Original Article
DOI: https://doi.org/10.70749/ijbr.v3i3.812

INDUS JOURNAL OF BIOSCIENCE RESEARCH
https://induspublishers.com/IJBR
ISSN: 2960-2793/ 2960-2807

Unraveling the Hidden Diversity of Gut Microbiota: A Genome Sequencing and Comparative Genome Analysis of Novel Bacterial Strains in Human Stool

Muhammad Imran Malik[1], Ayesha Iram[2], Abuzar Mehdi Khan[2], Muhammad Qasim[3], Waseem Abbas[4], Maha Munir[4]

[1]Department of Zoology, University of Education, Lahore, Multan Campus, Punjab, Pakistan.
[2]Department of Zoology, University of Education, Lahore, Faisalabad Campus, Punjab, Pakistan.
[3]Department of Zoology, The Islamia University of Bahawalpur, Punjab, Pakistan.
[4]Center for Applied Molecular Biology, University of The Punjab, Lahore, Punjab, Pakistan.

ARTICLE INFO

Keywords

PCR, Bacteria, Biochemical, Recombinant, Stool, Gene.

Corresponding Author: Muhammad Imran Malik,
MS Scholar, Department of Zoology, University of Education, Lahore, Multan Campus, Punjab, Pakistan.
Email: imranmalik8182@gmail.com

Declaration

Authors' Contribution: All authors equally contributed to the study and approved the final manuscript.
Conflict of Interest: No conflict of interest.
Funding: No funding received by the authors.

Article History

Received: 07-01-2025
Revised: 26-02-2025
Accepted: 12-03-2025

ABSTRACT

E. coli strains have been isolated and identified in Pakistan, but detailed genomic data specific to these strains is still lacking. This gap in knowledge prevents a full understanding of E. coli's genomics in the region. Research indicates that both phagocytic and nonphagocytic mammalian cells can use Shigella, Listeria, Salmonella and invasive E. coli as vectors for gene delivery. This study seeks to investigate the epidemiology and pathogenic features of E. coli in Pakistan. Stool samples were collected from various regions exhibiting severe symptoms and the bacteria were isolated, cultured and analyzed. The process included DNA extraction, 16s PCR analysis and genome sequencing. Key features of the Escherichia coli SWL-1 genome are a GC content of 50.8%, 5020 coding sequences, 61 RNA sequences and 4 contigs. The genome encompasses genes related to cofactors, vitamins, cell walls, capsules, virulence, disease defense, potassium metabolism and carbohydrates, among others. E. coli strains are categorized into six main types: ETEC, EIEC, EPEC, EHEC, EAEC and DAEC, each causing different levels of diarrhea. Urinary tract infections (UTIs) are the most common extraintestinal infections caused by uropathogenicE. coli (UPEC). Additionally, ExPECE. coli features adhesin operons specific to the antigen, including the virulence genes.

INTRODUCTION

Escherichia coli (*E. coli*) is a bacterium often seen in the human intestinal microbiota and is known for its role in quality control. Most of the bacteria used for transferring traits into arranged and unqualified phagocytes are facultative intracellular types designed to survive post-cell entry (Grillot-Courvalin et al., 2002). Recently, the genomes of many *E. coli* K-12 strains, which were significant in genetic research and recombinant DNA technology, have been sequenced (Oshima et al., 2008). *Escherichia coli* SE15 (O150:H5), a human commensal bacterium recently isolated from the feces of a healthy adult and classified into *E. coli* phylogenetic group B2 (which comprises most extraintestinal pathogenic *E. coli*), has had its complete and annotated genomic sequence published (Toh& Yun, 2010). Whole genome sequencing is becoming a routine method in bacteriological research, laying the groundwork for future investigations (Ng &Kirkness, 2010).Human gut bacterial genome datasets reveal genes for arsenic-active proteins produced by bacteria in human feces, which can biochemically convert arsenic in vitro (Delgado et al., 2006). While substantial work has been done on the phylogenetic and genomic analyses of *E. coli* globally, comprehensive genomic research on this bacterium in Pakistan remains limited. Understanding the pathotype and characterization of *E. coli* is crucial for comprehending the diseases it causes. Our primary objective was to delineate the range of diarrheal diseases caused by various *E. coli* strains.Our research delves into the genomics, epidemiology and pathogenicity of *E. coli* strains from the Pakistani region, providing essential insights for addressing the numerous diseases caused by

