Bangladesh J Microbiol, Volume 34, Number 2, December 2017, pp 97-101

Original Article

Genetic Relatedness among *Escherichia coli* Isolated from Clinical Samples

Sunjukta Ahsan[1*], Ruhul Amin[1], and Anindita Bhowmik[1]

[1]*Department of Microbiology, University of Dhaka, Dhaka 1000, Bangladesh*

A common member of the intestinal microbiota in humans and animals is *Escherichia coli*. Based on the presence of virulence factors, *E. coli* can be potentially pathogenic. The focus of this study was to determine genotypic diversity among clinical strains of *E. coli*. Test isolates were identified by biochemical, culture-based and molecular tests. Partial sequence of 16s rDNA was determined by Polymerase Chain Reaction using universal primers. The sequences were used as input in Basic Local Alignment Search Tool (BLAST) to reconfirm the identity of the isolates. Sequences were edited in Chromas and aligned in Bioedit. Afterwards, sequences were fed into Mega 5 and a dendogram was constructed. We used three sequences deposited in GenBank for comparison with our test sequences. It was found that two of our sequences showed similarity to the reference sequences. One of the isolates was similar to a Uropathogenic and a commensal *E. coli*. The other isolate showed similarity to a UPEC strain. It is possible that both of these isolates were commensal bacteria that led to opportunistic infection in the patients. The rest of the nine strains were different from the reference sequences. These isolates may be opportunistic pathogens that have undergone mutation events leading to sufficiently different rDNA sequences from the reference bacteria. Or they may also be pathogenic strains that caused infection in patients. This study indicated that clinical isolates of *E. coli* show variation in terms of their rDNA sequences. In order to study their phylogeny, other supporting tools such as investigation for phylogenetic markers, RAPD, ARDRA, ERIC-PCR, ribotyping, etc. may be used to study variation among isolates to a more significant extent.

Keyword: *E. coli*, Phylogenetic analysis, Molecular typing

Introduction

It is known that most strains of *Escherichia coli* are harmless and live in the intestines of healthy humans and animals; however some of the *E. coli* strains can cause a wide variety of intestinal and extra-intestinal diseases, such as diarrhea, urinary tract infections, septicemia, and neonatal meningitis[1]. The presence of *Escherichia coli* in environmental sample is implicit evidence for fecal contamination and represents a threat to human and environmental health[2].

The traditional identification of bacteria on the basis of phenotypic characteristics is generally not as accurate as identification based on genotypic methods. Comparison of the bacterial 16S rRNA gene sequence has emerged as a preferred genetic technique. 16S rRNA gene sequence analysis can better identify poorly described, rarely isolated, or phenotypically aberrant strains, can be routinely used for identification of bacteria, and can lead to the recognition of novel pathogens and non-cultured bacteria[3].

Phylogeny is the study of evolutionary relatedness among various groups of organism. Phylogenetic analyses have shown that *E. coli* strains fall into four main phylogenetic groups (A, B1, B2, andD)[4] and that virulent extra-intestinal strains belongmainly to group B2 and to a lesser extent, to group D[5,6,7], whereas most commensal strains belong to group A. Thesestudies also give a better understanding of how pathogenic strains acquire virulence genes[5]. Identification of the source of bacterial contamination may be useful for developing plans to resolve or reduce water contamination occurrences.

The sequencing step in sequence based typing is preceded by locus specific PCR amplification to generate templates for the sequencing step. The sequencing step requires single stranded DNA templates, DNA primers, a DNA polymerizing enzyme, deoxynucleotidephosphates (dNTPs) and fluorescently labelled di-deoxynucleotidephosphates (ddNTPs).Phylogenetic tree is prepared based on genome sequences.

In the present study, the genetic relatedness among clinical *E. coli* was investigated. As all our test isolates were uropathogenic *E. coli* (UPEC), we determined to what extent they are related at the level of the 16s rDNA sequence and compared them to the sequences of commensal and other UPEC strains deposited in gene bank to compare diversity.

Methods and Materials

Bacteria

A total of 19 *E. coli*isolated from urine samples at Medinova Medical service were used for the study, 11 of which were sequenced for dendogram construction.

