Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Seasonal Dynamics of diarrheagenic Escherichia coli pollution in black clam Villorita cyprinoides and its harvesting habitats of Vembanad Lake, India

Neenu Pavithran
Cochin University of Science and Technology
Parvathi Ammini
parvathi@cusat.ac.in

Cochin University of Science and Technology
Sanath Kumar
Central Institute of Fisheries Education
Aleena Paul
Cochin University of Science and Technology
Adithya Suresh
Cochin University of Science and Technology
Neeraja R
Cochin University of Science and Technology
Navaneed VP
Cochin University of Science and Technology
Shruthi PK
Cochin University of Science and Technology
Paul Mwangi
Cochin University of Science and Technology

Research Article

Keywords: Faecal pollution, Escherichia coli, Shellfish, Virulence genes, Serotypes, Vembanad lake

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11410970/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Seasonal Dynamics of diarrheagenic *Escherichia coli* pollution in black clam *Villorita cyprinoides* and its harvesting habitats of Vembanad Lake, India

Neenu Pavithran,[1,2], Parvathi Ammini [1]*, Sanath Kumar H[3], Aleena Paul [1], Adithya Suresh[1,4], Neeraja R.[1,4], Navaneed V.P.[1,5], Shruthi P.K.[1], Paul Mwangi[1].

[1]*Department of Biotechnology, Cochin University of Science and Technology, Cochin- 682 022, Kerala, INDIA*

[2]*Department of Marine Biology, Microbiology and Biochemistry, Cochin University of Science and Technology, Cochin – 682 016, Kerala, INDIA.*

[3]*Central Institute of Fisheries Education, Mumbai, INDIA.*

[4]*ICAR - Central Institute of Fisheries Technology (CIFT), Cochin- 682029, INDIA.*

[5]*Department of Microbiology, Kasturba Medical College, Manipal– 576104, Karnataka, INDIA.*

[6]*Department of Chemical Engineering, Konkuk University, Seoul 05029, Republic of Korea.*

*Corresponding author
Dr Parvathi A
Professor and Head
Department of Biotechnology
Cochin University of Science and Technology
Kochi-682022, Kerala, India
Tel: +91-484-2575396, Fax: +91-484-2577595
Email ID: parvathi@cusat.ac.in
ORCID ID: Parvathi Ammini (0000-0001-6916-2671)

ABSTRACT

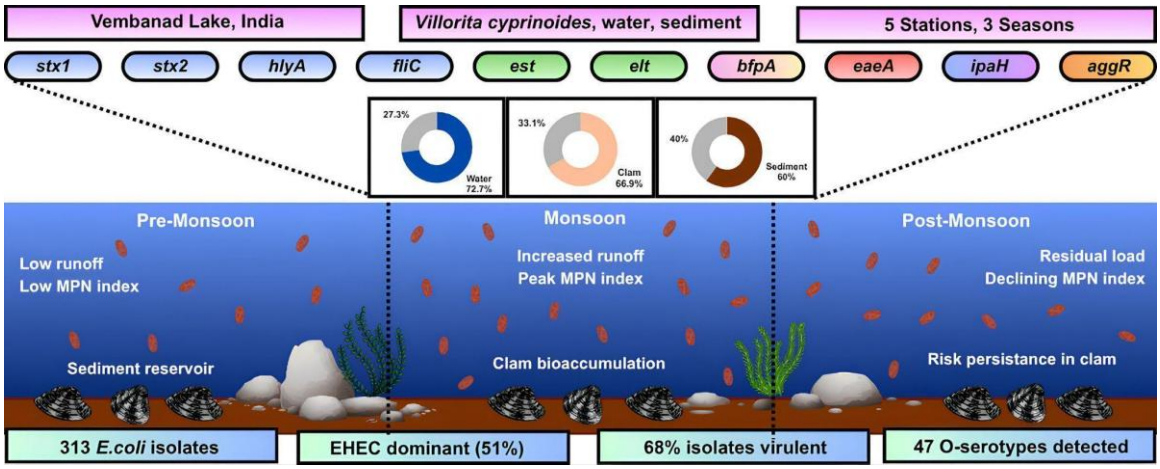
Microbial contamination of estuarine ecosystems is a major public health concern, particularly in the shellfish-harvesting areas where the accumulation of potential pathogens pose severe health risks for consumers. The study investigated the occurrence, seasonal dynamics, pathogroup distribution, and virulence gene profiles of *Escherichia coli* in black clams (*Villorita cyprinoides*), harvesting waters, and sediment at five major harvesting locations in Vembanad Lake, India. Samples were collected monthly over one year from February 2023 to January 2024. The Most Probable Number (MPN) method was implemented for the enumeration of faecal coliforms and *E. coli*, followed by biochemical identification and PCR-

based confirmation targeting the *uidA* gene and 11 virulence-associated markers. Elevated faecal indicator bacterial loads were detected across the sampling period, with significant seasonal variation, characterised by increased prevalence during the monsoon period, suggesting enhanced land-based runoff and anthropogenic inputs. Elevated MPN levels were recorded during the monsoon (MON) and post-monsoon (PSMON) seasons compared to the pre-monsoon (PRMON) period, indicating persistent faecal contamination and bacterial accumulation beyond the monsoon season. Among 313 confirmed *E. coli* isolates, 68.05% harboured at least one virulence gene. The prevalence of virulence associated gene was highest in water samples (72.72%) followed by clam (66.94%) and sediment (60%) samples. Of the 213 virulent isolates, 51.17% belonged to the enterohemorrhagic *E. coli* (EHEC) pathogroup, indicating dominance of toxigenic strains. A total of 47 distinct serotypes were detected, demonstrating considerable diversity among isolates. Prevalence of pathogenic *E. coli* in these different samples shows persistent faecal pollution, which is a potential threat to public health through consumption of contaminated shellfish. These findings further stress the importance of microbial surveillance and improved pollution management in tropical estuarine systems where shellfish harvesting occurs.

Key words

Faecal pollution, *Escherichia coli*, Shellfish, Virulence genes, Serotypes, Vembanad lake.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Aquatic ecosystems are complex interconnected systems in which biotic and abiotic factors interact to regulate nutrient cycling, fisheries production, and a range of ecosystem services. They also serve as important indicators of environmental health as water quality can directly reflect the effects of anthropogenic disturbances (Hu et al., 2025). However, increasing environmental pollution has significantly impaired the quality of aquatic ecosystems worldwide. Population explosion, inadequate infrastructure, sewage treatment, poor sanitation, and unsustainable utilization of water resources have led to the release of substantial quantities of unprocessed wastes into aquatic ecosystems (Tom et al., 2021). Stormwater runoff and heavy rainfall events further intensify contamination by mobilizing accumulated pollutants and facilitating the discharge of untreated sewage into the aquatic environment. Such inputs can substantially increase microbial loads and contribute to the deterioration of water quality and environmental degradation (Olds et al., 2018). Among many adverse effects of aquatic pollution, microbial contamination has become increasingly important due to its direct implication on public health and ecosystems (Ali et al., 2025). *Escherichia coli* is a frequently occurring bacterium residing in the digestive systems of humans and warm-blooded animals and is widely used as an indicator of recent faecal pollution in environmental samples (Gomes et al., 2024). Consequently, presence of *E. coli* may not reflect the potential health significance of contamination. Characterization of virulence-associated genes can provide information on pathogenic potential of environmental isolates *(*Antony et al., 2021; Al Qabili et al., 2022).

Vembanad Lake, the longest freshwater – brackish water ecosystem in India, is a Ramsar site of international ecological importance, and is increasingly being subjected to anthropogenic stress from urban effluents, agriculture runoff, aquaculture effluents, and domestic sewage. This has led to high levels of microbial contamination in the water column with the presence of faecal indicator bacteria (FIB) including *E. coli* in the water, sediments, and associated biota (Abdulaziz et al., 2023). Microbiological pollution, along with other environmental stressors such as heavy metals, eutrophication, and aquatic weed infestation also contributes to water quality deterioration and ecological instability (Kulk et al., 2021). Such multi-stressors may cause changes in water quality and may affect resident organisms, particularly filter-feeding species that are associated with suspended particles and sediment contaminants (Chinnadurai et al., 2020).

Among these ecologically significant organisms, *Villorita cyprinoides* (J. E. Gray, 1825), also commonly known as Indian black clam, is a native bivalve species found in the Cochin estuary and other estuaries of the Indian Peninsula. They contribute to more than 70% of clam fisheries

in India (Arun, 2009). *V. cyprinoides* is a pollution indicator, as bivalves are known to accumulate microbial pathogens and contaminants due to their filter-feeding life style (Don Xavier et al., 2021). This bioaccumulation ability creates a direct pathway between environmental contamination and potential trophic transfer and thus between environmental contamination and human exposure (Fehrenbach et al., 2024). Thus *V. cyprinoides* serve as an important link between environmental assessment of water quality and health risk assessment for humans.

In this context, characterization of pathogenicity-associated genes in *E. coli* isolated from water, sediment and filter-feeding *V. cyprinoides* can provide insight beyond conventional measurements of faecal indicator bacteria. The detection and distribution of virulence associated genes can help to evaluate the level of microbial contamination from its pathogenic potential (Singh et al., 2020). Examining these determinants across different environmental compartments is relevant because water, sediment and bivalve tissues represent interconnected but distinct reservoirs through which faecal bacteria may present, circulate, and accumulate. Such an integrated approach can also strengthen the application of bivalves as sentinel organism for assessing microbial contamination and potential exposure pathways within aquatic systems.

Considering the ecological importance of Vembanad Lake, the bioindicator potential of *V. cyprinoides,* and the emerging concern of pathogenic microbial contamination, the present study investigates the occurrence of *E. coli* in water, sediment and *V. cyprinoides* collected from the lake. The study assessed the prevalence of virulence associated genes and serotype of *E. coli* by integrating host associated and environmental matrices. This approach enables a comprehensive assessment of the distribution and prevalence of *E. coli* strains having pathogenic potential within the ecosystem. Furthermore, the study aims to elucidate the connection between anthropogenic contamination sources and microbial risk pathways, which could assist in evaluating environmental quality and the safety of seafood and public health risks.

2. MATERIALS AND METHODS

2.1. Study area and Sampling

Samples were collected from five stations in Vembanad Lake, Kerala, India, namely S1-Keltron Ferry, S2-Perumbalam, S3-Thanneermukkom, S4-Kumarakom, and S5-Muhamma (Fig.1). These stations are ecologically and hydrologically different from each other lying in

the southern part of the lake, which is the largest natural clam harvesting region. S1 and S2 stations are located in the southern brackish-water zone, influenced by tidal exchange and relatively higher salinity. On the other hand, S4 and S5 are located to the north of Thanneermukkom barrage, where freshwater predominates and tidal flushing is restricted, resulting in contrasting physicochemical conditions that may influence clam distribution and microbial dynamics.

Sampling was conducted monthly (for 12 months) between February 2023 to January 2024. For each sampling event, 15 samples were collected, comprising of clams (*Villorita cyprinoides,* n = 5), sediment (n = 5), and surface water (n = 5). Clam samples were harvested using the traditional *Kolli* method and placed in sterile bags. Sediment samples (~100 g) were collected using sterile scoops from a depth of ~5–10 cm. Surface water samples (~500 mL) were collected in sterile polyethylene bottles. All collected samples were kept chilled during transportation to the laboratory and were processed within 6 h of collection.

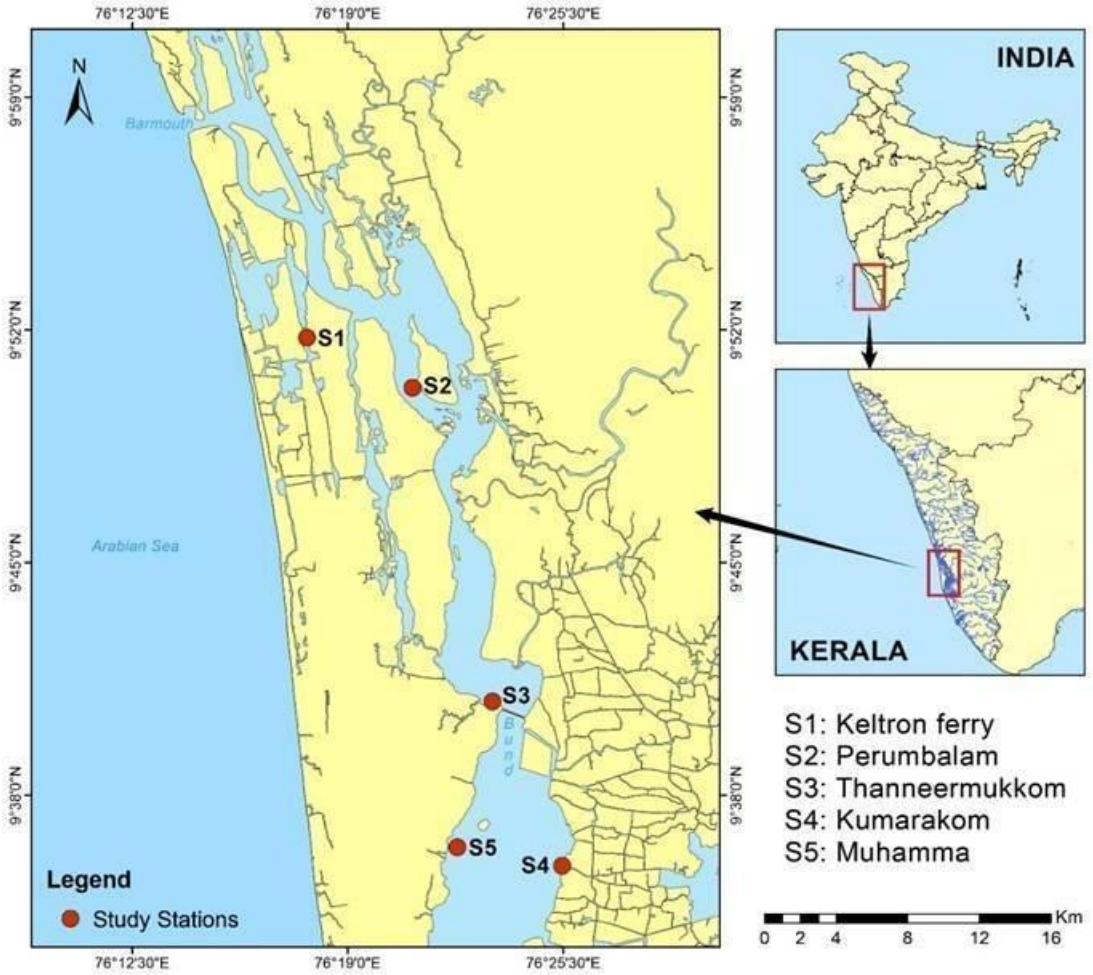


Fig. 1. Map showing the location of the studied area in Vembanad Lake

2.2. Faecal coliforms (FC) enumeration and *E. coli.* isolation

The Most Probable Number (MPN) method was used for quantifying the faecal coliform (FC) and *E. coli*. Lauryl Sulfate Tryptose Broth (LSTB) was used for the presumptive coliform detection, followed by confirmation in EC broth. For clam samples, 25 g of soft tissue was homogenized in 225 mL of sterile phosphate-buffered saline (PBS), serially diluted, and processed according to the Bacteriological Analytical Manual (BAM) of the U. S Food and Drug Administration (FDA) (Feng et al., 2011) with minor adaptations. The MPN counts were expressed in terms of MPN/g for clam and sediment samples and MPN/100 mL for water samples.