this versatile bacterium in the area. We focused on characterizing isolated *E. coli* strains through human stool samples from Pakistan, including DNA extraction and sequencing, with particular attention to the 16S gene sequence (Matushek et al., 1996). The assembled sequence revealed a single circular chromosome with a G+C content of 50.8%, 5,046 coding sequences and 114 distinct RNAs, organized into four contigs (Nash et al., 2010).Our research also explores the functional elements of *E. coli*. The *E. coli* SWL-1 genome includes pathways related to carbohydrate metabolism, DNA metabolism, cofactors, vitamins, prosthetic groups, pigments, cell walls, capsules, virulence factors, diseases, defense mechanisms, potassium metabolism, phages and prophages, transposable elements, plasmids, miscellaneous cellular processes and RNA metabolism (Gonzaga-Jauregui et al., 2012). Notably, certain adhesion operons were identified in the central region of specific pathotypes, shedding light on the complex virulence mechanisms and genetic diversity within pathogenic strains (Micciulla et al., 2019).Additionally, our study aimed to elucidate potential pathogenicity and virulence determinants. We investigated the six major *E. coli* pathotypes responsible for severe human diarrhea: enteroaggregative *E. coli* (EAEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC) and diffusely adherent *E. coli* (DAEC) (Ritchie et al., 2012). Our findings also highlighted that uropathogenic *E. coli* (UPEC) is the leading cause of urinary tract infections (UTIs), while newborn meningitis-associated *E. coli* (NMEC) is recognized as the pathotype that causes meningitis and sepsis (Chen et al., 2006).

MATERIALS AND METHODS

Molecular Characterization of Collection of Samples

The samples were collected using a plastic cap designed like a hat. To efficiently gather the samples, the collection device was positioned over a toilet bowl. We obtained human fecal samples from various locations in Islamabad, Pakistan, to analyze natural type bacteria present in the human stool.

Bacterial Strains Isolated and Cultivated Fresh Feces of a Healthy Adult

One gram of healthy adult feces was suspended in 9.0 mL of phosphate-buffered saline (pH 7.0).The suspension was repeatedly plated on DHL agar and incubated at 37°C for 24 hours. Single colonies were then isolated The eight red colonies from the DHL medium were grown twice on Luria-Bertani (LB) agar plates and identified as *E. coli* based on the following characteristics: Gram-negative, rod-shaped, aerobic and anaerobic growth, no spore formation, motility and gas/lactic acid generation from glucose/lactose. The isolates were then stored at -85°C for future analysis.

Extraction of DNA

The CTAB method was employed to extract DNA from isolated bacteria. The culture was mature for 72 hour at 27°C in 10 ml of nutrient broth. Following this 1.5 ml of the A particle was formed after centrifugation of liquid broth at 13,500 rpm for 5 minutes. The supernatant was discarded and the pellet was resuspended in 567 µl of TE buffer using a pipette. To confirm the concentration at 100 µg/ml, 3 µl of 10% SDS added. After complete mixing, the mixture was incubated at room temperature for an hour. Next, add 100 µl of 5M NaCl and mix gently. Then, add 80 µl of CTAB/NaCl solution. The mixture was incubated at 65°C for 30 minutes, with vortexing every ten minutes.

An equal amount of chloroform/isoamyl alcohol was added and the mixture was spun for 2 minutes at 13,500 rpm in a small centrifuge. The top layer was transferred to a sterile tube and 0.6 volume isopropanol was added. The tube was then kept on ice for 15 minutes. After shaking the tube until a white DNA precipitate formed, it was centrifuged at 12,000 Rpm for 10 minutes to pellet. The supernatant was carefully removed and the process was repeated once. The pellet was completely dry. The sample was dissolved in 100 µl TE buffer and allowed to settle for 5 minutes. The extracted DNA was then decontaminate ready for future usage. The quality of the template DNA was evaluated using marker electrophoresis.