Corresponding author:
Dr. Sunjukta Ahsan, Associate Professor, Dept. of Microbiology, University of Dhaka, Email: sunjukta@du.ac.bd, Phone: 8801948283983

Detection of 16s rDNA gene of E. coli sequences by PCR

PCR primers for universal 16s rDNA were used to detect the samples as bacterial isolates.Bacteria grown on Nutrient agar medium were used for preparation of template DNA. Sterile toothpick was used to pick a bacterial colony and mixed with 20 µl sterile distilled water. The mixture was heated at 100°C for 5 minutes in a PCR block (MJ Research,USA). Heat lysed cells were centrifuged for 3 mins in a short spin machine (ABC-M6 mini centrifuge, USA). Cell debris were pelleted at the bottom and the supernatant contained DNA. This was used as template DNA for subsequent PCR.Commercial Master mix (*Taq* 2x master mix, NEB, UK) was used in all PCR. It was two times (2x) concentrated and contained dNTPs, Taq polymerase, $MgCl_2$ and buffer. A sterile eppendorf was taken and the tube was labeled. Measured volume of water and $MgCl_2$ were taken in the eppendorf. Then the master mix was added. During addition of master mix the eppendorf was placed on icebox because Taq polymerase might lose its activity if it was not placed in ice. Though the master mix contained $MgCl_2$, extra $MgCl_2$ was added for greater efficiency of Taq polymerase present in the master mix. After that forward and reverse primers were added in appropriate volumes. Table 2 shows the reaction components in a 20 µl PCR reaction mixture. A total of 18µl reaction mixture was dispensed in each PCR tube. Two µl of template was added to each tube and mixed with a pipette tip.PCR reaction was carried out in a programmable Gradient cycler (MJ Research,USA) that involved an initial DNA denaturation, followed by a number of cycles of denaturation, primer annealing, and product extension. A final DNA extension step completed the reaction.The PCR for Universal 16s rDNAwas performed using an initial denaturation step at 95°C for 5 minutes, followed by 35 cycles each consisting of an initial denaturation step at 95°C for 30 seconds, primer annealing at 58°C for 30 seconds and extension at 68°C for 30seconds. A final extension step was carried out at 68°C for 10minutes. PCR amplicons were analyzed using agarose gel electrophoresis in 1% agarose (w/v). The gel was then stained in Ethidium bromide solution (40 µg/ml) for 30 min and destained in distilled water for 5 minutes. Then stained gel was observed with a UV transilluminator (Gel Doc, Bio-Rad, USA). Photographs were taken and bands were analyzed with "Quality One®" software (Bio-Rad, USA). The molecular mass marker, 1 kb DNA ladder marker (Promega, UK) was used for Universal 16s rDNA.

PCR primers

The sequences of primers used in this study are shown in Table 1.

DNA Purification

DNA was purified by Invisorb DNA CleanUp (STRATEC Molecular GmbH, Germany). It was used for purification of DNA fragments after PCR reactions and purification of contaminated DNA after DNA purification.

DNA sequencing

DNA sequencing is performed in CARS (Centre for Advanced Research in Sciences).

Genotypic characterization

The 16s rDNA sequences generated with universal forward and reverse primers were aligned in MUSCLE (www.ebi.ac.uk) with known sequence of *E. coli* (KR148984.1) deposited in GenBank (http://www.ncbi.nlm.nih.gov/) and edited manually. Genotypic characterization was done using MEGA 5.2 version software. The phylogenetic analysis was constructed using neighbor-joining algorithm[8,9].

Results

Amplification of 16s rDNA by using universal primer

Molecular genotypic technique involves the amplification of any phylogenetically informative target such as 16S rRNA gene for genera identification[3]. So amplification of 16S rRNA marker gene was performed to conform the test *E. coli*.

A 1kb amplicon corresponding to part of the 16s rDNA was obtained using PCR (Figure 1).

