ISSN Online: 2349-7785

Genome Analysis to Identify Virulence-Related Features Present in an Emerging Enteric Pathogen *Escherichia albertii* KF1 as Compared to that of the Prototypical *Escherichia coli* O127:H6 E2348/69

Shelly Sinha [1, 2], Sukanta Kumar Sen [1] and Bomba Dam [1]*

[1] Microbiology Laboratory, Department of Botany, (DST-FIST & UGC-DRS Funded), Institute of Science, Visva-Bharati (A Central University), Santiniketan, West Bengal 731235, India
[2] Department of Botany, Rabindra Mahavidyalaya, Champadanga, Hooghly, West Bengal 712401, India

Received 23 October 2017; accepted in revised form 30 November 2017

Abstract: Diarrhoea has a high impact on the rate of morbidity and mortality of infants less than five years in developing countries. Among different etiological agents, the most common causal organisms for persistent diarrhoea in children are enteric bacterial genera like *Escherichia*, *Shigella*, *Salmonella*, *Klebsiella* and *Enterobacter*. Out of them *E. coli* is the most dominant and have several pathotypes that are commonly referred as enteropathogenic *E. coli* (EPEC). Apart from the species *E. coli*, *Escherichia albertii* KF1 has recently been identified as an emerging enteropathogen associated with infant diarrhoea. In this study we have performed in silico analyses on the genome sequence of this strain using different bioinformatic platforms and tried to identify regions or genes related to virulence which are in common to the prototypical *Escherichia coli* O127:H6 E2348/69 and uniquely present in it. Both the genomes share 3374 CDS, with some conserved pathogenicity encoded features like the typical enteropathogenic type three secretion system (T3SS) associated with locus of enterocyte effacement (LEE) and non-LEE effectors. In addition, several singletons are also identified in the genome of *E. albertii* KF1 within eight potential genomic islands (GEIs). This includes genes encoding barrel domain-containing outer membrane with a putative type five secretion system (T5SS), multidrug export proteins (MATE) and a region coding for flagellar synthesis and a large number of transposases and hypothetical proteins of unknown functions. Therefore, the study is very relevant to the complete understanding of the importance of exclusive GEIs and virulence related singletons along with the conserved coding regions to determine the pathogenicity and emergence of *E. albertii* KF1 as an enteropathogen.

Key words: Comparative genomics, enteropathogen, *Escherichia coli* O127:H6 E2348/69, *Escherichia albertii* KF1, genomic islands.

Introduction

Diarrhoea is one of the threatening and most important disease that cause global morbidity and mortality in children. Globally, millions of children under five succumbed to death every year, due to recurrent occurrence of the infection (www.who.int/mediacentre/factsheets/fs330/). It is a syndrome caused by a combination of factors that include severe fluid loss, dehydration and blood dysentery along with septic bacterial infection [6]. Toxins produced upon infection lead to malabsorption and subsequent malnutrition with cognitive disability [27]. Among different etiological agents, like bacteria, virus and other parasitic organisms, enteropathogenic bacteria are the leading causes for persistent diarrheal infection in children [20]. The

*Corresponding author (Bomba Dam)
E-mail: < bomba.dam@visva-bharati.ac.in >

common enterobacteria known to cause the infection are members of the genera *Escherichia*, *Shigella*, *Salmonella*, *Klebsiella* and *Enterobacter* with *Escherichia coli* being the most dominant agent [24].

E. coli can vary in their pathogenicity due to a diverse set of genes present in them. These pathotypes taken together are collectively called diarrheagenic *E. coli* (DEC) [19]. Among them, enteropathogenic *E. coli* (EPEC) cause severe and fatal infantile diarrhea [7]. They are again divided into typical (tEPEC) and atypical (aEPEC) sub-groups based on presence or absence of *bfp* operon (encoding bundle forming pilus) [11]. During the series of sporadic outbreaks of infantile diarrhea in 1940s and 1950s, tEPEC isolates were found responsible in majority of them [5,10,23]. Apart from the most common species '*coli*' of genus *Escherichia*, other reported species with pathogenicity includes *fergusonii*, *albertii*, *vulneris*, *blattae* and *marmotae* [7].