DNA Gel Electrophoresis

To prepare 1.8% agarose gel, 1.8 g of powdered agarose was dissolved in 90 ml of 1X TAE buffer. The mixture was cooked in the microwave until clear, then cooled to 28°C. After cooling, the solution included 2 µl of ethidium bromide, a staining dye. to avoid bubble formation, combs were put in the gel tray and gently pressed into the liquid. The tray was then set at 25°C for around 20 minutes to firm up. Once the gel had set, the stoppers were carefully removed from the ends of the gel tray. The tray was then submerged in 1X TAE buffer, ensuring that the gel was fully covered by the buffer. The electrophoresis equipment was activated. For loading, 2 µl of 2X dye was mixed with 3 µl of sample DNA on a parafilm sheet (ensure thorough mixing by pipetting). The gel was then subjected to electrophoresis at 120 V for 45 minutes. After electrophoresis, the DNA fragments were visualized and photographed. The resulting DNA sample were prepare for use in PCR profile.

PCR Profiling

High-quality samples were chosen for PCR profiling using pathogen-specific primers.To confirm the occurrence of *E. coli* in the inaccessible samples, we performed 16S PCR following DNA extraction.

Gel electrophoresis of PCR Products

To evaluate the PCR data, 2% agarose gel electrophoresis was used. To make the gel, 2 g of

powdered agarose was dissolved in 100 cc of 1X TAE buffer to get 2% gel.

The solution was heated in the microwave until clear, then cooled to 28°C. After cooling, 3 µl of Ethidium bromide (a staining dye) was added. Combs were inserted into the gel tray and gently pressed into the liquid to prevent bubble formation. On a parafilm sheet, combine 4 µl of 6X loading dye with 10 µl of PCR product (pipette thoroughly). The prepared samples and 3 µl of a 1 kb ladder were put into the gel wells. The gel was then operated at 70 volts for two hours and thirty minutes. The PCR findings were visualized and photographed using a transilluminator.

The Sequencing of DNA

All DNA samples were transferred to Microbe (Korea) for library preparation and high-throughput DNA sequencing utilizing the Illumine NovaSeq 6000 sequencing machine from the Illumina platform. Next-generation sequencing techniques were employed to sequence the genome.

BLAST

BLAST (Basic Local Alignment Search Tool) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was utilized to assess sequence similarity. This program assesses the statistical importance of nucleotide or protein sequence matches against sequence databases. The 16S rRNA gene was sequenced and then searched against the NCBI database using the BLAST program to discover and compare comparable sequences.

Phylogenetic Analysis

Instead of using evolutionary biology A molecular-based phylogenetic analysis is employed. MEGA X (https://www.megasoftware.net/downloads/dload_win_gui) demonstrates how to assess. phylogenetic trees. Phylogenetic analysis of molecular data, such as DNA and amino acid sequences, is critical for evolutionary and molecular biology. With the growing use of DNA sequencing, comprehensive gene and protein sequence data are now available in public web databases. Given the large number of molecules with varying rates of evolution, identifying the right molecule (gene or protein) for a given lineage is critical for reliable phylogenetic analysis. If the incorrect molecule is used phylogenetic analysis may become less trustworthy. As computer technology progresses, so does the accuracy and efficiency of phylogenetic analysis software. As a result, there are several techniques and algorithms available for creating phylogenetic trees. This review discusses several popular phylogenetic analysis methodologies. Proper interpretation of the phylogenetic tree is critical, which includes evaluating a molecular-based phylogenetic analysis is used.

Databases Submissions

To receive an accession number, the 16S rRNA sequence was submitted to NCBI GenBank at(https://submit.ncbi.nlm.nih.gov/). The submission was completed using the BankIt tool.

