frontiers | Frontiers in Veterinary Science

TYPE Original Research
PUBLISHED 09 October 2025
DOI 10.3389/fvets.2025.1639243

OPEN ACCESS

EDITED BY
Klaas Dietze,
Friedrich-Loeffler-Institute, Germany

REVIEWED BY
Anna Szczerba-Turek,
University of Warmia and Mazury in Olsztyn, Poland
Sumeet Kumar Tiwari,
Quadram Institute, United Kingdom

*CORRESPONDENCE
Nicolás Galarce
✉ ngalarce@uchile.cl

RECEIVED 01 June 2025
ACCEPTED 19 September 2025
PUBLISHED 09 October 2025

CITATION
Martínez V, Cartajena JT, Méndez E, Dörner J, Méndez D, Arriagada G, Toledo J, Arancibia R, Pizarro N, Castro D, Luna D, Ramos R, Jorquera J, Escobar B, Kudva IT and Galarce N (2025) Comparative genomics of Shiga toxin-producing *Escherichia coli* reveals host-specific adhesiome adaptations in humans and cattle.
Front. Vet. Sci. 12:1639243.
doi: 10.3389/fvets.2025.1639243

Comparative genomics of Shiga toxin-producing *Escherichia coli* reveals host-specific adhesiome adaptations in humans and cattle

Víctor Martínez[1], José T. Cartajena[2,3], Estefanía Méndez[4], Jessica Dörner[1], Diego Méndez[1,2], Gabriel Arriagada[5], Jorge Toledo[6,7], Richard Arancibia[8], Nicolás Pizarro[9,10], Daniela Castro[11,12], Daniela Luna[1], Romina Ramos[13], Joaquín Jorquera[13], Beatriz Escobar[2], Indira T. Kudva[14] and Nicolás Galarce[2]*

[1]Departamento de Ciencia Animal, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile, [2]Departamento de Medicina Preventiva Animal, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile, [3]Programa de Magíster en Ciencias Animales y Veterinarias, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile, [4]Life Sciences Faculty, Universidad Andres Bello, Santiago, Chile, [5]Instituto de Ciencias Agroalimentarias, Animales y Ambientales, Universidad de O'Higgins, San Fernando, Chile, [6]Red de Equipamiento Científico Avanzado, Facultad de Medicina, Universidad de Chile, Santiago, Chile, [7]Departamento de Ciencias de la Salud, Universidad de Aysén, Coyhaique, Chile, [8]Departamento de Ciencias Clínicas, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile, [9]Instituto de Investigaciones Agropecuarias, Osorno, Chile, [10]Facultad de Medicina Veterinaria y Agronomía, Universidad de las Américas, Santiago, Chile, [11]Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias y Forestales, Universidad Católica del Maule, Curicó, Chile, [12]Escuela de Medicina Veterinaria, Facultad de Recursos Naturales y Medicina Veterinaria, Universidad Santo Tomás, Santiago, Chile, [13]Escuela de Medicina Veterinaria, Facultad de Ciencias de la Vida, Universidad Andrés Bello, Santiago, Chile, [14]Food Safety and Enteric Pathogens Research Unit, National Animal Disease Center, Agricultural Research Service, United States Department of Agriculture, Ames, IA, United States

Introduction: Shiga toxin-producing *Escherichia coli* (STEC) is a zoonotic pathogen responsible for severe human infections, with cattle recognized as the principal animal reservoir for human infection. Adhesion is a critical step in STEC colonization, facilitating persistence and transmission. While human-associated adhesion mechanisms have been extensively studied, those driving colonization in cattle remain less understood. In this study, we characterized the adhesiome of STEC strains isolated from Chilean cattle and compared them with a global collection to identify host-specific adhesion patterns and genetic adaptations.
Methods: A total of 948 fecal samples from Chilean cattle were screened, yielding 71 confirmed STEC isolates, which were analyzed alongside 546 publicly available genomes to compare host-specific adhesion patterns. The adhesiome was examined based on gene presence/absence patterns, followed by a genome-wide association study (GWAS) and variant effect analysis to identify host-specific adhesion genes and their functional implications.
Results: Adhesin gene analysis revealed distinct adhesion strategies between hosts. Several genes, including *ehaA*, *stgABC*, *yadLMN*, and *iha*, were significantly associated with cattle, while *eae*, *cah*, *ypjA*, and *paa* were more frequent in human-associated STEC. Functional enrichment analysis revealed differences in biological processes, including protein folding and fimbrial usher porin activity in cattle, and response to methylglyoxal in humans. GWAS identified *yeeJ*, *espP*, and *fimC* as strongly associated with cattle strains, whereas *clpV*, *ybgQ*, and sab were linked to human isolates. Variant analysis showed higher genetic diversity in human isolates, with *yadK*, *espP*, and *ybgP* exhibiting the highest variant densities. However, the functional effects of adhesin mutations were largely conserved across hosts, suggesting selective constraints on adhesion mechanisms.

Discussion: Our findings provide new insights into STEC host adaptation and highlight potential targets to reduce zoonotic transmission and improve pre-harvest food safety strategies. Future research should focus on functional validation of host-specific adhesin variants and their potential as preventive strategies.

KEYWORDS

STEC, *E. coli*, Shiga, adhesiome, WGS, GWAS, cattle, human

1 Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are a group of emerging zoonotic pathogens responsible for significant public health and economic burdens worldwide (1). These bacteria produce Shiga toxins (Stx), potent cytotoxins that can cause severe human diseases, including hemorrhagic colitis and hemolytic-uremic syndrome (HUS), especially in young children (2). STEC is primarily transmitted to humans through consumption of contaminated food, particularly beef (3).

Adult cattle are recognized as the primary reservoir of STEC, with reported prevalence rates in Latin America ranging from 14 to 90% (4, 5). Their widespread presence increases the risk of environmental contamination and zoonotic transmission, underscoring the need for effective livestock-based control strategies. Beyond human health concerns, STEC infections impose substantial economic costs related to healthcare, productivity loss, and outbreak management. In the United States alone, healthcare costs associated with STEC infections were estimated to exceed USD 311 million in 2018 (6).

STEC utilizes diverse adhesins—including surface-associated and secreted proteins—to colonize host tissues (7). A key adhesion determinant is the locus of enterocyte effacement (LEE), which encodes intimin (*eae*) and type III secretion system (T3SS) proteins, promoting intimate attachment to enterocytes and microvilli effacement (8). LEE-positive strains, such as O157:H7 and several non-O157 serotypes (e.g., O111:NM, O26:H11, O103:H2), are strongly associated with outbreaks and severe human disease (9). However, the emergence of LEE-negative STEC strains utilizing alternative adhesins, encoded in pathogenicity islands like the locus of adhesion and autoaggregation (LAA) and the locus of proteolysis activity (LPA), highlights additional colonization strategies (10, 11). Collectively, these adhesion determinants constitute the adhesiome, defined as the complete set of fimbrial and non-fimbrial adhesin genes that facilitate bacterial colonization of host tissues (12).

While human-associated adhesion mechanisms have been extensively characterized, those facilitating bovine colonization remain poorly understood. Since cattle serve as the main reservoir for human infection, elucidating these mechanisms is critical for designing targeted pre-harvest interventions. Newborn calves are typically colonized by STEC shortly after birth, acquiring the pathogen from maternal microbiota and the surrounding environment (13, 14). Once ingested, STEC can survive, persist, and colonize the gastrointestinal tract, particularly targeting the recto-anal junction (RAJ) as its primary site of colonization (15, 16). It has been shown that proteins encoded by the LEE play a critical role in STEC adherence to RAJ's stratified squamous epithelium (RSE) cells, including intimin (17, 18). However, other adhesins are involved in the colonization and persistence of LEE-positive STEC in cattle, such as EhaA (19), and H7 flagella (20), among others. Moreover, Kudva et al. (21) registered a similar LEE-positive adhesion pattern of LEE-negative STEC strains to RAJ cells, concluding that adhesins other than intimin are involved in this phenotype. Adhesion of STEC to bovine gut by specific virulence factors is of paramount importance since it allows its persistence and the successful transmission to other hosts (22). Therefore, understanding STEC colonization is the key step to controlling the infection.

The genetic diversity of STEC, including variations within the adhesiome, poses a major challenge for prevention strategies, such as vaccine development. The complexity of STEC adhesion mechanisms across hosts underscores the need for genomic surveillance to characterize circulating strains and their colonization traits (23, 24). This study aims to characterize the adhesiome of STEC strains isolated from cattle in Chile and compare them with global strains to identify host-specific adhesion patterns. Understanding these colonization mechanisms is fundamental for designing effective intervention strategies to mitigate STEC transmission at the livestock level. By elucidating key adhesins involved in bovine persistence, this research contributes to the development of targeted mitigation strategies, contributing to One Health-based strategies to mitigate transmission risks at the human-animal interface.

2 Materials and methods

2.1 Sample collection for STEC isolation

A total of 948 fecal samples were collected between June 2023 and March 2024 from abattoirs and farms located across seven regions, representing the majority of Chile's cattle population (25). The sampled animals included both juveniles and adults, with most being of mixed breed. Approximately 20 g of fecal material per animal was aseptically collected directly from the rectum by trained veterinarians and transported under refrigerated conditions in sterile flasks until laboratory processing. All sampling procedures were approved by the Institutional Committee of Care and Use of Animals, Universidad de Chile (Protocol No. 23658—VET—UCH).

2.2 Sample processing and STEC identification

Samples were processed following protocols from previous studies (26, 27). Briefly, 5 g of each fecal sample were enriched in 9 mL of tryptone soy broth (Becton Dickinson and Co., United States) and incubated overnight at 42 °C. A 25 µL aliquot of the enrichment culture was then plated onto MacConkey agar (Becton Dickinson and Co., United States) and incubated at 37 °C for 18–24 h.

Bacterial growth from confluent areas was resuspended in 500 µL of sterile nuclease-free water, subjected to heat treatment at 100 °C for

15 min, and centrifuged at 26,480 × g for 5 min. DNA concentration and purity (A260/A280 ratio) were determined using a NANO-400 micro-spectrophotometer (Hangzhou Allsheng Instruments Co., China). Samples with optimal purity (1.8–2.0) were stored at −20 °C for subsequent analyses.

The presence of *stx1* and *stx2* genes was confirmed by multiplex PCR (28) in a LifeECO® thermal cycler (Hangzhou Bioer Technology Co., China). The STEC97 strain [*stx1*-positive, *eae*-positive, *stx2*-positive (29)] was used as a positive control, while *E. coli* ATCC 25922 served as the negative control. Up to 30 colonies per positive sample were individually plated on MacConkey agar (Becton Dickinson and Co., United States) and CHROMagar™ STEC (CHROMagar Microbiology, France) (30, 31). After 24 h of incubation, colonies were screened by multiplex PCR to confirm the presence of *stx1* and/or *stx2* genes. PCR-confirmed colonies were then tested for the *uspA* gene, encoding the universal stress protein A, to verify *E. coli* species identity (32). A single confirmed isolate per sample was stored at −80 °C for further analysis.

2.3 Whole genome sequencing of STEC strains

Genomic DNA from all STEC strains was extracted using the Wizard Genomic DNA Purification Kit (Promega, United States), following the manufacturer's instructions. DNA concentration was measured using fluorometry with the Qubit® dsDNA BR Assay kit (Life Technologies, United States), and DNA quality was assessed with an Epoch microplate spectrophotometer (Biotek Instruments, United States). A total of 1 ng of DNA was used for library preparation using the Nextera XT DNA Library Prep Kit (Illumina, United States), following the manufacturer's protocol. The average fragment size of libraries was determined by capillary electrophoresis using the High Sensitivity NGS Fragment Analysis Kit (Advanced Analytical Technologies, United States). Libraries were quantified using the KAPA Library Quantification Kit (Kapa Biosystems, United States) on a Rotor-Gene Q platform (Qiagen, Germany). Whole-genome sequencing (WGS) was performed on a NovaSeq X Plus platform (Illumina, United States) with a 150-cycle paired-end reagent kit at Haplox Company, Hong Kong.

Additionally, two STEC strains isolated from clinical human stool samples, corresponding to serotypes O157:H7 and O26:H11 and provided by the Instituto de Salud Pública de Chile, were included in the analysis for comparison purposes. These strains were processed for WGS as described above. All genome sequences were deposited in GenBank under BioProject number PRJNA656305.

2.4 Publicly available sequence data

To enhance the comparative analysis, 568 publicly available *E. coli* genome sequences were retrieved from GenBank's Sequence Read Archive (SRA)[1] on December 15, 2024. Genomes were selected based on diverse geographical origins, host species, and serotypes to broadly represent the global diversity of STEC strains.