Among them, *E. albertii* KF1 is an emerging enteropathogen and has been isolated in 2003, from diarrheic children [14]. Several of its strains have been completely sequenced to understand the genomic features present therein [22]. The organism has been recently recognised as a close relative of *E. coli*, showing different types of pathogenesis [22]. The detailed genomic feature of this emerging enteropathogen, which distinguishes it from other *Escherichia* species, is still poorly understood. The strain have been earlier recognized as *Hafnia alvei* [2] but biochemical and molecular phylogenetic analysis placed it more close to *E. coli* and *Shigella boydii* [1, 25]. However, several variations have been reported in this new enteric pathogen as compared to typical enteropathogenic strain that separates it from *E. coli* pathogroups [8]. Some of the major variations include its biochemical inability to ferment lactose, D-sorbitol and D-xylose and immunologically by lacking invasion antigen H [1].

Availability of complete genome sequence empowered researchers to explore different genomes and interpret the genes present therein. A substantial number of genes in most of the genomes remained functionally unannotated or classified as 'putative' or 'hypothetical'. Therefore, re-annotation followed by a detailed bioinformatic characterization of relevant genomes would be rewarding in terms of identifying novel features.

In this study, the available genome sequences of two enteropathogenic strains, including the emerging *E. albertii* KF1, and the prototypical enteropathogen *E. coli* O127:H6 E2348/69 [15] are re-annotated and characterized in detail. The main focus was to identify genomic islands (GEIs) and other pathogenicity related regions particularly in *E. albertii* KF1 that would justify its development of enteropathogenicity.

Materials and methods

Data source

Genome sequences of *E. coli* O127:H6 E2348/69 (Accession no. FM180568) and *E. albertii* KF1 (Accession no. CP007025) are retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/).

Bioinformatic analyses of genome sequences

The tRNA coding genes are predicted with tRNA scan-SE 1.21 version [18]. RNAmmer 1.2 was used to predict rRNA genes [16]. Description of protein function from genomes sequences was accessed through SWISS-PROT database. Accessory regions of the strains were determined by finding genomic islands (GEIs) using Island Viewer 4 webserver [4]. Comparative analysis of respective genomes is done for the prediction and interactive visualization of GEIs (regions of probable horizontal origin) using IslandViewer 4 webserver (http://www.pathogenomics.sfu.ca/islandviewer). The predictions were made by four different software platforms namely, Islander method [13], IslandPick [17], SIGI-HMM [26], Island Path-DIMOB [12] and an integrated approach that combines the result of all four predictions. The IslandPath-DIMOB method was subsequently used for detailed prediction and analysis of the GEIs along with recognition of different genes responsible for virulence.

Accurate annotations of genes as virulence factors in enteropathogenic genomes were done through Virulence Factor Database (VFDB). A collection of manually curated antibiotic resistance genes as well as a predictive tool (Resistance Gene

Identifier) was performed using Comprehensive Antibiotic Resistance Database (CARD). Virulence factor coding genes were analysed using wg-VISTA alignment result of conserved region [9]. Genomes were re-annotated using RAST [3] platform to have a proper metabolic reconstruction. Pair wise alignment of the corresponding proteins was done using ClustalW, applying default parameters.

