International Journal of Medical Research Professionals

Original Article

Whole Genome Sequence of Multidrug Resistant Escherichia Coli from a Child with Acute Diarrhea

K Rajeshwari[1*], Beena Uppal[2], Rakesh Singh[3]

[1]Director Professor of Pediatrics, Maulana Azad Medical College, New Delhi, INDIA.

[2]Director Professor & Head Microbiology, University of Delhi, FMS, Maulana Azad Medical College, New Delhi, INDIA.

[3]Principal Scientitst, Biotechnology, NBPGR, New Delhi. INDIA.

Article History
Received: 25 Mar 2016
Revised: 27 Mar 2016
Accepted: 30 Mar 2016

***Correspondence to:**
Dr K.Rajeshwari
Director Professor of Pediatrics
Maulana Azad Medical College
New Delhi, India
Rajeshwari.dr@gmail.com

ABSTRACT

Background: Escherichia coli are a common cause of acute diarrheal illness especially in developing countries. Enteropathogenic E.coli (EPEC) commonly affects infants and cause significant morbidity. An EPEC strain O146 was recovered from the stool of an infant with acute diarrhea from a medical college hospital in Delhi, India and was found resistant to many commonly used antimicrobials and was characterized here.

Results: We sequenced the genome of Escherichia coli O146 using Illumina NextSeq 500 platform with 2 x 150 pair-end read chemistry. Approximately ~1 Gb of high quality reads of the sample were assembled using velvet software. We have obtained 188 contigs which comprising of ~4.9Mb genome. It has several genes that are related to virulence and pathogenicity that are responsible for causing disease.

Conclusion: This E.coli O146 is a multiple drug resistant isolate recovered from a child with acute diarrhea .This is the first description of a whole genome sequence of a drug resistant EPEC strain in India.

KEYWORDS: Escherichia coli, EPEC strain O146, Genome, Diarrhea.

INTRODUCTION

Every year more than four billion diarrheal illnesses occur each year and of this majority occur in developing countries[1]. In children less than 5 years it is an important cause of sickness and significantly contributes to mortality. The deaths caused by diarrheal illnesses have decreased worldwide and in India also due to increasing awareness and administration of oral rehydration solutions. The number of deaths due to this disease has declined from over 2 million in the year 2001 to about 1.5 million in 2012[2].The clinical spectrum of EPEC diarrhea is characterised by acute watery diarrhea and occasionally bloody stools are seen. They are implicated as a major causative factor leading to dehydration in less than 12 month olds especially in developing countries. The organisms are transmitted due to ingestion of contaminated water and certain meat products. In a study on acute diarrhea from Niger in 28% of cases one pathogenic bacterium could be identified in children with watery diarrhea and EPEC was isolated in 11% cases among 416 cases of acute diarrhea[3].

Studies from India focusing on EPEC alone in children with acute diarrhea are far and few. In a study from Kolkata in diarrheal stools, multiplex PCR was used to characterise *E.coli* isolates and enteropathogenic *Escherichia coli* was isolated in about 2% cases only.[4] In other studies isolation rates of EPEC was 4.0% in Somalia and 5% in Thailand[5, 6]. To the best of our knowledge there are no published genome sequences from India representing antibiotic resistant EPEC isolates. Here we describe the genome sequence of E.coli O146 isolate from a infant with acute diarrhea admitted to Lok Nayak Hospital and associated Maulana Azad Medical Colege, New Delhi resistant to It was resistant to norfloxacin, ciprofloxacin, ofloxacin, ceftriaxone, cefotaxime, nalidixic acid and amoxicillin.

MATERIALS AND METHODS

Whole genome sequencing was done on a isolate from a one year old male child resident of East Delhi, who was admitted with history of diarrhea of one day duration, was passing 6-12 stools per day, watery, no blood or mucus, vomiting for 24 hours and moderate grade fever of 100-102*F. There were no signs of dehydration. Systemic examination was normal.

Stool examination of the sample revealed a grossly liquid yellow stool and microscopic examination revealed no pus cells, red blood cells, ova, cyst or parasite. No budding yeast cells seen. For isolation of E. *coli*, stool specimen was plated on MacConkey (Hi Media), followed by incubation for 16–18 hrs at 37°C. Five typical, lactose fermenting pink colour colonies per sample, were selected confirmed as *E.coli* by their motility and standard biochemical reactions.