Aliquots were streaked onto Eosin Methylene Blue (EMB) agar from positive EC broth tubes. Colonies showing a metallic green sheen were presumptively identified as *E. coli*. Representative isolates were purified and maintained on nutrient agar slants at 4°C. Biochemical confirmation was performed using IMViC tests (Indole, Methyl Red, Voges–Proskauer, and Citrate utilization test), and isolates showing the characteristic pattern (++--) were confirmed as *E. coli* (Divya et al., 2016).

2.3. Molecular Confirmation Using *uidA* Gene

Genomic DNA was isolated by boiling, and the supernatant was used as template DNA. PCR amplification targeted a 147 bp fragment of the species-specific *uidA* gene, which encodes for the β-D-glucuronidase using the primers UAL-754 (5' AAAACGGCAAGAAAAAGCAG-3) and UAR-900 (5'ACGCGTGGTTACAGTCTTGCG 3') (Antony et al., 2021). Thermal cycling included an initial denaturation at 94°C for 2 min; 25 cycles of 94°C for 1 min (denaturation), 58°C for 1.5 min (annealing), and 72°C for 2 min (extension) followed by a final extension at 72°C for 5 min. Amplified products were electrophoresed on a 1.5% w/v agarose gel in 1× TAE buffer (Himedia, India), stained with ethidium bromide (Genei, India), and imaged using a Gel Documentation System (Bio Rad Gel Doc EZ Imager, USA).

2.4. Virulence gene detection in *E. coli*

PCR was performed to detect virulence genes using previously published primers (Table 1). A multiplex PCR assay was conducted for the simultaneous detection of Shiga toxin genes (*stx1* and *stx2*), *eaeA*, and *hlyA* (Antony et al., 2021). The PCR amplification was carried out under the following conditions: initial denaturation at 94 °C for 5 minutes; 30 cycles of denaturation at 94 °C for 30 seconds, annealing at 50 °C for 30 seconds, and extension at 72 °C for

1.5 minutes; followed by a final extension at 72 °C for 5 minutes. Singleplex PCRs were performed for additional virulence genes, including *elt, est, bfpA, fliC, aggR, ial,* and *ipaH*. PCR products were visualised on 2.0% agarose gel stained with ethidium bromide.

2.5. Serotyping of *E. coli* isolates

Confirmed *E. coli* strains were serotyped at the National Salmonella and Escherichia Centre, Central Research Institute, Kasauli, India. Based on the results, strains were classified into specific serotypes or classified as untypeable (UT) or rough (R).

Table 1. Primer sequences and PCR conditions for virulence gene detection

Target Gene	Primer Sequence (5′–3′)	Amplicon Size (bp)	Annealing Temp (°C)	Reference
stx1	F:ACACTGGATGATCTCAGTGG R: CTGAATCCCCCTCCATTATG	180	58	(Antony et al., 2021)
stx2	F:CCATGACAACGGACAGCAGTT R: CCTGTCAACTGAGCAGCACTTTG	255	58	(Antony et al., 2021)
eaeA	F: GACCCGGCACAAGCATAAGC R: CCACCTGCAGCAACAAGAGG	384	58	(Antony et al., 2021)
hlyA	F: GCATCATCAAGCGTACGTTCC R: AATGAGCCAAGCTGGTTAAGCT	534	58	(Antony et al., 2021)
elt	F:GGCGACAGATTATACCGTGC R: CGGTCTCTATATTCCCTGTT	450	55	(Prakasan et al., 2022)
est	F:ATTTTTMTTTCTGTATTRTCTT R:CACCCGGTACARGCAGGATT	190	56	(Prakasan et al., 2022)
bfpA	F:AATGGTGCTTGCGCTTGCTGC	326	58	(Prakasan et al., 2022)

	R: GCCGCTTTATCCAACCTGGTA			
fliC	F:GCGCTGTCGAGTTCTATCGAGC R: CAACGGTGACTTTATCGCCATTCC	625	60	(Prakasan et al., 2022)
aggR	F:GTATACACAAAAGAAGGAAGC R: ACAGAATCGTCAGCATCAGC	254	58	(Prakasan et al., 2022)
ial	F:CTGGATGGTATGGTGAGG, R:GGAGGCCAACAATTATTTCC	320	56	(Nave et al., 2016)
ipaH	F:GTTCCTTGACCGCCTTTCCGATACCGTC R:GCCGGTCAGCCACCCTCTGAGAGTAC	600	58	(Prakasan et al., 2022)

2.6. Determination of antibiotic resistance profiles of *E. coli*

Antimicrobial susceptibility was assessed using the Kirby–Bauer disk diffusion technique on Mueller–Hinton agar, following the method described by Bauer et al. (1966). A panel of antibiotics representing 14 different classes were tested on plates including Tetracycline (30 mcg), Amphenicol (Chloramphenicol (C 30 mcg)), Fluroquinolones (Norfloxacin (NX 10 mcg), Ofloxacin (OF 2mcg), Ciprofloxacin (CIP 10 mcg), Nalidix acid (NA 30 mcg)), Asamycin (Rifampicin (RIF 30 mcg)), Carbapenem (Meropenem (MRP 30 mcg)), Macrolide (Erythromycin (E 15 mcg)), Nitoimidazole (Metronidazole (MET 10 mcg)), Aminocoumarin (Novobiocin (NB 5 mcg)), Sulfonamide (Co-trimoxazole (COT 25 mcg)), Aminoglycosides (gentamycin (GEN 10 mcg), Tobramycin (TOB 10 mcg), Streptomycin (S 25 mcg)), Penicillin-Beta lactamase inhibitors (Amoxyclav (AMC 20/10 mcg), Piperaciliin-Tazobactam (PIT 100/10 mcg), Cephalosporins (Cefaclor (CF 30mcg), Ceftriazone (CTR 30 mcg), Cefotaxime (CTX 30 mcg), Cefazolin (CZ 30mcg), Cefepime (CPM 30mcg)), Polymyxins (Polymyxin B (PB 50 mcg)) and Lipopeptides (Colistin (CL 10 mcg)). Clinical and Laboratory Standards Institute 32 (CLSI, 2023), findings on the intrinsic resistance of Enterobacteriaceae and the significance of antibiotic classes in the treatment of human and animal diseases were taken into consideration while selecting the antimicrobial range.

2.7. Statistical analysis

Statistical analyses were performed using Python (version 3.x), and nonparametric tests were performed using functions from the SciPy and pandas libraries (Virtanen et al., 2020). Differences between sample matrices were evaluated using the Kruskal–Wallis test, followed by Dunn's post-hoc test. A *p*-value below 0.05 was considered statistically significant.

2.8. Distance-based redundancy analysis

To explore associations between faecal indicator bacteria and selected environmental variables (temperature, rainfall, salinity, dissolved oxygen, and pH), a distance-based redundancy analysis (dbRDA) ordination was conducted. This analysis utilized the most parsimonious distance-based linear model (DistLM) applied to log (X+1)-transformed data, with Bray–Curtis similarity as the distance measure. The analysis was performed using PRIMER v7 (Clarke & Gorley, 2015). Seasonal mean values for each parameter across the study period were used as input data, and analysis was based on Pearson's correlation matrix to identify the major environmental gradients influencing bacterial distribution across seasons.

3. RESULTS

3.1. Spatial and temporal variation in faecal coliform (FC) counts across water, sediment and clam

Results showed there were significant seasonal variations in FC concentrations across the water, sediment and clam samples collected in the study ($P<0.001$). Increased FC levels were observed during the monsoon (MON) and post-monsoon (PSMON) seasons compared to the pre-monsoon (PRMON) season, suggesting a pronounced seasonal pattern in microbial contamination. During the pre-monsoon season, FC concentration ranged from 1.36 to 3.04 log MPN/100 mL in water samples, whereas in the clam and sediment samples, it ranged from 0.47 to 3.04 log MPN/g. During the monsoon season, the concentration of FC in clams increased significantly reaching up to 3.04 log MPN/g. The FC MPN indices for sediment and water samples were 0.96 to 3.04 log MPN/g and 2.66 to 3.04 log MPN/100 mL, respectively during the monsoon season. FC concentrations in water samples were as high as 3.04 log MPN/100 mL during the post-monsoon season, while in clam and sediment samples, the FC concentrations where low ranging from 1.96 to 3.04 log MPN/g and 1.63 to 3.04 log MPN/g, respectively. High FC levels during this period suggest increased input of faecal bacteria, through runoff and freshwater inflow.

Similar seasonal trend was observed in clam samples (Fig 2). The highest concentration of FC was recorded during the MON and PSMON seasons compared with the PRMON period. In the PRMON season, FC concentration ranged from 0.47 to 3.04 log MPN/g, whereas during the

monsoon season, concentrations were consistently high reaching up to 3.04 log MPN/g. In the post monsoon season, FC concentration ranged from 1.96 to 3.04 log MPN/g. The higher FC concentrations in clams during the monsoon season is attributed to their filter-feeding behaviour which facilitate the uptake and bioaccumulation of faecal bacteria from the surrounding water column. A similar trend was seen in sediment samples. The concentration of FC during the pre-monsoon season ranged from 0.47 to 3.04 log MPN/g, which further increased to 1.63 to 3.04 log MPN/g and 0.96 to 3.04 log MPN/g during monsoon and post monsoon, respectively. This suggests that the sediments serve as reservoirs for faecal bacteria, particularly during the period of increased inflow, runoff and deposition. Water samples also exhibited seasonal variation in FC concentration with high values during post-monsoon season, reaching up to 3.04 log MPN/100 mL. During monsoon, the concentrations ranged from 2.66 to 3.04 log MPN/100 mL, while during the pre-monsoon season, the concentration ranged from 1.36 to 3.04 log MPN/100 mL. The consistently elevated FC levels in water during the monsoon season highlighted its role as the primary medium for bacterial transport and distribution. In general, the concentration of FC showed a clear seasonal pattern with significantly high levels during the PSMON and MON seasons in sediment, clams and water samples. The result of this study suggests that there are strong impacts of rainfall driven runoff, freshwater inputs and sediment processes on the contamination of shellfish harvesting areas with microbes.

Statistical analysis using Kruskal Wallis test to compare the differences in faecal coliform levels of each sample type with seasons, showed that there were significant seasonal differences (Pre-monsoon: $p = 0.01$; Monsoon: $p = 0.001$; Post-monsoon: $p = 0.001$), and significant variations between clam, sediment and water matrices. Further, sediment samples showed significant differences between clam and water in all seasons with post-hoc Dunn's test ($p < 0.05$), and no significant differences were found between clam and water samples. Highest variations were seen during the monsoon season indicating increased spatial heterogeneity in faecal coliform levels across sample types.

