Molecular detection of Enterotoxigenic *Escherichia coli* Toxins and Colonization Factors from Diarrheic Children in Pediatric Teaching Hospital, Sulaymaniyah, Iraq

Hezhan Faeq Rasul, Sirwan Muhsin Muhammed, Huner Hiwa Arif, Paywast Jamal Jalal

Department of Biology, College of Science, University of Sulaymaniyah, Sulaymaniyah, Iraq

**ABSTRACT**

Enterotoxigenic *Escherichia coli* (ETEC) is one well-established causative agent of diarrhea in the developing countries among young children. This prospective study was performed at Laboratories of University of Sulaimani (in Sulaymaniyah City/Iraq) from September to October 2021 which aimed to determine the prevalence of ETEC among children and the most prevalence colonization factor (CFA/I) among ETEC. One hundred and twenty-five fresh stool samples were collected from hospitalized – children with diarrhea at Dr. Jamal Ahmed Rashid’s Pediatric Teaching Hospital. The collected samples were cultured on MacConkey and eosin methylene blue agar as selective and differential media for Gram- negative bacteria. Colonies were identified through Gram staining and biochemical tests including: Indole, methyl red, and catalase reaction test. Vitek-2 machine was depended to test some obtained isolates. Most of isolates (60%) showed positive results for *E. coli* – out of this percentage, 14 (18.66%) were positive for ETEC using polymerase chain reaction assay identifying stable and labile toxins (LTs). It was noticed that all of the ETEC isolates were stable toxin producer isolates whereas LT producer isolates were not identified. Colonization factor 5 (CS5) has been detected among three ETEC isolates (21.42%), meanwhile, 11 isolates (78.57%) have not expressed colonization factors at all.

**Index Terms:** *Escherichia coli*, Enterotoxigenic *E. coli*, Stable toxin, Labile toxin

1. INTRODUCTION

*Bacterium coli commune* was initially reported as a commensal Gram-negative rod from the healthy individual’s intestinal flora by Theodor Escherich, a German pediatrician, in 1885, and in his honor, these rods were named *Escherichia coli* [1]. The genus *Escherichia coli* is distributed widely and is the most common facultative anaerobe found among humans and warm-blooded animals’ – large intestine [2]. Depending on the number of virulence determinants found, specific combinations were created, determining the currently known *E. coli* pathotypes, which are generally recognized as diarrheagenic *E. coli* (DEC) [3]. DEC pathotypes are classified into enteropathogenic *E. coli*, enterotoxigenic *E. coli* (ETEC), EIEC is for enteroinvasive *E. coli*, and Shiga stands for enterohemorrhagic *E. coli*, enteroaggregative *E. coli* (toxin-producing *E. coli*) is another type of *E. coli*. They pathotypes of *E. coli* vary widely in terms of preferred host colonization locations, virulence mechanisms, and clinical symptoms and out [4]-[6].

In the developing countries, ETEC is still one of the most common causes of infectious diarrhea in travelers and children [7]. Watery diarrhea, vomiting, stomach cramps, and,
ETECs ability to stick to and colonize intestinal epithelium, it is critical for pathogenicity. In addition, its ability to produce heat-labile labile toxin (LT) enterotoxin and/or heat-stable stable toxin (ST) enterotoxin, both of which can produce diarrhea. ST is a limited peptide made up of 18–19 amino acid residues, whereas LT is a high-molecular-weight (84 kDa) enterotoxin with an active alpha subunit surrounded by five identical binding B subunits [11]. The two main genotypes of ST are STa and STb; typically, ETEC strains isolated from people produce STa (STI or STh), which is encoded by the estA gene, whereas STb (STII orSTP) is primarily produced by animal ETEC strains which is encoded by the eltB [12]. The LTs that ETEC strains produce are likewise a diverse category of toxins. There are two main LT families known as LT-I and LT-II [13]. LT genes eltA and eltB produce LT-I and LT-II, respectively. The ST genes are possible to express independently or in tandem with the LT genes eltA and eltB [13]. ETEC strains can express seven different toxin combinations: STh, STp, STh/LT, STp/LT, LT, and less typically, STh/STp and STh/STp [14]. The existence of colonization factors (CFs) on membrane of a bacterial cell, which normally form pili, also known as fimbriae, is necessary for colonization [15]. Depends on antigenic specificity and/or the N-terminal amino-acid sequence of the main subunit, different forms of colonization factor antigens (CFA) and putative colonization factors have been identified(pilin) such as CFA/I, CS1, CS2, CS3, CS4, CS5, CS6, CS7, CS14, CS17, CF19, CF21, CF22 [16]. There is a limited research available on ETEC colonization and prevalence of diarrheagenic ETEC among human, particularly children below the age of six in this region. To fill this gap, this study have conducted to determine E. coli, ETEC toxins (ST and LT) producers, and colonization factors from children under 6 years suffering from watery diarrhea in the Pediatric Teaching Hospital, in Sulaimani City.