For identification of EPEC, slide agglutination with antisera to common EPEC O antigens, was carried out. E. coli strain grown on a nutrient agar plate, was suspended in normal saline solution, autoclaved for 15 minutes and then examined by slide agglutination using commercially available antisera, in a kit identified as "Pathogenic E. coli Antisera" (Denka Seiken Co.,Ltd.,Tokyo, Japan). Serotyping showed it to be O146.

DNA isolation was done using MB505 HiPurA™ Bacterial and Yeast Genomic DNA Miniprep Purification Spin Kit procured from HiMedia Labs, Mumbai. Genomic DNA was stored at -80˚C and used for sequencing.

ANTIBIOGRAM TESTING

The isolate was subjected to antibiotic sensitivity testing using Kirby Bauer's disc diffusion method on Mueller Hinton Plates as per CLSI guidelines. It was carried out with standard discs Amikacin (30mcg), Ampicillin (10mcg), Azithromycin (15mcg), Cefotaxime (30mcg), Ceftriaxone (30mcg),Ciprofloxacin (5mcg), Gentamicin (10mcg), Nalidixic Acid (30 mcg), Nitrofurantoin (300 mcg), Norfloxacin (10mcg), Ofloxacin (5mcg).

The time to recovery was 24 hours. Child received intravenous fluids and injection ceftriaxone. As there was no response to ceftrixone, injection amikacin was added and child subsequently improved. No complications were seen and the child had an uneventful hospital stay and subsequently discharged.

GENOME SEQUENCING

Qualitative and quantitative analysis of gDNA

Quality was checked on 1% agarose gel (loaded 5 µl) for the single intact band. The gel was run at 110 V for 30 mins. 1 µl of each sample was loaded in Nanodrop 8000 for determining $A_{260/280}$ ratio and 1 µl of each sample was used for determining concentration using Qubit® 2.0 Fluorometer.

Preparation of 2 x150 NextSeq library

The paired-end sequencing library was prepared using Illumina TruSeq Nano DNA HT Library Preparation Kit. 200ng gDNA was fragmented by Covaris to generate a mean fragment distribution of 550bp. Covaris shearing generates dsDNA fragments with 3' or 5' overhangs. The fragments were then subjected to end-repair. This process converts the overhangs resulting from fragmentation into blunt ends using End Repair Mix. The 3' to 5' exonuclease activity of this mix removes the 3' overhangs and the 5' to 3' polymerase activity fills in the 5' overhangs. A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment.

This strategy ensures a low rate of chimera (concatenated template) formation. Indexing adapters were ligated to the ends of the DNA fragments, preparing them for hybridization onto a flow cell. The ligated products were purified using SP beads supplied in the kit. The size-selected product was PCR amplified as described in the kit protocol.

Quantity and quality check (QC) of library on Bioanalyzer

The amplified library was analyzed in Bioanalyzer 2100 (Agilent Technologies) using High Sensitivity (HS) DNA chip as per manufacturer's instructions.

Cluster Generation and Sequencing

After obtaining the Qubit concentration for the library and the mean peak size from Bioanalyser profile, library was loaded onto NextSeq for cluster generation and sequencing.

Paired-End sequencing allows the template fragments to be sequenced in both the forward and reverse directions on NextSeq. The kit reagents were used in binding of samples to complementary adapter oligos on paired-end flow cell. The adapters were designed to allow selective cleavage of the forward strands after resynthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment.

The library was prepared using illumina TruSeq Nano DNA HT Sample Preparation Kit. The mean fragment size of the library is 980. The next generation sequencing run for Sample was performed using Paired end (PE) 2x150bp library on Illumina NextSeq to generate 1 GB data.

Quality check

The raw data was filtered using Trimmomatic v0.30 and generate the high quality data for further assembly.

Parameters considered for filtration are as follows:

- Adapter trimming
- SLIDINGWINDOW: Perform a sliding window trimming of 20 bp, cutting once the average quality within the window falls below a threshold of 20.
- LEADING: Cut bases off the start of a read, if below a threshold quality of 20.
- TRAILING: Cut bases off the end of a read, if below a threshold quality of 20.
- MINLENGTH: Drop the read if it is below 50 bp length.

After the quality filtration (mean quality score >=20) and adaptor trimming using Trimmomatic as described above, the high quality reads are used for assembly. High quality reads of Sample-535 was assembled using Velvet with kmer-131, exp_cov auto and cov_cutoff auto parameters. A total of 188 validated contigs were obtained for the sample. The list of software used for assembly are shown in table 1.