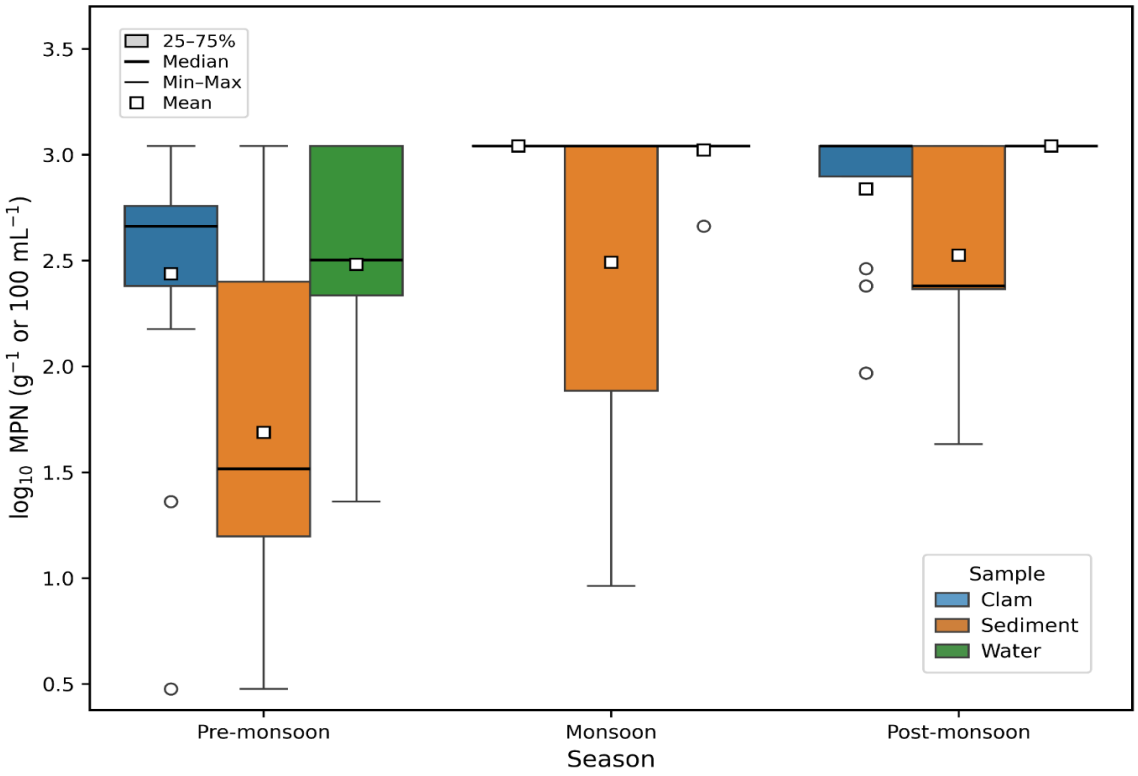


Fig. 2. Seasonal changes in concentrations of FC in sediment, clam and water samples from clam-harvesting sites in Vembanad Lake during the PRMON, MON, and PSMON periods.

3.2. Station-wise distribution of faecal coliforms in water, sediment and clam.

The boxplot (Fig. 3) shows the distribution of FC concentrations across the five sampling stations in Vembanad Lake for clam, sediment, and water samples. FC concentrations were detected across all the environmental compartments, but the extent and magnitude of contamination varied seasonally and spatially. Water samples showed the greatest responsiveness to seasonal variation, followed by sediment and clam samples. At stations S2 ($p = 0.0279$), S4 ($p = 0.0273$), and S5 ($p = 0.0273$), significant seasonal variation in FC concentration were observed in the water column. This seasonal variability may be associated with monsoon-driven runoff, freshwater influx, and anthropogenic discharges. Sediment samples showed significant seasonal variations in stations, S1 ($p = 0.0380$) and S2 ($p = 0.0196$). The constantly high faecal coliform concentration in sediments could possibly indicate their role as persistent environmental reservoirs of microbial contamination. In clam samples, significant seasonal variation was detected only at stations S2 ($p = 0.0052$) and S5 ($p = 0.0273$), suggesting that the influence of seasonality on bioaccumulation was site-specific rather than

uniform in the estuary. Post-hoc Dunn's test revealed that statistically significant pairwise differences were mainly restricted to station S2. In clam samples from S2, faecal coliform levels during the PRMON season differed significantly from both MON and PSMON seasons ($p < 0.05$), indicating enhanced microbial accumulation during MON months. Similarly, sediment samples at S2 showed a significant difference between PRMON and MON seasons ($p < 0.05$). Even though the Kruskal–Wallis test was significant, no other stations showed significant pairwise differences after Bonferroni correction, indicating potential overlaps in seasonal distribution at those locations.

Station S2 exhibited significant differences in FC concentrations in the sediment, clam and water samples and thus could be considered as a potential contamination hotspot in the study area. This could be due to more anthropogenic influence, increased runoff or a lower flushing efficiency at this station. The station S3, on the other hand, did not show any significant changes from one season to another in any of the compartments indicating relatively stable environmental conditions and low seasonal variation in faecal contamination. The increased faecal coliform counts in the clam and sediment samples compared with water samples at several stations further support the role of shellfish as long-term reservoirs of faecal contamination in tropical estuarine ecosystems.

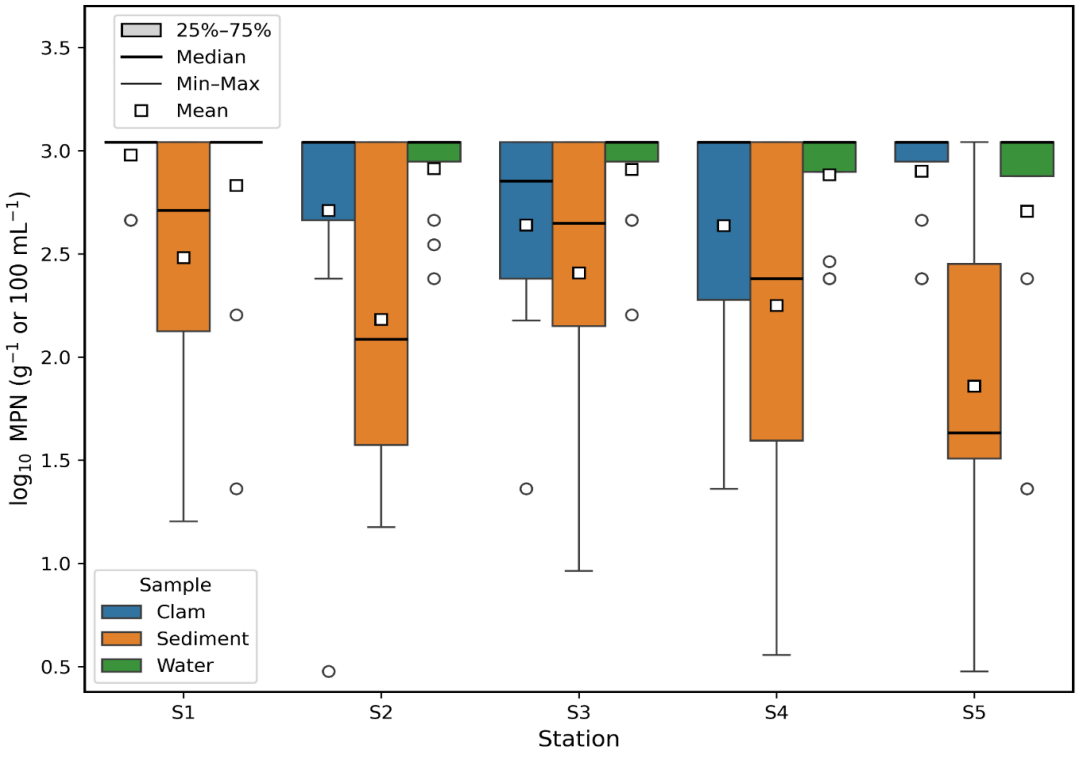


Fig. 3. Station-specific changes of faecal coliforms in sediment, clam and water.

3.3. E. coli virulence genes occurrence in various seasons and stations

Enumeration and isolation of *E. coli* were carried out through the MPN method, followed by selective plating and biochemical characterization with molecular confirmation. In total, 943 presumptive *E. coli* isolates were obtained on EMB agar. Among these, 398 isolates exhibited characteristic IMViC reaction typical of *E. coli*. These presumptive isolates were further subjected to molecular characterization using amplification of *uidA* gene. Around 313 *uidA* isolates were positive and included for further analysis. The distribution of *uidA* confirmed isolates varied across seasons and sample types. A higher number of confirmed isolates was obtained during the PRMON season, whereas comparatively fewer isolates were detected during the MON and PSMON seasons. Among sample types, water yielded the highest number of isolates, followed by clam and sediment samples. Of the total 313 *uidA* positive isolates, 121 were obtained from clams, 132 from water and 60 from sediments.

Among the *uidA* confirmed isolates, at least one virulence-associated gene was detected in 68.05% of isolates. The prevalence of virulence genes was highest among water isolates (72.72%), followed by clam (66.94%) and sediment (60%) isolates. These findings indicate widespread distribution of pathogenic *E. coli* across all the three environmental matrices, (Fig 4).

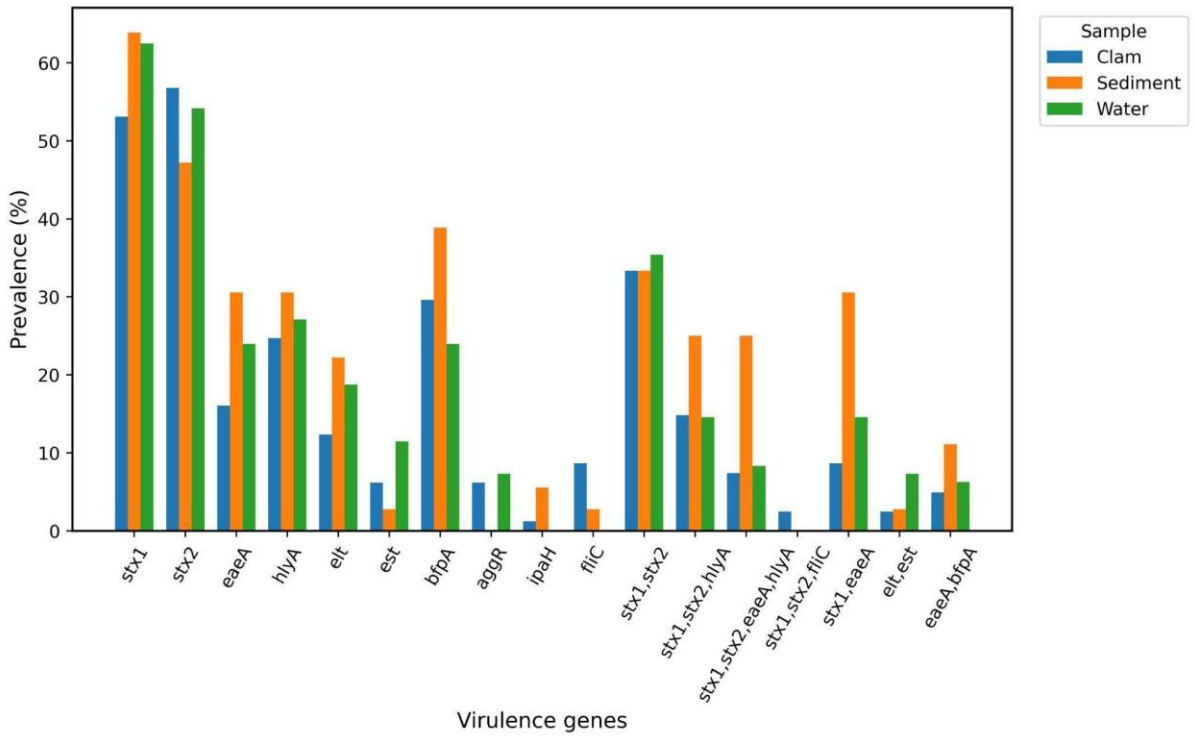


Fig. 4. Distribution of virulence genes in *E. coli* isolates recovered from sediment, clam, and water samples.

Among the Shiga toxin-associated genes examined, *stx1* showed the highest detection frequency, occurring in 45.46% of water isolates, 38.34% of sediment isolates, and 34.13% of clam isolates. Similarly, *stx2* was highly prevalent, particularly in water isolates (39.4%), followed by clam (36.51%) and sediment (28.34%). The hemolysin gene (*hlyA)* was detected more in water (19.7%) and sediment isolates (18.34%) than in clam isolates (15.88%). The *eaeA* gene, associated with the attaching and effacing lesion, was detected predominantly in sediment isolates (18.34%), followed by water (17.42%) and clam isolates (10.32%).

Enterotoxigenic *E. coli* (ETEC)- associated genes, *elt* and *est*, were frequently detected across all sample types. *The elt* gene prevalence was high in water (13.63%), and sediment isolates (13.34%), but was lower in clam isolates (7.94%). The *est* had lower prevalence in water (8.34%), clam (3.97%), and sediment isolates (1.67%).

The *bfpA* gene (typical of enteropathogenic *E. coli (EPEC))* was significantly higher in sediment isolates (23.34%), followed by clam (19.04%) and water isolates (17.43%). The *aggR* gene, indicative of enteroaggregative *E. coli (EAEC),* was more frequently detected in water (5.31%) than in clams (3.97%), with no detection in sediment. *ipaH*, a virulence marker of enteroinvasive *E. coli (EIEC),* was observed in 3.34% of sediment and 0.8% in clam isolates.

The *ial* gene was absent across all the samples. The *fliC* gene, a flagellar marker, was minimally detected in the clam and sediment isolates (5.56% and 1.67%, respectively).