Annotation, Genome and Assemblage

Genome assembly and marginal note were performed using various computational methods. To determine functional categories, the following databases were utilized The following resources are availableKyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups (COG), Non-Redundant Proteins, Gene Ontology (GO), Evolutionary Genealogy of Non-Supervised Orthologous Groups, Virulence Factors of Pathogenic Bacteria (VFDB), Antibiotic-Resistant Genes and Carbohydrate-Active Enzymes.

RAST is used to Identify ORFs.

RAST (Rapid Annotation with Subsystem Technology) was used to annotate genomic characteristics such protein-coding genes and RNA. RAST was used to detect ORFs (Open Reading Frames) via its web platform. (https://rast.nmpdr.org/).

Pathogenicity and Virulence Factors have been Identified

RAST method was utilized to identify potential pathogenicity and virulence determinants. For this purpose, a dataset of reference genes known to be associated with pathogenicity was employed. These reference genes are cataloged in the VFDB (Virulence Factor Database), which provides researchers with up-to-date information on virulence factors from different bacterial illnesses. Each reference gene in the dataset was analyzed using BLAST and RAST against the reference sequences.

RESULTS

Molecular Characterization of Sample Assortment

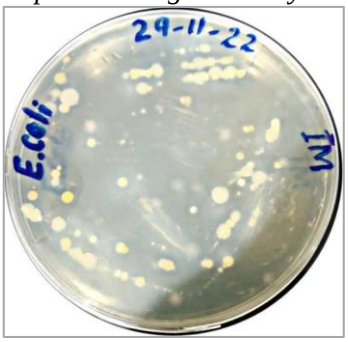
The fecal samples are collected using a plastic cap designed like a hat. To facilitate quick collection, the device is placed over a toilet bowl. Human fecal samples were gathered from multiple locations Islamabad, Pakistan, for the purpose to analyze Bacteria of the wild kind are found in human excrement.

Isolation of Pathogen

Pink cultures appear after 24 to 27 hours of incubation observed and selected for further investigation.

Figure 1

Depicts Pathogen Colony

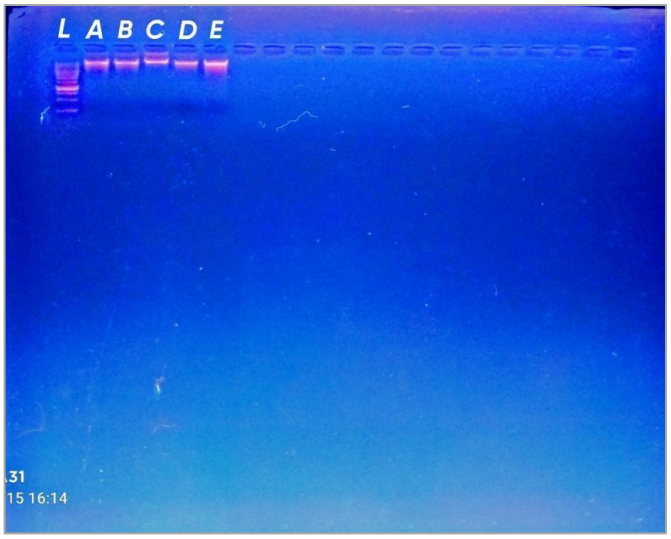


Qualitative examination of extracted DNA

After extraction from the samples, the whole DNA's quality was evaluated using electrophoresis on a 2% agarose gel.The gel was subsequently examined using gel documentation methods that detect UV light. The results, shown in Figure 3.2, indicate that a sufficient quantity of DNA was present, with distinct bands corresponding to the 16S rRNA genes.

Figure 2

Lane L (marker, leader) on a 1% agarose gel; Extracted DNA is shown.