Results and discussions

Basic description of the two genomes

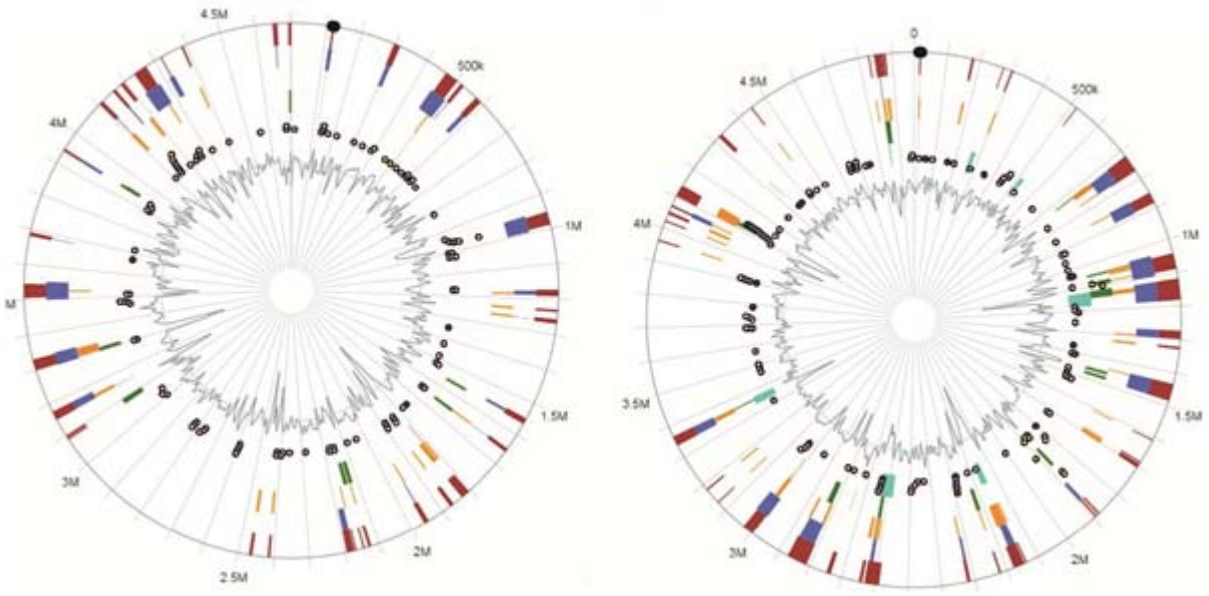
The basic features of the genomes of the prototypical *E. coli* O127:H6 E2348/69 and the emerging enteropathogen *E. albertii* KF1 are enlisted in Table 1. The two genomes are referred here after in the manuscript as *E. coli* O127 and *E. albertii* KF1 or simply strains O127 and KF1. The genome size of *E. albertii* KF1 is 4.7 Mb, and that of *E. coli* O127 is 5.0 Mb (Table 1). Although having a ~300 Kb smaller genome, *E. albertii* KF1 encode sufficiently high number of total CDS (4823) as compared to *E. coli* O127 (4886), thereby increasing its overall gene density (1.026 genes/Kb). Both the genomes have seven rRNA operons, but vary in the number of tRNA coding genes with higher numbers in strain O127. The number of pseudogenes is higher (391) in the genome of *E. albertii* KF1 that after re-annotation on RAST platform reduced to 121. However, the same in *E. coli* O127 remain unchanged to 149 even after re-annotation. The number of protein coding genes in *E. albertii* KF1 and *E. coli* O127 are 4312 and 4548 respectively, with 3374 sequences being shared between the two genomes. Large number of genes encoding hypothetical or conserved hypothetical protein with unknown functions is identified in *E. albertii* KF1. The average nucleotide identity (ANI) percentage between the two genomes is low (89 %) thereby suggesting high divergence. Even the observed core / pan genome ratio is 60 % that represents high variability among them.

Genomic islands

Genomes evolve through varied processes including mutation, rearrangement of segments or horizontal gene transfer. Pathogenic bacterial strains evolve by acquiring virulence factors and organizing them in specific regions of their genome that are referred as genomic islands (GEI). Presence of GEIs in both the genome was predicted using the web server IslandViewer 4 that have four different bioinformatic prediction methods, namely Islander, IslandPick, SIGI-HMM, IslandPath-DIMOB methods and an integrated platform combining all four predictions. The total GEIs count identified in the two genomes using different platforms varied significantly (Fig. 1). While, Islander predicted six GEIs in *E. coli* O127, none is detected in *E. albertii* KF1. IslandPick identified seven and eight putative GEIs respectively in the two genomes. Similarly, SIGI-HMM that analysed individual genes of codon usage, anticipated 33 GEIs for strain O127 and 18 for strain KF1. However, IslandPath-DIMOB, predicted 14 and 20 GEIs associated regions in genomes of strains O127 and KF1. The involvement

Table 1. Complete genome features of *E. albertii* KF1 and *E. coli* O127:H6 E2348/69

General features	*E. albertii* KF1	*E. coli* O127:H6 E2348/69
Accession number	CP007025	FM180568
Size (Mb)	4.7	5.0
No. of genes	4823	4886
No. of proteins	4312	4548
Gene density (gene/Kbp)	1.026	0.977
GC content	49.7	50.6
No. of tRNA	86	92
No. of rRNA operons	07	07
No. of pseudogenes	391	149
No. of singletons	778	961



(A) (B)

Fig. 1. Graphical circular view of the genomes of (A) *E. albertii* KF1 and (B) *E. coli* O127:H6 E2348/69 with their predicted GEIs, using IslandViewer 4 platform. Size of the genome is marked outside. Visualization of predicted GEIs by different methods, from inside to outside: Islander (turquoise), IslandPick (green), SIGI-HMM (orange), IslandPath-DIMOB (blue) and Integrated (dark red). Small filled circles represent genes coding for virulence (purple), antimicrobial resistance (pink) and pathogenicity associated factors (yellow). The innermost radiating circle shows the GC skew of the genome.