Combinations of multiple virulence genes were also assessed. The co-occurrence of *stx1* + *stx2* was most frequent in the water (25.75%) followed by clam (21.42%) and sediment isolates (20%). The presence of *stx1* + *stx2* + *hlyA* was higher in sediment isolates (15%), followed by water (10.61%) and clam isolates (9.53%). The presence of *stx1* + *stx2* + *eaeA* + *hlyA* was significantly higher in sediment isolates (15%), followed by water (6.07%) and clam isolates (4.76%). The combination *stx1* + *stx2* + *fliC* was found exclusively in clam (1.66%). The combination *stx1* + *eaeA* was found higher in sediment (18.34%), followed by water (10.61%) and clam isolates (5.76%). The co-occurrence of *elt* + *est* was most frequent in the water isolates (5.31%) followed by clam and sediment isolates (1.66% and 1.67%, respectively). The combination *eaeA* + *bfpA* was found to be higher in sediment (6.67%), followed by water (4.55%) and clam isolates (3.31%)

Overall, the detection of multiple virulence associated genes across water, sediment and clam indicates the widespread occurrence of potentially pathogenic *E. coli* in the study area. The frequent detection of *stx1* and *stx2*, either individually or in combination with other virulence associated genes is particularly noteworthy from a food safety perspective because of potential exposure of shellfish consumers to toxigenic *E. coli.*

3.4. Occurrence of *E. coli* virulence genes across the three sample matrices across the five stations

The hierarchical heat map (Fig. 5) illustrates the distribution and relative abundance of virulence genes among *E. coli* isolates obtained from sediment, clam and water samples across the five harvesting stations in Vembanad Lake. The pattern of clustering reveals heterogeneity in the occurrence of virulence genes among the different samples and stations. The presence of Shiga toxin-producing genes *stx1* and *stx2* were most frequently detected virulence gene across all environmental matrices. Clam and water samples from stations S2, S3, and S4 shared particularly high abundances of *stx1* indicating that toxigenic strains are widely distributed. Similarly, *stx2* showed relatively higher prevalence, particularly in clam and water samples suggesting that the estuarine environment is a key reservoir for pathogenic *E. coli*.

Other virulence genes such as *bfpA*, *eaeA*, and *hlyA* were also detected across stations and sample types. The *eaeA* gene, which is known to be related to bacterial adhesion and intestinal

colonization, was present in a large number in clam, sediment, and water isolates, indicating that environmental isolates may have high pathogenic potential. The prevalence of *hlyA* gene was noticeable in water and clam samples, suggesting the presence of virulent EHEC-associated lineages in the harvesting environments. Other genes including *elt*, *ipaH*, *fliC*, *aggR*, and *est* were found at lower frequencies, and were only present in few stations or in a limited number of samples, indicating the uneven distribution of diarrheagenic *E. coli* pathotypes within the estuary.

The water and clam samples from stations S2 and S4 showed similar clustering patterns reflecting the similarities in their virulence gene profiles. Hierarchical clustering revealed that sediment samples are subclustered, confirming the potential for sediment to hold pathogenic bacteria consistently over time. Across all the sample matrices, station S2 showed higher abundance of virulence genes, which confirms the hotspot status of this station in the Vembanad lake for microbial contamination.

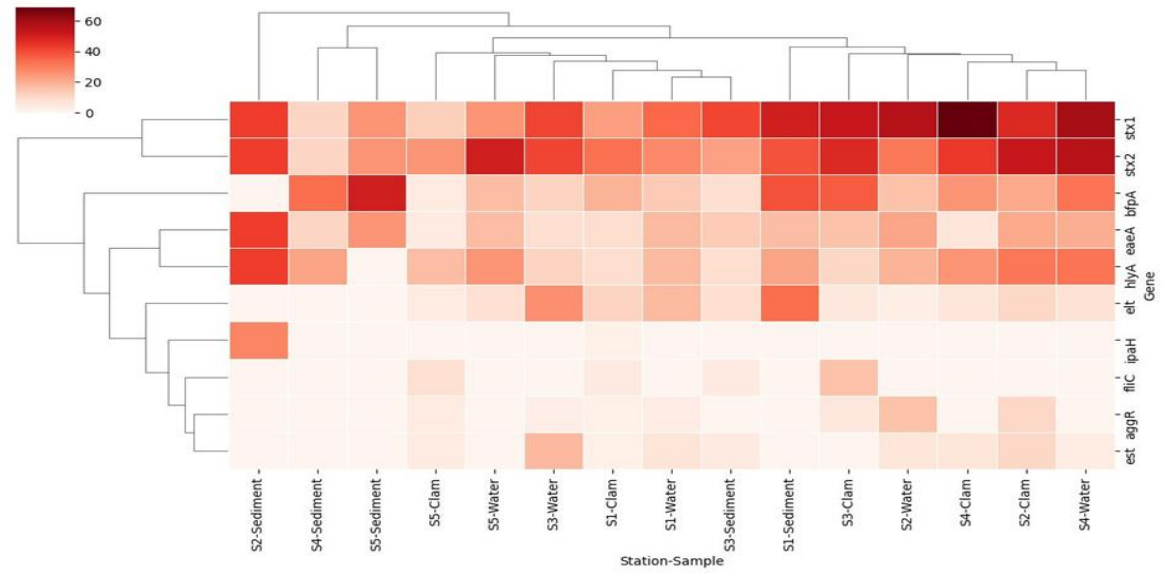


Fig.5. Hierarchical heat map showing the variation of virulence genes in the three sample matrices across the five stations

3.5. Distribution of *E. Coli* virulence genotypes and pathogroups

A total of 313 *uidA* confirmed *E. coli* isolates obtained from sediment, clam, and water samples were examined for virulence genes linked to different diarrheagenic *E. coli* pathogroups. Genotypic analysis demonstrated the presence of several virulence genes, with a predominance of EHEC associated virulence markers. The most frequently detected genotype was *stx1* in 24 isolates (11.26%), followed by *stx2* in 20 isolates (9.38%). The *hlyA* and *fliC* genes were present in few isolates. Gene combinations of EHEC pathogroup *stx1+stx2* occurred in 19 isolates (8.92%) followed by *stx1+stx2+ eaeA+hlyA* in 16 isolates (7.51%) and *stx1+stx2+ hlyA* in 11 isolates (5.16%). Other EHEC gene combinations, such as *eaeA+stx2* and *hlyA+stx2,* were less frequently observed, occurring only in 2 isolates (0.94%). These results indicate a substantial representation of EHEC pathotype (51.17%), suggesting a high proportion of Shiga toxin-producing *E. coli* strains among the isolates obtained from the study area.

The *eaeA* gene encodes intimin and is associated with the attaching-effacing phenotype, which was present in 4 isolates (1.87%), and *bfpA,* an adherence marker associated with typical of Enteropathogenic *E. coli* (EPEC), was present in 8 isolates (3.75%). These findings suggest presence of typical EPEC in lower frequencies than EHEC- associated genotypes. Enterotoxigenic *E. coli* (ETEC) marker such as *elt* was found in 7 isolates (3.28%). However, the *elt + est* combination occurred in only 1 isolate (0.46%), suggesting limited occurrence of classic ETEC strains (4.22%). Similarly, enteroinvasive *E. coli* (EIEC) marker *ipaH* was detected at very low levels in 3 isolates (1.40%).

The EAEC-associated gene *aggR* was present in only 1 isolate (0.46%), indicating less occurrence of enteroaggregative strains**.** These findings suggest that EHEC was the most prevalent and diverse pathogroup, with a considerable number of isolates harboring multiple

virulence factors such as *stx*, *eaeA*, and *hlyA*, followed by EPEC and ETEC.

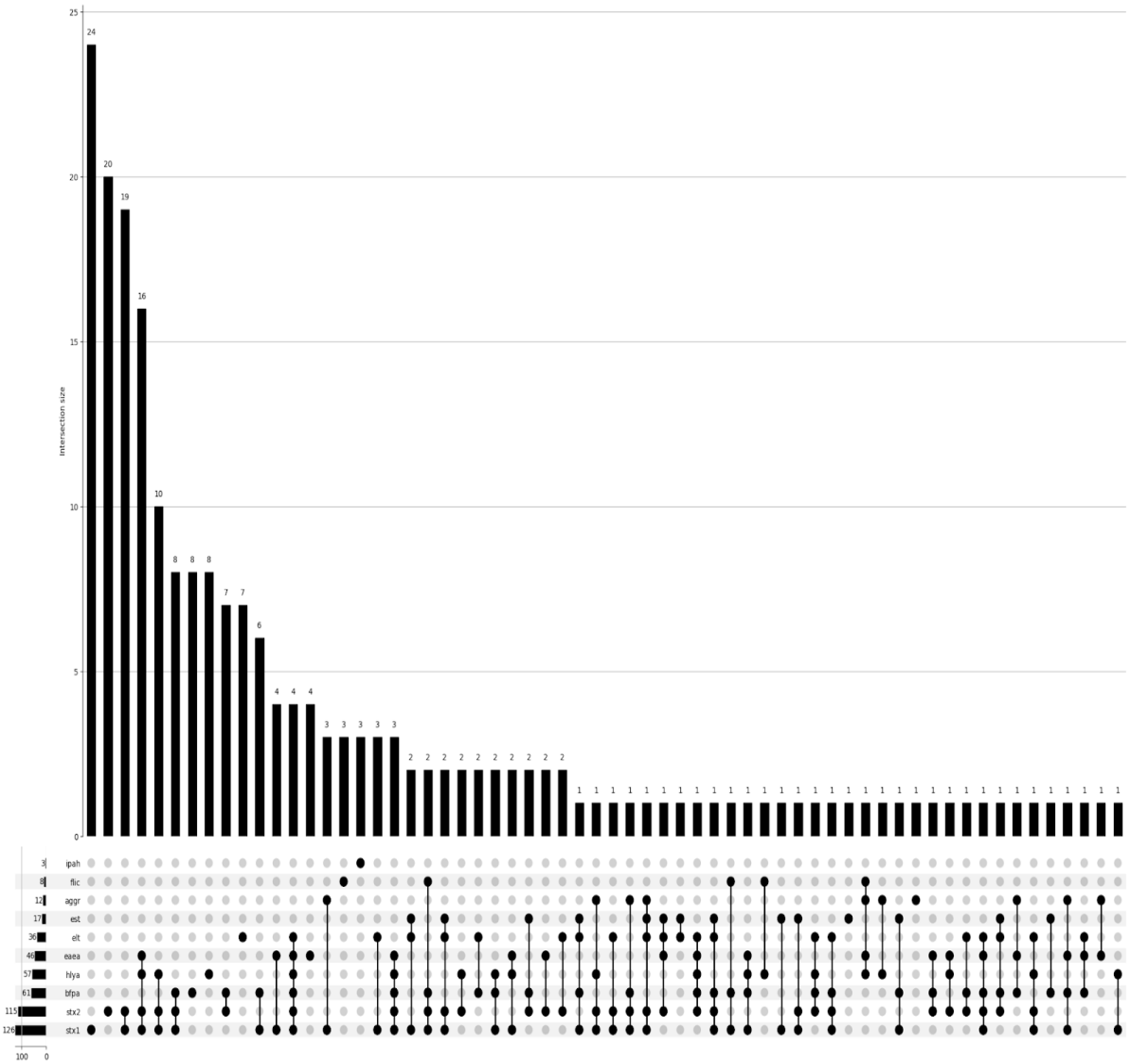


Fig. 6. UpSet plot showing the frequency of main virulence gene combination in *E. coli* isolates obtained from sediment, clams, and water samples in Vembanad Lake (n = 213). Above each bar, the number denotes how many isolates exhibited the specific combination of virulence genes indicated by the filled dots

Table 2. The table represents the total number and percentage of positive isolates for each genotype, including combinations of multiple virulence genes and their classification into diarrheagenic *E. coli* pathogroups (EHEC, ETEC, EPEC, EAEC, EIEC) among environmental isolates (n=213).

3.6. Distribution of *E. coli* serotypes across sediment, clam and water samples

Analysis of serotypes identified multiple *E. coli* serotypes among isolates obtained from

GENOTYPE	PATHOGROUP	NO. OF *E. coli* ISOLATES	No. (%) isolates belonging to specific pathogroups
stx1	EHEC	24	109 (51.17)
stx2		20	
hlyA		8	
fliC		3	
stx1,stx2		19	
stx1,stx2,hlyA		11	
stx1,stx2,eaeA,hlyA		16	
stx1,eaeA		4	
eaeA,stx2		2	
hlyA,stx2		2	
elt	ETEC	7	9 (4.22)
est		1	
elt, est		1	
bfpA	EPEC	8	8 (3.75)
eaeA	Atypical EPEC	4	4 (1.87)
ipaH	EIEC	3	3 (1.408)
aggR	EAEC	1	1 (0.469)

sediment, clam and water samples (Fig. 7). Out of 313 *E. coli* isolates, 249 isolates were serotyped, and 64 isolates were untyped. In terms of overall prevalence, serotype O11 was the most dominant (5.75%), followed by O8 (5.11%). Serotypes O7 and O25 accounted for 4.15% of isolates. Of 121 *E. coli* strains serotyped from clams, 26% could be serotyped. O11 was the most prevalent serotype (9.92%), followed by O25 and O8 (6.62% each). Of 60 strains isolated

from sediment, O7 showed the highest prevalence (10%), and serotypes O109, O11, O15, O169, O23, and O66 showed prevalence of 3.34% each. Among 132 strains isolated from harvested water, serotype O18 had the highest prevalence of 8.34% (Table 3).