2. MATERIALS AND METHODS

2.1. Sample Collection
During the period from September to October, 125 sample stools were collected from children <6 years at Dr. Jamal Ahmed Rashid’s Pediatric Teaching Hospital and SMART Private Hospital in Sulaymaniya City. Both sexed were included (63 females and 62 males). The necessary information about the patients were taken from the hospitals, and the collected samples were transferred from the hospitals to the Advanced Bacteriology Laboratory from Biology Department of University Sulaimani within less than 3 hours in an ice box to culture them.

2.2. Bacterial Cultivation and Characterization
All samples were preliminary cultured on differential and selective media for presumptive isolation of Gram-negative enteric rods. These included MacConkey agar for first isolation and eosin methylene blue for confirmation as described by [17], [18]. All lactose fermenting, deeply pink, circular, medium in size, colonies were subcultured on the medium (eosin methylene blue) agar (Neogene, UK). All plates were incubated at 37°C for 18–20 h. Colonies showing green metallic sheen on EMB agar were considered as E. coli strains. E. coli samples utilized in the present study were identified by Gram staining [19] and initial biochemical tests including indole [20], methyl red [21], catalase test [22], and other bacteriologic characterization using Vitek-2 system (VITEK®2 GN ID card) by Vitek machine (BioMerieux, France) were performed for some of them [23].

2.3. DNA Extraction and Purification
The DNA of isolates under test was isolated and purified using and following the directions.

2.3.1. Colony extraction
It was performed by transferring two colonies from fresh bacterial culture then mixed with 40 μl of ddH₂O and preheated at 95°C for 10 min using the thermo cycler and purified DNA obtained by centrifugation at 12,000 rpm for 1 min. The supernatant was used as a polymerase chain reaction (PCR) template [24].

2.3.2. DNA extraction with kit
Overnight fresh colonies from nutrient broth (Neogene, UK) utilized. Genomic DNA from E. coli isolates was extracted and purified using a DNeasy kit (AddPrep Genomic, Korea) according to manufacturer protocol.

2.4. PCR Method
PCR mixture contained the DNA template, forward/reverse primer (Macrogen, Korea), and Master Mix (Taq Master (2 × conc)/addbio. Korea) deionizing water (Accumax, Korea).
2.4.1. 16S rRNA
PCR was performed for 75 samples E. coli to identify 16S rRNA using this reaction included: Initial denaturation for 5 min at 94°C, followed by 35 cycles of amplification (1 min at 94°C, 1 min at 56–58°C and 1 min at 72°C), and finally finished with 7 min at 72°C [25].

2.4.2. ST and LT
The 96-well plates were used to amplify stable and LT genes. The PCR procedure included pre-incubation at 95°C for 1 min, followed by 35 cycles of (1 min at 95°C, 1.10 min at 45°C, 1.30 min at 72°C), final incubation at 72°C for 5 min [26]. The products have run on 2% agarose gel (TransGen, China).