of the accessory genomic region of *E. coli* O127 in virulence and resistance has already been reported [15]. Thus, the genome of *E. albertii* KF1 is explored further using IslandPath-DIMOB that scrutinizes the genome based on the nucleotide biasness and presence of mobility genes (Fig. 2). Out of the 20 selected regions in this emerging pathogen, eight have been considered as potential candidate GEIs based on their size and functions (Table 2). They span a large region of the genome (23 Kb to 38 Kb), except GEI6 and GEI7, being smaller in size. Pathogenicity and virulence related genes are identified in all GEIs. Moreover, out of the eight GEIs, five are related with potential virulence and antibiotic resistance coding genes while remaining three have putative virulence related genes.

Genes encoded within GEIs

The GEIs include the accessory gene pool and contain common features such as insertion sequences (IS), resistance and virulence factors, prophages and small mobile units (Table 2). Apart from the virulence related factors, they also contain some basic metabolism related genes that encodes, polysaccharide export protein Wza, protein-tyrosine-phosphatase, protein-tyrosine kinase, cytochrome d ubiquinol oxidase subunit 2 protein, C4-dicarboxylate ABC transporter, nitrite extrusion protein 2 and SsrA-binding protein, ubiquinone-binding protein and AMP nucleosidase. The virulence related gene identified in the different GEI code for type III secretion system (T3SS), membrane proteins, outer membrane pore InvG, effectors NleG and enterobacterial Ail/Lom family protein. Additional shared genes include those encoding membrane transport, cell division and iron uptake.

Transposases of at least two broad IS3 and HTH superfamilies are noticed in both the genomes, with little similarity among themselves (Table 3). One interesting observation in the GEI2 of *E. albertii* KF1 is the presence of a fully conserved and complete gene cluster of the flagella biosynthesis op-