RESULTS

The total number of contigs were 188, total contig size of 4954160 bp, average contig size 26351 bp, contig N50 92999 and a maximum contig size of 293318 bp. The contigs were used to identify the coding region using Prodigal on default parameters to identify the coding sequences from assembled contigs of the sample. The CDS statistics of the sample are shown in table 2.

Table1: List of Software used for Assembly

Software	Version
Trimmomatic	0.30
Velvet	1.2.10
BLAST2GO	2.7.0
Misa	Perl scripts

Table 2: CDS statistics of the Sample

Description	Sample
Number of CDS	4,538
Total CDS Size	4,288,497
Average CDS Size	945
CDS N50	1,173
Max CDS Size	9,492

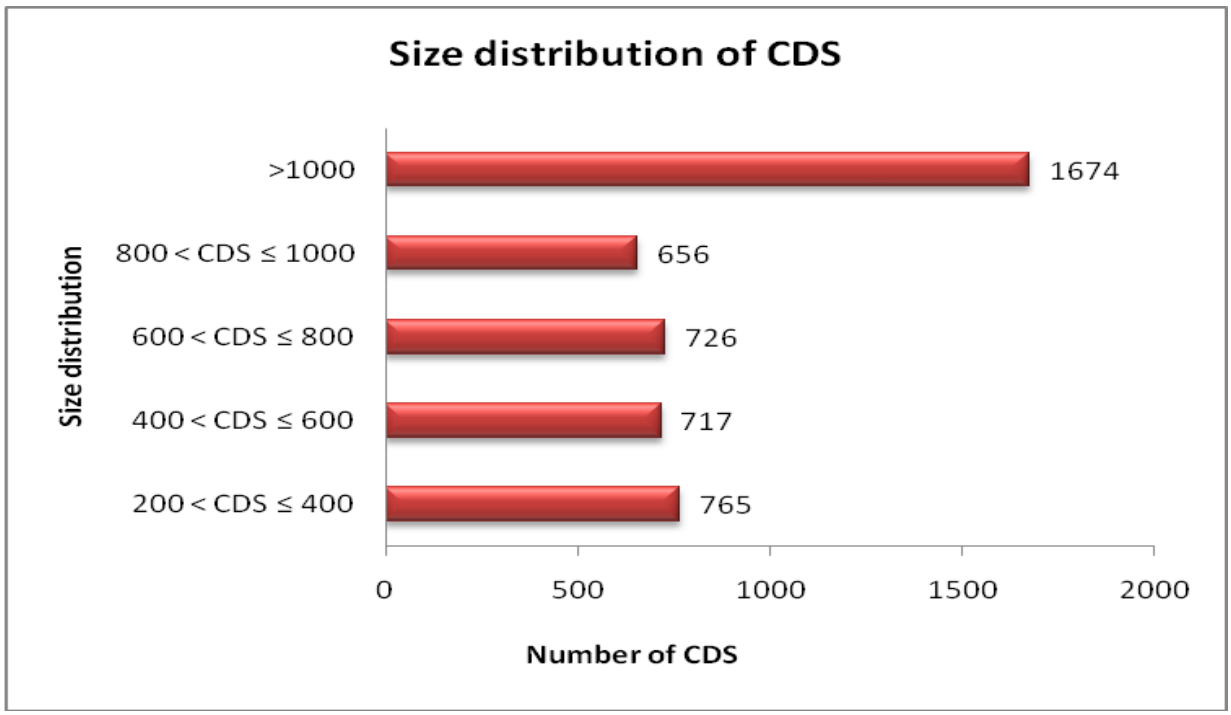
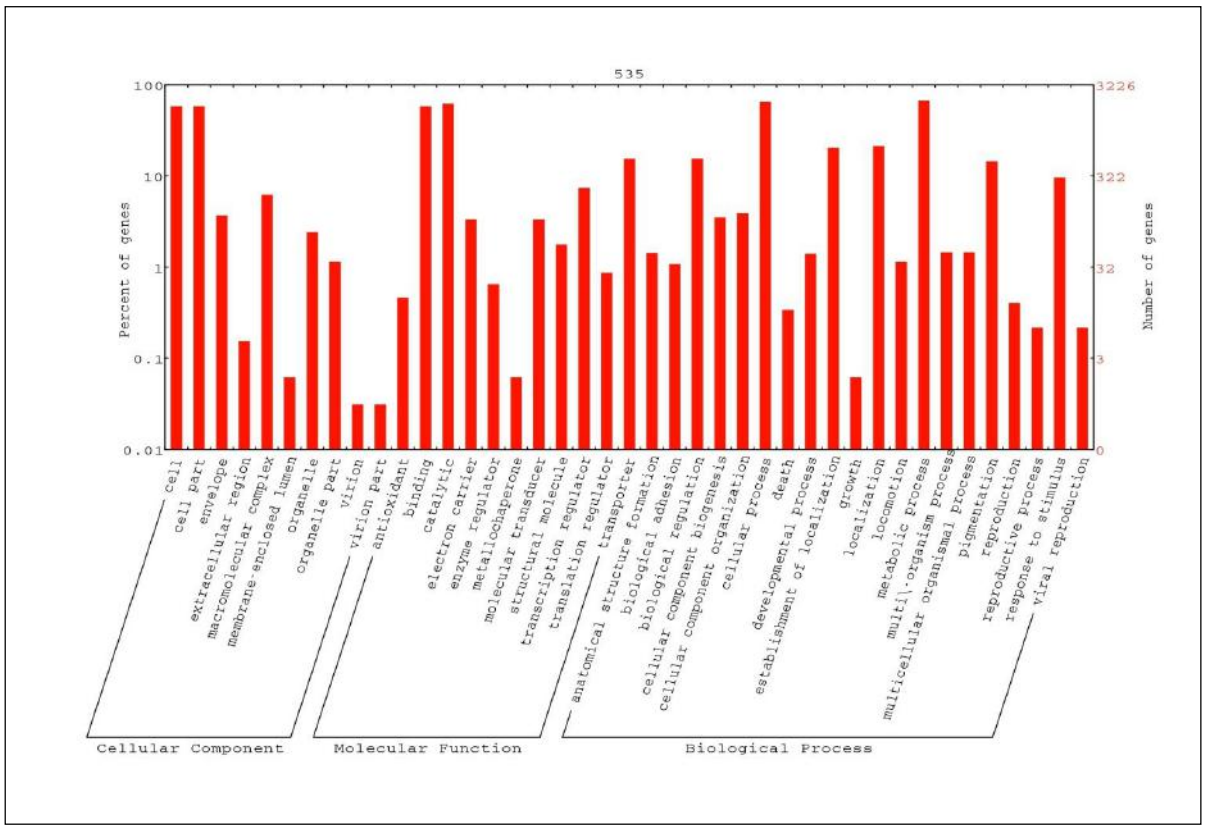