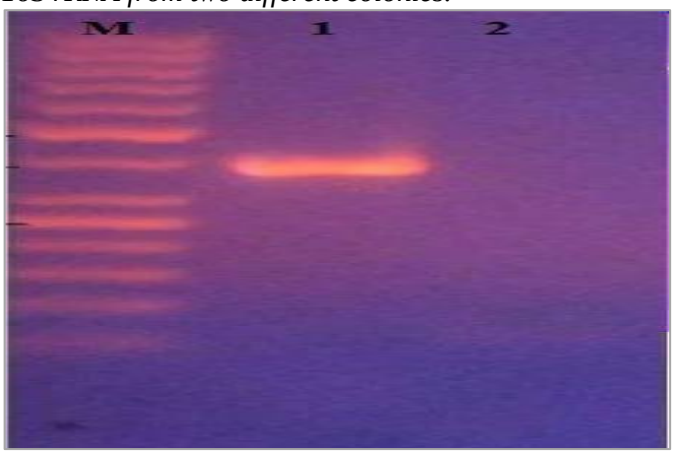


Gel electrophoresis for PCR Products

The PCR was conducted in a 20 µl reaction volume. To amplify the 16S rRNA gene from selected bacterial colonies, 16S rRNA primers were used. The denaturation temperature was set at 94°C, the annealing temperature at 52°C *and* the extension temperature at 72°C. A 10 kb PCR marker was employed. After obtaining The PCR findings were analyzed using agarose gel electrophoresis, as shown in the figure. The bands were examined under ultraviolet light before being removed for gene purification.

Figure 3

Under ultraviolet light, a 1% agarose gel displays Lane L (marker/Leader), Lane 1 and two PCR products of 16S rRNA from two different colonies.



BLAST

To examine the series, Blast (Basic Local Alignment Search Tool) was performed. This program assesses the statistical significance of nucleotide and protein sequence matches across databases. The sequence was BLAST and confirmed to be completely similar to *Escherichia coli*. Figure 3.4 shows the BLAST findings, including the scientific name, Evalue, query reporting, identification, consent Number OP985050.

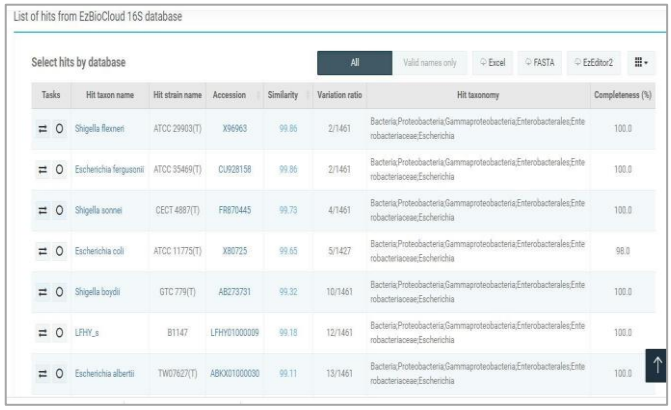
Figure 4

BLAST Results

This research is conducted using Chun Lab's EzBio Cloud database. EzBio Cloud focuses on Bacterial and Archaeal taxonomy, ecology, genetics, metagenomics and microbiome research. A search was carried out in this database using the 16S rRNA gene sequence. Figure 3.5 displays the findings, including the matched taxon name, hit strain name and accession number.

Figure 5

EzBio Cloud results



Phylogenetic Analysis

The phylogenetic tree was constructed using the Neighbor-Joining method with TreeBEST, based on the core-pan gene. The standard nucleotide identity (ANI) for all genome sequences in this research will be determined using J Species WS, using both The modes are one-to-one and all-to-all. The 16S rRNA genetic material series for *Escherichia coli* is shown under.