All genomes were screened for the presence of *stx* subtypes by mapping reads against reference sequences using BWA (33). After filtering, a final dataset comprising 546 confirmed STEC genomes was obtained for downstream analysis. Metadata, including host origin, country of isolation, and year, were retrieved using the eSearch tool from the EMBOSS suite (34).

2.5 Epidemiological typing and phylogenomic analysis

All FASTQ Illumina reads were assembled *de novo* using SPAdes (v.3.15.2) with default parameters (35). Genome assembly quality was evaluated using CheckM2 (36), which estimates genome completeness and contamination based on machine learning models trained with lineage-specific marker sets available in the DIAMOND[2] database.

The program was not run using a specific model; a cosine similarity calculation was performed to determine the appropriate completeness model for each isolate. The program predicts protein sequences to annotate all the genomes with DIAMOND. Finally, in all cases the Neural network contamination model was used.[3] This quality control step was critical to ensure the reliability of assemblies, particularly for genomes retrieved from the SRA. Only genomes with a completeness score greater than 99.6%, as estimated by CheckM2, were included in the analysis.

Prediction of *stx* subtypes was performed on ABRicate (v.0.8.13).[4] Sequence types (STs) of all STEC strains were predicted by Achtman's multilocus sequence typing (MLST) scheme using the GitHub platform.[5] The housekeeping genes used included *adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*. Additionally, SerotypeFinder 2.0[6] was used to determine serotype (37).

2.6 Adhesiome analysis

Adhesin gene analysis was conducted using AdhesiomeR,[7] an R-based tool that executes BLASTn searches against a curated database of 427 adhesin genes identified across diverse *E. coli* strains. Adhesin sequences were classified according to sequence identity thresholds: highly similar (>95%), moderately similar (75–95%), and unrelated (<75%) (12). In order to enhance confidence in adhesin identification, a strict mode with gene-specific bit score threshold was applied. Comparative analyses were based on presence/absence matrices (1 = present; 0 = absent) and clustering profiles to detect host-specific adhesin signatures between human- and cattle-associated STEC isolates. Differences in the detection rates of adhesin genes between cattle- and human-associated isolates were analyzed using a *Z*-test for two proportions. The analysis was performed in Microsoft Excel

1 www.ncbi.nlm.nih.gov/genbank/

2 https://github.com/bbuchfink/diamond

3 https://github.com/chklovski/CheckM2/blob/main/README.md

4 https://github.com/tseemann/abricate

5 https://github.com/tseemann/mlst

6 https://cge.food.dtu.dk/services/SerotypeFinder/

7 https://github.com/ksidorczuk/adhesiomeR.git

(Microsoft Office 365, version 16.100.2), with statistical significance set at $p < 0.05$.

To further explore genetic variability, we performed genome-wide association analysis (GWAS) using two complementary approaches. First, we used the complete genome of *E. coli* K-12 (ASM584v2) as the reference, which contains the majority of core adhesin genes and allowed variant calling and functional annotation at the whole-genome level. Second, recognizing that several adhesin genes were absent from the K-12 genome, we conducted an additional GWAS focused on the adhesiome. For this, we extracted all 427 adhesin gene sequences identified in our adhesiome analysis—including those shared between human- and cattle-associated strains—to ensure a more comprehensive assessment. These sequences were retrieved from AdhesiomeR[8] and used as templates for further analysis. The complete genome sequences of STEC strains isolated from stool samples ($n = 158$ from cattle and $n = 205$ from humans) were mapped to the *E. coli* K-12 genome and to the adhesiome sequences using BWA (38). After sorting using BamTools (39), the PCR duplicates were removed using Sambamba (40). Single nucleotide polymorphisms (SNPs) and short insertions/deletions (indels) were identified using FreeBayes and jointly using all the samples in the analysis (41).

2.7 Genome-wide association analysis

GWAS was conducted based on SNPs and indels identified from variant call format (VCF) files generated by FreeBayes. Quality control of variants was performed using VCFtools, applying a minimum quality threshold (MinQ) of 30 and a mean depth filter of 100× (42). A logistic regression model was fitted to the data, using host species (cattle = 1, human = 0) as the binary outcome, following the approach described by Pérez-Enciso et al. (43).

We used Pyseer (44) to discover variants significantly associated with the cattle host, being humans deemed as controls in the standard association analysis. To account for potential confounding by population structure, we conducted a principal component analysis (PCA) on variant call format (VCF) files using PLINK [-pca command (45)]. The model included the first five principal components as fixed effects, as eigenvalues showed no substantial drop beyond this threshold. The pairwise distance matrix was generated using Mash (46). Following the Benjamini and Hochberg correction, the *p*-values were adjusted for multiple comparisons using the false discovery rate (FDR) method, implemented through the p.adjust function in R (FDR $\leq$0.05). Model coefficients were also presented as odds ratios (OR), with values greater than one indicating an association of the alternative allele with the cattle host.

Annotation and functional impact prediction of host-associated variants were performed using snpEff (47). For GWAS analyses based on the *E. coli* K-12 genome, the pre-compiled *E. coli* str. K-12 substr. MG1655 database provided by SnpEff was employed. For the adhesiome sequences that belonged to different bacterial strains, we constructed a custom SNPeff database. The adhesiome sequences were annotated using AUGUSTUS (48), with the *E. coli* K-12 database selected as the reference strain to guide the annotation (obtained as a GFF file format). This annotation file was then converted to GTF format using AGAT to develop the snpEff custom database of the adhesiome. Predicted coding sequences (CDS) and protein sequences of the identified genes were obtained using the script getAnnoFasta.pl (available at https://github.com/nextgenusfs/augustus.git). All these files were used to compile the snpEff database.

3 Results

3.1 STEC identification in Chilean cattle

Of the 948 fecal samples collected from cattle, 71 (7.5%) were confirmed as STEC-positive, harboring *stx1* and/or *stx2* genes along with the *uspA* marker. Among these, 70.4% harbored *stx2*, 26.8% carried both *stx1* and *stx2*, and only 2.8% were positive for *stx1* alone. The geographical distribution of STEC varied across Chile, with regional differences in detection rates (Figure 1).

3.2 Epidemiological typing and phylogenomic analysis

For comparative purposes, 546 publicly available STEC genomes were retrieved from NCBI, encompassing isolates from humans (50.0%), food (23.8%), and cattle (19.4%). These genomes spanned a broad temporal range (1978–2021) and were predominantly collected from Germany (43.8%), Chile (16.1%), and France (13.2%).

Among the 619 genomes analyzed (71 Chilean isolates plus 546 public genomes), we identified 29 distinct *stx* subtypes profiles, with *stx2a* (20.0%) and *stx1a* (18.9%) being the most prevalent. Serotype analysis indicated O157:H7 (10.8%), O26:H11 (7.1%), and O130:H11 (6.0%) as the most frequently detected. MLST classification identified 141 STs, with ST11 (13.4%) and ST297 (8.7%) being the most common. Complete metadata is provided in Supplementary Table 1. Figures 2, 3 depict the distribution of serotypes and STs according to geographical origin and host.

Among the 71 STEC strains isolated in this study, eight *stx* subtypes profiles were identified, with *stx2d* (38.0%), *stx2c* (26.8%), and *stx1a* + *stx2a* (12.7%) being the most prevalent. Among the 12 predicted serotypes, O130:H11 (42.3%), O185:H7 (21.1%), and O113:H21 (8.5%) were the most common. Likewise, the most frequently detected STs were ST297 (49.3%), ST2387 (20.5%), and ST223 (9.6%).

Due to the high diversity in serotypes, *stx* profiles, and STs, we examined the genomic structure of STEC isolates from food, cattle, and humans ($n = 572$) to assess whether genomic similarity clusters corresponded to host origin or geographic region. Based on STs, we generated MinHash sketches of draft whole-genome assemblies using k-mers of length 31 and a sketch size of 100,000 in Sourmash. This analysis revealed substantial genomic variability between groups, potentially enhancing STEC adaptability to diverse environments and supporting a wide range of virulence strategies (Figure 4). In parallel, genetically homogeneous clades were observed within specific STEC STs, suggesting associations with ecological factors or adaptive niches, further supported by the multidimensional scaling analysis (Supplementary Figure 1).

8 https://adhesiomer.quadram.ac.uk/app/adhesiomeR

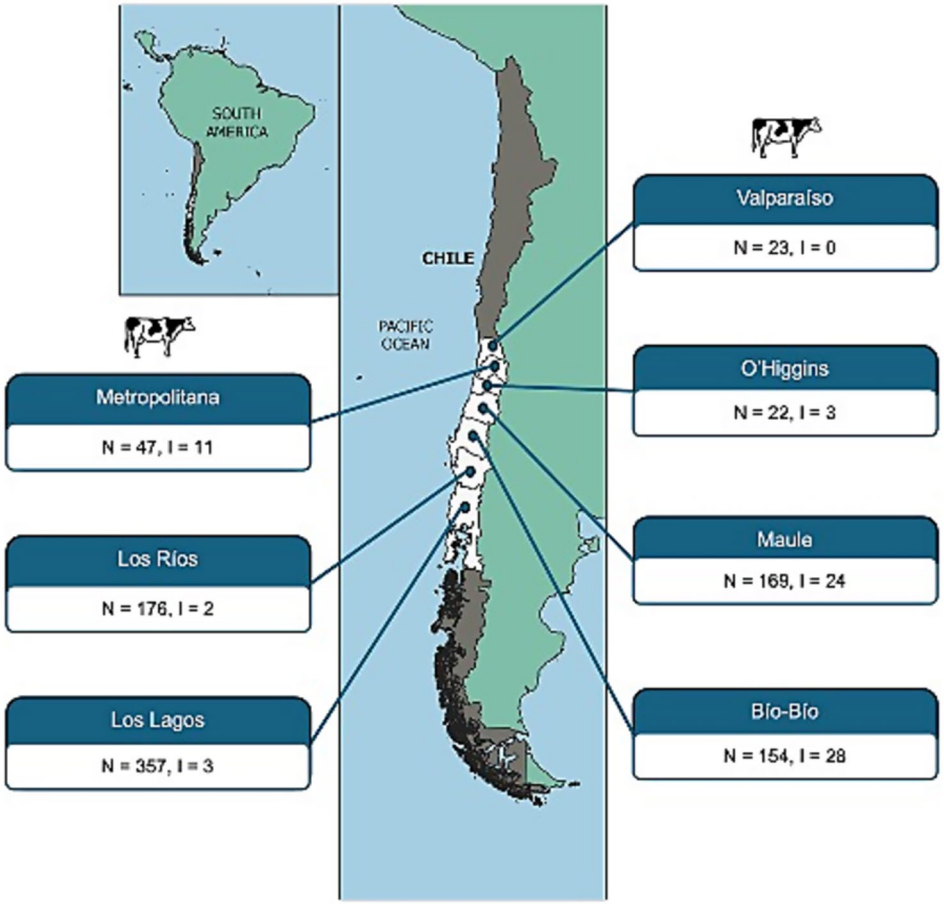

FIGURE 1
Geographical distribution of STEC strains isolated from cattle feces across Chilean regions, based on official administrative divisions. For each region, the number of fecal samples collected (N) and the number of STEC-positive isolates recovered (I) are indicated.

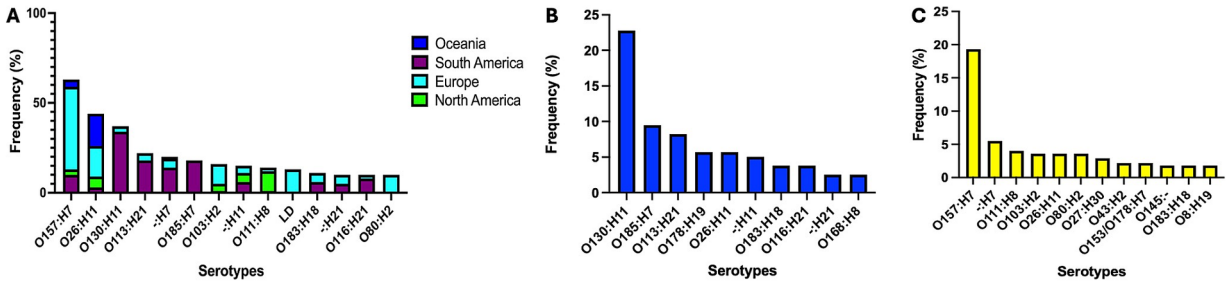

FIGURE 2
Distribution of major STEC serotypes across continents **(A)** and by host species (**B**, cattle; **C**, humans). Selected continents represent 96.7% of analyzed genomes. LD, low discrimination between predicted serotypes.