Figure 1: Distribution of CDS according to their length

The functional annotation was performed using the 4,538 CDS by aligning those CDS to non-redundant database of NCBI using BLASTX with an e-value cutoff of 1e-06.BLASTX resulted in the annotation of 4,521 CDS and 17 CDS had no significant blast hit for Sample-535. The majority of hits (87%) were found to be against *Escherichia coli.*

Gene Ontology analysis helps us in specifying all the annotated nodes comprising of GO functional group such as Biological Process, Molecular Function and Cellular Component. The GO distribution of CDS for the Sample was done and Biological Process (2,784), Molecular Function (2,876) and Cellular Component (1,875) were identified (Fig 2).

Mapping of CDS to the biological pathways were performed blast against the KEGG database using KAAS server. A total of 2,266 CDS mapped in to metabolic pathways, majority of which are falling in to carbohydrate metabolism (313), membrane transport (266), carbon metabolism (217), amino acid metabolism (141) etc. The pathway analysis of the CDS has been provided in Table 3.



Various genes causing resistance to betalactam antibiotics due to production of betalactamases by various genes were identified in the sample. They are listed below.

Drug resistance		
00312 beta-Lactam resistance [PATH:ko00312]	**Gene name**	**function**
CDS_881_contig_24	K09476 ompF	outer membrane pore protein F
CDS_3686_contig_165	K09476 ompF	outer membrane pore protein F
CDS_2273_contig_75	K09475 ompC	outer membrane pore protein C
CDS_1133_contig_30	K08218 ampG	MFS transporter, PAT family, beta-lactamase induction signal transducer AmpG
CDS_124_contig_2	K01207 E3.2.1.52, nagZ	beta-N-acetylhexosaminidase [EC:3.2.1.52]
CDS_3145_contig_141	K01467 ampC	beta-lactamase class C [EC:3.5.2.6]
CDS_672_contig_18	K15580 oppA, mppA	oligopeptide transport system substrate-binding protein
CDS_1771_contig_45	K15580 oppA, mppA	oligopeptide transport system substrate-binding protein
CDS_1772_contig_45	K15581 oppB	oligopeptide transport system permease protein
CDS_1773_contig_45	K15582 oppC	oligopeptide transport system permease protein
CDS_1774_contig_45	K15583 oppD	oligopeptide transport system ATP-binding protein
CDS_1775_contig_45	K10823 oppF	oligopeptide transport system ATP-binding protein
CDS_1032_contig_27	K17836 penP, blaZ	beta-lactamase class A [EC:3.5.2.6]
CDS_2127_contig_59	K18148 rtcB	release factor H-coupled RctB family protein
CDS_1173_contig_30	K03585 acrA, mexA, adeI, smeD	membrane fusion protein
CDS_1172_contig_30	K18138 acrB, mexB, adeJ, smeE	multidrug efflux pump
CDS_955_contig_25	K12340 tolC	outer membrane channel protein
CDS_4421_contig_186	K05366 mrcA	penicillin-binding protein 1A [EC:2.4.1.- 3.4.-.-]
CDS_1311_contig_33	K05515 mrdA	penicillin-binding protein 2
CDS_1957_contig_52	K03587 ftsI	cell division protein FtsI (penicillin-binding protein 3)

Table 3: Genes involved in functional pathways

Pathway	No of CDS
METABOLISM	
Carbon metabolism	217
Carbohydrate metabolism	313
Energy metabolism	160
Lipid metabolism	69
Nucleotide metabolism	115
Amino acid metabolism	210
Metabolism of other amino acids	70
Glycan biosynthesis and metabolism	53
Metabolism of cofactors and vitamins	161
Metabolism of terpenoids and polyketides	39
Biosynthesis of other secondary metabolites	19
Xenobiotics biodegradation and metabolism	46
GENETIC INFORMATION PROCESSING	
Transcription	4
Translation	78
Folding, sorting and degradation	52
ENVIRONMENTAL INFORMATION PROCESSING	
Replication and repair	58
Membrane transport	266
Signal transduction	159
CELLULAR PROCESSES	
Transport and catabolism	8
Cell motility	54
Cell growth and death	12
Immune system	1
Endocrine system	9
Digestive system	2
Excretory system	2
Nervous system	4
ORGANISMAL SYSTEM	
Environmental adaptation	6
Cancers	12
Immune diseases	2
Neurodegenerative diseases	6
Endocrine and metabolic diseases	4
Infectious diseases	35
Drug resistance	20

DISCUSSION & CONCLUSION

Diarrheagenic *E.coli* include enteropathogenic *E.coli* (EPEC), enterotoxigenic *E.coli* (ETEC), enteroinvasive *E.coli* (EIEC), enteroaggregative *E.coli* (EAEC) and shiga toxin producing *E.coli* (STEC)[7]. EPEC, in addition to their ability to induce attaching and effacing (AE) lesions, also possesses a large EPEC adherence factor (EAF) plasmid and the cluster of genes that encode the bundle-forming pili (BFP)[8,9]. The EPEC strains which carry the eae gene but lack the EAF plasmid and the stx gene are described as atypical EPEC. This bacterium had eaea gene and bfpA gene suggesting it to be typical EPEC. It also had the espA gene. Inaddition we have identified several virulence related genes such as toxin, adhesion, invasion and capsule biosynthesis from the sample.

Whole Genome analysis of the Sample was sequenced on Illumina NextSeq platform with 2 x 150 pair-ends read chemistry. Approximately ~1 Gb of high quality reads of sample-535 were assembled using velvet software.

We have obtained 188 contigs which comprising of ~4.9Mb genome. Out of 188 contig, we have predicted 4,538 coding sequences (CDS). We obtained BLAST hits of 99% of CDS against NCBI database. From the Gene Ontology analysis, a total number of 3,226 GO terms were assigned to CDS.

The whole genome has been named MAMC 18 after the institution and ward of isolation and has been deposited with the National Centre For Biotechnology Information, Bethesda, Maryland, US, where it has been granted an accession number. This is the first ever instance of a whole genome sequenced by the Department of Microbiology, Maulana Azad Medical College.

Nucleotide Sequence Accession Number.

This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JWHN00000000 and is available for complete viewing.

Antibiotic Resistance Profile

It was sensitive to nitrofurantoin, azithromycin, gentamicin, amikacin. It was resistant to norfloxacin, ciprofloxacin, ofloxacin, ceftriaxone, cefotaxime, nalidixic acid, amoxicillin. To better understand the genetic basics of antibiotic resistance the genome was searched for specific genes which conferred antibiotic resistance. A list of these genes is provided as additional file 1.

