VIRULENCE
2026, VOL. 17, NO. 1, 2672206
https://doi.org/10.1080/21505594.2026.2672206

Taylor & Francis
by informa

PATHOGEN GENOMICS

OPEN ACCESS Check for updates

Phylogenomic framework and virulence gene boundaries of emerging Shiga toxin-producing *Escherichia coli* O118 informed by the comprehensive profiling of 359 O118 genomes

Irvin Rivera [a,b], Sara S.K. Koenig[a,b], Armando L. Rodriguez[c], Joseph M. Bosilevac[d], and Mark Eppinger[a,b]

[a]Department of Molecular Microbiology & Immunology, University of Texas at San Antonio, San Antonio, TX, USA; [b]South Texas Center for Emerging Infectious Diseases (STCEID), San Antonio, TX, USA; [c]Research Computing Support Group, University of Texas at San Antonio, San Antonio, TX, USA; [d]U.S. Department of Agriculture (USDA), Agricultural Research Service (ARS), U.S. Meat Animal Research Center, Clay Center, NE, USA

ABSTRACT

Non-O157 Shiga toxin-producing *Escherichia coli* (STEC), particularly the O118 serogroup, are emerging pathogens linked to severe foodborne illnesses, which are characterized by the production of a potent phage-borne cytotoxin. This study explores the genomic landscape, virulence factors, and resistance traits of O118 STEC. We analyzed 357 publicly available O118 genomes, all but three available only in draft stage spanning ten different H-antigen types. To enhance the availability of high-quality reference genomes, we additionally included two O118:H16 STEC strains from our collection that were sequenced to closure. Multilocus sequence typing based on the 4,160 shared genes revealed phylogenetic clustering by H-type and delineated distinct STEC-phylogroups, alongside relationships to non-STEC pathovars such as uropathogenic *E. coli* (UPEC), enteropathogenic *E. coli* (EPEC), and enterotoxigenic *E. coli* (ETEC). Identified STEC-phylogroups encompass H6, H12, H16, and H2 strains with diverse Shiga toxin profiles (stx_{1a}, stx_{2a}, stx_{2b}, stx_{2c}, stx_{2f}). A subset of H2-STEC lacked *stx*, suggesting potential secondary phage loss. Most STEC groups carried the locus of enterocyte effacement (LEE). Further, a strong correlation was observed between H-antigens and *eae* subtypes, with specific pairings such as H6/*eae*-ι, H16/*eae*-β, and H2/*eae*-ε. Horizontally acquired pathogenicity islands, including O-island 122 in H16 strains and a novel pathogenicity-associated island carrying antibiotic resistance, along with other loci related to colonization and interbacterial competition, further enhance their virulence potential. Our findings underscore the genetic diversity and virulence potential of O118 STEC. Understanding such phylogroup-linked virulence and resistance traits is crucial for effective surveillance and public health interventions.

ARTICLE HISTORY
Received 2 October 2025
Revised 19 February 2026
Accepted 4 May 2026

KEYWORDS
Shiga toxin-producing *Escherichia coli* (STEC); serogroups O118 and O151; whole genome sequencing; multi-locus sequence typing; comparative phylogenomics

Introduction

Escherichia coli O118 is an emerging foodborne Shiga toxin-producing *E. coli* (STEC) serogroup with a bovine zoonotic reservoir [1]. Infections can cause severe human gastrointestinal disease and may progress to life-threatening complications, such as hemorrhagic colitis and hemolytic uremic syndrome (HUS) [2,3]. Vulnerable populations, such as children and the elderly, are particularly affected [4–6]. This underscores the importance of monitoring this emerging serogroup in both veterinary and public health contexts and assessing its virulence potential and phylogenetic boundaries [4,7]. *E. coli* are historically classified by their somatic O- and flagellar H-antigens [8–10]. Serotype O157:H7 is the dominant causative agent of STEC disease in the U.S [11,12]. However, the incidence of non-O157 infections has steadily increased in the U.S. and globally [13–16]. Emerging serogroups O26, O45, O103, O111, O121, and O145 account for most clinical non-O157 STEC infections in the US and are colloquially referred to as the "Big Six" [13,17,18]. These accounted for 83% of non-O157 STEC infections in the United States from 2000 to 2010 [19–22]. As diagnostic testing increased, non-O157 STEC, namely the "Big Six," surpassed the previously dominant O157:H7 lineage in 2014 [15,18]. Between 2021 and 2023, non-O157 infections accounted for 168k annual cases, compared to 96k O157 cases [19,23]. The first reported human infection with stx_1-positive O118:H2 STEC occurred in 1996 during a school outbreak in Japan [24] linked to salad [25]. Since then, strains of O118 (including its related O151 subgroup) [26,27] have been recognized as an emerging and human pathogenic STEC lineage, many of which possess Shiga toxins (*stx*), intimin (*eae*), and

CONTACT Mark Eppinger mark.eppinger@utsa.edu
Supplemental data for this article can be accessed online at https://doi.org/10.1080/21505594.2026.2672206.