3.3 Adhesiome analysis

3.3.1 Analysis of the pangenome of the adhesiome in host related strains

To characterize the adhesiome of STEC genomes from cattle and human origins, we used the AdhesiomeR tool to identify adhesin clusters, which defined fimbrial and non-fimbrial adhesins function. The distribution of adhesin-related gene clusters among STEC genomes is presented in Figure 5. Adhesin clusters (Figure 5A) revealed a predominance of the A-A cluster among cattle-associated strains (88.5%), while human-associated strains exhibited greater diversity, with A-G as the most prevalent cluster (44.2%). Fimbrial clusters (Figure 5B) were predominantly represented by the F-C cluster in cattle (83.9%), whereas human-associated strains showed a broader distribution, with F-C and F-F being the most frequent (24.4 and 18.6%, respectively). Similarly, among non-fimbrial clusters

10.3389/fvets.2025.1639243

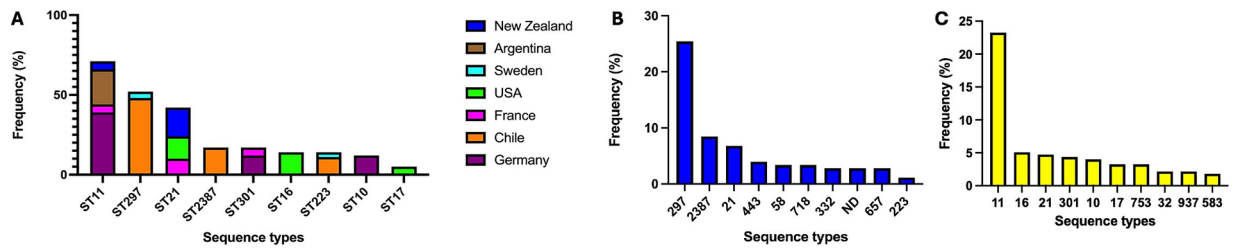

FIGURE 3
Distribution of the most frequent STEC sequence types (STs) across countries **(A)** and host species (**B**, cattle; **C**, humans). Selected countries represent 96.7% of analyzed genomes.

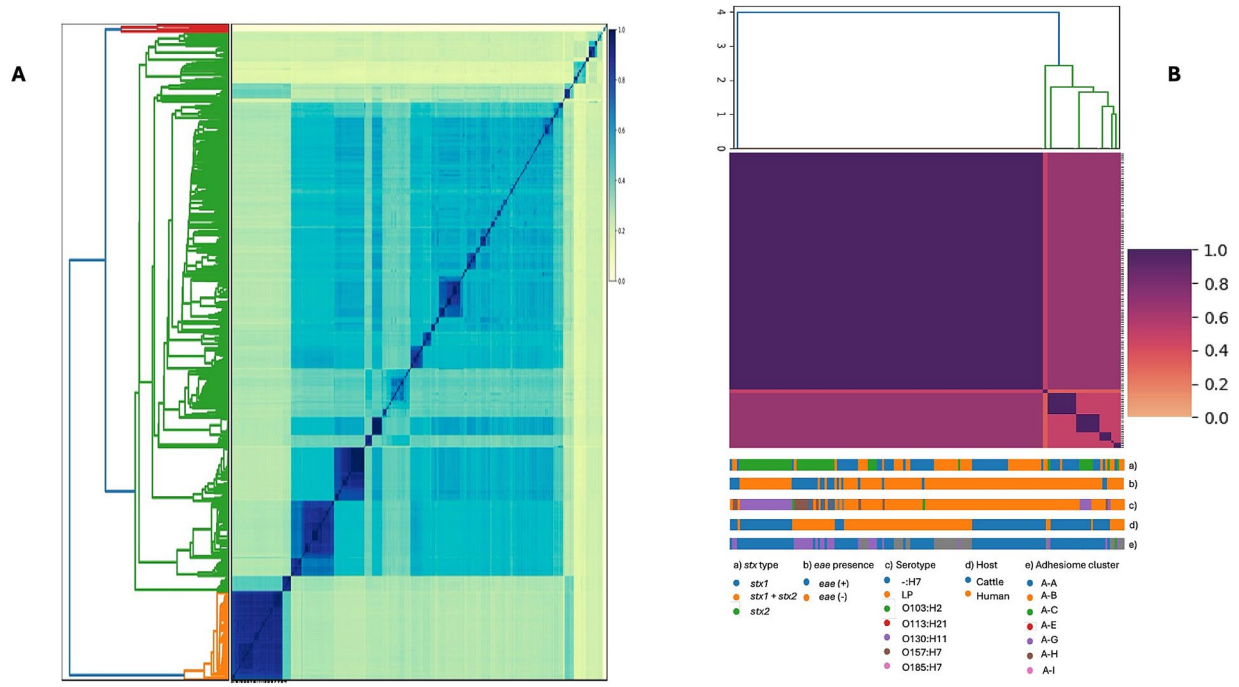

FIGURE 4
Genome similarity among STEC isolates based on whole-genome sequences. **(A)** Pairwise genome similarity (Jaccard similarity index, JSI) among 572 isolates from cattle, humans, and food. **(B)** Genetic diversity of the subset of isolates included in the adhesiome analysis, restricted to human and cattle stool samples. Colored tracks indicate *stx* type (a), *eae* presence (b), serotype (c), host (d), and adhesiome cluster (e). In both panels, darker colors represent higher similarity (JSI close to 1).

(Figure 5C), N-B predominated in cattle strains (67.8%), while human strains were more diverse, with higher proportions of N-A and N-D clusters (17.4 and 14%, respectively).

These findings suggest that cattle-associated STEC strains exhibit a more conserved adhesin profile, whereas human-associated strains display greater diversity, possibly reflecting adaptation to host-specific selective pressures and the diversity of the strains analyzed.

Next, we compared the genes identified in the adhesin clusters according to their host (Figure 6). This analysis revealed a combination of conserved core adhesins and host-associated adhesion profiles. Many adhesin genes, particularly those encoded within the *fim*, *csg*, and *ecp* operons, were highly conserved across cattle- and human-associated strains (>87% prevalence), indicating that certain adhesion mechanisms are fundamental for bacterial colonization, irrespective of the host species.

Nevertheless, differences were observed for specific adhesins. For example, *ehaA* was more prevalent in cattle (97.6%) compared to humans (74.2%). Similarly, *stgA*, *stgB*, and *stgC* were more common in cattle-associated strains, suggesting a greater reliance on alternative fimbrial adhesins for colonization. Additionally, the *yadM*, *yadL*, *yadN*, and *iha* genes were detected more frequently in cattle strains, supporting their potential role in host-specific adaptation. Conversely, STEC strains from humans showed a higher frequency of *eae*, *cah*, *ypjA*, and *paa*, all of which have been implicated in epithelial attachment and virulence.

Using the gene presence/absence data from the adhesiome, we conducted a *Z*-test for differences in proportions to identify host-specific functional enrichments (Supplementary Table 2). Several biological processes related to pilus biology, including pilus formation (GO:0009289), pilus assembly (GO:0009297), and pilus organization (GO:0043711), as well as cell adhesion involved in single-species

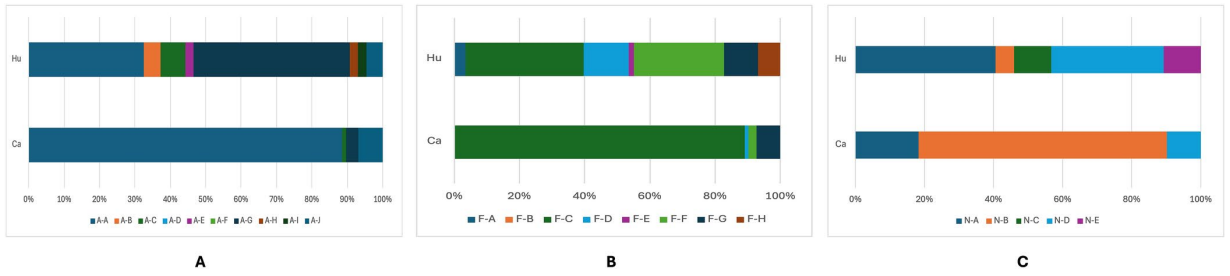

FIGURE 5
Frequency of adhesin **(A)**, fimbrial **(B)**, and non-fimbrial **(C)** clusters among STEC genomes from humans (Hu) and cattle (Ca). Bars represent the number of genomes identified in each cluster. Different letter combinations denote distinct clusters, following the AdhesiomeR nomenclature.

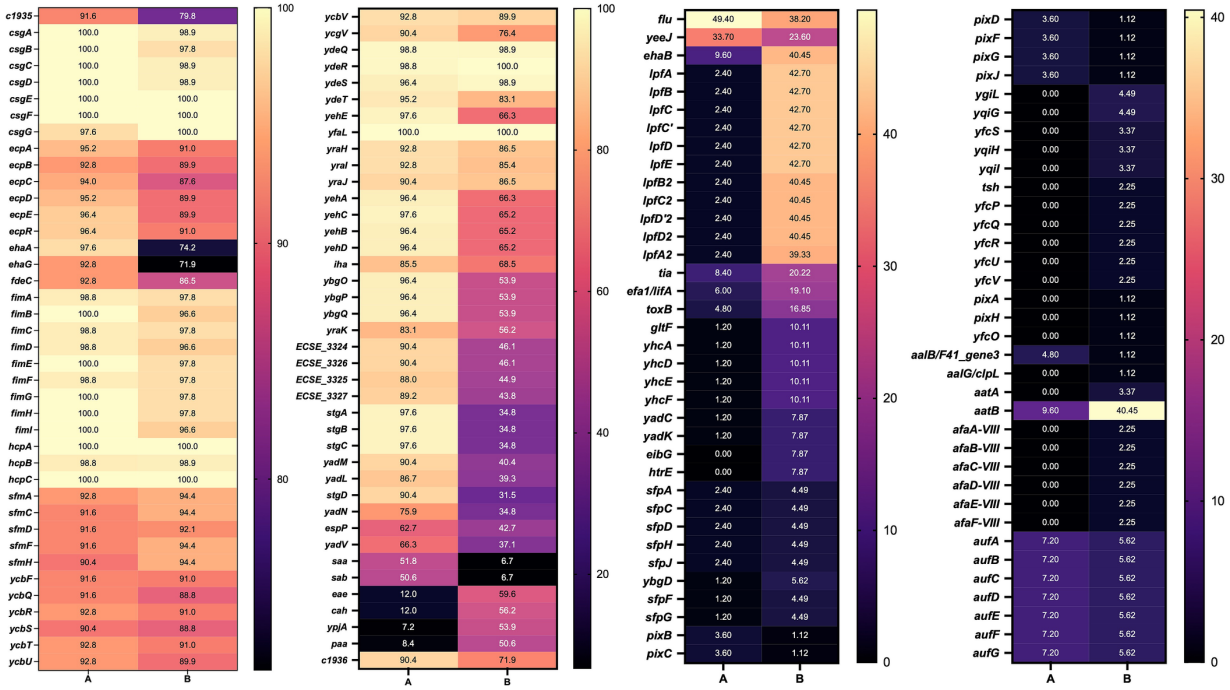

FIGURE 6
Heatmap of genes present in STEC genomes from cattle and human sources. A: cattle; B: human, numbers indicate the percent prevalence of individual genes.

biofilm formation (GO:0043709), were significantly enriched in both cattle- and human-associated strains.

In cattle, processes such as protein folding (GO:0006457) and fimbrial usher porin activity (GO:0015473) were significantly enriched. In contrast, response to methylglyoxal (GO:0051595) was significantly enriched in humans (Supplementary Table 3). These findings support the hypothesis that STEC strains exhibit host-specific adhesion strategies, with cattle-associated strains relying more on fimbrial and curli-mediated mechanisms, while human-associated strains may have evolved broader stress response capabilities to persist within the human gastrointestinal environment.

3.3.2 Genome wide association analysis using *Escherichia coli* K-12

The GWAS using the *E. coli* K-12 reference genome is presented in Figure 7A. Gene enrichment analysis revealed that cattle-associated genes were significantly linked to trehalose transport (GO:0015771) and protein-phosphocysteine-trehalose phosphotransferase system transporter activity (GO:0090589), involving the *bglF* and *ascF* genes. A complete list of enriched genes and associated pathways is provided in Supplementary Table 4.

Significant variants were predominantly associated with adhesin genes. Genes such as *yadK*, *ybgP*, *yfcS*, *flu*, and *ybgQ* exhibited association signals surpassing the significance thresholds at the whole-genome level. ORs derived from regression coefficients indicated stronger associations of *yadK* and *ybgP* with cattle isolates, while *flu*, *yfcS*, and *ybgQ* were more closely linked to human-associated strains (Figure 7B).