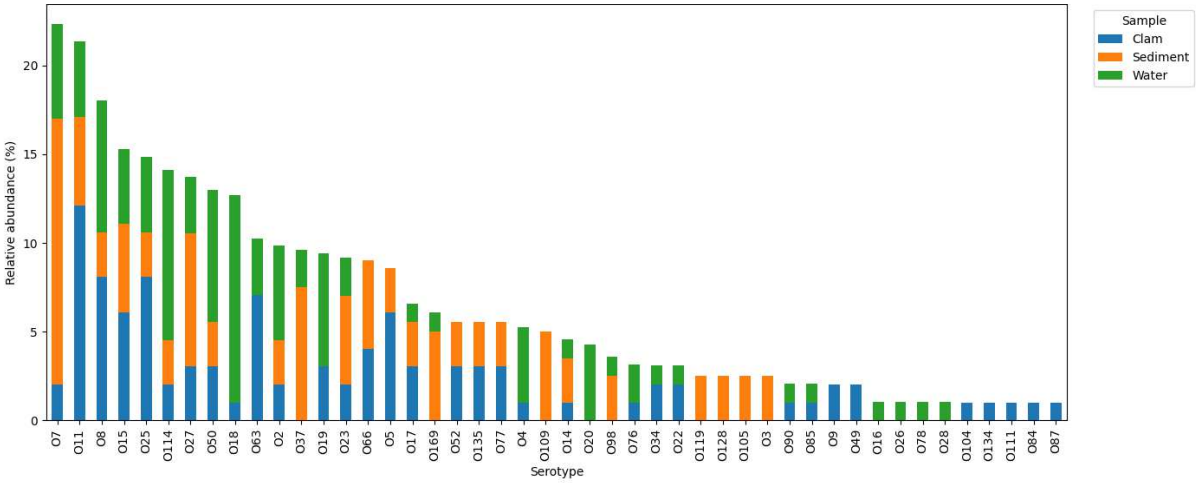


Fig. 7. Relative abundance (%) of serotypes identified among *E. coli* isolates recovered from sediment, clam, and water samples.

Table 3. Serotype prevalence among *E. coli* isolates obtained from sediment, clam, and water samples collected from Vembanad Lake.

serotypes	Overall serotype status (n=313) Serotyped (n=249) Untyped (n=64)	Distribution of serotypes by sample type		
		Clams (n=121)	Sediment (n=60)	Water (n=132)

	No. of isolates	% of sample	No. of isolates	% of clams	No. of isolates	% of sediment	No. of isolates	% of water
O104	1	0.32	1	0.83	0	0	0	0
O105	1	0.32	0	0	1	1.67	0	0
O109	2	0.64	0	0	2	3.34	0	0
O11	18	5.75	12	9.92	2	3.34	4	3.04
O111	1	0.32	1	0.83	0	0	0	0
O114	12	3.83	2	1.66	1	1.67	9	6.82
O119	1	0.32	0	0	1	1.67	0	0
O128	1	0.32	0	0	1	1.67	0	0
O134	1	0.32	1	0.83	0	0	0	0
O135	4	1.28	3	2.48	1	1.67	0	0
O14	3	0.96	1	0.83	1	1.67	1	0.76
O15	12	3.83	6	4.96	2	3.34	4	3.04
O16	1	0.32	0	0	0	0	1	0.76
O169	3	0.96	0	0	2	3.34	1	0.76
O17	5	1.6	3	2.48	1	1.67	1	0.76
O18	12	3.83	1	0.83	0	0	11	8.34
O19	9	2.88	3	2.48	0	0	6	4.55
O2	8	2.56	2	1.66	1	1.67	5	3.79
O20	4	1.28	0	0	0	0	4	3.04

O22	3	0.96	2	1.66	0	0	1	0.76
O23	6	1.92	2	1.66	2	3.34	2	1.52
O25	13	4.15	8	6.62	1	1.67	4	3.04
O26	1	0.32	0	0	0	0	1	0.76
O27	9	2.88	3	2.48	3	5	3	2.28
O28	1	0.32	0	0	0	0	1	0.76
O3	1	0.32	0	0	1	1.67	0	0
O34	3	0.96	2	1.66	0	0	1	0.76
O37	5	1.6	0	0	3	5	2	1.52
O4	5	1.6	1	0.83	0	0	4	3.04
O49	2	0.64	2	1.66	0	0	0	0
O5	7	2.24	6	4.96	1	1.67	0	0
O50	11	3.51	3	2.48	1	1.67	7	5.31
O52	4	1.28	3	2.48	1	1.67	0	0
O63	10	3.19	7	5.79	0	0	3	2.28
O66	6	1.92	4	3.31	2	3.34	0	0
O7	13	4.15	2	1.66	6	10	5	3.79
O76	3	0.96	1	0.83	0	0	2	1.52
O77	4	1.28	3	2.48	1	1.67	0	0
O78	1	0.32	0	0	0	0	1	0.76
O8	16	5.11	8	6.62	1	1.67	7	5.31

O84	1	0.32	1	0.83	0	0	0	0
O85	2	0.64	1	0.83	0	0	1	0.76
O87	1	0.32	1	0.83	0	0	0	0
O9	2	0.64	2	1.66	0	0	0	0
O90	2	0.64	1	0.83	0	0	1	0.76
O98	2	0.64	0	0	1	1.67	1	0.76
ROUGH	16	5.11	2	1.66	4	6.67	10	7.58
UT	64	20.45	20	16.53	16	26.67	28	21.22

UT= Untypeable

3.7. Distribution of *E. coli* isolates across pathogroups and serotypes

The heatmap (Fig. 8) presents the distribution of *E. coli* serotypes across pathogroups. Several serotypes, like O11, O18, and O7, exhibited high isolate counts with the EHEC pathogroup. In contrast, several serotypes showed low isolate numbers or were absent in many pathogroups. Thus, pathogroup-wise distribution revealed that EHEC comprised the highest number of isolates across multiple serotypes, whereas fewer serotypes and lower isolate counts were represented in EPEC, ETEC, EIEC, and EAEC.

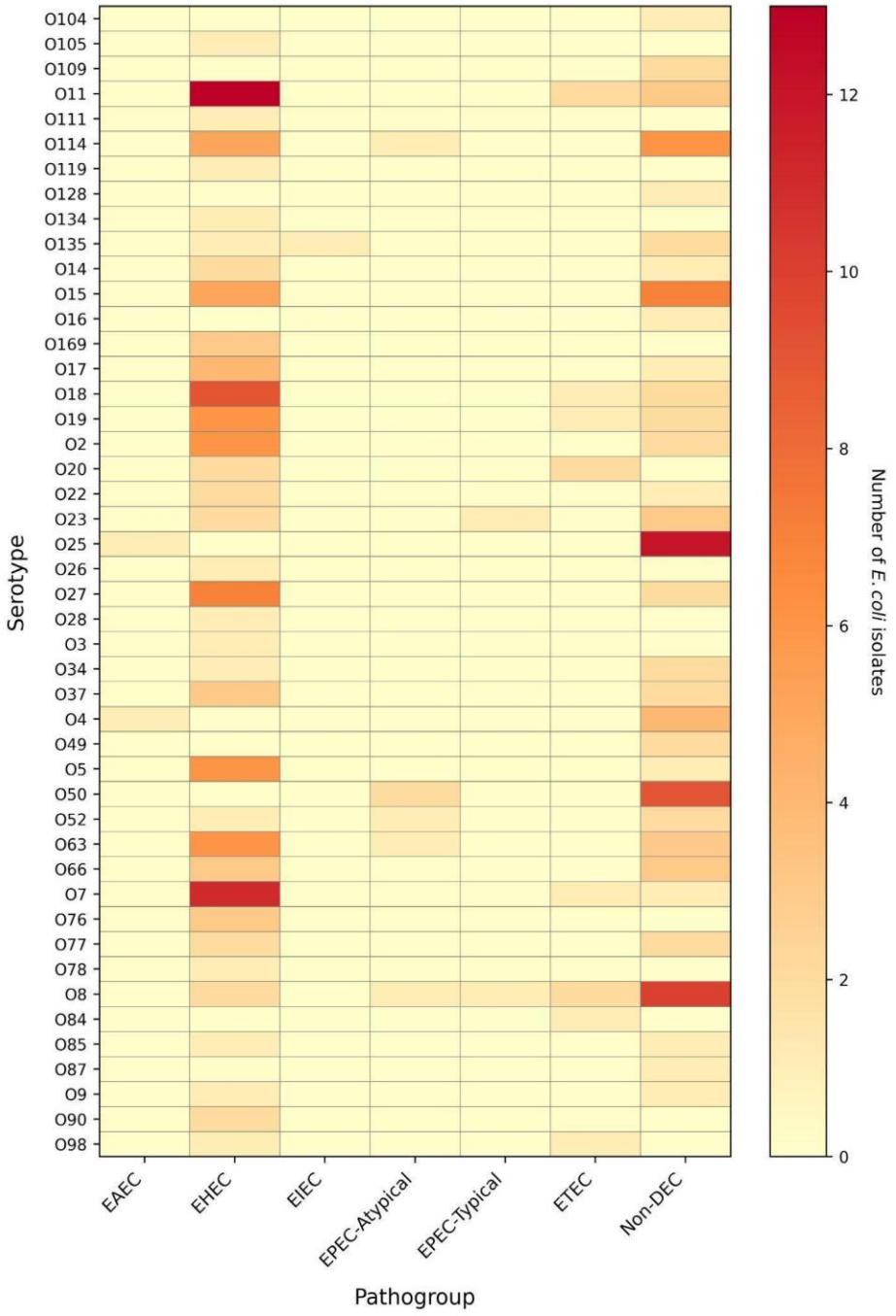


Fig. 8. The heatmap depicts various serotype-pathogroup combinations of *E. coli*, with colour intensity representing isolate counts ranging from 0 to 12.

3.8. Abundance of *E. coli* serotypes among various virulence genes

Different clustering patterns and varying connectivity between virulence genes and *E. coli* serotypes were identified using a network diagram **(**Fig. 9). *Stx1*, being the most prevalent gene, showed high abundance connections with serotypes O11 and O18. In contrast, low-abundance association patterns were observed with serotypes O23, O5, O90, O22, O20, O2. *aggR*, while being the least prevalent virulence gene identified in the study, still exhibited an association pattern with serotypes O6 and O76. Certain *E. coli* serotypes, including O11, O114, O14, O15, O17, O2, O23, O25, O27, O50, O7, and O8, were ubiquitously detected across clam, sediment, and water samples, which shows the widespread distribution of diarrheagenic *E. coli* (Table 4).

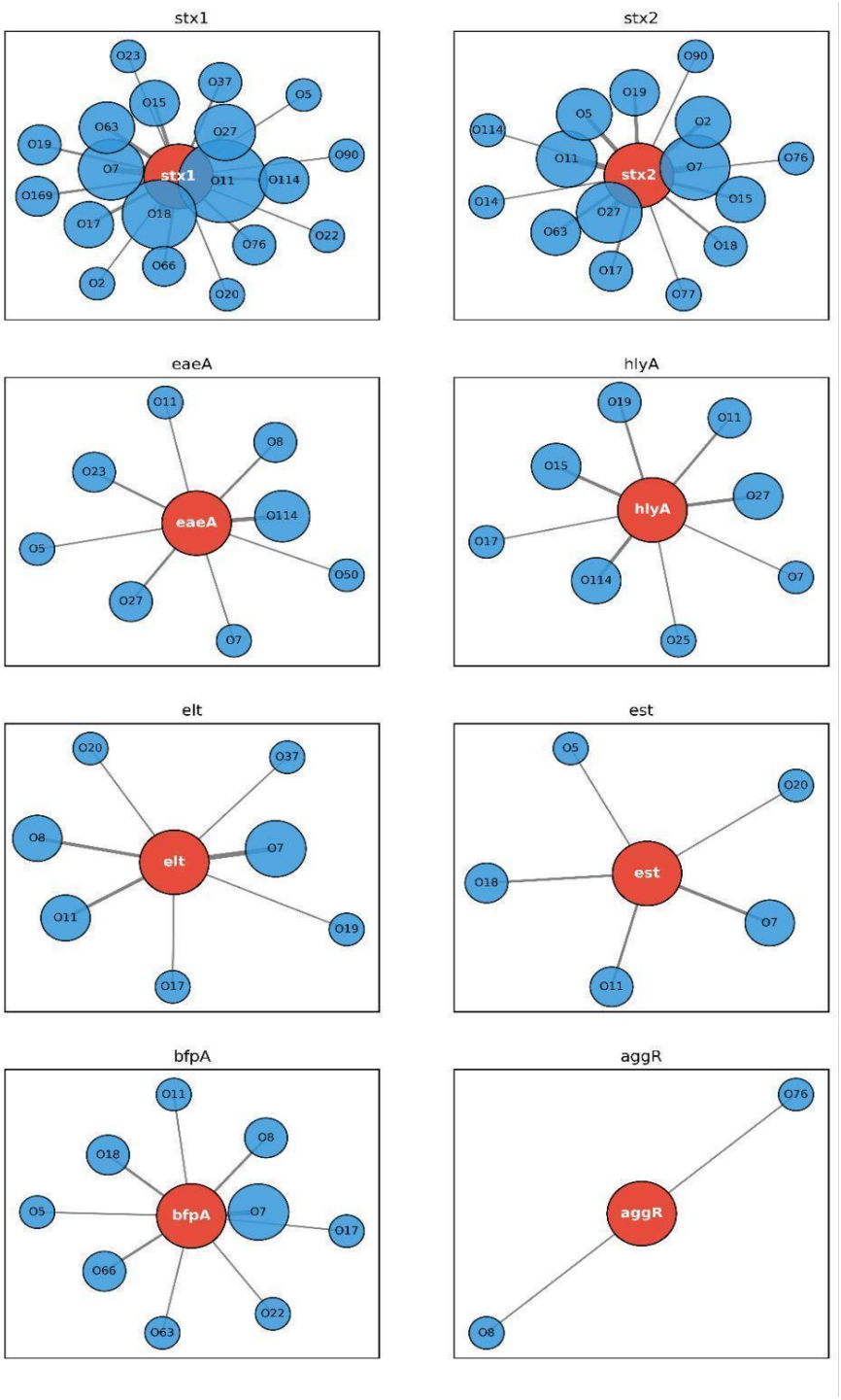


Fig. 9. The abundance of *E. coli* serotypes among virulence genes, illustrating a network plot

Table 4. Pathgroup-wise distribution of different serotypes of *E. coli* isolated in this study

Pathogroup	*E. coli* serovars identified	Virulence genotypes
EHEC	O11	*stx1* (4); *stx1, stx2* (2) ; *stx1,stx2,hlyA* (1) *; stx1, stx2, eaeA, hlyA (2)*
	O114	*stx1, eaeA* (1)*; stx1, stx2, hlyA*(1)
	O14	*stx2, eaeA, hlyA* (1)
	O15	*stx1* (1)*; stx2* (1)*; hlyA* (3)*; stx1, stx2* (2)*; stx1, stx2, hlyA* (1)
	O17	*stx1* (1)*;* stx1, stx2 (1)
	O2	*stx1* (1)*; stx2* (4)*; stx1, stx2, hlyA* (1)
	O23	*stx1, eaeA* (1)*; stx1, stx2, eaeA, hlyA* (1)
	O25	*fliC* (1)
	O50	*fliC* (1)
	O7	*stx1* (2)*; stx2* (1)*; fliC* (1)
	O8	*hlyA* (1)
	O27	*stx1, stx2, eaeA, hlyA* (3)
EPEC	O114	*eaeA* (1)
	O50	*eaeA* (2)
ETEC	O11	*elt* (1)*; est* (1)

3.9. Antibiotic susceptibility testing of *E. coli*

The incidence and antibiotic resistance profile of virulent *E. coli* strains (n=213) were determined using antibiotics belonging to various classes (Fig. 10). The highest resistance was observed for Cefotaxime with 83.2%, followed by Rifampicin with 81.7%, Erythromycin (68.1%) and Tetracycline (65.4%).