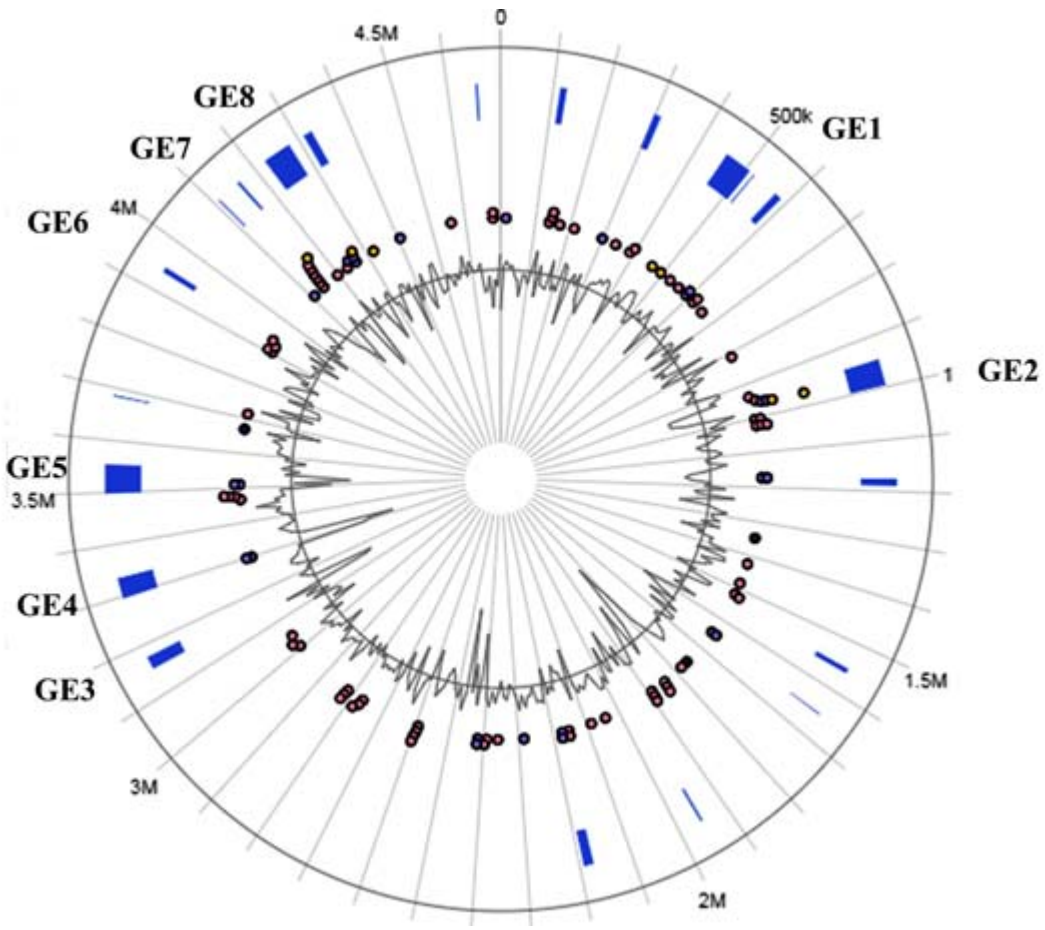


Fig. 2. Prediction of eight candidate GEIs using IslandPath-DIMOB in the genome of *E. albertii* KF1. The size of the genome is shown besides the circular view. The GEIs are shown in blue and are marked as GE1 to GE8. Out of the 12 identified GEIs eight regions are considered with potential candidate GEIs for this study, as they encode the genes associated with pathogenicity. The innermost radiating circle shows the GC skew of the genome. Virulence coding genes (purple), antimicrobial resistance coding genes (pink) and pathogenicity associated genes (yellow) presented as smaller circles.

Table 2. List of GEIs identified in *E. albertii* KF1 genome with their encoded genes and function

GEI	Character*	Size#	Start	End	Functional components$
GEI1	Virulence associated	38131	463080	501211	Polysaccharide export protein Wza (3) Protein-tyrosine-phosphatase (2) Protein-tyrosine kinase (2) Phosphoanhydride phosphorylase (2) Membrane protein (1) Cytochrome d ubiquinol oxidase subunit (2) Cytochrome bd-II ubiquinol oxidase subunit (2) Hydrogenase expression/formation

table 2. (continued).

GEI	Character*	Size#	Start	End	Functional components$
					protein (2) Hydrogenase-1 operon protein HyaE (2) Hydrogenase 1 maturation protease (2) Hydrogenase 1 b-type cytochrome subunit (2)
GEI2	-	37011	952416	989427	Protease (3) Transcriptional regulator (2) Flagella related *flh/fli* (1) Type III secretion system protein (1) Phage tail protein (1) Transposase (1)
GEI3	Virulence associated	23736	3152315	3176051	C4-dicarboxylate ABC transporter (2) Nitrite extrusion protein (2) Toxin B (1) TatD family hydrolase (1)
GEI4	Virulence associated	30952	3300717	3331669	T3SS outer membrane pore InvG (1) T3SS inner membrane ring protein (1) T3SS protein InvA (1)
GEI5	-	39395	3506011	3545406	Tail assembly (1) Late control (4) SsrA-binding protein (2) Ubiquinone-binding protein (2) RnfH family protein (3) Outer membrane protein assembly factor BamE (1)
GEI6	Virulence associated	7039	3939556	3946595	DUF2138 domain-containing protein (4) T3SS effector NleG (1) Transposase (1)
GE17	-	9264	4155516	4164780	Transposase (4) Protein ninE (4) Enterobacterial Ail/Lom family protein (1)
GEI8	Virulence associated	38854	4232349	4271203	MATE family multidrug exporter (1) AMP nucleosidase (2) MFS transporter (1) Metal-binding protein ZinT (1) Sulfoxide reductase heme-binding subunit YedZ (1) T3SS effector protein NleD (1)

Functions of the proteins are indicated by numbers indicated in brackets at the end.
(1) virulence linked; (2) metabolism related; (3) role in both metabolism and virulence; (4) unknown function
*The GEI2, GEI5 and GEI7 do not have genes that are linked with virulence directly, but have proteins which are putatively linked with pathogenicity
$A large number of hypothetical proteins are present within every genomic island
Except GEI6 and GEI7, which are small in size (7 and 9 Kb respectively), rest of the GEIs mentioned above are of size ranging within 23Kb to 38Kb.