3.3.3 Genome wide association analysis using the adhesiome

We performed an association analysis using the complete adhesiome, acknowledging that not all adhesin genes are represented

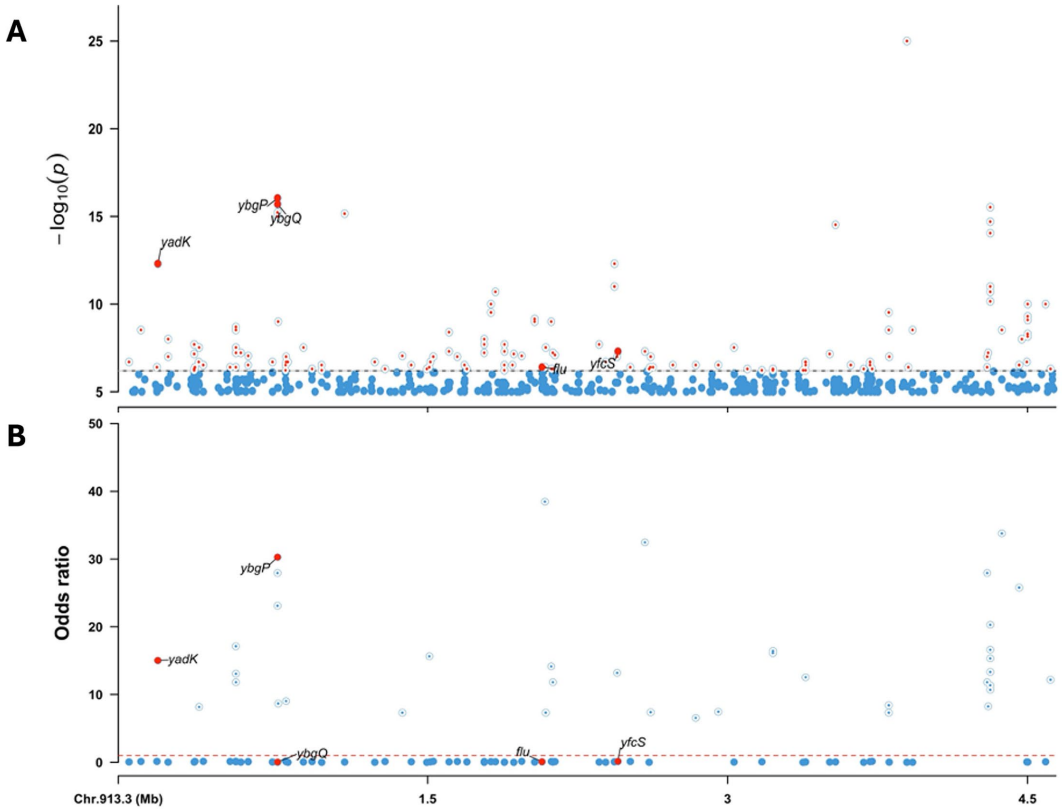

FIGURE 7
Genome-wide association analysis (GWAS) of human- and cattle-associated STEC strains using the *E. coli* K-12 genome as reference. **(A)** Manhattan plot showing the -log10(*p*) values from the association analysis. Significant variants in the adhesiome (e.g., *yadK*, *ybgP*, *ybgQ*, *flu*, *yfcS*) are indicated by large red dots. **(B)** Odds ratios (OR) derived from regression coefficients, with OR > 1 indicating association with cattle strains and OR < 1 indicating association with human strains. The dashed line at OR = 1 marks the neutral threshold. Together, these results highlight adhesiome genes significantly associated with host specificity.

in the *E. coli* K-12 reference genome. This analysis yielded similar results to the whole-genome association, with significant signals for genes shared between K-12 and the STEC adhesiome (*yadK*, *ybgP*, *ybgQ*, *flu*, and *yfcS*).

Additionally, other adhesin genes absent from the core K-12 genome showed specific associations with the cattle host. Variants in *yeeJ*, *espP*, and *fimC* exhibited the strongest statistical associations (lowest *p*-values) and the highest OR for the cattle host (Figure 8), suggesting roles in cattle-specific colonization and adaptation under host-driven selective pressures.

Conversely, although significant, *clpV*, *ybgQ*, and *sab* displayed ORs below one, indicating stronger associations with the human host. These genes exhibited low variant density, possibly reflecting conserved functions in host interaction or membrane integrity.

Overall, these results underscore distinct adaptation strategies shaped by selective pressures, promoting STEC colonization in cattle and highlighting adhesins as potential targets for intervention strategies.

3.3.4 Prediction of variant effects using the whole genome data and the adhesiome sequences

We used snpEff to predict the effects of variants identified in the GWAS analysis, utilizing the pre-compiled *E. coli* K-12 genome database. For the adhesiome sequences, a custom snpEff database was generated by annotating adhesin genes extracted from the adhesiome sequence data. Out of 433 adhesin sequences analyzed, 285 genes were successfully annotated.

The number of variants per gene was significantly higher in isolates from humans compared to cattle (4,780 vs. 3,507), suggesting a greater overall genetic diversity among human-associated STEC strains. However, the functional impact of these mutations was similar across hosts (Supplementary Table 6).

Notably, mutations with moderate-to-high predicted impacts were particularly enriched near loci associated with significant GWAS signals both in the K-12 genome and adhesiome sequences, suggesting potential adaptive selection for alternative alleles favoring host-specific colonization (Table 1). For example, an indel in *ybgQ* at position 749,777 (K-12 reference) introduced a frameshift mutation with a high predicted impact on protein function.

In cattle-associated strains, alternative alleles with moderate-to-high predicted impacts were often observed at near-fixation frequencies, particularly in *yadK*, *espP*, and *ybgP*. Conversely, some genes, such as *ybgP*, exhibited moderate-to-high impact variants with low reference allele frequencies in cattle, challenging detection in GWAS analyses.

Variants occurring at intermediate frequencies and exhibiting high linkage disequilibrium with moderate-to-high impact mutations were more likely to reach significance. See Table 1 for detailed information on variants located within 100 bp of significant GWAS signals.

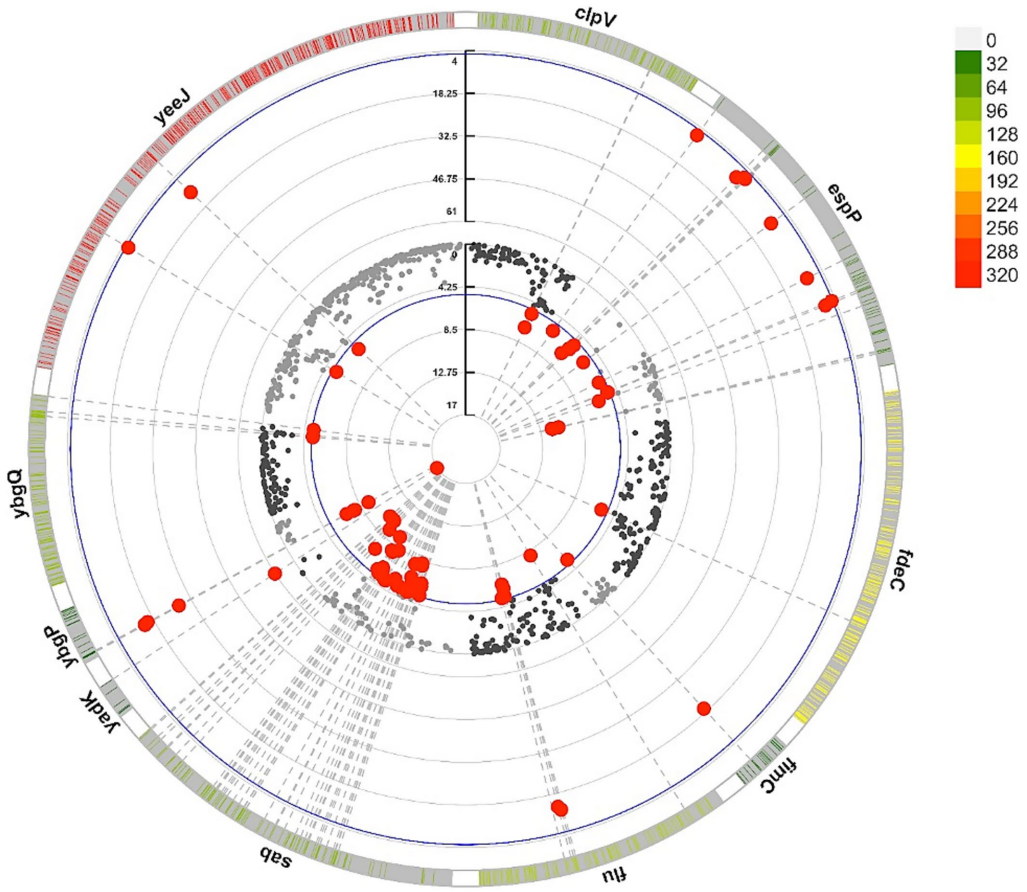

FIGURE 8
Circular representation of adhesin-associated variants in STEC genomes from cattle and human. The outermost layer shows the density of variants across adhesin genes with at least one variant significantly associated with the host, with higher color intensity indicating greater variant frequency. The middle and inner layers display the odds ratios (OR) and statistical significance (−log10(*p*)) from the GWAS, respectively, with red dots highlighting significant variants associated with cattle strains. All significant genes are listed in Supplementary Table 5.

4 Discussion

Commensal bovine-adapted *E. coli* strains are considered the evolutionary precursors of diarrheagenic pathotypes, including STEC (49), with cattle-associated STEC potentially acting as a bridge to human infection (49, 50). Upon transmission, STEC must adapt to new environments and host-specific factors such as diet, hygiene, and antimicrobial exposure, all of which may shape genomic evolution (51). Therefore, identifying molecular markers that differentiate cattle- and human-associated STEC strains is essential for understanding transmission dynamics and designing targeted preventive strategies aimed at reducing bacterial carriage in livestock.

4.1 STEC diversity

In this study, 71 STEC strains (7.5%) were identified among the collected stool samples. This detection rate is lower than in previous reports (27, 52, 53) and may reflect several factors, including improved farm biosecurity, differences in feeding practices, and environmental conditions. Seasonal and geographical variation in STEC prevalence has been widely documented, with higher recovery rates during spring and summer (54–56). Ambient temperature, rainfall, and vector abundance have been suggested as drivers of these seasonal trends, since warmer conditions may favor STEC persistence outside the host and increase exposure opportunities. Feeding practices can also influence bacterial shedding; animals fed forage typically shed fewer STEC than grain-fed cattle (22, 57). In our study, most samples were collected during winter and early spring, a period characterized by lower temperatures and high rainfall, which may reduce environmental persistence and transmission. Moreover, in the regions analyzed, most animals were pasture-fed. Together, these factors may at least in part explain the relatively low prevalence observed. Future longitudinal studies will be required to further elucidate these factors. Importantly, despite this lower detection rate, the strains recovered are consistent with those reported in previous studies (26, 27, 58) and are representative of the STEC populations currently circulating in Chilean cattle.

Among the isolates recovered, the majority carried *stx2* subtypes, which are strongly associated with severe human disease (59). From a global perspective, *stx2* was also the predominant subtype among all analyzed STEC genomes, reinforcing concerns about the threat posed by circulating STEC strains, as Stx2 exhibits both greater cytotoxic activity and higher affinity for host ribosomes (60).

10.3389/fvets.2025.1639243

TABLE 1 Description of the effect of variants significantly associated with the host phenotype in the GWAS analysis of the *E. coli* K-12 strain or when considering the adhesiome sequences, and respectively with their allele frequencies.