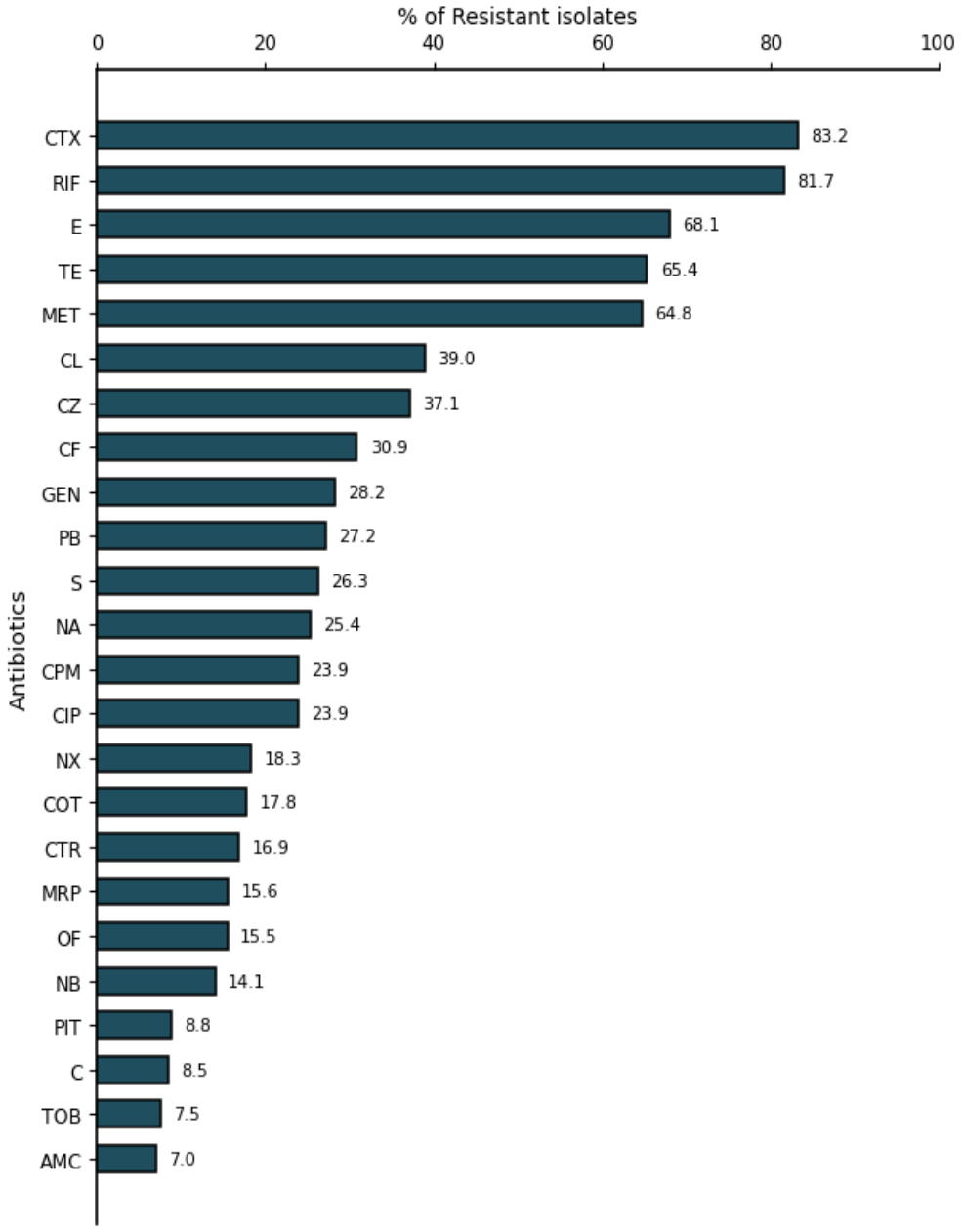


Fig. 10. Antimicrobial susceptibility profiles of pathogenic *E. coli* isolates (n= 213)

3.10. Environmental Drivers of Bacterial Contamination

Clear seasonal separation in bacterial contamination patterns was evident across sampling periods, from dbRDA ordination obtained from DistLM (Fig. 11). Monsoon (MON) and post-monsoon (PSMON) samples clustered on the positive side of the ordination whereas pre-monsoon (PRMON) samples were mainly distributed on the negative side. This indicates a strong seasonal variation in distribution of bacterial contamination. The dbRDA axes explained a substantial proportion of the variation in the dataset. The first axis (dbRDA1) accounted for 34.6% of the fitted variation corresponding to 38.5% of total variation, while the second axis (dbRDA2) explained 22.9% of the fitted variation corresponding to 25.5% of total variation. Thus, the first two axis collectively represented 57.5% of fitted variation and 64% of total variation indicating that the ordination captured a considerable proportion of the bacterial relationships.

The environmental variables and bacterial indicators showed distinct directional relationships within the ordination space. Samples located along the positive axis of dbRDA1 were strongly associated with *E. coli* concentrations in water, sediment, and clam, as well as rainfall ($p \leq 0.001$). The above observations suggest a close association of the bacterial load with rainfall and runoff, as surface runoff and enhanced transport of potentially facilitated faecal pollution in the water bodies. In contrast, salinity and dissolved oxygen (DO) ($p \leq 0.001$) were all on the negative side of dbRDA1, suggesting an inverse relationship with the bacterial contamination. The higher concentration of these physicochemical parameters would decrease the chances of bacterial survival. Overall, the results identify monsoon-driven rainfall and associated runoff as the primary environmental drivers of *E. coli* contamination, with pronounced seasonal variability observed during MON and PSMON periods.

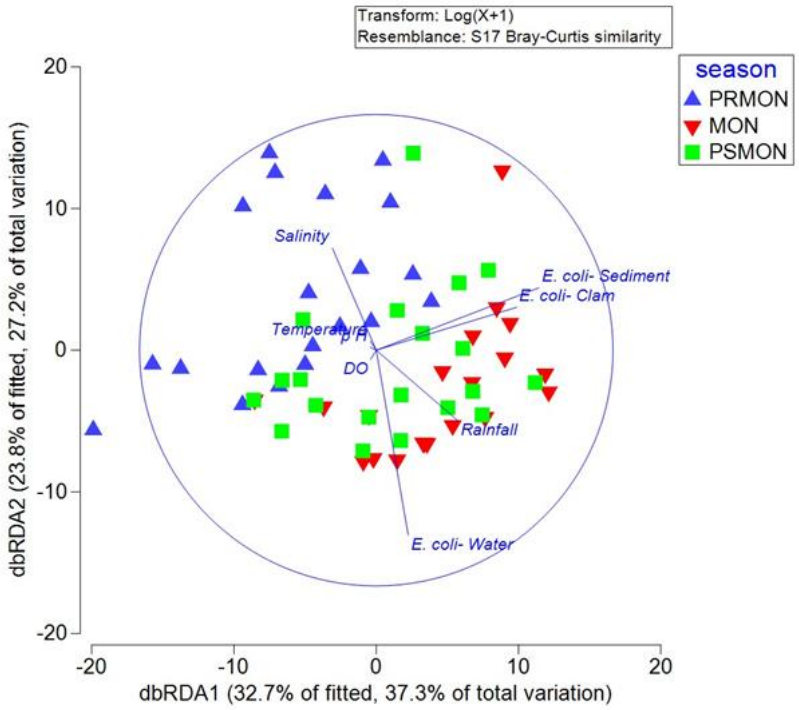


Fig. 11. dbRDA plot showing the relationship with environmental parameters and bacteriological data in different shellfish harvesting areas in Vembanad Lake

4. DISCUSSION

The study assesses *E. coli* contamination in water, sediment and clam tissues collected from five shellfish-harvesting stations in Vembanad Lake, India's largest Ramsar estuarine hotspot of aquatic biodiversity in India. However, increasing anthropogenic activities, including land reclamation, eutrophication and untreated anthropogenic activities have contributed to ecological degradation of this ecosystem (Asha et al., 2016). Assessment of the ecosystem health revealed that over 30% of Vembanad Ramsar coastal zone is currently under fair to poor health category emphasizing the need for integrated biological monitoring (Padua et al., 2023).

The study results indicate the presence of potentially pathogenic and virulent *E. coli* in clam, sediment and water of the Vembanad Lake. Our study demonstrates faecal contamination and potential public health implications due to microbial pollution across this ecosystem. The distribution of these organisms across all the environmental compartments is relevant to clam safety, as these organisms are filter feeders, they accumulate microorganisms from surrounding water and sediment, potentially acting as vehicles for human exposure. Their presence reflects both environmental pollution and public health concerns associated with the consumption of contaminated clams. Aquaculture effluents, agricultural runoff, municipal sewage discharge

and tourism have deteriorated water quality in Vembanad Lake. In our study, faecal coliform level was highest during MON and PSMON season than during the PRMON season. Similar findings have been reported in other studies where enhanced surface runoff, sewage discharge and resuspended sediments during intense periods of prolonged rain increased faecal bacterial numbers in this aquatic system (Chinnadurai et al., 2020, Antony et al., 2021).

Increased surface runoff during monsoon months mobilizes faecal contamination from human settlements, agricultural and drainage channels etc. into water bodies. Enhanced freshwater influx, and sediment resuspension may further redistribute bacterial contaminants. Elevated levels during the post-monsoon also occurred due to prolonged rainy period and delayed clearance of microbial contamination from these environmental compartments. Due to rainfall and freshwater influx, salinity and dissolved oxygen (DO) levels were found to be low during monsoon, which was negatively correlated with FC. Reduced salinity enhances the survival of enteric bacteria and decreasing DO limits oxidative stress, thereby increasing faecal bacterial number (Desta et al., 2024). Of the total 313 *E. coli* isolates detected, most isolates were from PRMON followed by MON and PSMON seasons. The PRMON season had favourable and stable environmental conditions for bacterial survival, while intense flushing during monsoon distributed bacterial populations to various compartments from various land based polluted areas.

The hydrological characterization of this lake, particularly Thaneermukkam barrage provides an important context for understanding the spatial and seasonal variation in bacterial contamination (Asha et al., 2016). Thanneermukkam barrage was established in 1976 across the backwater system primarily to prevent mixing of seawater into the southern part of lake and facilitate paddy culture. Station S3 located near the barrage divides the southern basin of Vembanad Lake into oligohaline and estuarine segments. This barrage has altered the natural salinity gradient and hydrological connectivity between the fresh water dominated and estuarine region of lake (Cleetus et al., 2024, Asha et al., 2016). While the barrage's construction has successfully curbed saline water intrusion into upstream areas, the hydrological manipulation, combined with the monsoon-driven influx of pollutants has affected the distribution and retention patterns around stations S3, S4 and S5 (Kumarakom, Muhamma, Thanneermukkom). Changes in freshwater inflow, saline water intrusion and water exchange may influence bacterial transport, dispersal and persistence differently on both sides of the barrage. The presence of Thaneermukkam barrage, therefore represents an important

hydrological feature to consider while interpreting the spatial distribution of *E. coli* and fecal contamination in Vembanad Lake.

The major strength of our investigation is the simultaneous assessment of faecal contamination and *E. coli* virulence across three distinct environmental compartments such as harvesting water, sediment, and bivalve tissue and its spatiotemporal variation. The studies that concurrently assessed water, sediment, and bivalves within the same framework remain comparatively less, and the extent to which bioaccumulation reflects environmental patterns across compartments is limited (Garcés-Ordóñez et al., 2022). The detection of a diverse range of virulence genes such as *stx1, stx2, eaeA, hlyA, elt, est, bfp, aggR, ipaH, ial, and fliC* in environmental and shellfish isolates confirms the presence of several diarrheagenic *E. coli* pathotypes in clams and their associated environment. Conditions like haemorrhagic colitis and haemolytic uremic syndrome (HUS) have been associated with the presence of such multi-virulence genotypes (Karmali et al., 2010). Studies have reported that among various virulence genes detected, Shiga toxin-producing *E. coli* (STEC) strains that carry *stx1* were more prevalent when compared to *stx2* (Bennani et al 2011). However, Antony et al. (2021) reported a higher prevalence of the *stx2* gene than the *stx1* gene in *E. coli* isolated from black clams in the Cochin estuary. Nevertheless, the detection of the virulence gene should be interpreted as evidence of pathogenic potential rather than confirmation of toxin expression or clinical virulence.