Table 3. Comparative analysis of virulence coding regions of *E. albertii* KF1 and *E. coli* O127:H6 E2348/69

Virulence factor	*E. albertii* KF1			*E. coli* O127:H6 E2348/69			Similarity (%)	Function
	Start	End	Protein id	Start	End	Protein id		
Porin	295994	314521	WP_025237139	265574	266629	WP_001339261	90	Membrane transport
Glycine zipper family protein	308418	308762	WP_025237144	1329912	1330256	WP_000726974	97	Membrane transport
Translocated intimin receptor Tir	1807892	1809634	WP_025237816	4113830	4115482	WP_001339882	96	Membrane transport
Antitermination protein	489567	490388	WP_025237241	1058356	1058979	WP_001235460	90	Membrane protein
Intimin	1804393	1807221	WP_025237814	4110343	4113162	WP_000627890	90	Membrane protein
Intimin-like SinH	-	-	-	2860904	2863108	WP_024201073	-	Membrane protein
Temperature sensitive hemoglutinin	2179155	2182946	WP_252379741	-	-	-	-	Outer membrane protein
Protease	311484	312431	WP_025237145	51793	51881	WP_001201835	<40	Outer membrane protein
Holin	463080	501211	WP_000839572	743713	743919	WP_000411802	96	Fimbrilin protein
EscD/YscD/HrpQ T3SS inner membrane ring	1802904	1804124	WP_025237813	-	-	-	-	Secretion system
Outer membrane autotransporter barrel domain	303386	307813	WP_025237143	-	-	-	-	T5SS
Outer membrane PhoE	3990809	3993397	WP_025238698	-	-	-	-	T5SS
Isomerase/hydrolase	300053	300712	WP_025237142	1335346	1336005	WP_000284277	96	Oxidative stress resistance
Septum formation inhibitor	301205	301894	WP_002462028	3667188	3667881	WP_000203095	92	Inhibits septum formation
Septum site-determining protein MinD	301918	302730	WP_000101044	1331816	1332628	WP_000101055	96	Cell division
Cell division topological specificity factor	302734	303000	WP_001185665	1331546	1331812	WP_001185665	97	Transcription regulation
Transposase	309802	310167	WP_000567766	270073	270375	WP_001374976	<40	IS3 components
Transposase	310125	311030	WP_000376504	769506	769754	WP_000567696	<40	HTH superfamily
MATE / AcrB	4232349	4271203	WP_024164769	441731	444880	YP_002327973	<40	Defence
Enterobactin (EntF)	918225	922156	WP_025237424	545224	549105	WP_000777601	99	Iron uptake

eron, *flhA*/*B*/*C*/*D* along with the gene for transcriptional regulator, and *fliD* that encode flagellar filament capping protein (Table 2). This is in contradiction to previous finding that *E. albertii* KF1 is nonmotile [21]. These genes if present as functional units could help the strain to sustain unfavourable conditions and eventually help in its newly designed job as an enteropathogen.

In addition, GEIs of both the genomes also harbour singletons with higher number in *E. coli* O127 (961). The 778 singletons that are exclusively identified in *E. albertii* KF1 include several insertion sequences (IS), like IS2, IS66, IS64, IS62 and TnpB transposases and prophage like regions (Table 2). Unique virulence-related genes identified in *E. albertii* KF1 include the outer membrane autotransporter barrel domain-containing protein, phage tail protein, inner membrane ring protein, outer membrane assembly protein BamE and PhoE plus MATE family multidrug exporter protein. Although the strain O127 have a multidrug efflux pump subunit *acrB*, it is different from the MATE family multidrug exporter (Table 3). Even the proteases encoded within both the genomes show very less homology. Strain KF1 also encodes several hypothetical proteins within each GEI. These proteins of unknown or putative function might also provide some unique and/or novel virulence factor to its host.

Secretion system

A significant array of advanced secretion system machineries to export various virulence factors across the cell membrane have been found in bacteria. Among them, T3SS is one of the conserved transmembrane complex systems encoded by locus of enterocyte effacement, LEE [19]. Besides LEE, the secretion system also involves the association of non LEE effectors [7]. These are present in both the EPEC and enterohemorrhagic (EHEC) pathotypes of *E. coli* and both are recognised as DEC. Earlier studies reported that different strains of *E. albertii* could secrete T3SS and harbour a conserved region encoding LEE [22]. *E. albertii* KF1, like all other EHEC and EPEC strains including strain *E. coli* O127 is found to possess core pathogenomic components of the non LEE-encoded intimin (*eae*) secretion system that include two conserved proteins viz. key intimin (90 % similarity), and a putative coding region for translocator intimin effector Tir (96 % similarity) (Table 3). LEE-encoded and T3SS-dependent effectors are also observed in *E. albertii* KF1 with NleG, NleD and putative NleG8 proteins (Table 2).