Gene	Position	−log10(value)	Reference allele freq. in cattle	Reference allele freq. in humans	Effect of the variant	Type of change	Moderate or high in regions near the actual significant variant (100 bp)	Origin
espP	88	5.8	0.13	0.84	Moderate	intergenic_region	Yes	Adhesiome
espP	849	6.9	0.12	0.86	Low	synonymous_variant	Yes	Adhesiome
espP	924	6.1	0.12	0.83	Low	synonymous_variant	Yes	Adhesiome
espP	938	6.1	0.12	0.83	Moderate	missense_variant	Yes	Adhesiome
espP	948	6.1	0.12	0.83	Moderate	missense_variant	Yes	Adhesiome
espP	963	5.5	0.04	0.51	Low	synonymous_variant	Yes	Adhesiome
espP	1,625	5.9	0.18	0.92	Moderate	missense_variant	Yes	Adhesiome
espP	2,512	5.6	0.19	0.92	Moderate	missense_variant	Yes	Adhesiome
espP	2,920	5.2	0.37	0.74	Moderate	missense_variant	Yes	Adhesiome
espP	2,945	5.2	0.40	0.89	Moderate	missense_variant	Yes	Adhesiome
fimC	459	5.5	0.79	0.99	Low	synonymous_variant	Yes	Adhesiome
flu	1,950	5.2	0.71	0.96	Moderate	missense_variant	Yes	Adhesiome
flu	1,983	5.3	0.71	0.98	Low	synonymous_variant	Yes	Adhesiome
yadK	462	17.0	NA	0.98	Moderate	intergenic_region	Yes	Adhesiome
ybgP	99	7.8	0.03	0.82	Low	synonymous_variant	Yes	Adhesiome
ybgP	105	7.6	0.03	0.82	Low	synonymous_variant	Yes	Adhesiome
ybgP	111	9.4	0.05	NA	Low	synonymous_variant	Yes	Adhesiome
ybgP	117	6.9	0.08	NA	Low	synonymous_variant	Yes	Adhesiome
yeeJ	1,827	5.4	0.61	0.84	Low	synonymous_variant	Yes	Adhesiome
yeeJ	2,919	5.8	0.70	0.96	Low	synonymous_variant	Yes	Adhesiome
flu	2,072,135	5.1	0.67	0.22	Low	synonymous_variant	Yes	K-12
yadK	151,138	12.3	0.01	0.96	Low	synonymous_variant	Yes	K-12
ybgP	749,591	15.0	0.01	0.88	Low	synonymous_variant	Yes	K-12
ybgP	749,597	16.1	0.01	0.88	Low	synonymous_variant	Yes	K-12
ybgQ	749,777	15.8	0.15	0.61	High	frameshift_variant	Yes	K-12
yfcS	2,451,806	7.3	0.88	0.90	Low	synonymous_variant	Yes	K-12

In Chile, the predominant STEC serotypes identified were O130:H11 (ST297), O185:H7 (ST2387), and O113:H21 (ST223), consistent with previous reports (26, 58), suggesting the persistence of specific strains within cattle reservoirs. In contrast, global data showed a higher prevalence of O157:H7 (ST11) and O26:H11 (ST21), along with O130:H11 (ST297). Our results underscore the considerable genomic variability in STs and serotypes among circulating STEC strains.

Notably, O130:H11 (ST297) predominated among cattle isolates, whereas O157:H7 (ST11) was more common among human isolates, consistent with global patterns (26, 61–64). O157:H7, which carries the LEE pathogenicity island and is strongly associated with the *stx2* gene, is the serotype most frequently linked to severe human disease (59, 65). Conversely, O130:H11, a LEE-negative serotype, has been implicated in sporadic outbreaks (66–68). Despite lacking the LEE locus, O130:H11 strains harbor virulence factors such as *ehxA*, *saa*, *sab*, *lpfA*, and *iha*, which may contribute to severe human disease (62, 69, 70).

Although seropathotype classification (based on ST or serotype) has historically been employed to assess the association of STEC strains with HUS and human outbreaks (71), its reliability is limited by the high genetic plasticity of STEC. The distribution of diverse serotypes across regions and hosts complicates epidemiological surveillance and strain tracking (72). Given these challenges, prioritizing molecular markers related to virulence and host adaptation over traditional serotyping could improve monitoring protocols and control strategies. This strategy would strengthen epidemiological surveillance, veterinary health measures, and public health interventions, ultimately enabling more effective and targeted control of STEC.

The high genomic variability of STEC facilitates its adaptation to diverse hosts through mechanisms such as recombination and horizontal gene transfer, impacting both virulence and transmission dynamics (49, 73, 74). However, despite this genetic plasticity, the existence of genetically homogeneous clades within specific STs suggests that certain STEC lineages have undergone niche specialization, likely driven by host- or environment-specific selective pressures. For example, at the bovine RAJ, selective pressures such as predation by bacterivorous protozoa may have favored strains encoding Stx and LEE genes, which inhibit protozoal grazing and promote persistence (50).

4.2 Adhesiome analyses

The evolutionary strategies of STEC may account for the greater adhesin diversity observed in human-associated strains compared to cattle-associated strains (49). Our analysis, combined with the AdhesiomeR cluster classification, confirmed that *ehaA*, *ehaG*, *stgA-C*, *yadL-N*, and *iha* were significantly more prevalent in cattle strains. EhaA, an autotransporter protein, facilitates rapid cell aggregation, biofilm formation, and adhesion to bovine RAJ epithelial cells when overexpressed in *E. coli* K-12 (19). Similarly, EhaG enhances autoaggregation, biofilm formation, and binding to collagens I–V, laminin, fibronectin, and fibrinogen (75). Although Stg adhesins have been implicated in adhesion to human and avian epithelial cells (76), their specific role in STEC gut colonization remains unclear. The Yad fimbrial autotransporter family modulates gene expression and virulence in STEC O157:H7 (77), potentially contributing to RAJ colonization (78). Additionally, Iha, a dual-function adhesin and siderophore receptor, is widely distributed among LEE-positive and LEE-negative STEC strains, and is often carried on mobile genetic elements, facilitating horizontal dissemination (11, 79, 80). Collectively, these findings suggest that cattle-adapted STEC strains rely on a relatively conserved adhesin repertoire optimized for stable colonization, biofilm formation, extracellular matrix adhesion, and iron acquisition.

In contrast, human-associated STEC strains exhibited greater diversity in adhesins, including fimbrial and non-fimbrial types, suggesting a more flexible colonization strategy adapted to heterogeneous environments. These strains showed a significantly higher representation of *eae*, *cah*, *ypjA*, and *paa* genes. Eae is a key virulence factor in LEE-positive STEC, leading to attaching and effacing lesions (81). Similarly, *cah* (calcium-binding antigen 43 homologous) is predominantly found in LEE-positive strains (82), promoting autoaggregation, biofilm formation, and persistence in STEC O157:H7 (81, 83). YpjA (homologue of EhaD) has been shown to enhance biofilm formation in STEC O157:H7, reinforcing its contribution to intestinal colonization (19). Meanwhile, *paa* (porcine attaching and effacing-associated protein) has been associated with *eae*-positive strains from both animals and human origins, suggesting a synergistic role with intimin-mediated adhesion (49, 84). These findings indicate that human-associated STEC strains rely on a broader adhesin repertoire than cattle strains, which may reflect adaptation to a more variable intestinal environment and stronger immune pressures.

Differences in adhesin gene abundance between cattle- and human-associated STEC strains align with distinct functional enrichments revealed by gene ontology analysis, supporting the notion of host-specific adaptation strategies. In cattle-associated strains, enriched biological processes point to strong selection for biofilm formation and attachment mechanisms, particularly involving fimbrial and non-fimbrial adhesins (85), autotransporter adhesins such as EhaA (86), chaperone-usher adhesins like Yad (87), and the adhesin-siderophore Iha (79). Conversely, human-associated strains exhibit significant enrichment in methylglyoxal detoxification, a response to a toxic glycolysis byproduct (88), potentially indicating a metabolic shift toward glyoxylate cycle activation. This finding suggests a potential metabolic shift toward glyoxylate cycle activation, favoring carbon conservation and survival under nutrient-limited conditions, typical of the human intestine (89). Moreover, methylglyoxal detoxification mechanisms also contribute to the neutralization of reactive oxygen species (ROS), enhancing bacterial survival against host immune defenses (90). Altogether, these observations indicate that adaptation to oxidative and metabolic stress represents a key selective force shaping the genomic and functional landscape of human-associated STEC strains.

4.3 GWAS

GWAS represent powerful tools for identifying genetic variants associated with bacterial adaptations, including adhesion, by linking genotype to phenotype while accounting for confounding factors such as population structure (91). Applying GWAS to pathogenic bacteria improves our understanding of virulence mechanisms and host-pathogen interactions; however, relatively few studies have explored the

population structure and adaptation of STEC using this approach. Some investigations have focused on associating genomic features with pathogenic properties in STEC. For instance, Matussek et al. (92) analyzed 238 STEC genomes from patients with and without HUS to identify genetic predictors of disease severity. Their integrative approach, combining serotyping, *stx* subtyping, virulence profiling, phylogenomics, and a pangenome-wide association study (PWAS), showed that O157:H7 clade 8 strains and *stx2a* or *stx2a* + *stx2c* subtypes were strongly associated with HUS, while *stx1a* was more frequent in non-HUS cases. Virulence genes related to adherence (*eae*, *tir*, *paa*), toxins (*toxB*, *astA*), and type III secretion system proteins were enriched in HUS isolates, and hundreds of accessory genes were linked to severe disease, including adhesins (*yfcP*, *yehD*, *elfG*, *sfmA*) and regulators, although many encoded hypothetical proteins. The authors concluded that severe human disease results from the interplay between canonical virulence determinants, accessory genetic elements, and host–pathogen interactions. Similarly, Peroutka-Bigus et al. (93) compared genomic and phenotypic features of three human outbreak-associated and one cattle-derived STEC O157:H7 isolates to assess host adaptation. Despite differences in virulence gene expression, adherence, and Stx production among outbreak isolates, no significant differences were detected in cattle colonization or shedding compared with the cattle-associated strain. This highlights that genomic and phenotypic variation in STEC O157:H7 does not necessarily correspond to host-specific adaptation, emphasizing that host specificity cannot be inferred solely from genetic or phenotypic traits. More recently, Espadinha et al. (94) performed a PWAS of 531 STEC isolates, identifying associations between the development of HUS and the presence of *stx2a*, *stx1a* + *stx2a*, or *stx1a* + *stx2c*, as well as the co-occurrence of genes such as *ygiW* (stress-induced protein) with group_5720 (transcriptional regulation) and *pfkA* (6-phosphofructokinase-1) with *fieF* (Zn^{2+}/Fe^{2+}/Cd^{2+} efflux transporter), among other epidemiological factors. In the same context, Marques Da Silva et al. (95) explored the genomic determinants of cattle colonization by comparing the genomes of STEC O22:H8 and O157:H7 strains. They identified 28 virulence-associated genes unique to O22:H8, primarily involved in adherence (e.g., *cfaA*, *cfaB*, *cfaC*, *cfaD*/*cfaE*, *sisA*, *lesP*, *hes*, *pagC*, *tpsA*, and *tpsB*), autotransporters (*ag43*), and invasion (*tia*). These findings highlight the complexity of adhesin gene distribution and function in STEC, reinforcing the need for further research to elucidate the genetic basis of host-specific colonization.

Beyond the host-specific differences between cattle- and human-associated STEC strains, our GWAS comparing these isolates with the non-pathogenic *E. coli* K-12 genome provides additional evidence for host-driven selection, highlighting potential adaptive traits that distinguish STEC from commensal *E. coli* strains. This analysis revealed significant enrichment of trehalose transport and protein-phosphocysteine-trehalose phosphotransferase system activities, associated with the *bglF* and *ascF* genes. It is well established that *E. coli* can use trehalose as an alternative carbon source and synthesize it intracellularly to counteract osmotic stress by stabilizing membrane integrity (96). The BglF and AscF proteins, key components of the phosphotransferase system, mediate the transport and phosphorylation of β-glucosides, including trehalose, thus regulating its metabolism and contributing to the bacterial stress response (97). These metabolic adaptations may enhance the resilience of STEC strains to osmotic and environmental stresses encountered in the bovine gastrointestinal environment.

Our GWAS comparing the adhesiome of cattle- and human-associated STEC with *E. coli* K-12 revealed distinct host-specific adhesion strategies. In cattle strains, *yadK* and *ybgP* were significantly associated. YadK, a chaperone-usher adhesin, enhances acid stress resistance, biofilm formation, and epithelial attachment, promoting persistence in the bovine gastrointestinal tract (98, 99). YbgP facilitates adhesion to epithelial and abiotic surfaces, promoting environmental persistence (87, 100). In contrast, human-associated strains exhibited significant enrichment of *ybgQ*, *yfcS*, and *flu*. *ybgQ* encodes the usher protein necessary for YbgP fimbriae assembly, suggesting a role in epithelial colonization (101). YfcS enhances biofilm formation and bacterial aggregation (87, 100), while *flu* (also known as *agn43*), an autotransporter adhesin, promotes microcolony formation, biofilm stability, and immune evasion in both LEE-positive and LEE-negative STEC strains (11, 102). These findings suggest that cattle-associated STEC strains prioritize adhesion mechanisms suited for long-term gut colonization, whereas human-associated strains exhibit a broader adhesin repertoire, likely reflecting selective pressures favoring host invasion, immune evasion, and adaptation to the intestinal niche.