hemolysins [25,28]. The carried ΦStx prophages feature different combinations of *stx*-subalelles [11,29] that can also form hybrid toxins [30]. Both stx_1 and stx_2 are potent translation inhibitors [25,29,31–33] and their alleles have been documented in O118 [6,28]. Stx_2, particularly $stx_{2a,}$ and $stx_{2d,}$ have been associated with elevated cytotoxicity [34,35] compared to the morbidity caused by stx_{1a+} STEC [8,14,36]. However, Stx_1 has been shown to exhibit increased cytotoxicity in Vero cells compared to Stx_{2a} [34,37,38]. Adding further complexity, Stx_1 has been reported to reduce the overall toxicity of Stx_{2a}, potentially through competitive binding or interference at the receptor level [34]. Besides the pathovar-defining Shiga toxin, another notable virulence factor includes the locus of enterocyte effacement (LEE). This pathogenicity island encodes a type III secretion system (T3SS) and effectors, the outer membrane adhesin intimin (*eae*), and its translocated receptor (*tir*) [9,25]. The latter is also carried by enteropathogenic *E. coli* (EPEC) [39]. Antimicrobial (AR) and multidrug-resistant (MDR) isolates are described for O118 STEC [25,28,40,41]. STEC also carry serogroup-specific virulence plasmids and an array of other plasmids with diverse functions [17,42,43]. Isolates belonging to serogroup O118 occur across multiple *E. coli* pathovars, including EPEC, which encode intimin and form characteristic attaching and effacing (A/E) lesions [4,24]; Enteroaggregative *E. coli* (EAEC) which form biofilm-like aggregates [4,13]; and extraintestinal pathogenic E. coli (ExPEC) that are capable of causing severe infections outside the gastrointestinal tract, such as urinary tract infections (UTIs) and bloodstream infections [24,44]. In this study, we provide a comprehensive framework for O118 *E. coli* with a focus on emerging STEC O118. The gathered information on virulence gene content and genome makeup provide a foundation to delineate evolutionary and pathovar boundaries of these emerging and human pathogenic non-O157 lineages.

Materials and method

Bacterial strains analyzed in this study

For this study, we analyzed a total of 359 *E. coli* O118 genomes: 357 publicly available genomes retrieved from the NCBI Pathogen Detection Isolate Browser (https://www.ncbi.nlm.nih.gov/pathogens/isolates/) as of July 2024, which included three closed genomes; and further two clinical O118 strains from the Nebraska Public Health Laboratory that we sequenced to closure to provide high-quality references. Strain-associated metadata and genome statistics are presented in Table S1.

Genome sequencing, assembly, and annotation

Strains were cultured overnight at 37°C with shaking at 220 rpm in lysogeny broth (LB) (Thermo Fisher Scientific, Asheville, NC, USA). To maximize total genomic DNA (gDNA) yields, bacterial overnight cultures were diluted to OD_{600} of 0.03 in fresh LB medium and grown at 37°C with shaking at 220 rpm to mid-log phase (OD_{600} ~0.5). Total gDNA was extracted using the Monarch HMW DNA Extraction Kit (New England Biolabs, Ipswich, MA, USA). Genomes were sequenced to closure using Nanopore long-read technology (Oxford Nanopore, UK). Sequencing libraries were prepared with Rapid Barcoding Kit (RBK-114) according to the manufacturer's instructions and sequenced on the PromethION platform on R10.4.1.Flow Cell (FLO-MIN114). Reads in the fastq format were imported into Galaxy v.22.05 [45]. Default parameters were used for all software unless specified otherwise. Fastq reads were QC-ed using FastQC (v.0.74+Galaxy0) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and assembled with Flye (v.2.9.3) [46]. The resulting contigs were evaluated with QUAST (v.5.2.0 + Galaxy1). The chromosomal *dnaA* and plasmid *repA* genes were designated as the zero point of the closed molecules prior to annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [47]. Genome assemblies were profiled with SeqKit v2.8.2 to obtain genome statistics [48,49].

Phylogenomic analyses and MLST-schemas

The O118 genomes, along with the sequence of K-12 substrain MG1655 [50,51], were imported into Ridom SeqSphere+ (v.8.3) (Ridom GmbH, Münster, Germany) for core genome (cg) and targeted Multilocus Sequence Typing (MLST) [52–54]. The Sequence Type (ST) was determined according to the EnteroBase schema [54]. Allele sequences for the Achtman scheme, targeting seven housekeeping genes (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*), were accessed on the EnteroBase website (https://enterobase.warwick.ac.uk/species/ecoli/download_7_gene). A cgMLST schema was developed using the closed chromosome of K-12 substrain MG1655 [51] as seed as previously described [55]. Core and accessory MLST targets were identified according to the inclusion/exclusion criteria of the SeqSphere+ Target Definer. The allele information from the targeted seven-gene schema and the defined core genome gene of the panel strains were used to

establish phylogenetic hypotheses using the minimum-spanning method with default settings [56,57].