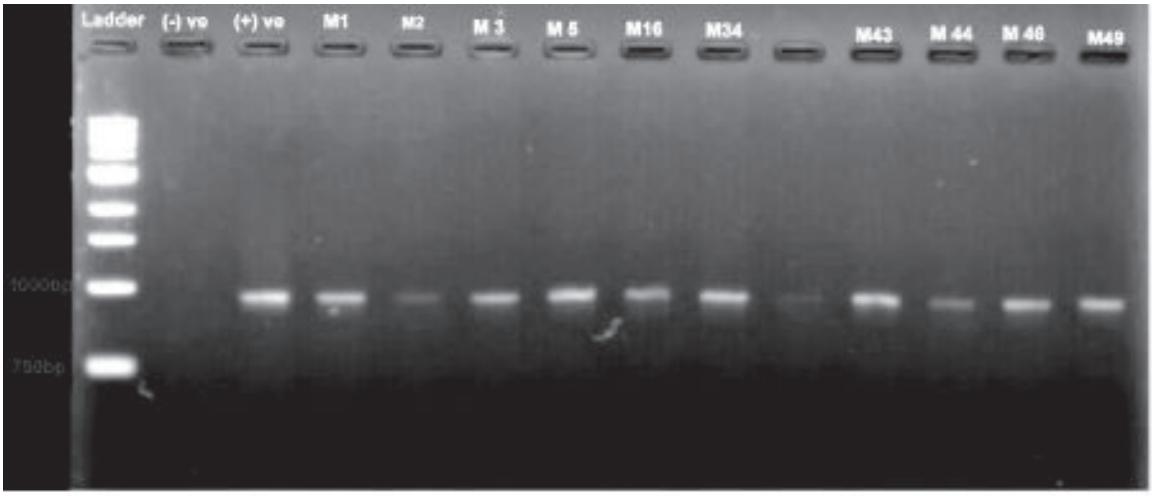

Figure 1. *PCR analysis of 16S rRNA universal for Escherichia coli.* Samples were designated with M followed by a number. *E. coli DH5± was used as a positive (+) control. (+) ve and (-) ve denote positive and negative controls respectively. The ladder is a 1kb ladder purchased from Promega(USA).*

Genetic Relatedness among test isolates

The genetic relatedness among the test isolates was determined by plotting a dendogram based on 16s rDNA partial sequences using Neighbour-joining method (Figure 2). The evolutionary history was inferred using the Neighbor-Joining method[8]. The optimal tree with the sum of branch length = 21.73223416 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method[9] and are in the units of the number of base substitutions per site. The analysis involved 14 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 474 positions in the final dataset. Evolutionary analyses were conducted in MEGA5[10].

In this study, sequences of 16s rDNA of two uropathogenic *E. coli* (GenBank IDs: KX816955.1, strain DS-04 and JX290084.1, strain KVP104) isolated in India and a commensal *E. coli* (GenBank ID: AB269763.1) isolated in Japan were used for comparison with the test isolates. From the dendogram, it is observed that isolate M34 showed similarity to uropathogenic and commensal *E. coli*. On the other hand, isolate M6 was in the same cluster as uropathogenic *E. coli* strain KVP104. The rest of the isolates showed sequence variation with the reference bacteria used.