Notable differences emerged when analyzing the complete adhesiome of cattle- and human-associated STEC strains compared to the analysis based solely on the *E. coli* K-12 genome. In cattle-associated strains, *yeeJ*, *espP*, and *fimC* exhibited the highest significance values and OR, suggesting a prominent role in host-specific colonization and adaptation, likely driven by selective pressures within the bovine gastrointestinal tract. *yeeJ* encodes an autotransporter protein with structural similarity to intimin, facilitating biofilm formation and enhancing STEC O157:H7 binding to eukaryotic cells (103, 104). EspP, a serine protease autotransporter of *Enterobacteriaceae* family, encoded on STEC virulence plasmids, promotes adhesion to bovine rectal epithelial cells, intestinal colonization, and supports biofilm formation and HeLa cell adherence (105, 106). Its proteolytic activity targets extracellular matrix and mucus proteins, facilitating tissue penetration and receptor exposure for other adhesins. *fimC*, a key component of the *fim* operon, plays a critical role in the assembly of type 1 fimbriae, which promote epithelial adhesion and biofilm formation in *E. coli*. In STEC O157:H7, type 1 fimbriae have been implicated in RAJ cell colonization, highlighting their potential contribution to cattle adaptation and long-term persistence (107, 108).

Conversely, *clpV*, *ybgQ*, and *sab* exhibited a highly significant OR below one, suggesting their association with human-host adaptation. ClpV, a cytosolic ATPase and essential component of the type VI secretion system (T6SS), facilitates bacterial competition by delivering effector proteins—such as peptidoglycan hydrolases, phospholipases, and DNases—into target cells (109, 110). In the STEC O157:H7 strain EDL933, ClpV mediates the translocation of catalase into macrophages, promoting immune evasion; notably, deletion of *clpV* reduces lethality in murine infection models (111). Furthermore, ClpV has been associated with HUS-producing STEC strains, supporting its potential role in virulence (112, 113), although its precise contribution to intestinal colonization remains unclear. Similarly, Sab, a plasmid-encoded autotransporter, enhances adherence to human epithelial cells and biofilm formation in LEE-negative STEC strains (114), potentially facilitating intestinal colonization and persistence.

In addition to the gene presence/absence patterns identified through GWAS, variant analysis offered deeper insights into

host-specific selective pressures shaping adhesin functionality. Human-associated STEC exhibited greater genetic diversity, although the predicted functional impacts on adhesins were largely conserved across hosts. A key observation was the presence of SNPs with moderate-to-high effects on protein sequences. In particular, given *ybgQ*'s role in adhesion and outer membrane integrity, structural disruptions caused by SNPs may significantly affect bacterial adherence and survival within the human gut. Conversely, cattle-associated strains exhibited alternative mutations with moderate-to-high effects, especially in *yadK*, *espP*, and *ybgP*, where near-fixation frequencies suggest strong host-driven selection favoring these variants for bovine colonization. Notably, different *espP* alleles exhibit distinct biological activities: EspPα and EspPγ are secreted and enzymatically active, whereas EspPβ and EspPδ display reduced or absent proteolytic function (115). Among them, EspPγ specifically cleaves pepsin and coagulation factor V in humans (116), while EspPα is more frequently found in human isolates. In contrast, other EspP variants are predominantly associated with animal reservoirs and environmental sources (117). These host-driven selective pressures likely induce functional modifications that may either enhance or attenuate virulence. Further experimental studies, including *in vitro* and *in vivo* infection models, are necessary to clarify the impacts of these alternative variants on colonization efficiency and adhesin functionality.

A potential limitation of our GWAS is the uneven geographic distribution of publicly available STEC genomes, with an over-representation of isolates from countries such as Chile, Germany, and France. However, our analyses were focused on host origin rather than country of isolation, and GWAS methods corrected for population stratification were applied to mitigate this potential bias. In addition, the dataset encompassed diverse serotypes and lineages, which supports the robustness of the host-specific associations identified.

Targeting adhesins in STEC offers a promising avenue for developing effective intervention strategies. Given the essential role of adhesins in host colonization, strategies such as anti-adhesin antibodies, competitive inhibitors, or adhesin-based vaccines could significantly reduce bacterial adherence and gut colonization, ultimately lowering transmission risk. These approaches could be particularly beneficial in pre-harvest control programs designed to decrease STEC carriage in cattle. Future research should focus on identifying adhesins with high conservation across STEC strains and evaluating their potential as preventive targets, including efficacy assessments through animal model studies.

5 Conclusion

This study provides new insights into the genetic diversity and functional roles of adhesins in STEC host adaptation. We identified adhesin-encoding genes strongly associated with cattle strains (e.g., *yadK*, *espP*, *fimC*) and others in human-associated strains (*clpV*, *ybgQ*, *sab*) whose presence and polymorphisms may reflect host-specific selective pressures. Importantly, these genes are not only markers of host origin but also have potential functional implications, where cattle-associated strains may display enhanced colonization and persistence at the bovine RAJ, while human-associated strains may have facilitated immune evasion and adhesion to epithelial cells, increasing their virulence in the human host. However, their precise roles in host colonization require further investigation, particularly to evaluate their potential as intervention targets. Future functional studies, including adhesion assays and transcriptomic profiling under host-mimicking conditions, will be crucial to elucidate their contribution to bacterial fitness, biofilm dynamics, and immune evasion strategies. Moreover, variant analysis revealed greater genetic diversity in adhesin genes among human-associated strains, although functional effects remained comparable across hosts, suggesting selective constraints that preserve key adhesion mechanisms. High-impact mutations should be further examined through protein structure modeling and functional assays to assess their influence on adhesion efficiency and pathogenicity. Collectively, such studies may contribute to the identification of conserved adhesin targets for vaccine development and other intervention strategies aimed at reducing STEC transmission at the livestock level.

Data availability statement

The whole-genome sequences of STEC strains that support the findings of this study are openly available in GenBank at https://www.ncbi.nlm.nih.gov/genbank/, reference number PRJNA656305.

Ethics statement

The animal studies were approved by Comité Institucional de Cuidado y Uso de Animales of the Universidad de Chile. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

VM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JC: Data curation, Investigation, Methodology, Writing – original draft. EM: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. JD: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. DM: Formal analysis, Investigation, Methodology, Writing – original draft. GA: Formal analysis, Investigation, Writing – original draft. JT: Investigation, Methodology, Writing – original draft. RA: Investigation, Resources, Writing – original draft. NP: Investigation, Resources, Writing – original draft. DC: Investigation, Resources, Writing – original draft. DL: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. RR: Formal analysis, Investigation, Methodology, Writing – original draft. JJ: Data curation, Formal analysis, Investigation, Writing – original draft. BE: Formal analysis, Investigation, Writing – original draft. IK: Data curation, Investigation, Visualization, Writing – review & editing. NG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project

administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Grant Number 1230776. IK was supported by the USDA-ARS CRIS Project 5030-32000-225-00D.

Acknowledgments

We sincerely thank Dr. Juan Carlos Hormazábal from the Instituto de Salud Pública de Chile for kindly providing the human-origin STEC isolates analyzed in this study. We also extend our gratitude to Dr. Luis Altamirano, and the abattoir professionals for their invaluable collaboration in sample collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets.2025.1639243/full#supplementary-material

References

1. Rivas M, Chinen I, Guth BEC. Enterohemorrhagic (Shiga toxin-producing) *Escherichia coli* In: AG Torres, editor. *Escherichia coli* in the Americas. Cham: Springer (2016). 97–123.

2. Cody EM, Dixon BP. Hemolytic uremic syndrome. *Pediatr Clin*. (2019) 66:235–46. doi: 10.1016/j.pcl.2018.09.011

3. Kintz E, Brainard J, Hooper L, Hunter P. Transmission pathways for sporadic Shiga-toxin producing *E. coli* infections: a systematic review and meta-analysis. *Int J Hyg Environ Health*. (2017) 220:57–67. doi: 10.1016/j.ijheh.2016.10.011

4. Fernández D, Sanz ME, Parma AE, Padola NL. Characterization of Shiga toxin-producing *Escherichia coli* isolated from newborn, milk-fed, and growing calves in Argentina. *J Dairy Sci*. (2012) 95:5340–3. doi: 10.3168/jds.2011-5140

5. Gonzalez AGM, Cerqueira AMF. Shiga toxin-producing *Escherichia coli* in the animal reservoir and food in Brazil. *J Appl Microbiol*. (2020) 128:1568–82. doi: 10.1111/jam.14500

6. Economic Research Service. (2025). Cost estimates of foodborne illnesses. Available online at: https://www.ers.usda.gov/data-products/cost-estimates-of-foodborne-illnesses. (Accessed March 16, 2025)

7. McWilliams BD, Torres AG. Enterohemorrhagic *Escherichia coli* adhesins In: Enterohemorrhagic *Escherichia coli* and other Shiga toxin-producing *E. coli*. Hoboken, NJ: John Wiley & Sons (2015). 131–55.

8. Stevens MP, Frankel GM. The locus of enterocyte effacement and associated virulence factors of enterohemorrhagic *Escherichia coli* In: Enterohemorrhagic *Escherichia coli* and other Shiga toxin-Producing *E. coli*. Hoboken, NJ: John Wiley & Sons (2015). 97–130.

9. CDC. (2025). BEAM dashboard. Available online at: https://www.cdc.gov/ncezid/dfwed/BEAM-dashboard.html. (Accessed March 16, 2025)

10. Schmidt H, Zhang W-L, Hemmrich U, Jelacic S, Brunder W, Tarr PI, et al. Identification and characterization of a novel genomic island integrated at selC in locus of enterocyte effacement-negative, Shiga toxin-producing *Escherichia coli*. *Infect Immun*. (2001) 69:6863–73. doi: 10.1128/iai.69.11.6863-6873.2001

11. Montero DA, Velasco J, Del Canto F, Puente JL, Padola NL, Rasko DA, et al. Locus of adhesion and autoaggregation (LAA), a pathogenicity island present in emerging Shiga toxin-producing *Escherichia coli* strains. *Sci Rep*. (2017) 7:7011. doi: 10.1038/s41598-017-06999-y

12. Sidorczuk K, Burdukiewicz M, Cerk K, Fritscher J, Kingsley RA, Schierack P, et al. adhesiomeR: a tool for *Escherichia coli* adhesin classification and analysis. *BMC Genomics*. (2024) 25:609. doi: 10.1186/s12864-024-10525-6

13. Besser TE, Richards BL, Rice DH, Hancock DD. *Escherichia coli* O157:H7 infection of calves: infectious dose and direct contact transmission. *Epidemiol Infect*. (2001) 127:555–60. doi: 10.1017/S095026880100615X

14. Fairbrother JM, Nadeau É. Colibacillosis In: JJ Zimmerman, LA Karriker, A Ramirez, KJ Schwartz, GW Stevenson and J Zhang, editors. Diseases of swine. *11th* ed. Hoboken, NJ: John Wiley & Sons (2019). 807–34.

15. Naylor SW, Low JC, Besser TE, Mahajan A, Gunn GJ, Pearce MC, et al. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. *Infect Immun*. (2003) 71:1505–12. doi: 10.1128/iai.71.3.1505-1512.2003

16. Mir RA, Brunelle BW, Alt DP, Arthur TM, Kudva IT. Supershed *Escherichia coli* O157:H7 has potential for increased persistence on the rectoanal junction squamous epithelial cells and antibiotic resistance. *Int J Microbiol*. (2020) 2020:2368154. doi: 10.1155/2020/2368154

17. Dziva F, van Diemen PM, Stevens MP, Smith AJ, Wallis TS. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. *Microbiology*. (2004) 150:3631–45. doi: 10.1099/mic.0.27448-0

18. Naylor SW, Roe AJ, Nart P, Spears K, Smith DGE, Low JC, et al. *Escherichia coli* O157:H7 forms attaching and effacing lesions at the terminal rectum of cattle and colonization requires the LEE4 operon. *Microbiology*. (2005) 151:2773–81. doi: 10.1099/mic.0.28060-0

19. Wells TJ, Sherlock O, Rivas L, Mahajan A, Beatson SA, Torpdahl M, et al. EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ Microbiol*. (2008) 10:589–604. doi: 10.1111/j.1462-2920.2007.01479.x

20. Mahajan A, Currie CG, Mackie S, Tree J, McAteer S, McKendrick I, et al. An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157:H7 with bovine intestinal epithelium. *Cell Microbiol*. (2009) 11:121–37. doi: 10.1111/j.1462-5822.2008.01244.x

21. Kudva IT, Hovde CJ, John M. Adherence of non-O157 Shiga toxin-producing *Escherichia coli* to bovine recto-anal junction squamous epithelial cells appears to

be mediated by mechanisms distinct from those used by O157. *Foodborne Pathog Dis*. (2013) 10:375–81. doi: 10.1089/fpd.2012.1382