2.4.3. Colonization genes
To identify genes of colonization factors of CFA/I, CS1, CS2, CS3, CS4v, CS5, CS6, CS14, CS17. The same pre-mentioned procedure was depended with different primer for each gene. The genes were amplified by an initial denaturation at 94°C (1 min), followed by 35 cycles of amplification (94°C for 30 s, 52°C for 30 s, and 72°C for 1 min), finally, 5 min at 72° [27]. The amplicon was separated with 3% agarose gels by gel electrophoresis (Cleaver-CS-300v, UK) and then visualized by ethidium bromide (TransGen, China). The specificity of the primers was tested by both BLAST search and is illustrated in Table 1.

2.5. DNA Sequencing
The sequencing was performed for 10 samples with amplified 16S rRNA -F and 16S rRNA-R (forward and reversed primers (10 pmol). DNA sequencing was achieved by Sanger sequencing/ABI 3500, Macrogen Genome Center, Korea using BigDye kit.

2.6. Phylogenetic Tree
Evolutionary analysis was conducted by MEGA7 program. The evolutionary history was deduced using the Kimura 2-parameter model and the maximum likelihood technique [28]. It is shown the tree with the greatest log likelihood (−505.1852). Next to the branch is the proportion of trees where the related taxa clustered together. The starting tree(s) for the heuristic search were automatically generated by applying the neighbor-join and BioNJ algorithms to a matrix of pairwise distances calculated using the maximum composite likelihood technique and thereafter picking the topology with the best log likelihood value. To represent evolutionary rate differences across sites (5 categories [+G, parameter = 0.0500]), a distinct gamma distribution was utilized. The branch distances are calculated by the number of replacement per location, as well as the tree is depicted to scale. A total of 18 nucleotide sequences were examined. The codon locations were included 1st + 2nd + 3rd + non-coding. All positions containing gaps and missing data were eliminated. There were a total of 320 positions in the final dataset [29].

3. RESULTS AND DISCUSSION
3.1. Detection of DEC
It was appeared that out of 125 tested samples, 83 (66.4%) were Gram-negative bacteria after they were cultured on