Virulence gene profiling showed that the EHEC-associated gene sets were the most common and being present in 51.17% of virulent *E. coli* isolates. The higher prevalence of *stx1* may be due to the persistence of *stx* carrying lysogenic phage, which can facilitate horizontal transfer and environmental persistence of *stx* gene among susceptible populations (Schmidt & Muniesa, 2023). However, in this study the phage associated mechanism or horizontal gene transfer of *stx* genes are not examined and hence this interpretation could be regarded as a possible explanation.

The ETEC (*elt*, *est*) genes-containing strains known for causing traveller's diarrhoea and watery diarrhoea, particularly in children accounted for 4.22% of the total isolates. In contrast, EPEC and atypical EPEC genes were present in only 3.75% and 1.87% of isolates, respectively, followed by EAEC (*aggR*) and EIEC (*ipaH*), indicating a lower presence of these pathogroups in the stations. Thus, simultaneous detection of diverse pathogroups reflects considerable genetic diversity within environmental isolates of *E. coli*. This also suggests contamination from multiple anthropogenic and environmental stressors. Municipal sewage, domestic waste,

agricultural discharges, and tourism-associated activities may provide multiple routes for the introduction of faecal bacteria into this environment. Such differences highlight the spatial heterogeneity of environmental isolates of *E. coli* and may reflect differences in sources of pollution and environmental contamination.

Water acts as a primary medium for transport and dispersal of planktonic bacteria, which explains the higher detection in water samples (Maugeri et al., 2004). Our study revealed a higher prevalence of toxigenic genes in water (45.08%), followed by clam (38.03%), and sediment (16.91%). In contrast, sediments serve as a longer-term reservoir of microbial cells and free DNA, which can bind to particles and perhaps become less detectable using molecular assays (Sabatino et al., 2025). Pathogen loads and related virulence genes in the water column rapidly increased post flooding in 2018. It has demonstrated vulnerability of aquatic environment to large-scale contamination from sewage, sediment, and runoff. Floodwaters transport bacteria and pathogenic *E. coli* from terrestrial and land sources into surface waters. Shortly after such occurrences, this "flush-in" process causes greater concentrations of virulence determinants in water (Nwaubani & Sherchan, 2026).

Serotyping demonstrated huge diversity among *E. coli* isolates in clam, sediment, and water samples. A total of 47 serotypes from various pathogroups, including O11, 018, and O7, were isolated from various samples (n = 313), highlighting the health significance of environmental contamination. Ecologically, high serotype diversity found in this study indicates several potential contamination sources and complex ecological dynamics in the estuarine environment. Although, other studies have reported a lower diversity of *E. coli*, with 17 serotypes (Prakasan et al., 2022) and 25 serotypes (Sukumaran et al., 2012). In contrast, Chandran et al. (2008) reported 43 *E. coli* serotypes. Among the 47 serotypes identified, a widespread occurrence of O11, O114, O14, O15, O17, O2, O23, O25, O27, O50, O7, O8 were observed in our study across multiple environmental compartments. Globally, serotype O2 causes bloodstream infections (Mamani et al., 2019), while O8 is linked to urinary-tract infections (Mohammed et al., 2021; El Halfawy et al., 2026); O15 predominates in Lithuanian uropathogenic *E. coli* (Prakapaite et al., 2019); and O25 is the most common extraintestinal pathogenic *E. coli* (ExPEC) serotype worldwide, including ESBL-producing strains (Ciesielczuk et al., 2016). Previous literature on infection by the serotypes O11, O114, O14, O17, O23, O27, O50 and O7 have not been reported yet. These serotypes are highly pathogenic, highlighting the possible public health dangers posed by lakes acting as reservoirs for pathogenic pathogens (Abhirosh et al., 2011).

The serotype distribution of *E. coli* from the study revealed a habitat-specific pattern, with serotype O11 dominating the clam samples, O7 dominating the sediment samples, and O18 dominating the water samples. To the best of our knowledge, no such sample-specific predominance of the mentioned serotypes of *E. coli* have been reported in Vembanad shellfish harvesting environments. Earlier studies in estuarine and coastal environments, reported the presence of various *E. coli* serotypes in water, sediment, and shellfish samples, but the distribution of the serotypes was shown to overlap rather than to exhibit dominance of serotypes in particular environments. The results of the current study, therefore, add new knowledge to the *E. coli* serotype distribution in environments associated with clam harvesting. Further studies should focus on the exploration of the factors that may influence the *E. coli* serotype distribution in such environments.

Global public health and environmental security are threatened by the dissemination of antibiotic-resistant bacteria (ARB) in aquatic ecosystems. In our study, the antibiotic resistance pattern showed high resistance to cefotaxime (83.2%), followed by rifampicin and erythromycin with 81.7% and 68.1%, respectively. Divya & Hatha (2019) reported a high prevalence (40%) of multi drug resistance, particularly to ampicillin, tetracycline, and nalidixic acid. According to Sukumaran et al. (2012), the Cochin estuary serves as a habitat for resistant bacteria. Vaiyapuri et al. (2021) also reported a higher resistance of *E. coli* isolates to cefotaxime, which is a third-generation cephalosporin drug. Multi-drug resistance (MDR) was common, particularly showing high resistance to ampicillin and tetracyclines, which are widely used in human and veterinary medicine. The widespread occurrence of multidrug-resistant isolates reflects selective pressure associated with antimicrobial use in human medicines, agriculture, and aquaculture (Ajayi et al., 2024; Apreja et al., 2022).

The dbRDA results indicated strong seasonal influence on bacterial contamination patterns, with monsoon and post-monsoon periods showing higher associations with faecal indicator bacteria in water, sediment, and clam. This phenomenon is consistent with higher surface runoff inflow of untreated sewage and sediment resuspension in the monsoon and post-monsoon period, promoting the transport and persistence of pathogenic bacteria in the aquatic environment. The rainfall induced influx of freshwater directly upturns the bacterial loading and alters key physicochemical parameters such as salinity and dissolved oxygen (Baliarsingh et al., 2021) which were negatively correlated with faecal indicators. Reduced salinity in

estuarine waters has been found to promote the survival of enteric bacteria, and lower DO levels can support bacterial persistence by reducing oxidative stress and microbial competition.

The strong correlation between sediment-associated bacteria and rainfall highlights the role of sediments as a sink and secondary reservoir of faecal bacteria (Perkins et al., 2014). During monsoon period, increased water flow and sediment resuspension can redistribute sediment-associated bacteria into the water column again, sustaining contamination even after peak rainfall events (Jamieson et al., 2005). These findings highlight the need for integrated management strategies that address both in-situ water quality parameters and external factors for the survival of bacteria during the rainy season.

The results of this study from a One Health point of view indicate that environmental pollution, ecosystem health, and human health risk are interlinked. The occurrence of harmful bacteria, their antibiotic resistance, and persistence in sediment, water and biota emphasize the importance of integrated management approaches. To prevent further degradation and to protect the ecological and public health value of this vulnerable wetland ecosystem, it is critical to improve wastewater treatment infrastructure, regulate agricultural-aquaculture effluents, restore natural hydrological regimes, and implement long-term environmental monitoring (Wang et al., 2024). Despite providing important baseline data, this study has certain limitations. The analysis was limited to an analysis of selected virulence genes and whole genome sequencing could yield greater insights into evolutionary relationships and resistance mechanisms. More research into the genomic, metagenomic and resistome analyses, is required to better understand how ecological persistence of pathogenic *E. coli* occurs in this tropical estuarine environment.

5. CONCLUSION

Molecular screening of *E coli* isolates revealed significant diversity in prevalence and distribution of shiga toxin genes *(stx1* and *stx2)* across three environmental compartments. *stx1* was more frequently detected in water isolates (42.6%), whereas *stx2* were common to clams (33.3%). Other related genes to adhesion and toxin production namely *eaeA, hlyA* and *bfpA* were also detected especially with sediments and water samples. Most water sources were positive for combinations of virulence genes, including *stx1, stx2, eaeA* and *hlyA* suggest that water may be a reservoir and vehicle of the transmission of some of the pathogenic *E. coli*. Similarly, the *ipaH* and *aggR* detection denotes presence of some classes such as enteroinvasive and enteroaggregative *E. Coli* in these environments. These findings therefore, support the need for regular monitoring of microbial and virulence-associated markers in

sediments, shellfish harvesting waters and bivalves. At the same time, it is also important to implement shellfish sanitation and post-harvest handling practices. Overall, the study provides baseline information for assessing microbial contamination in various environmental compartments of the Vembanad Lake and highlights the importance of integrating molecular surveillance methods such as virulence gene marker detection for public health risk assessment.

Acknowledgements

The authors acknowledge the Head of the Department, Department of Biotechnology, Cochin University of Science and Technology, for providing the necessary research facilities. The authors also express their gratitude to the Vice Chancellor of CUSAT for institutional support. The authors graciously thank Dr. Jomy Joseph of the Microbial Ecology Lab, Department of Biotechnology, CUSAT, for providing the standard strain. Additionally, the authors extend their gratitude to the Council of Scientific and Industrial Research (CSIR), New Delhi, for providing a research fellowship grant for Neenu P to assist with the preparation of this manuscript.

Funding

The author Neenu P (ID No: 35101695) received a research fellowship grant from Council of Scientific and Industrial Research (CSIR), New Delhi to assist with the preparation of this manuscript.

Authors and Affiliations

Department of Biotechnology, Cochin University of Science and Technology, Cochin- 682 022, Kerala, INDIA

Neenu Pavithran, Parvathi Ammini, Aleena Paul, Shruthi P.K., Paul Mwangi

Central Institute of Fisheries Education, Mumbai, INDIA.

Sanath Kumar H

ICAR- Central Institute of Fisheries Technology (ICAR-CIFT), Willington Island, Cochin-692029, Kerala, India

Adithya Suresh

Department of Microbiology, Kasturba Medical College, Manipal– 576104, Karnataka,

INDIA.

Neeraja R.

Department of Chemical Engineering, Konkuk University, Seoul 05029, Republic of Korea.

Navaneed V.P.

Author contributions

Neenu Pavithran: Conceptualization, data curation, investigation, methodology, formal analysis, writing-original draft, software and visualization.

Parvathi A: Conceptualization, writing- review and editing, validation, funding acquisition, supervision, methodology.

Adithya Suresh: Writing-original draft, methodology, software and visualization.

Aleena Paul: Writing-original draft and investigation.

Neeraja R. : Data curation, Investigation,

Navaneed V.P. : Data curation, Investigation

Shruthi P.K.: Investigation

Paul Mwangi: Investigation

All authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

REFERENCE

1. Abdulaziz, A., Sathyendranath, S., Vijayakumar, S. K., Menon, N., George, G., Kulk, G., Raj, D., Krishna, K., Rajamohananpillai, R., & Tharakan, B. (2023). The distribution of fecal contamination in an urbanized tropical lake and incidence of acute diarrheal disease. *Acs Es&t Water*, *3*(6), 1561–1573. https://doi.org/10.1021/acsestwater.2c00255
2. Abhirosh, C., Sherin, V., Thomas, A., Hatha, A., & Mazumder, A. (2011). Potential public health significance of faecal contamination and multidrug-resistant Escherichia coli and Salmonella serotypes in a lake in India. *Public Health*, *125*(6), 377–379. DOI: 10.1016/j.puhe.2011.03.015
3. Ajayi, A. O., Subulade, A. A., Adejoh, O. P., & Popoola, O. A. (2024). Investigation of Antibiotic-Resistant Enterococci in Farm Animals and Vegetables: A case study of Selected Towns in Ekiti State, Nigeria. *Researchers Journal of Science and Technology*, *4*(5), 85–100. https://doi.org/10.83080/rejost.vol4no5.141
4. Ali, N. M., Khan, M. K., Mazhar, B., & Mustafa, M. (2025). Impact of water pollution on waterborne infections: emphasizing microbial contamination and associated health hazards in humans. *Discover Water*, *5*(1), 19. https://doi.org/10.1007/s43832-025-00198-x
5. Al Qabili, D. M. A., Aboueisha, A.-K. M., Ibrahim, G. A., Youssef, A. I., & El-Mahallawy, H. S. (2022). Virulence and antimicrobial-resistance of shiga toxin-producing E. coli (STEC) Isolated from edible shellfish and its public health significance. *Archives of Microbiology*, *204*(8), 510. DOI: 10.1007/s00203-022-03114-2
6. Antony, A. C., Silvester, R., Divya, P., Aneesa, P., Francis, B., Hussain, M. S., Umesh, B., George, J., & Abdulla, M. H. (2021). Faecal contamination and

prevalence of pathogenic E. coli in shellfish growing areas along south-west coast of India. *Regional Studies in Marine Science*, *44*, 101774. 10.1016/j.rsma.2021.101774