22. Sapountzis P, Segura A, Desvaux M, Forano E. An overview of the elusive passenger in the gastrointestinal tract of cattle: the Shiga toxin producing *Escherichia coli*. *Microorganisms*. (2020) 8:877. doi: 10.3390/microorganisms8060877

23. Starks CM, Miller MM, Broglie PM, Cubbison J, Martin SM, Eldridge GR. Optimization and qualification of an assay that demonstrates that a FimH vaccine induces functional antibody responses in women with histories of urinary tract infections. *Hum Vaccin Immunother*. (2021) 17:283–92. doi: 10.1080/21645515.2020.1770034

24. Gadar K, McCarthy RR. Using next generation antimicrobials to target the mechanisms of infection. *npj Antimicrob Resist*. (2023) 1:1–14. doi: 10.1038/s44259-023-00011-6

25. Instituto Nacional de Estadísticas. (2025). Censo Agropecuario. Available online at: http://www.ine.gob.cl/estadisticas/economia/agricultura-agroindustria-y-pesca/censos-agropecuarios. (Accessed March 16, 2025)

26. Galarce N, Sánchez F, Escobar B, Lapierre L, Cornejo J, Alegría-Morán R, et al. Genomic epidemiology of Shiga toxin-producing *Escherichia coli* isolated from the livestock-food-human interface in South America. *Animals*. (2021) 11:1845. doi: 10.3390/ani11071845

27. Galarce N, Escobar B, Sánchez F, Paredes-Osses E, Alegría-Morán R, Borie C. Virulence genes, Shiga toxin subtypes, serogroups, and clonal relationship of Shiga toxin-producing *Escherichia coli* strains isolated from livestock and companion animals. *Animals*. (2019) 9:733. doi: 10.3390/ani9100733

28. Cebula TA, Payne WL, Feng P. Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. *J Clin Microbiol*. (1995) 33:248–50. doi: 10.1128/jcm.33.1.248-250.1995

29. Borie CF, Monreal Z, Martinez J, Arellano C, Prado V. Detection and characterization of enterohaemorrhagic *Escherichia coli* in slaughtered cattle. *J Vet Med B*. (1997) 44:273–9. doi: 10.1111/j.1439-0450.1997.tb00973.x

30. Arancia S, Babsa S, Brambilla G, Chiani P, Ferreri C, Galati F, et al. (2017). Report of the 19th Inter-Laboratory Study on the Detection of Shiga toxin-producing *E. coli* (STEC) in Sprout Spent Irrigation Water (PT19). Rome, Italy: Istituto Superiore di Sanità.

31. Jenkins C, Perry NT, Godbole G, Gharbia S. Evaluation of chromogenic selective agar (CHROMagar STEC) for the direct detection of Shiga toxin-producing *Escherichia coli* from faecal specimens. *J Med Microbiol*. (2020) 69:487–91. doi: 10.1099/jmm.0.001136

32. Chen J, Griffiths MW. PCR differentiation of *Escherichia Coli* from other gram-negative bacteria using primers derived from the nucleotide sequences flanking the gene encoding the universal stress protein. *Lett Appl Microbiol*. (1998) 27:369–71. doi: 10.1046/j.1472-765X.1998.00445.x

33. Jung Y, Han D. BWA-MEME: BWA-MEM emulated with a machine learning approach. *Bioinformatics*. (2022) 38:2404–13. doi: 10.1093/bioinformatics/btac137

34. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. (2000) 16:276–7. doi: 10.1016/S0168-9525(00)02024-2

35. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. (2012) 19:455–77. doi: 10.1089/cmb.2012.0021

36. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods*. (2023) 20:1203–12. doi: 10.1038/s41592-023-01940-w

37. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol*. (2015) 53:2410–26. doi: 10.1128/jcm.00008-15

38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. (2009) 25:2078–9. doi: 10.1093/bioinformatics/btp352

39. Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics*. (2011) 27:1691–2. doi: 10.1093/bioinformatics/btr174

40. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. (2015) 31:2032–4. doi: 10.1093/bioinformatics/btv098

41. Garrison E, Marth G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*. Available online at: https://doi.org/10.48550/arXiv.1207.3907. [Epub ahead of preprint]

42. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. (2011) 27:2156–8. doi: 10.1093/bioinformatics/btr330

43. Pérez-Enciso M, de los Campos G, Hudson N, Kijas J, Reverter A. The 'heritability' of domestication and its functional partitioning in the pig. *Heredity*. (2017) 118:160–8. doi: 10.1038/hdy.2016.78

44. Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. Pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics*. (2018) 34:4310–2. doi: 10.1093/bioinformatics/bty539

45. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. (2007) 81:559–75. doi: 10.1086/519795

46. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. (2016) 17:132. doi: 10.1186/s13059-016-0997-x

47. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*. (2012) 6:80–92. doi: 10.4161/fly.19695

48. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res*. (2004) 32:W309–12. doi: 10.1093/nar/gkh379

49. Arimizu Y, Kirino Y, Sato MP, Uno K, Sato T, Gotoh Y, et al. Large-scale genome analysis of bovine commensal *Escherichia coli* reveals that bovine-adapted *E. coli* lineages are serving as evolutionary sources of the emergence of human intestinal pathogenic strains. *Genome Res*. (2019) 29:1495–505. doi: 10.1101/gr.249268.119

50. Koudelka GB, Arnold JW, Chakraborty D. Evolution of STEC virulence: insights from the antipredator activities of Shiga toxin producing *E. coli*. *Int J Med Microbiol*. (2018) 308:956–61. doi: 10.1016/j.ijmm.2018.07.001

51. Moraïs S, Winkler S, Zorea A, Levin L, Nagies FSP, Kapust N, et al. Cryptic diversity of cellulose-degrading gut bacteria in industrialized humans. *Science*. (2024) 383:eadj9223. doi: 10.1126/science.adj9223

52. Vidal R, Corvalán L, Vivanco S. Caracterización de cepas de *Escherichia coli* productor de Shigatoxina (STEC) aisladas desde cerdos y bovinos sanos, faenados en la Región Metropolitana. *Avances Cienc Vet*. (2012) 27:41.

53. Díaz L, Gutierrez S, Moreno-Switt AI, Hervé LP, Hamilton-West C, Padola NL, et al. Diversity of non-O157 Shiga toxin-producing *Escherichia coli* isolated from cattle from central and southern Chile. *Animals*. (2021) 11:2388. doi: 10.3390/ani11082388

54. Fernández D, Rodríguez EM, Arroyo GH, Padola NL, Parma AE. Seasonal variation of Shiga toxin-encoding genes (*stx*) and detection of *E. coli* O157 in dairy cattle from Argentina. *J Appl Microbiol*. (2009) 106:1260–7. doi: 10.1111/j.1365-2672.2008.04088.x

55. Cobbold R, Desmarchelier P. A longitudinal study of Shiga-toxigenic *Escherichia coli* (STEC) prevalence in three Australian dairy herds. *Vet Microbiol*. (2000) 71:125–37. doi: 10.1016/S0378-1135(99)00173-X

56. Stanford K, Johnson RP, Alexander TW, McAllister TA, Reuter T. Influence of season and feedlot location on prevalence and virulence factors of seven serogroups of *Escherichia coli* in feces of Western-Canadian slaughter cattle. *PLoS One*. (2016) 11:e0159866. doi: 10.1371/journal.pone.0159866

57. Jacob ME, Callaway TR, Nagaraja TG. Dietary interactions and interventions affecting *Escherichia coli* O157 colonization and shedding in cattle. *Foodborne Pathog Dis*. (2009) 6:785–92. doi: 10.1089/fpd.2009.0306

58. Gutiérrez S, Díaz L, Reyes-Jara A, Yang X, Meng J, González-Escalona N, et al. Whole-genome phylogenetic analysis reveals a wide diversity of non-O157 STEC isolated from ground beef and cattle feces. *Front Microbiol*. (2021) 11:622663. doi: 10.3389/fmicb.2020.622663

59. Fruth A, Lang C, Größl T, Garn T, Flieger A. Genomic surveillance of STEC/EHEC infections in Germany 2020 to 2022 permits insight into virulence gene profiles and novel O-antigen gene clusters. *Int J Med Microbiol*. (2024) 314:151610. doi: 10.1016/j.ijmm.2024.151610

60. Basu D, Li X-P, Kahn JN, May KL, Kahn PC, Tumer NE. The A1 subunit of Shiga toxin 2 has higher affinity for ribosomes and higher catalytic activity than the A1 subunit of Shiga toxin 1. *Infect Immun*. (2015) 84:149–61. doi: 10.1128/iai.00994-15

61. Masana MO, D'Astek BA, Palladino PM, Galli L, Del Castillo LL, Carbonari C, et al. Genotypic characterization of non-O157 Shiga toxin-producing *Escherichia coli* in beef abattoirs of Argentina. *J Food Prot*. (2011) 74:2008–17. doi: 10.4315/0362-028X.JFP-11-189

62. Fernández D, Krüger A, Polifroni R, Bustamante A, Sanso AM, Etcheverría AI, et al. Characterization of Shiga toxin-producing *Escherichia coli* O130:H11 and O178:H19 isolated from dairy cows. *Front Cell Infect Microbiol*. (2013) 3:9. doi: 10.3389/fcimb.2013.00009

63. Torres AG, Amaral MM, Bentancor L, Galli L, Goldstein J, Krüger A, et al. Recent advances in Shiga toxin-producing *Escherichia coli* research in Latin America. *Microorganisms*. (2018) 6:100. doi: 10.3390/microorganisms6040100

64. Mussio P, Martínez I, Luzardo S, Navarro A, Leotta G, Varela G. Phenotypic and genotypic characterization of Shiga toxin-producing *Escherichia coli* strains recovered from bovine carcasses in Uruguay. *Front Microbiol*. (2023) 14:1130170. doi: 10.3389/fmicb.2023.1130170

65. Byrne L, Adams N, Jenkins C. Association between Shiga toxin-producing *Escherichia coli* O157:H7 *stx* gene subtype and disease severity, England, 2009–2019. *Emerg Infect Dis*. (2020) 26:2394–400. doi: 10.3201/eid2610.200319

66. Elliott EJ, Robins-Browne RM, O'Loughlin EV, Bennett-Wood V, Bourke J, Henning P, et al. Nationwide study of haemolytic uraemic syndrome: clinical, microbiological, and epidemiological features. *Arch Dis Child*. (2001) 85:125–31. doi: 10.1136/adc.85.2.125

67. Galli L, Miliwebsky E, Irino K, Leotta G, Rivas M. Virulence profile comparison between LEE-negative Shiga toxin-producing *Escherichia coli* (STEC) strains isolated from cattle and humans. *Vet Microbiol*. (2010) 143:307–13. doi: 10.1016/j.vetmic.2009.11.028

68. Jinnerot T, Tomaselli ATP, Johannessen GS, Söderlund R, Urdahl AM, Aspán A, et al. The prevalence and genomic context of Shiga toxin 2a genes in *E. coli* found in cattle. *PLoS One*. (2020) 15:e0232305. doi: 10.1371/journal.pone.0232305

69. Shen J, Zhi S, Guo D, Jiang Y, Xu X, Zhao L, et al. Prevalence, antimicrobial resistance, and whole genome sequencing analysis of Shiga toxin-producing *Escherichia coli* (STEC) and Enteropathogenic *Escherichia coli* (EPEC) from imported foods in China during 2015–2021. *Toxins*. (2022) 14:68. doi: 10.3390/toxins14020068

70. Osek J, Wieczorek K. Isolation and molecular characterization of Shiga toxin-producing *Escherichia coli* (STEC) from bovine and porcine carcasses in Poland during 2019–2023 and comparison with strains from years 2014–2018. *Int J Food Microbiol*. (2025) 428:110983. doi: 10.1016/j.ijfoodmicro.2024.110983

71. Karmali MA, Mascarenhas M, Shen S, Ziebell K, Johnson S, Reid-Smith R, et al. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J Clin Microbiol*. (2003) 41:4930–40. doi: 10.1128/jcm.41.11.4930-4940.2003

72. Franz E, Delaquis P, Morabito S, Beutin L, Gobius K, Rasko DA, et al. Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing *Escherichia coli* (STEC) in global food production systems. *Int J Food Microbiol*. (2014) 187:57–72. doi: 10.1016/j.ijfoodmicro.2014.07.002

73. Coombes BK, Gilmour MW, Goodman CD. The evolution of virulence in non-O157 Shiga toxin-producing *Escherichia coli*. *Front Microbiol*. (2011) 2:90. doi: 10.3389/fmicb.2011.00090

74. Blankenship HM, Mosci RE, Dietrich S, Burgess E, Wholehan J, McWilliams K, et al. Population structure and genetic diversity of non-O157 Shiga toxin-producing *Escherichia coli* (STEC) clinical isolates from Michigan. *Sci Rep*. (2021) 11:4461. doi: 10.1038/s41598-021-83775-z