Another important feature that distinguishes EPEC from EHEC is the inability to produce Shiga toxins or cause hemorrhagic colitis or hemolytic uremic syndrome. In fact, other strains of *albertii* are also known to produce the stx2a shiga toxin which shows its affinity with EHEC strains [22]. Although strain KF1 does not encode this toxin, the non-LEE effectors associated to shiga toxin production are putatively identified in its genome.

Re-annotation of the genomes on RAST platform clearly demarcated some pathogenesis related genes like those coding for endolysin, capsid and Hok/Gef family protein in both the strains. Even, the *de novo* analysis provided new insights in their membrane transport mechanisms. It includes EscC/YscC/HrcC family T3SS outer membrane ring protein, outer membrane usher protein, envelope stress response membrane protein PspB, PspC, EscJ/YscJ/HrcJ family. The T3SS inner membrane ring protein and outer membrane protein BamE were exclusively identified in strain KF1.

Production of lesions through the T3SS finally helps in the adherence of the organism. Attachment of enteropathogens is further facilitated through EPEC adherence factor (EAF), type IV bundle-forming pili (*bfp*) and virulence regulator *per*. Although, *E. albertii* KF1 genome does not possess EAF, a specific region (1394236 - 1395438) is identified that showed homology to pilin biosynthesis and could provide a similar function as the type IV pili of strain O127. BLAST analysis also confirmed a conserved domain of PulF superfamily helping cell motility and intercellular trafficking.

The barrel domain-containing outer membrane autotransporter possess a distinct secretion mechanism in *E. albertii* KF1 having similarity with T5SS. These autotransporters carry virulence factors that help the strain to translocate the effector proteins and help in adherence by promot-

ing actin like function resulting in bacterial mobility. As *E. albertii* is non motile [20], this feature adds its fitness ability in unfavourable conditions. On the contrary, the T3SS of type 2 (or ETT2) like genomic islands that provide dynamics in the pathogenicity of the harbouring strains like *E. coli* strain O127 and other *E. albertii* strains [21], is absent in *E. albertii* KF1. However, *E. albertii* KF1 harbours barrel domain autotransporter and MATE proteins that enhances its virulence potential in spite of lacking the ETT2 system.

Thus, in conclusion, the comparative genomics of the prototypical and emerging enteropathogen identified a proportion of genes that are shared between them with unique singletons. The exclusive singletons linked with virulence and resistance are localized in five GEIs of *E. albertii* KF1. This includes, *phoE* and *bamH* (encodes outer membrane protein) coupled with T5SS, type IV pili biosynthesis, a conserved region with genes encoding for flagella, MATE proteins, abundant transposases (with different superfamilies) and several hypothetical proteins. Even, some of the encoding genes show very less similarity with prototypical strain. Presence of these exclusive features makes *E. albertii* KF1 to be a successful candidate as an emerging enteropathogen.

Acknowledgements

This work is financially supported by the DST-SERB Young Scientist Award [Project No.SB/YS/LS-16/2013].

References

1. **Abbott, S.L., Connor, J.O., Robin, T., Zimmer, B.L., Janda, J.M. (2003).** Biochemical properties of a newly described *Escherichia* species, *Escherichia albertii*. J. Clin. Microbiol. 41(10): 4852-4854.
2. **Albert, M.J., Faruque, S.M., Ansaruzzaman, M., Islam, M.M., Haider, K., Alam, K., Kabir, I., Robins-Browne, R. (1992).** Sharing of virulence-associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. J. Med. Microbiol. 37: 310-314.
3. **Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Kevin Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., LOsterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O. (2008).** The RAST server: rapid annotations using subsystems technology. BMC Genomics. 9: 75.
4. **Bertelli, C., Laird, M.R., Williams, K.P., Lau, B.Y., Hoad, G., Winsor, G.L., Brinkman, F.S.L. (2017).** IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets, Nucleic Acids Res. 45(3): W30-W35.
5. **Bray, J. (1945).** Isolation of antigenically homogenous strains of *Bact. coli neapolitanum* from summer diarrhoea in infants. J. Pathol. Bacteriol. 57: 239-247.
6. **Casburn-Jones, A.C. (2004).** Management of infectious diarrhoea. Gut. 53(2): 296-305.
7. **Croxen, M.A., Law R.J., Scholz, R., Keeney, K.M., Wlodarska, M, Finlay, B.B. (2013).** Recent advances in understanding enteric pathogenic *Escherichia coli*. Clin. Microbiol. Rev. 26 (4): 822-880.
8. **Fiedoruk, K., Daniluk, T., Swiecicka, I., Murawska, E., Sciepuk, M., Leszczynskaa, K. (2014).** First complete genome sequence of *Escherichia albertii* strain KF1, a new potential human enteric pathogen. Genome Announc. 2(1): e00004-e00014.
9. **Frazer K.A., Pachter L., Poliakov A., Rubin E.M., Dubchak I. (2004).** VISTA: computational tools for comparative genomics. Nucleic Acids Res. 1: W273-279.
10. **Ghosh, P. K., Ali, A. (2010).** Isolation of atypical enteropathogenic *Escherichia coli* from children with and without diarrhoea in Delhi and the National Capital Region, India. J. Med. Microbiol. 59: 1156-1162.