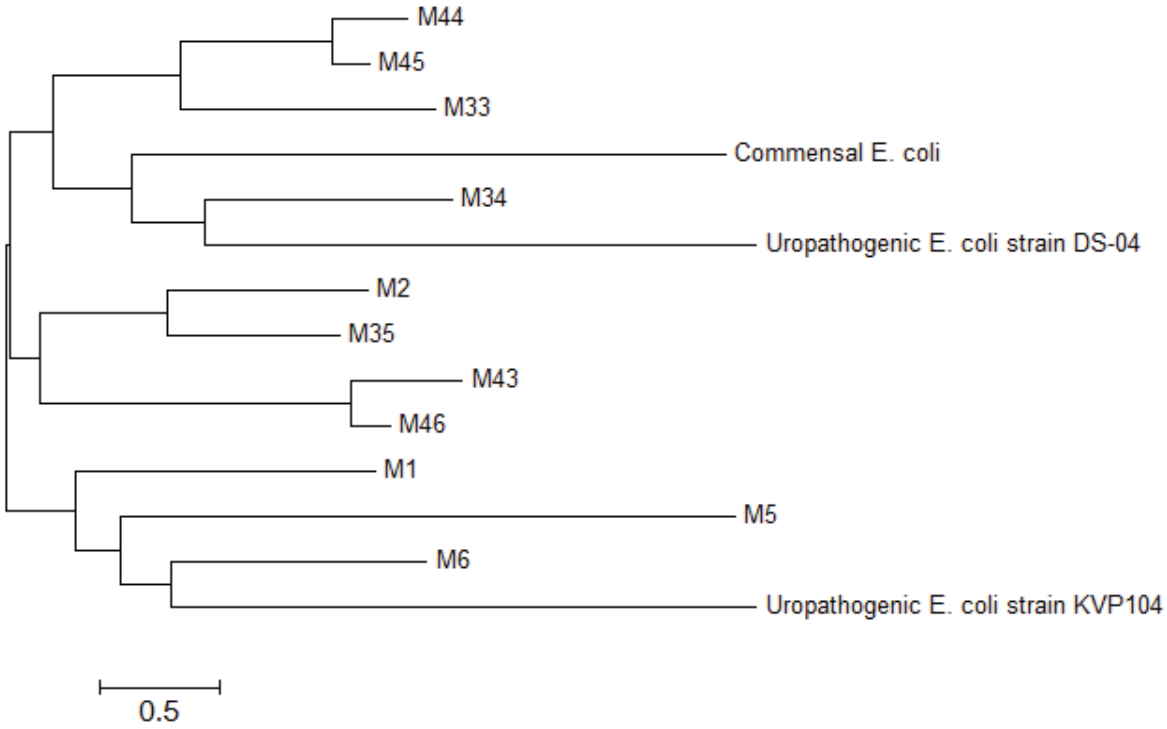

Figure 2. *Evolutionary relationships of studied taxa*

Table 1.*Primers for 16s rDNA used in this study*

Primer	Sequence (5´to 3´)	Tm (°C)	Amplicon size (bp)	Reference
Universal forward	5´AGAGTTTGATCCTGGCTCAG3´	54.3	1000	3
Universal reverse	5´CTTGTGCGGGCCCCCGTCAATTC3´	65.1	1000	3

Table 2.*Volume of Reagents for preparation of PCR Reaction mixture*

Reagent	Volume(µl)	Stock solution	Final concentration
Autoclaved distilled H_2O	04	–	–
Primer 1 forward	01	10 µM	0.5 µM
Primer 2 reverse	01	10 µM	0.5 µM
$MgCl_2$	01	25mM	2.0 mM
DMSO	01	–	–
Master mix	10	2x	1x
Template	02	–	–

Discussion

This study was undertaken to investigate genetic relatedness among Uropathogenic *E. coli* isolated from human patients. With the invention of polymerase chain reaction (PCR) and automated DNA sequencing, the genome of some bacteria has been sequenced completely. A comparison of the genomic sequences of bacterial species showed that the 16S ribosomal RNA (rRNA) gene is highly conserved within a species and among species of the same genus and, hence, can be used as the new gold standard for the specification of bacteria[11]. To study bacterial phylogeny and taxonomy, the 16S rRNA gene sequences are very useful. With the gene presence in almost all bacteria, often existing as a multigene family, or operons, the function of the 16S rRNA gene over time has not changed, suggesting that random sequence changes are a more accurate measure of time and the 16S rRNA gene (1500 bp) is large enough for informatics purposes[12]. Using 16S rRNA sequences, numerous bacterial genera and species have been reclassified and renamed; classification of uncultivable bacteria has been made possible, phylogenetic relationships have been determined, and the discovery and classification of novel bacterial species have been facilitated[11]. This method has been successful in identifying *Enterobacteriaceae* species from a bone marrow transplant recipient[12], and the use of this method to identify or discover novel bacteria in clinical microbiology laboratories has successfully been reported[11,13].

Sequence analysis of the test isolates indicated sequence difference among them in terms of 16s rDNA sequences. Hence, the technique was able to differentiate among the isolates. In earlier studies[12] 16S rRNA gene sequencing was successfully used for bacterial pathogen identification in the clinical laboratory. 16S rRNA gene sequencing has been used for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories[11], and 16S rRNA was developed as a PCR target for detection of *E. coli* in Rainbow Trout[14].*E. coli* isolated from human and animal sources were successfully discriminated using 16s rDNA sequences by earlier studies[15]. Furthermore, it was reported that the use of 16S rRNA gene sequence to study bacterial taxonomy has been used widely for a number of reasons[12]. These reasons include (i) its presence in almost all bacteria, often existing as a multigene family or operons; (ii) the fact that the function of the 16S rRNA gene over time has not changed, suggesting that random sequence changes are a more accurate measure of time (evolution); and (iii) the fact that the 16S rRNA gene (1, 500 bp) is large enough for informatics purposes.