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>426</td>
</tr>
<tr>
<td>16S-F</td>
<td>5′GACGTA\text{CTGCCAGATAAGC-3′}</td>
</tr>
<tr>
<td>16S-R</td>
<td>5′-TTAGTCTT\text{CGACGGTACG-3′}</td>
</tr>
<tr>
<td>ST toxin</td>
<td>186</td>
</tr>
<tr>
<td>STF</td>
<td>5-TCT GTA TTG TCT TTT TCA CC-3, 5-T\text{TAATA GCA CCC GGT ACA AGC-3,}</td>
</tr>
<tr>
<td>STR</td>
<td></td>
</tr>
<tr>
<td>LT toxin</td>
<td>273</td>
</tr>
<tr>
<td>LTF</td>
<td>5-\text{ACG GCCGTTAC TAC TCTC-3}</td>
</tr>
<tr>
<td>LTR</td>
<td>5-\text{TGG TCTCGG TCGATATG-3}</td>
</tr>
<tr>
<td>CFA/I</td>
<td>170</td>
</tr>
<tr>
<td>CS1</td>
<td>243</td>
</tr>
<tr>
<td>CS3</td>
<td>100</td>
</tr>
<tr>
<td>CS5</td>
<td>226</td>
</tr>
<tr>
<td>CS4</td>
<td>198</td>
</tr>
<tr>
<td>CS6</td>
<td>152</td>
</tr>
<tr>
<td>CS14</td>
<td>162</td>
</tr>
<tr>
<td>CS17</td>
<td>130</td>
</tr>
<tr>
<td>CS2</td>
<td>368</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>5′-GACGTA\text{CTGCCAGATAAGC-3′}</td>
<td></td>
</tr>
<tr>
<td>16S-F</td>
<td>5′-TTAGTCTT\text{CGACGGTACG-3′}</td>
<td></td>
</tr>
<tr>
<td>16S-R</td>
<td></td>
<td>[25]</td>
</tr>
<tr>
<td>ST toxin</td>
<td>5-TCT GTA TTG TCT TTT TCA CC-3, 5-T\text{TAATA GCA CCC GGT ACA AGC-3,}</td>
<td></td>
</tr>
<tr>
<td>STF</td>
<td></td>
<td>[26]</td>
</tr>
<tr>
<td>STR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT toxin</td>
<td>5-\text{ACG GCCGTTAC TAC TCTC-3}</td>
<td></td>
</tr>
<tr>
<td>LTF</td>
<td>5-\text{TGG TCTCGG TCGATATG-3}</td>
<td></td>
</tr>
<tr>
<td>LTR</td>
<td>5-\text{ACG GCCGTTAC TAC TCTC-3}</td>
<td></td>
</tr>
<tr>
<td>CFA/I</td>
<td>5-\text{GCTAT TCTCCCGCATACAA-3}</td>
<td></td>
</tr>
<tr>
<td>CS1</td>
<td>243</td>
<td></td>
</tr>
<tr>
<td>CS3</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>CS5</td>
<td>226</td>
<td></td>
</tr>
<tr>
<td>CS4</td>
<td>198</td>
<td></td>
</tr>
<tr>
<td>CS6</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>CS14</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>CS17</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>CS2</td>
<td>368</td>
<td></td>
</tr>
<tr>
<td>CTA\text{GGTCAGATCCGTTTG-3}</td>
<td></td>
<td>[27]</td>
</tr>
<tr>
<td>CTG\text{AGCAAATCCGTTTC-3}</td>
<td></td>
<td>[27]</td>
</tr>
<tr>
<td>TTG\text{GCGACTCTGCAAGTGGT-3}</td>
<td></td>
<td>[27]</td>
</tr>
<tr>
<td>TTG\text{GCGACTCTGCAAGTGGT-3}</td>
<td></td>
<td>[27]</td>
</tr>
<tr>
<td>TTG\text{GCGACTCTGCAAGTGGT-3}</td>
<td></td>
<td>[27]</td>
</tr>
</tbody>
</table>
EMB and MacConkey agar, they showed metallic shine and pink color respectively. Seventy-five (60%) of them were rod shape purple color, when they were grown in peptone water, they produced forming pink ring color at top of tubes after addition of Kovac's reagent. The color of the broth cultured changed to red after adding methyl red indicator to tube during performing methyl red test. H2O2 was added to fresh colonies, bubble formation indicated positive catalase test. The percentage of appeared *E. coli* similarity to an ideal *E. coli* by Vitek-2 test was done for the samples of 70, 23, 60, and 35 which were 99%, 93%, 87%, and 94%, respectively. Seventy-five isolates have given positive for 16S rRNA-based PCR, as shown in Fig. 1.

Our study explored that the most diarrheagenic pathogens among Gram negative in Sulaimani are *E. coli*, which is compatible with the results reported in a local study by Hasan et al. (2020) done in Dhok city [30], Shatub et al. (2021) found similar results (61.3%) [31], whereas Khalil (2015) in Baghdad reported lower positive rates (38.6%) [32] as well as other investigators who showed lower positive results [33]-[35]. Several studies from worldwide revealed varying DEC detection rates in *E. coli* among children under 5 years old, ranging between 4% and 87% in Africa including 22.9%, 7.4%, 55.9%, and 86.5%, Asia (45.2%, 4.7%, 6.82%), and America 5.5%. These variations could be related to changes in DEC pathotype distribution from one region to the others, also between countries in the same region [36]. According to many reports around the world, various factors may be the primary causes of diarrheal outbreaks including; traveling to tropical zones, consuming contaminated, and lack of personal hygiene [35]. However, considering that prior studies have focused on certain aspects such as geographical conditions, sampling period, study population, hygienic level of region, and detection technologies [37].

The proportion of infected males 44 (58.6%) was relatively higher than females 31 (41.3%), the infected males higher than females were like to result reported by Hasan et al. (2020) who reported 87.4% among males and 87.0 among females [30]. The current observations were agreed with results reported by the result mentioned by Amir et al. (2020) in Iran who showed 53.01% for male and 46.99% for female [35]. Similarly, our observations were parallel to the results concluded by Ochien and Atieno (2021) in West Kenya (55.9% and 44.1%) male and female, respectively [38], whereas the current results were not agreed with the result of Abbasi in Iran (2020) who reported higher rates among females than males [34].