ATTGAAAGTTTGATCATGGCTCAGATTGAACGC
TGGCGGCAGGCCTAACACATGC
AAGTCGAACGGTAACAGGAAGAAGCTTGCTCT
TTGCTGACGAGTGGCGGACGGG

TGAGTAATGTCTGGGAAACTGCCTGATGGAGG
GGGATAACTACTGGAAACGGTA
GCTAATACCGCATAACGTCGCAAGACCAAAGA
GGGGGACCTTCG
GGCCTCTTGCCATCGGATGTGCCCAGATGGGAT
TAGCTAGTAGGTGGGGTAACG
GCTCACCTAGGCGACGATCCCTAGCTGGTCTGA
GAGGATGACCAGCCACACTGG
AACTGAGACACGGTCCAGACTCCTACGGGAGG
CAGCAGTGGGGAATATTGCACA
ATGGGCGCAAGCCTGATGCAGCCATGCCGCGT
GTATGAAGAAGGCCTTCGGGTT
GTAAAGTACTTTCAGCGGGGAGGAAGGGAGTA
AAGTTAATACCTTTGCTCATTG
ACGTTACCCGCAGAAGAAGCACCGGCTAACTC
CGTGCCAGCAGCCGCGGTAATA
CGGAGGGTGCAAGCGTTAATCGGAATTACTGG
GCGTAAAGCGCACGCAGGCGGT
TTGTTAAGTCAGATGTGAAATCCCCGGGCTCAA
CCTGGGAACTGCATCTGATACT
GGCAAGCTTGAGTCTCGTAGAGGGGGGTAGAA
TTCCAGGTGTAGCGGTGAAATG
CGTAGAGATCTGGAGGAATACCGGTGGCGAAG
GCGGCCCCCTGGACGAAGACTG
ACGCTCAGGTGCGAAAGCGTGGGGAGCAAACA
GGATTAGATACCCTGGTAGTCC
ACGCCGTAAACGATGTCGACTTGGAGGTTGTG
CCCTTGAGGCGTGGCTTCCGGAG
CTAACGCGTTAAGTCGACCGCCTGGGGAGTAC
GGCCGCAAGGTTAAAACTCAAA
TGAATTGACGGGGGCCCGCACAAGCGGTGGAG
CATGTGGTTTAATTCGATGCAA
CGCGAAGAACCTTACCTGGTCTTGACATCCACG
GAAGTTTTCAGAGATGAGAATG
TGCCTTCGGGAACCGTGAGACAGGTGCTGCAT
GGCTGTCGTCAGCTCGTGTTGTG
AAATGTTGGGTTAAGTCCCGCAACGAGCGCAA
CCCTTATCCTTTGTTGCCAGCGG
TCCGGCCGGGAACTCAAAGGAGACTGCCAGTG
ATAAACTGGAGGAAGGTGGGG
ATGACGTCAAGTCATCATGGCCCTTACGACCA
GGGCTACACACGTGCTACAATGG
CGCATACAAAGAGAAGCGACCTCGCGAGAGCA
AGCGGACCTCATAAAGTGCGTC
GTAGTCCGGATTGGAGTCTGCAACTCGACTCCA
TGAAGTCGGAATCGCTGTAATC
GTGGATCAGAATGCCACGGTGAATACGTTCCC
GGGCCTTGTACACACCGCCCGTC
ACACCATGGGAGTGGGTTGCAAAAGAAGTAGG
TAGCTTAACCTTCGGGAGGGCG
CTTACCACTTTGTGATCATGACTGGGGTGAAGT
CGTAACAAGGTAACCGTAGGGG
AACCTGCGGTTGGATCACCTCCTT

The *E. coli* strain SWL-1 is most closely related to *E. coli* ATCC 11775. The figure below displays other *E. coli* strains associated with strain SWL-1.

Figure 6

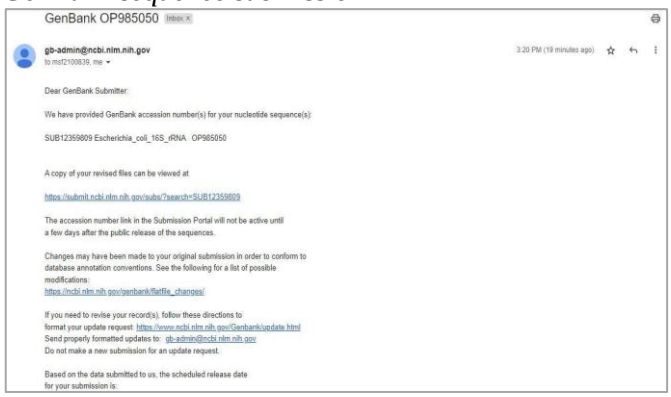
Phylogenetic Study

The Database Submission

The 16s rRNA sequence submitted to GenBank is shown below, along with information and entry number.