75. Totsika M, Wells TJ, Beloin C, Valle J, Allsopp LP, King NP, et al. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl Environ Microbiol*. (2012) 78:2179–89. doi: 10.1128/AEM.06680-11

76. Lymberopoulos MH, Houle S, Daigle F, Léveillé S, Brée A, Moulin-Schouleur M, et al. Characterization of stg fimbriae from an avian pathogenic *Escherichia coli* O78:K80 strain and assessment of their contribution to colonization of the chicken respiratory tract. *J Bacteriol*. (2006) 188:6449–59. doi: 10.1128/jb.00453-06

77. Gonyar LA, Sauder AB, Mortensen L, Willsey GG, Kendall MM. The Yad and Yeh fimbrial loci influence gene expression and virulence in enterohemorrhagic *Escherichia coli* O157:H7. *mSphere*. (2024) 9:e00124-24. doi: 10.1128/msphere.00124-24

78. Nawrocki EM, Kudva IT, Dudley EG. Investigating the adherence factors of *Escherichia coli* at the bovine recto-anal junction. *Microbiol Spectr*. (2024) 12:e01270-24. doi: 10.1128/spectrum.01270-24

79. Tarr PI, Bilge SS, Vary JC, Jelacic S, Habeeb RL, Ward TR, et al. Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect Immun*. (2000) 68:1400–7. doi: 10.1128/iai.68.3.1400-1407.2000

80. Toma C, Martínez Espinosa E, Song T, Miliwebsky E, Chinen I, Iyoda S, et al. Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing *Escherichia coli*. *J Clin Microbiol*. (2004) 42:4937–46. doi: 10.1128/jcm.42.11.4937-4946.2004

81. Farfan MJ, Torres AG. Molecular mechanisms that mediate colonization of Shiga toxin-producing *Escherichia coli* strains. *Infect Immun*. (2012) 80:903–13. doi: 10.1128/iai.05907-11

82. Biscola FT, Abe CM, Guth BEC. Determination of Adhesin gene sequences in, and biofilm formation by, O157 and non-O157 Shiga toxin-producing *Escherichia coli* strains isolated from different sources. *Appl Environ Microbiol*. (2011) 77:2201–8. doi: 10.1128/AEM.01920-10

83. Carter MQ, Brandl MT, Kudva IT, Katani R, Moreau MR, Kapur V. Conditional function of autoaggregative protein Cah and common Cah mutations in Shiga toxin-producing *Escherichia coli*. *Appl Environ Microbiol*. (2017) 84:e01739-17. doi: 10.1128/AEM.01739-17

84. Ho TD, Davis BM, Ritchie JM, Waldor MK. Type 2 secretion promotes enterohemorrhagic *Escherichia coli* adherence and intestinal colonization. *Infect Immun*. (2008) 76:1858–65. doi: 10.1128/iai.01688-07

85. Sabate R, de Groot NS, Ventura S. Protein folding and aggregation in bacteria. *Cell Mol Life Sci*. (2010) 67:2695–715. doi: 10.1007/s00018-010-0344-4

86. van Ulsen P. Protein folding in bacterial adhesion: secretion and folding of classical monomeric autotransporters In: D Linke and A Goldman, editors. Bacterial adhesion: chemistry, biology and physics. Dordrecht: Springer (2011). 125–42.

87. Wurpel DJ, Beatson SA, Totsika M, Petty NK, Schembri MA. Chaperone-usher fimbriae of *Escherichia coli*. *PLoS One*. (2013) 8:e52835. doi: 10.1371/journal.pone.0052835

88. Ferguson GP, Tötemeyer S, MacLean MJ, Booth IR. Methylglyoxal production in bacteria: suicide or survival? *Arch Microbiol*. (1998) 170:209–18. doi: 10.1007/s002030050635

89. Iacometti C, Marx K, Hönick M, Biletskaia V, Schulz-Mirbach H, Dronsella B, et al. Activating silent glycolysis bypasses in *Escherichia coli*. *BioDesign Res*. (2022) 2022:9859643. doi: 10.34133/2022/9859643

90. Ahn S, Jung J, Jang I-A, Madsen EL, Park W. Role of glyoxylate shunt in oxidative stress response. *J Biol Chem*. (2016) 291:11928–38. doi: 10.1074/jbc.M115.708149

91. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. *PLoS Comput Biol*. (2018) 14:e1005958. doi: 10.1371/journal.pcbi.1005958

92. Matussek A, Mernelius S, Chromek M, Zhang J, Frykman A, Hansson S, et al. Genome-wide association study of hemolytic uremic syndrome causing Shiga toxin-producing *Escherichia coli* from Sweden, 1994–2018. *Eur J Clin Microbiol Infect Dis*. (2023) 42:771–9. doi: 10.1007/s10096-023-04600-1

93. Peroutka-Bigus N, Nielsen DW, Trachsel J, Mou KT, Sharma VK, Kudva IT, et al. Phenotypic and genomic comparison of three human outbreak and one cattle-associated Shiga toxin-producing *Escherichia coli* O157:H7. *Microbiol Spectr*. (2024) 12:e04140-23. doi: 10.1128/spectrum.04140-23

94. Espadinha D, Brady M, Brehony C, Hamilton D, O'Connor L, Cunney R, et al. Case-control study of factors associated with hemolytic uremic syndrome among Shiga toxin-producing *Escherichia coli* patients, Ireland, 2017–2020. *Emerg Infect Dis*. (2025) 31:728–40. doi: 10.3201/eid3104.240060

95. Marques Da Silva W, Smith LY, Aburjaile FF, Larzabal M, Alves SIA, Farace P, et al. Whole-genome sequence reveals genetic determinants of the colonization of Shiga toxin-producing *Escherichia coli* O22:H8 isolates in cattle. *Curr Microbiol*. (2025) 82:207. doi: 10.1007/s00284-025-04185-2

96. Elbein AD, Pan YT, Pastuszak I, Carroll D. New insights on trehalose: a multifunctional molecule. *Glycobiology*. (2003) 13:17R–127R. doi: 10.1093/glycob/cwg047

97. Chen Q, Postma PW, Amster-Choder O. Dephosphorylation of the *Escherichia coli* transcriptional antiterminator BglG by the sugar sensor BglF is the reversal of its phosphorylation. *J Bacteriol*. (2000) 182:2033–6. doi: 10.1128/jb.182.7.2033-2036.2000

98. Spurbeck RR, Stapleton AE, Johnson JR, Walk ST, Hooton TM, Mobley HLT. Fimbrial profiles predict virulence of uropathogenic *Escherichia coli* strains: contribution of Ygi and Yad fimbriae. *Infect Immun*. (2011) 79:4753–63. doi: 10.1128/iai.05621-11

99. Chingcuanco F, Yu Y, Kus JV, Que L, Lackraj T, Lévesque CM, et al. Identification of a novel adhesin involved in acid-induced adhesion of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology*. (2012) 158:2399–407. doi: 10.1099/mic.0.056374-0

100. Korea C-G, Badouraly R, Prevost M-C, Ghigo J-M, Beloin C. *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone–usher fimbriae with distinct surface specificities. *Environ Microbiol*. (2010) 12:1957–77. doi: 10.1111/j.1462-2920.2010.02202.x

101. Stubenrauch C, Belousoff MJ, Hay ID, Shen H-H, Lillington J, Tuck KL, et al. Effective assembly of fimbriae in *Escherichia coli* depends on the translocation assembly module nanomachine. *Nat Microbiol*. (2016) 1:1–8. doi: 10.1038/nmicrobiol.2016.64

102. Wallecha A, Oreh H, van der Woude MW, deHaseth PL. Control of gene expression at a bacterial leader RNA, the *agn*43 gene encoding outer membrane protein Ag43 of *Escherichia coli*. *J Bacteriol*. (2014) 196:2728–35. doi: 10.1128/jb.01680-14

103. Martinez-Gil M, Goh KGK, Rackaityte E, Sakamoto C, Audrain B, Moriel DG, et al. YeeJ is an inverse autotransporter from *Escherichia coli* that binds to peptidoglycan and promotes biofilm formation. *Sci Rep*. (2017) 7:11326. doi: 10.1038/s41598-017-10902-0

104. Moreau MR, Kudva IT, Katani R, Cote R, Li L, Arthur TM, et al. Nonfimbrial adhesin mutants reveal divergent *Escherichia coli* O157:H7 adherence mechanisms on human and cattle epithelial cells. *Int J Microbiol*. (2021) 2021:8868151. doi: 10.1155/2021/8868151

105. Dziva F, Mahajan A, Cameron P, Currie C, McKendrick IJ, Wallis TS, et al. EspP, a type V-secreted serine protease of enterohaemorrhagic *Escherichia coli* O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells. *FEMS Microbiol Lett*. (2007) 271:258–64. doi: 10.1111/j.1574-6968.2007.00724.x

106. Xicohtencatl-Cortes J, Saldaña Z, Deng W, Castañeda E, Freer E, Tarr PI, et al. Bacterial macroscopic rope-like fibers with cytopathic and adhesive properties. *J Biol Chem*. (2010) 285:32336–42. doi: 10.1074/jbc.M110.162248

107. Katani R, Kudva IT, Srinivasan S, Stasko JB, Schilling M, Li L, et al. Strain and host-cell dependent role of type-1 fimbriae in the adherence phenotype of super-shed *Escherichia coli* O157:H7. *Int J Med Microbiol*. (2021) 311:151511. doi: 10.1016/j.ijmm.2021.151511

108. Edison LK, Kudva IT, Kariyawasam S. Comparative transcriptome analysis of Shiga toxin-producing *Escherichia coli* O157:H7 on bovine rectoanal junction cells and human colonic epithelial cells during initial adherence. *Microorganisms*. (2023) 11:2562. doi: 10.3390/microorganisms11102562

109. Navarro-Garcia F, Ruiz-Perez F, Cataldi Á, Larzábal M. Type VI secretion system in pathogenic *Escherichia coli*: structure, role in virulence, and acquisition. *Front Microbiol*. (2019) 10:1965. doi: 10.3389/fmicb.2019.01965

110. Vazquez-Lopez J, Navarro-Garcia F. *In silico* analyses of Core proteins and putative effector and immunity proteins for T6SS in enterohemorrhagic *E. coli*. *Front Cell Infect Microbiol*. (2020) 10:195. doi: 10.3389/fcimb.2020.00195

111. Wan B, Zhang Q, Ni J, Li S, Wen D, Li J, et al. Type VI secretion system contributes to enterohemorrhagic *Escherichia coli* virulence by secreting catalase against host reactive oxygen species (ROS). *PLoS Pathog*. (2017) 13:e1006246. doi: 10.1371/journal.ppat.1006246

112. Aas CG, Drabløs F, Haugum K, Afset JE. Comparative transcriptome profiling reveals a potential role of type VI secretion system and fimbriae in virulence of non-O157 Shiga toxin-producing *Escherichia coli*. *Front Microbiol*. (2018) 9:1416. doi: 10.3389/fmicb.2018.01416

113. Bai X, Ylinen E, Zhang J, Salmenlinna S, Halkilahti J, Saxen H, et al. Comparative genomics of Shiga toxin-producing *Escherichia coli* strains isolated from pediatric patients with and without hemolytic uremic syndrome from 2000 to 2016 in Finland. *Microbiol Spectr*. (2022) 10:e0066022. doi: 10.1128/spectrum.00660-22

114. Herold S, Paton JC, Paton AW. Sab, a novel autotransporter of locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* O113:H21, contributes to adherence and biofilm formation. *Infect Immun*. (2009) 77:3234–43. doi: 10.1128/iai.00031-09

115. Navarro-Garcia F. Serine proteases autotransporter of *Enterobacteriaceae*: structures, subdomains, motifs, functions, and targets. *Mol Microbiol*. (2023) 120:178–93. doi: 10.1111/mmi.15116

116. Brockmeyer J, Bielaszewska M, Fruth A, Bonn ML, Mellmann A, Humpf H-U, et al. Subtypes of the plasmid-encoded serine protease EspP in Shiga toxin-producing *Escherichia coli*: distribution, secretion, and proteolytic activity. *Appl Environ Microbiol*. (2007) 73:6351–9. doi: 10.1128/AEM.00920-07

117. Khan AB, Naim A, Orth D, Grif K, Mohsin M, Prager R, et al. Serine protease *espP* subtype α, but not β or γ, of Shiga toxin-producing *Escherichia coli* is associated with highly pathogenic serogroups. *Int J Med Microbiol*. (2009) 299:247–54. doi: 10.1016/j.ijmm.2008.08.006