### 3.1.1. DNA sequencing
All accession numbers have shown in Fig. 2. AY342058.1 is the accession number of a ST gene which sequencing was performed.

A phylogenetic tree based on the 16S ribosomal RNA sequences was extracted from a representative set of 10 *Enterobacteriaceae* genomes and compared to some other different strains (Fig. 2). All strains referred to one clad, the clad of *E. coli* was more similar to *Shigella flexneri*. The numbers (close individual) clustered and bootstrap percentage of 100 replications. Some of the tree nodes are uncertainly predicted. It has concluded that the analysis of variable genes identifies interstrain relationships that may be correlated to the lifestyle of the organisms [39].

![Fig. 1. 16S rRNA gene PCR of *Escherichia coli*: M; is 100 bp ladder 2–6 were 16S rRNA gene.](image1)

![Fig. 2. Molecular phylogenetic analysis by maximum likelihood method.](image2)
3.2. Identification of ETEC Toxins

The ETEC characterization using PCR among isolates was done. For the tested children who suffered from non-bloody, acute diarrhea due to \( E. coli \) (\( n = 75 \)). All \( E. coli \) isolates were evaluated by PCR monoplex for ST gene and LT gene. A total of 14 (18.66%) were proved ETEC. The majority toxin profile among the selected strains were ST, whereas LT was not recorded in the present study, as shown in Fig. 3. While all other pathotype of \( E. coli \) were 81.33%.

Results of the present study were agreed with observations reported in studies done in other parts of Kurdistan and Iraq. PCR-based studies detected showed different percentage rates of ETEC in stool samples ranging from 18% to 26% [30], [32], [36], [40], [41]. Results of this study were not parallel to conclusions mentioned by other investigators who reported lower percentage rates of positive results [34], [37].

ETEC is more common among low- and middle-income individual states, where it is a prominent pathogenic strain in travelers’ diarrhea, with a large burden on these countries [42]. ETEC was recognized as a major pathotype among children below 5 years old. The high percentage rates of ETEC- positive results could be attributed to the family’s poor hygiene and artificial feeding [30], [41], [43]. Variations in our results with other researches could be related to changes in primers utilized, geographical considerations, population targeted, and sample size [44]. In nine of the 12 research conducted in Africa, 22 of 34 investigations in Asia, and three of six studies in Latin America and the Caribbean, ETEC was being the first or second most often isolated pathogen [45].

The proportion of infected male and female with ETEC was different, in our study, 9 (64.3%) ETEC was among males, while the rest 5 (35.7%) were among females which is parallel to results mentioned by other studies from West Kenya [38] whereas other investigators [46] found higher percentage positive results among females which are not in agreement with our results.

3.3. Distribution of ETEC among Children According to the Age

The ETEC pathotype was identified in all children according to their ages, with a slightly increasing number of infected children with ETEC under 12 months [47]. Our result highlighted the significant of ETEC like a cause of childhood diarrhea between the ages of 1 and 12 months, as shown in Fig. 4. The current observations were agreed to results reported by Shatub \textit{et al.} (2021) from Tikrit/Iraq [31] whereas our results were far with outcome of Khalil (2015) [32]. Our findings are backed up by a review article that looked at ETEC infection from 1984 to 2005, stratified infants by age, and found that the peak incidence occurs after 6 months and can last until 18 months [48]. This might be related to duration of breastfeeding, the source of drinking water, cleanliness, sanitation, age, and the level of maternal educational. In contrast, in the finding by Abdul-Hussein \textit{et al.} (2018) in Wasit/Iraq, the most ETEC prevalence was found among (3–24 months [33].

3.4. Stable and LT

In our result, it was noticed that 14 (18.66%) ST genes were present among isolates while there had no any identified LT solely and ST-LT toxin as shown in Fig. 3. The prevalence of ST was the most common toxin gene. In the other region of Kurdistan and the rest of the world, there are several reports showing differences in the prevalence of ETEC pathotype. The present study differed from a study in Duhok city by
Hasan et al. (2020), by which LT toxin solely identified in 37% of the cases [30], whereas seven strains positive for LT gene and three strains positive for ST were identified by Khalil (2015) in Bagdad [32]. Consequently, our study was close to a study by Shabazi et al. (2021) that found three ST and one LT gene in their results. Besides, the prevalence of LT and ST gene in several studies was presented as follows. For instance, a study by Alizade et al. (2014) found 11.97% ST and 9.86% LT, Saka et al. (2019) concluded with 19 ST and 14 LT, and Nazarian et al. (2014) with 4 ST and 1 LT [36], [37], [46], [49].