Figure 7

GenBank sequence submission



Genome Assembly and Explanation

General features

Table 1 provides a summary of key features of the *Escherichia coli* genome. The genome consists of a single circular chromosome with a length of 5,064,331 base pairs and shows no evidence of autonomous plasmids. The G + C content of the *E. coli* genome is 50.8 percent. The genome includes 5,046 coding sequences and 114 distinct types of RNA. It is organized into four contigs and contains 380 subsystems. In comparison, *E. coli*'s closest relative, *Escherichia coli* W3110, has 448 subsystems.

Table 1

General features of *Escherichia coli*

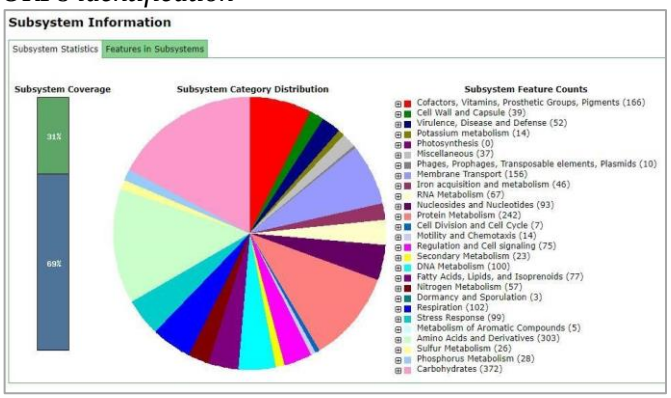
Length	5,064,331
G + C content %	50.8
Number of coding sequences	5046
Number of Contigs	4
Number of subsystems	380
N50	4870284
L50	1
Number of RNAs	114

ORFs Classification by using RAST

RAST is a software program used to identify genes in the genomes of microorganisms such as bacteria, archaea and viruses. It uses three-period nonhomogeneous Markov models to forecast coding sequences. Since its beginnings, RAST has helped to analyze the genomes of thousands of bacteria, archaea and viruses throughout the world. RAST, created as The Institute for Genomic Research's (TIGR) first microbial gene analysis tool, divides genes into functional groupings denoted by different colors. The distribution of gene categories includes 372 related carbohydrates, 100 to DNA metabolism, 166 to cofactors, vitamins, prothetic groups and pigments, 39 to cell walls and capsules, 52 to virulence, diseases and defense, 14 to potassium metabolism, 0 to photosynthesis, 10 to phages, prophages, transportable elements and plasmids, 37 miscellaneous categories, 156 to transferable membranes, 67 to RNA metabolism and 93 to nucleosides and nucleotides.

Figure 8

ORFs identification



Identifying Pathogenicity and Virulence Determinants

E. coli strains linked to varied degrees of severe human diarrhea are divided into six categories: enteroinvasive*E. coli* (EIEC), enteropathogenic*E. coli* (EPEC), enterohemorrhagic*E. coli* (EHEC), enteroaggregative*E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC). Furthermore, uropathogenic*E. coli* (UPEC) are the leading cause of urinary tract infections (UTIs), the most prevalent extraintestinal*E. coli* infection.

E. coli is also the most common Gram-negative bacteria responsible for meningitis, particularly in newborns. The pathotype of neonatal meningitis and sepsis is known as neonatal Meningitis-Associated *E. coli* (NMEC)