In the present study, comparison of test sequences with reference sequences indicated that one isolate M34 showed similarity to Uropathogenic (GenBank IDs: KX816955.1) as well as commensal *E. coli* (GenBank ID: AB269763.1), whereas M6 showed similarity to UPEC (GenBank ID: JX290084.1). The rest of the isolates were in different clusters. This result indicates that the isolate which was similar to commensal *E. coli* was most likely opportunistic. The rest of the isolates might have accumulated sufficient mutations to be different from others. On the other hand, they might have been phylogenetically pathogenic in origin possessing virulence traits. These results could be confirmed if further phylogenetic tests were done. For example, the isolates could have been tested for the presence of phylogenetic markers.

Conclusions

In the present investigation, the genetic relationship among 11*E. coli* isolated from human was studied. Comparison of the test isolates with uropathogenic and commensal *E. coli* indicated that whereas two isolates showed similarity to the reference strains the rest were in different clusters. This could indicate to accumulation of mutation in the sequences or differences in phylogeny. Further confirmation regarding the origin could be obtained by supporting tests, e.g. detection of phylogenetic markers. Other tests such as ERIC PCR, Rep PCR, Ribotyping, etc. could also be performed for studying variation among the isolates. Increasing the number of test isolates would increase the statistical validity of such results.

Acknowledgement

This study was supported by the Department of Microbiology ofUniversity of Dhaka, Bangladesh and a grant (2012-2013)provided by the University Grants Commission, Government ofBangladesh.

Conflict of interest

Authors declare no conflict of interest.

References

1. Orskov F and Orskov I. 1992. *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol*. **38**: 699–704.
2. Hartstra AV, Bouter KEC, Bäckhed F and Nieuwdorp M. 2015. Insights into the role of the microbiome in obesity and type 2 diabetes. *Diabetes Care.***38**: 159–165.
3. Clarridge JE. 2004. Impact of 16s rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases.*Clin Microbiol Rev.* **17**(4): 840-62.
4. Musser JM, Bemis DA, Ishikawa H and Selander RK. 1987. Clonal diversity and host distribution in *Bordetella bronchiseptica*. *J Bacteriol.***169**: 2793–2803
5. Bingen E, Picard B, Brahimi N, Mathy S, Desjardins P, Elion J and Denamur E. 1998. Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent b2 group strains. *J Infect Dis.***177**(3): 642-50.
6. Johnson JR and Stell AL. 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise.*J Infect Dis*.**181**(1): 261-72.
7. Picard B, Garcia JS, Gouriou S, Duriez P, Brahimi N, Bingen E, Elion J and Denamur E. 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun*.**67**: 546–553.
8. Saitou N and NeiM. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol* .**4(**4): 406–425.

9. Tamura K, Dudley J, Nei M and Kumar S. 2007. Mega 4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol*. **24**(8): 1596–1599.

10. Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.***28**: 2731-2739.

11. Woo PCY, Leung PKL, LeungKW and Yuen KY.2000. Identification by 16s ribosomal RNA gene sequencing of an *Enterobacteriaceae* species from a bone marrow transplant recipient. *J Clin. Pathol*. **53(**4): 211–215.

12. Patel JB. 2001. 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Mol. Diag*. **6**(4): 313–321.

13. Boudewijns M, Bakkers JM, Sturm PDJ and Melchers WJG. 2006. 16S rRNA gene sequencing and the routine clinical microbiology laboratory: a perfect marriage? *JClin Microbiol.***44:** 3469-3470.

14. Fattahi F, Mirvaghefi A, Farahmand H, Rafiee G, and Abdollahi A. 2013. Development of 16s rRNA targeted PCR method for the detection of *Escherichia coli* in rainbow trout (Oncorhynchus mykiss). *Iranian J Pathol*. **8**(1): 36–44.

15. Suardana IW. 2014. Analysis of Nucleotide Sequences of the 16S rRNA Gene of Novel *Escherichia coli* Strains Isolated from Feces of Human and Bali Cattle. *J Nucleic Acids*. 10.1155/2014/475754.