ST, a peptide with a molecular weight of 18 or 19 amino acids, combined with a carrier protein and then will be antigenic. As a result, after infected with ST-producing ETEC, immunological responses to ST are not produced. The percentage of strains that produce LT alone, ST alone, or LT/ST varies by geographic region; in general, 30–50% of clinical ETEC isolates appear to produce ST solely [9], [45].

For pathogenic strains, such as ETEC, it was shown that specific conditions of host might increase or decrease bacterial virulence. The impact of glucose and bile on the gene expression and protein level of ST generated by different ETEC isolates was investigated by Joffre et al. (2016), and he discovered that there are unique STa amino acid variations that respond differently to environmental signals such as bile [50].

A substantial amount of literature highlights the effect of different seasons on the prevalence of ETES-associated infections, by which in the late spring and entire summer, this type of infection was repeatedly identified [9], [51]-[54]. However, in our study, the incidence of ETES-associated infection is lower than the expected rate when compared to other studies. This may be due to the period of data collection that was performed in September and October.

### 3.5. Detection of Colonization Factors

Among the 14 ETEC produced ST isolates, nine primers were chosen for most common CF, among them, only 3 (21.42%) ETEC isolates shown CF, 11 (78.57%) ETEC isolates without CF. In our result, there was only 1 isolate (7%) posed CFA/I, 1 (7%) showed CS4, and also 1 (7%) with CS5 as shown in Fig. 5.

In among clinically important ETEC strains, over 22 antigenically different CFs were identified, but only a handful are frequently present in diarrheic patient samples [9]. Current results are compatible with the result reported by Peruski Jr et al. (1999), among the ST-positive strain CFA/1, CS4, and CS5 were reported, 77% isolates were failed to express any of 9 CF [55]. Our result is close to the finding of Shaheen et al. (2004) CFA/1 (9.7%), CS4 (2 Strain), and CS5 (2 strain) Shaheen et al. (2004) [56] and Kipkirui et al. (2021) [44].

The discrepancy between our findings and those of other studies could be attributable of variation in CF expression by ETEC in different geographical regions, as well as differences in laboratory methods/primers used to identify CFs [46]. In addition, due to sub cultured repeatedly or storage for long term, the plasmid containing the CF genes has been lost [44]. Decreasing the expression of CF genes, a mutation inside the genetic locus, and expression of a CF not covered by the primers used in the PCR panel [6], [57], [58]. Antigens for CF are only created in vivo, or a small percentage of strains do not generate CFs [59]. The CF antigens are differed from the 9 CFs screened for in this study. Decreasing of CF has been reported to be associated to LT strains [41], [60], which similar our results where lack LT toxin. However, some studies have reported that CFs are almost equally associated with LT- and ST-positive ETEC strains [61].

### 4. CONCLUSION

This study illustrated ETEC in children below 6 years old with acute non-bloody diarrhea. PCR-based detection of ETEC revealed that all isolated ETECs found with ST toxin-producing gene with no any ETEC -LT identified gene isolate. From the overall of ETEC isolates, only three of them showed CF which are CS4, CS5, and CFA/1 on the different strains. This finding can provide an evidence on the
prevalent of ETECs pathotype in this region, furthermore, creation a platform for vaccine development can be adapted from this finding.

ACKNOWLEDGMENT

We thank to the staff of Dr. Jmale Ahmed Rashid’s Pediatric Teaching Hospital for helping us and we thank to all persons contributed in this study.

REFERENCES


Rasul, et al.: Molecular detection of Enterotoxigenic Escherichia coli Toxins from Diarrheic Children in, Sulaymaniyah