DISCUSSION

Our research focuses on the genom characterization of pathogenic *Escherichia coli* strains isolated from human stool samples in Punjab, specifically Multan, Islamabad *and* Sahiwal. This region was chosen due to the lack of genomic data despite its relevance to epidemiology and pathogenicity studies. The aim was to improve isolation techniques for *E. coli* strains causing symptoms such as diarrhea, severe stomach pains, bloody diarrhea and vomiting. *and* to understand their behavior and evolution. After isolating the strains, we extracted DNA and performed sequencing. Phylogenetic analysis and ORF identification using RAST helped pinpoint pathogenic and virulence factors. BLAST results confirmed a 100% similarity to *E. coli*. The genome assembly revealed a circular chromosome of 4,641,652 base pairs, a G + C content of 50.8%, 4,527 coding sequences *and* 110 different RNAs. The genome is divided into 371 subsystems and 1 contig. Our study also highlighted the significant role of genomic surveillance and the need for targeted therapeutic interventions to address *E. coli*-related diseases in the region.

REFERENCES

1. Chen, S. L., Hung, C., Xu, J., Reigstad, C. S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R. R., Ozersky, P., Armstrong, J. R., Fulton, R. S., Latreille, J. P., Spieth, J., Hooton, T. M., Mardis, E. R., Hultgren, S. J., & Gordon, J. I. (2006). Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli* : A comparative genomics approach. *Proceedings of the National Academy of Sciences*, *103*(15), 5977-5982. https://doi.org/10.1073/pnas.0600938103
2. Delgado, S., Suárez, A., & Mayo, B. (2006). Identification of dominant bacteria in feces and colonic mucosa from healthy Spanish adults by culturing and by 16S rDNA sequence analysis. *Digestive Diseases and Sciences*, *51*(4), 744-751. https://doi.org/10.1007/s10620-006-3201-4
3. Gonzaga-Jauregui, C., Lupski, J. R., & Gibbs, R. A. (2012). Human genome sequencing in health and disease. *Annual Review of Medicine*, *63*(1), 35-61. https://doi.org/10.1146/annurev-med-051010-162644
4. Grillot-Courvalin, C., Goussard, S., & Courvalin, P. (2002). Wild-type intracellular bacteria deliver DNA into mammalian cells. *Cellular Microbiology*, *4*(3), 177-186. https://doi.org/10.1046/j.1462-5822.2002.00184.x
5. Matushek, M. G., Bonten, M. J., & Hayden, M. K. (1996). Rapid preparation of bacterial DNA for pulsed-field gel electrophoresis. *Journal of Clinical*

Microbiology, *34*(10), 2598-2600. https://doi.org/10.1128/jcm.34.10.2598-2600.1996

6. Micciulla, S., Gerelli, Y., & Schneck, E. (2019). Structure and conformation of wild-type bacterial Lipopolysaccharide layers at air-water interfaces. *Biophysical Journal*, *116*(7), 1259-1269. https://doi.org/10.1016/j.bpj.2019.02.020
7. Nash, J. H., Villegas, A., Kropinski, A. M., Aguilar-Valenzuela, R., Konczy, P., Mascarenhas, M., Ziebell, K., Torres, A. G., Karmali, M. A., & Coombes, B. K. (2010). Genome sequence of adherent-invasive escherichia coli and comparative genomic analysis with other E. coli pathotypes. *BMC Genomics*, *11*(1). https://doi.org/10.1186/1471-2164-11-667
8. Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. *Methods in Molecular Biology*, 215-226. https://doi.org/10.1007/978-1-60327-367-1_12
9. Oshima, T., Okuno, T., Arai, N., Suzuki, N., Ohira, S., & Fujita, S. (2008). Vertical solar-blind deep-ultraviolet Schottky photodetectors based on β-Ga2O3 substrates. *Applied Physics Express*, *1*(1), 011202. https://doi.org/10.1143/apex.1.011202
10. Ritchie, J. M., Rui, H., Zhou, X., Iida, T., Kodoma, T., Ito, S., Davis, B. M., Bronson, R. T., & Waldor, M. K. (2012). Inflammation and disintegration of intestinal villi in an experimental model for vibrio parahaemolyticus-induced diarrhea. *PLoS Pathogens*, *8*(3), e1002593. https://doi.org/10.1371/journal.ppat.1002593
11. Toh, K. C., & Yun, S. (2010). An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pacific Journal of optimization*, *6*(615-640), 15.