11. **Hernandes, R.T., Elias, W.P., Vieira, M.A.M., Gomes, T.A.T. (2009).** An overview of atypical enteropathogenic *Escherichia coli*. FEMS Microbiol. Lett. 297(2): 137-149.
12. **Hsiao, W., Wan, I., Jones, S.J., Brinkman, F.S.L. (2003).** IslandPath: aiding detection of genomic islands in prokaryotes. Bioinformatics. 19: 418-420.
13. **Hudson, C.M., Lau, B.Y., Williams, K.P. (2015).** Islander: a database of precisely mapped genomic islands in tRNA and tmRNA genes. Nucleic Acids Res.43: D48-D53.
14. **Huys, G., Cnockaert, M., Janda, J.M., Swings, J. (2003).** *Escherichia albertii sp.* nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. Int. J. Syst. Evol. Microbiol. 53: 807-810.
15. **Iguchi, A., Thomson, N. R., Ogura, Y., Saunders, D., Ooka, T., Henderson, I. R., Harris, D., Asadulghani, M., Kurokawa, K., Dean, P., Kenny, B., Quail, M.A., Thurston, S., Dougan, G., Hayashi, T., Parkhill, J., Frankel, G. (2009).** Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J. Bacteriol. 191(1): 347-354.
16. **Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W. (2007)**. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35: 3100-3108.
17. **Langille, M.G.I., Hsiao, W.W.L., Brinkman, F.S.L. (2008).** Evaluation of genomic island predictors using a comparative genomics approach. BMC Bioinformatics. 9: 329.
18. **Lowe, T.M., Eddy, S.R. (1997).** tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25: 955-964.
19. **Nataro, J.P., Kaper, J.B. (1998).** Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. 11(1): 142-201.
20. **Nataro, J.P. (2013).** Diarrhea among children in developing countries. Adv. Exp. Med. Biol. 764: 73-80.
21. **Nimri, L.F. (2013).** *Escherichia albertii,* a newly emerging enteric pathogen with poorly defined properties. Diagn Microbiol Infect Dis. 77(2): 91-95.
22. **Ooka, T., Ogura, Y., Katsura, K., Seto, K., Kobayashi, H., Kawano, K., Tokuoka, E., Furukawa, M., Harada, S., Yoshino, S., Seto, J., Ikeda, T., Yamaguchi, K., Murase, K., Gotoh, Y., Imuta, N., Nishi, J., Gomes, T.A., Beutin, L., Hayashi, T. (2015).** Defining the genome features of *Escherichia albertii*, an emerging enteropathogen closely related to *Escherichia coli*. Genome Biol. Evolu. 7(12): 3170-3179.
23. **Rajeshwari, K., Beena, U., Singh, R., Tiwari, G., Ajay, K.M. (2015).** Multi drug resistant enteropathogenic *E. coli* diarrhea in children. Am. J. Res. Commun. 3(9): 27-48.
24. **Robins-Browne, R.M. (1987).** Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. Rev. Infect. Dis. 9(1): 28-53.
25. **Stock, I., Rahman, M., Sherwood, K.J., Wiedemann, B. (2005).** Natural antimicrobial susceptibility patterns and biochemical identification of *Escherichia albertii* and *Hafnia alvei* strains. Diagn. Microbiol. Infect. Dis. 51: 151-163.
26. **Waack, S., Keller, O., Asper, R., Brodag, T., Damm, C., Fricke, W.F., Surovcik, K., Meinicke, P., Merkl, R. (2006).** Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics. 7: 142.
27. **Walker-Smith, J.A. (2001).** Post-infective diarrhoea. Curr. Opin. Infect. Dis. 14: 567-571.