheikh J, Hicks S, Dall'Agnol M, Phillips AD, Nataro JP. Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. Molecular microbiology. 2001 Sep 1;41(5):983-97.

50. Sjöling Å, Sadeghipoorjahromi L, Novak D, Tobias J. Detection of major diarrheagenic bacterial pathogens by multiplex PCR panels. Microbiological research. 2015 Mar 31;172:34-40.

51. Sjöling Å, Wiklund G, Savarino SJ, Cohen DI, Svennerholm AM. Comparative analyses of phenotypic and genotypic methods for detection of enterotoxigenic *Escherichia coli* toxins and colonization factors. Journal of clinical microbiology. 2007 Oct 1;45(10):3295-301.

52. SWISS-2DPAGE(http://world-2dpage.expasy.org/swiss-2dpage/ retrieved on 20th April 2017)

53. TCA/Acetone Protein Precipitation Protocol. Proteomics and Mass Spectrometry Facility, BRC, Cornell University Version 3.0

54. The treatment of diarrhea. A manual for physicians and other senior health workers (http://www.who.int/maternal_child_adolescent/documents/9241593180/en/)

55. Thiem VD, Sethabutr O, von Seidlein L, Van Tung T, Chien BT, Lee H, Houng HS, Hale TL, Clemens JD, Mason C, Trach DD. Detection of Shigella by a PCR assay targeting the ipaH gene suggests increased prevalence of shigellosis in Nha Trang, Vietnam. Journal of clinical microbiology. 2004 May 1;42(5):2031-5.

56. UNICEF data: Monitoring the situation of children and women. Diarrhoea remains a leading killer of young children, despite the availability of a simple treatment solution; 2016.

57. Vanbogelen RA, Sankar P, Clark RL, Bogan JA, Neidhardt FC. The gene-protein database of *Escherichia coli*: Edition 5. Electrophoresis. 1992 Jan 1;13(1):1014-54.

58. Wagner PL, Livny J, Neely MN, Acheson DW, Friedman DI, Waldor MK. Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. Molecular microbiology. 2002 May 1;44(4):957-70.

59. Weiss KF, Chopra N, Stotland P, Riedel GW, Malcolm S. Recovery of fecal coliforms and of *Escherichia coli* at 44.5, 45.0 and 45.5 C. Journal of Food Protection®. 1983 Mar 1;46(3):172-7.

60. Wilkins M. Proteomics data mining. Expert review of proteomics. 2009 Dec 1;6(6):599-603.

61. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Molecular microbiology. 2006 Jun 1;60(5):1136-51

APPENDIX

Acrylamide gel

Components	Separating/Resolving gel		Stacking gel
	12%(10ml)	12.5% (10ml)	(4%)(5ml)
30%acrylamide mixture	4ml	4.15ml	0.66ml
Tris HCl , 1.5 M	2.5ml	2.5ml	0.63ml
Distilled water	3.3ml	3.15ml	3.6ml
10% SDS	100 µl	100µl	50µl
10% APS	100µl	100µl	50µl
TEMED	5µl	5µl	5µl

Destaining Solution 500ml

Sr. No.	Ingredients	Volume /Quantity	Concentration
1.	Methanol	200ml	40% V/V
2.	Acetic acid	50ml	10V/V
3.	MQ water	250ml	50% V/V

Equilibration buffer I (100ml)

Sr. No.	Ingredient	Volume/Quantity	Concentration
1.	Urea	36.034gm	6M
2.	SDS	2gm	2% W/V
3.	Tris-Cl pH 8.8	4.543gm	0.375M
4.	Glycerol	20ml	20% V/V
5.	DTT	2gm	130mM
Add traces of bromophenol blue and final volume makeup to 100ml with d/w. DTT should be added just before use.			

Equilibration buffer II (100ml)

Sr. No.	Ingredient	Volume/Quantity	Concentration
1.	Urea	36.034gm	6M
2.	SDS	2gm	2% W/V
3.	Tris-Cl pH 8.8	4.543gm	0.375M
4.	Glycerol	20ml	20% V/V
6.	Iodoacetamide	2.4gm	135mM
Add traces of bromophenol blue and final volume makeup to 100ml with d/w. IAA should be added just before use.			

Laemmli's buffer or SDS Sample buffer

Sr. No.	Ingredients	Quantity/1000ml	Concentration
1.	Tris (pH 6.8)	15.14gm	0.125M
2.	SDS	40gm	4%W/V
3.	Glycerol	200ml	20%V/V
4.	β-mercaptoethanol	100ml	10%V/V

Luria broth (Himedia)

Sr. No.	Ingredients	Quantity/1000ml
1.	Casein enzymic hydrolysate	10gm
2.	Yeast extract	5gm
5.	Sodium chloride	5gm

SDS-PAGE running buffer (10X in 1000ml)

S. No.	Ingredients	Quantity	Concentration
1	Glycine	144.13gm	192mM
2	SDS	10gm	0.1%W/V
3	Tris pH 8.3	30.285gm	25mM
Make Final Volume Up to 1000ml with d/w			

MacConkey agar (Himedia)

Sr. No.	Ingredients	Quantity/1000ml
1.	Peptones (meat and casein)	3gm
2.	Pancreatic digest of gelatin	17gm
3.	Lactose monohydrate	10gm
4.	Bile salts	1.5gm
5.	Sodium chloride	5gm
6.	Crystal violet	0.001gm
7.	Neutral red	0.030gm
8.	Agar	13.500

Staining Solution 500ml

Sr. No.	Ingredients	Volume /Quantity	Concentration
1.	Methanol	200ml	40% V/V
2.	Acetic acid	50ml	10V/V
3.	MQ water	250ml	50% V/V
4.	CBB R-250	1.25gm	0.25% W/V

Rehydration buffer 100ml

Sr. No.	Ingredient	Quantity/100ml	Concentration
1.	Urea	48gm	8M
2.	CHAPS	2gm	2% W/V
3.	DTT	0.75gm	50mM
4.	Biolytes	0.20ml	0.2%
Add traces of bromophenol blue and final volume makeup to 100ml with d/w. Store at -80°C. DTT must be added just before use.			