7. Apreja, M., Sharma, A., Balda, S., Kataria, K., Capalash, N., & Sharma, P. (2022). Antibiotic residues in environment: Antimicrobial resistance development, ecological risks, and bioremediation. *Environmental Science and Pollution Research*, *29*(3), 3355–3371. 10.1007/s11356-021-17374-w
8. Arun, A. U. (2009). An assessment on the influence of salinity in the growth of black clam (Villorita cyprinoides) in cage in Cochin Estuary with a special emphasis on the impact of Thanneermukkom salinity barrier. *Aquaculture, Aquarium, Conservation & Legislation*, *2*(3), 319–330.
9. Asha, C. V., Retina, I. C., Suson, P. S., & Bijoy Nandan, S. (2016). Ecosystem analysis of the degrading Vembanad wetland ecosystem, the largest Ramsar site on the South West Coast of India—Measures for its sustainable management. *Regional Studies in Marine Science*, *8*, 408-421. https://doi.org/10.1016/j.rsma.2016.06.003
10. Baliarsingh, S. K., Lotliker, A. A., Srichandan, S., Basu, A., Nair, T. B., & Tripathy, S. K. (2021). Effect of tidal cycle on Escherichia coli variability in a tropical estuary. *Bulletin of Environmental Contamination and Toxicology*, *106*(4), 622–628. 10.1007/s00128-021-03106-w
11. Bauer, A. W., Kirby, W. M., Sherris, J. C., & Turck, M. (1966). Antibiotic susceptibility testing by a standardized single disk method. *American Journal of Clinical Pathology*, *45*(4_ts), 493–496. https://doi.org/10.1093/ajcp/45.4_ts.493
12. Bennani, M., Badri, S., Baibai, T., Oubrim, N., Hassar, M., Cohen, N., & Amarouch, H. (2011). First detection of Shiga toxin-producing Escherichia coli in

shellfish and coastal environments of Morocco. *Applied Biochemistry and Biotechnology*, *165*(1), 290–299. 10.1007/s12010-011-9251-x

13. Chandran, A., AA, M. H., Varghese, S., & Sheeja, K. M. (2008). Prevalence of multiple drug resistant Escherichia coli serotypes in a tropical estuary, India. *Microbes and Environments*, *23*(2), 153–158. 10.1264/jsme2.23.153
14. Chinnadurai, S., Campos, C. J., Geethalakshmi, V., Sharma, J., Kripa, V., & Mohamed, K. (2020). Microbiological quality of shellfish harvesting areas in the Ashtamudi and Vembanad estuaries (India): Environmental influences and compliance with international standards. *Marine Pollution Bulletin*, *156*, 111255. 10.1016/j.marpolbul.2020.111255
15. Ciesielczuk, H., Jenkins, C., Chattaway, M., Doumith, M., Hope, R., Woodford, N., & Wareham, D. W. (2016). Trends in ExPEC serogroups in the UK and their significance. *European Journal of Clinical Microbiology & Infectious Diseases*, *35*(10), 1661–1666. 10.1007/s10096-016-2707-8
16. Clarke, K. R., & Gorley, R. N. (2015). Getting started with PRIMER v7. *PRIMER-e: Plymouth, Plymouth Marine Laboratory*, *20*(1), 300.
17. Cleetus, R. I., Asha, C. V., Nandan, S. B., Suson, P. S., & Jayachandran, P. R. (2024). Ecological impact of artificial barrage on calanoid copepods in the Vembanad Lake, Southwest Coast of India. *Regional Studies in Marine Science*, *80*, 103865. 10.1016/j.rsma.2024.103865
18. CLSI, 2023. Performance Standards for Antimicrobial Susceptibility Testing. CLSI supplement M100, 33rd ed. Clinical and Laboratory Standards Institute, Wayne, PA. ISBN: 978-1-68440-170-3.
19. Desta, B. N., Tustin, J., Sanchez, J. J., Heasley, C., Schwandt, M., Bishay, F., Chan, B., Knezevic-Stevanovic, A., Ash, R., & Jantzen, D. (2024). Environmental

predictors of Escherichia coli concentration at marine beaches in Vancouver, Canada: A Bayesian mixed-effects modelling analysis. *Epidemiology & Infection*, *152*, e38. 10.1017/S0950268824000311

20. Divya, P., Paul, S., Fathima, P., & Abdulla, M. H. (2016). Comparative evaluation of EMB agar and hicrome E. coli agar for differentiation of green metallic sheen producing non E. Coli and typical E. Coli colonies from food and environmental samples. *Journal of Pure and Applied Microbiology*, *10*(4), 2863–2870. https://doi.org/10.22207/JPAM.10.4.48
21. Divya, S. P., & Hatha, A. M. (2019). Screening of tropical estuarine water in south-west coast of India reveals emergence of ARGs-harboring hypervirulent Escherichia coli of global significance. *International Journal of Hygiene and Environmental Health*, *222*(2), 235–248. 10.1016/j.ijheh.2018.11.002
22. Don Xavier, N., Nandan, S. B., Jayachandran, P., Neethu, K., Mohan, D., & Marigoudar, S. (2021). Eliciting heavy metal contamination on selected native organisms from Cochin estuary using contemporary biomarker approach. *Journal of Earth System Science*, *130*(3), 174. 10.1007/s12040-021-01676-1
23. El Halfawy, N. M., Gouda, M. K., Elgayar, F. A., & Badran, A. A. (2026). Genomic characterization of multidrug-resistant Escherichia coli strains identified from patients with urinary tract infection in Egypt. *Scientific Reports*. https://doi.org/10.1038/s41598-026-40536-0
24. Fehrenbach, G. W., Murphy, E., Pogue, R., Carter, F., Clifford, E., & Major, I. (2024). Comprehensive analysis and assessment of exposure to enteric viruses and bacteria in shellfish. *Marine Environmental Research*, *196*, 106404. https://doi.org/10.1016/j.marenvres.2024.106404

25. Feng, P., Weagant, S., & Jinneman, K. (2011). *Diarrheagenic Escherichia coli. Chapter 4A in Bacteriological Analytical Manual (BAM)*.

26. Garces-Ordonez, O., Saldarriaga-Velez, J. F., Espinosa-Díaz, L. F., Patiño, A. D., Cusba, J., Canals, M., ... & Thiel, M. (2022). Microplastic pollution in water, sediments and commercial fish species from Ciénaga Grande de Santa Marta lagoon complex, Colombian Caribbean. *Science of the Total Environment*, *829*, 154643. https://doi.org/10.1016/j.scitotenv.2022.154643

27. Gomes, L., Bordalo, A. A., & Machado, A. (2024). Characterization of Escherichia coli isolates in recreational waters: Implications for public health and one health approach. *Water*, *16*(18), 2695. https://doi.org/10.3390/w16182695

28. Hu, J., Han, G., & Zhang, Q. (2025). Impacts of environmental change and human activities on aquatic ecosystems. In *Water* (Vol. 17, Number 11, p. 1669). MDPI. https://doi.org/10.3390/w17111669

29. Jamieson, R. C., Joy, D. M., Lee, H., Kostaschuk, R., & Gordon, R. J. (2005). Resuspension of sediment-associated Escherichia coli in a natural stream. *Journal of Environmental Quality*, *34*(2), 581–589. 10.2134/jeq2005.0581

30. Karmali, M. A., Gannon, V., & Sargeant, J. M. (2010). Verocytotoxin-producing Escherichia coli (VTEC). *Veterinary Microbiology*, *140*(3–4), 360–370. 10.1016/j.vetmic.2009.04.011

31. Kulk, G., George, G., Abdulaziz, A., Menon, N., Theenathayalan, V., Jayaram, C., Brewin, R. J., & Sathyendranath, S. (2021). Effect of reduced anthropogenic activities on water quality in Lake Vembanad, India. *Remote Sensing*, *13*(9), 1631. https://doi.org/10.3390/rs13091631

32. Mamani, R., Flament-Simon, S. C., Garcia, V., Mora, A., Alonso, M. P., Lopez, C., Garcia-Menino, I., Diaz-Jimenez, D., Blanco, J. E., & Blanco, M. (2019).

Sequence types, clonotypes, serotypes, and virotypes of extended-spectrum β-lactamase-producing Escherichia coli causing bacteraemia in a Spanish hospital over a 12-year period (2000 to 2011). *Frontiers in Microbiology*, *10*, 1530. https://doi.org/10.3389/fmicb.2019.01530

33. Maugeri, T. L., Carbone, M., Fera, M. T., Irrera, G. P., & Gugliandolo, C. (2004). Distribution of potentially pathogenic bacteria as free living and plankton associated in a marine coastal zone. *Journal of Applied Microbiology*, *97*(2), 354–361. 10.1111/j.1365-2672.2004.02303.x

34. Mohammed, E. J., Allami, M., Sharifmoghadam, M. R., & Bahreini, M. (2021). Relationship between antibiotic resistance patterns and O-serogroups in uropathogenic Escherichia coli strains isolated from Iraqi patients. *Jundishapur J Microbiol*, *14*(8), e118833. https://doi.org/10.5812/jjm.118833

35. Nave, H. H., Mansouri, S., Moghadam, M. T., & Moradi, M. (2016). Virulence gene profile and multilocus variable-number tandem-repeat analysis (MLVA) of enteroinvasive Escherichia coli (EIEC) isolates from patients with diarrhea in Kerman, Iran. *Jundishapur Journal of Microbiology*, *9*(6), e33529. 10.5812/jjm.33529

36. Nwaubani, D. A., & Sherchan, S. P. (2026). The Impact of Climate Change on the Transmission of Enteric Diseases Caused by Enteric Bacteria and Viruses: A Systematic Review. *Total Environment Microbiology*, 100056. https://doi.org/10.1016/j.temicr.2026.100056

37. Olds, H. T., Corsi, S. R., Dila, D. K., Halmo, K. M., Bootsma, M. J., & McLellan, S. L. (2018). High levels of sewage contamination released from urban areas after storm events: A quantitative survey with sewage specific bacterial indicators. *PLoS Medicine*, *15*(7), e1002614. 10.1371/journal.pmed.1002614

38. Padua, S., Kripa, V., Prema, D., Mohamed, K. S., Jeyabaskaran, R., Kaladharan, P., ... & Babu, A. (2023). Assessment of ecosystem health of a micro-level Ramsar coastal zone in the Vembanad Lake, Kerala, India. *Environmental Monitoring and Assessment*, *195*(1), 95. 10.1007/s10661-022-10692-7

39. Perkins, T. L., Clements, K., Baas, J. H., Jago, C. F., Jones, D. L., Malham, S. K., & McDonald, J. E. (2014). Sediment composition influences spatial variation in the abundance of human pathogen indicator bacteria within an estuarine environment. *PloS One*, *9*(11), e112951. 10.1371/journal.pone.0112951

40. Prakapaite, R., Saab, F., Planciuniene, R., Petraitis, V., Walsh, T. J., Petraitiene, R., Semoskaite, R., Baneviciene, R., Kalediene, L., & Kavaliauskas, P. (2019). Molecular characterization of uropathogenic Escherichia coli reveals emergence of drug resistant O15, O22 and O25 serogroups. *Medicina*, *55*(11), 733. 10.3390/medicina55110733

41. Prakasan, S., Lekshmi, M., Ammini, P., Balange, A. K., Nayak, B. B., & Kumar, S. H. (2022). Occurrence, pathogroup distribution and virulence genotypes of Escherichia coli from fresh seafood. *Food Control*, *133*, 108669. 10.1016/j.foodcont.2021.108669

42. Sabatino, R., Sivalingam, P., Di Nezio, F., Borgomaneiro, G., Rogora, M., Corno, G., Fontaneto, D., & Di Cesare, A. (2025). Intra-and extracellular DNA as carriers of antibiotic resistance genes and class 1 integrons in river waters. *Hydrobiologia*, *852*(20), 5135–5146. https://doi.org/10.1007/s10750-025-05906-1

43. Schmidt, H., & Muniesa, M. (2023). The Prophage and Us—Shiga Toxin Phages Revisited. In *Pathogens* (Vol. 12, Number 2, p. 232). MDPI. https://doi.org/10.3390/pathogens12020232

44. Singh, A. S., Nayak, B. B., & Kumar, S. H. (2020). High prevalence of multiple antibiotic-resistant, extended-spectrum β-lactamase (ESBL)-producing Escherichia coli in fresh seafood sold in retail markets of Mumbai, India. *Veterinary Sciences*, *7*(2), 46. 10.3390/vetsci7020046

45. Sukumaran, D. P., Durairaj, S., & Abdulla, M. H. (2012). Antibiotic resistance of Escherichia coli serotypes from Cochin estuary. *Interdisciplinary Perspectives on Infectious Diseases*, *2012*(1), 124879. 10.1155/2012/124879

46. Tom, A. P., Jayakumar, J. S., Biju, M., Somarajan, J., & Ibrahim, M. A. (2021). Aquaculture wastewater treatment technologies and their sustainability: A review. *Energy Nexus*, *4*, 100022. 10.1016/j.nexus.2021.100022

47. Vaiyapuri, M., Sebastian, A. S., George, I., Variem, S. S., Vasudevan, R. N., George, J. C., Badireddy, M. R., Sivam, V., Peeralil, S., & Sanjeev, D. (2021). Predominance of genetically diverse ESBL Escherichia coli identified in resistance mapping of Vembanad Lake, the largest fresh-cum-brackishwater lake of India. *Environmental Science and Pollution Research*, *28*, 66206–66222. 10.1007/s11356-021-15110-y. Epub 2021 Jul 30.

48. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., & Bright, J. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. 10.1038/s41592-019-0686-2.

49. Wang, F., Xiang, L., Leung, K. S.-Y., Elsner, M., Zhang, Y., Guo, Y., Pan, B., Sun, H., An, T., & Ying, G. (2024). Emerging contaminants: A one health perspective. *The Innovation*, *5*(4). https://doi.org/10.1016/j.xinn.2024.100612