This bacterium is chacracterised by genes responsible for multidrug efflux system, multidrug transporter genes, putative multidrug efflux transporter MdtA, acrD multidrug efflux pump, MATE family multidrug exporter and anerobic multidrug efflux transporter.

Also present is gene encoding Tol C, the bacterial protein playing a important role in the expulsion of various diverse molecules like protein toxins and efflux of various antibacterial drugs from inside the cell. TolC is a trimeric twelve strand barrel forming a transperiplasmic tunnel[10]. This establishes a long single pore dissimilar to other membrane proteins and forms an exit duct to various substrates. TolC is recruited by a substrate laden translocase and then substrate passage is facilitated from cytosol to external to the external cellular environment. This explains the presence of multidrug resistance in the sample. Further the lack of response to ceftriaxone seen in this child could be explained by the presence of these various genes and hence amikacin was added for clinical response.

To conclude this bacterium was a typical EPEC having both eaea and bfpA genes .It also had many genes encoding betalactamases and other genes causing multiple drug resistance. The increasing antibiotic resistance seen among E.coli isolates from developing countries can thus be explained by the presence of multiple genes. Drugs need to be developed which can block the expression of these genes and thus result in clinical response in children affected with acute diarrhea.

Only recently the vast extent of ecological disaster due to ESBL producing *Enterobacteriaceae* in Asia is the focus of intense research. It is possible that some of the successful ESBL-producing strains originate from Asia. Deficient sewage systems and very poor quality of potable drinking water, in combination with a lack of strict control over prescription and sales of antibiotics are the major factors that have promoted the development of drug resistance.

REFERENCES

1. Lakshminarayanan S, Jayalakshmy R. Diarrheal diseases among children in India: Current scenario and future perspectives. J Nat Sc Biol Med 2015; 6:24-8.

2. Bhan MK. Accelerated progress to reduce under-5 mortality in India. Lancet Glob Health 2013; 1:e172-3.

3. Langendorf C, Le Hello S, Moumouni A, Gouali M, Mamaty AA, Grais RFet al.Enteric bacterial pathogens in children with diarrhea in Niger: diversity and antimicrobial resistance. PLoS One. 2015 Mar 23; 10(3):e0120275. doi: 10.1371/journal.pone.0120275. eCollection 2015.

4. Dutta S, Guin S, Ghosh S, Pazhani GP, Rajendran K,Bhattacharya MK et al. Trends in the Prevalence of Diarrheagenic Escherichia coli among Hospitalized Diarrheal Patients in Kolkata, India. PLoS ONE 2013;8(2): e56068. doi:10.1371/journal.pone.0056068.

5. Casalino M, Yusuf MW, Nicoletti M, Bazzicalupo P, Coppo A,Colonna B et al. A two-year study of enteric infections associated with diarrheal diseases in children in urban Somalia. Trans R Soc Trop Med Hyg 1988;82: 637–641.

6. Echeverria P, Orskov F, Orskov I, Knutton S, Scheutz F, Brown JE. Attaching and effacing enteropathogenic Escherichia coli as a cause of infantile diarrhea in Bangkok. J Infect Dis 1991; 164: 550–554.

7. Nataro, Kaper JB. Diarrhoeagenic Escherichia coli. Clin Microbiol Rev 1998; 11:142-201.

8. Clarke SC, Haigh RD, Freestone PE, Williams PH. Virulence of ente-ropathogenic Escherichia coli, a global pathogen. Clin Microbiol Rev 2003; 16:365-78.

9. Nunes BE, Saridakis OH, Irino K. Genotypic and phenotypic charac-terization of attaching and effacing Escherichia coli (AEEC) which were isolated from children with and without diarrhoea Londrina, Brazil. J Med Microbiol 2003; 52:499-504.

10. Koronakis V, Eswaran J, Hughes C. S. Structure and function of TolC: the bacterial exit duct for proteins and drugs.Annu Rev Biochem2004;73:467-89.

Source of Support: Nil.

Conflict of Interest: None Declared.

Cite this article as: K Rajeshwari, Beena Uppal, Rakesh Singh. Whole Genome Sequence of Multidrug Resistant Escherichia Coli from a Child with Acute Diarrhea. Int J Med Res Prof. 2016, 2(2); 274-79.