Pathogenome makeup and visualization

Chromosomes were compared with progressive Mauve [58] and visualized in pyGenomeViz-pgv-mauve [59] (https://github.com/moshi4/pyGenomeViz) using seaborn [60]. Chromosomes and carried plasmids were compared and visualized in Blast Ring Image Generator (BRIG) (v.0.95) [61]. Serotypes were determined *in silico* using *E. coli* Typer (ECTyper) (https://github.com/phac-nml/ecoli_serotyping) (Galaxy 1.0.0) in Galaxy v.22.05. If the *wzy* gene was reported absent by ECTyper, BLASTn was used to reevaluate its presence and fragmentation status. The ECTyper database was downloaded as a JSON file from GitHub (https://github.com/phac-nml/ecoli_serotyping) and converted into a blast database. The Average Nucleotide Identity (ANI) between all-vs-all 359 genomes was computed with skani [62], the distance matrix clustered by the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method and visualized using ANIclustermap (62) (https://github.com/moshi4/ANIclustermap). Virulence and antibiotic resistance genes (ARGs) were cataloged with Virulence Factor Database (VFDB) [63] and ResFinder (https://cge.cbs.dtu.dk/services/ResFinder/) [64], respectively. Pathovars were inferred *in silico* through BLASTp query [65] against a curated VFDB database, flagging *E. coli* pathovar-defining virulence genes as compiled from [63,66,67], e.g. EPEC (*eae*), STEC (stx_1 or stx_2, ±*eae*), enterotoxigenic *E. coli* (ETEC) (*eltA*, *estA*, or *cfaB*), and uropathogenic *E. coli* (UPEC) (*fimH*). Analyzed O118 strains carry enterohemorrhagic *E. coli* (EHEC)-associated virulence factors (such as stx_2 and *eae*). However, not all EHEC-genotype strains actually cause EHEC disease. We refer to them as STEC in this study to distinguish genetic potential from confirmed clinical disease manifestation, such as reported through patient data or outbreak reports. Ribosomal RNAs and CRISPR/Cas systems were detected with Basic Rapid Ribosomal RNA Predictor (Barrnap) [68] (v.0.7) and CRISPRCasFinder [69] (v4.2.20) in Proksee [70] (v.1.1.0). Anti-phage systems were detected using DefenseFinder [71] (Galaxy v.1.3.0+galaxy0). Prophages were compared and visualized in Easyfig (v.2.2.2) [72]. Boundaries and locations of intact, partial, or remnant prophages were identified using PHASTEST [73], followed by manual curation of the ΦStx-prophage and core genome borders by Mauve alignment of the genome to *E. coli* K-12 substrain MG1655 (GenBank accession U00096) [50] in Geneious Prime (v.2024.0.5). Direct repeats caused by phage integration were identified by BLASTn self-alignment and the Repeat Finder plugin (v1.0.1) in Geneious Prime. Replicase subtyping was performed to characterize the ΦStx_{1a}-encoding prophages, following the typing scheme introduced by [74]. Replicases were typed with EHEC phage replication unit (*eru*) schema [75,76]. The subtypes of phage-borne *stx* were determined by BLASTn against a curated *stx*-suballele database, as previously described [17,77,78]. A hierarchical classification approach was implemented to categorize phage-associated genes from the PHASTEST JSON output. All present phage-associated genes were further curated processing type and name descriptors, along with reported Gene Ontology (GO), for more in-depth functional classification. Genomic islands (GI) were detected with IslandViewer4 [79–81]. LEE islands and boundaries were curated, guided by LEE1-LEE4 operon genes *espG and espF*, respectively, and visualized in EasyFig (v.2.2.2) [72]. Intimin subtypes were determined by BLASTn against a curated database of published *eae*-subtype sequences, as reported previously [17]. Transposable elements and Insertion Sequence (IS) elements, along with inverted repeats, were cataloged through BLASTn [82] against the Transposon Central (TnCentral) database [83], comprised of TnCentral +Integrall+ISFinder entries [84,85] and ISEScan (v.1.7.2.3+Galaxy0) [86].

Plasmid makeup and visualization

Plasmids were visualized and decorated in BRIG (v.0.95) [61]. Plasmid mobility and incompatibility groups were recorded with Mobilome Typer (MOB-Typer) (v.3.0.3+Galaxy0) [87] and ABRicate (Galaxy v.1.0.1; https://github.com/tseemann/ABRicate) with options –minid 95 –mincov 90 using the PlasmidFinder database [82,88] (Carattoli A, Zankari E, Garcia-Fernandez A, Volby Larsen M, Lund O, Villa L, Aarestrup FM, Hasman H. Antimicrob. Agents Chemother. 2014. April 28th.). Using a MASH-based search strategy, phylogenetically related plasmids were identified by comparison against the PLSDB database (v.2024.05.31.v2) [89,90]. Serotypes of the plasmid-associated chromosomes molecule were determined *in silico* using ECTyper (https://github.com/phac-nml/ecoli_serotyping) (Galaxy v1.0.0) in Galaxy v.22.05. Virulence genes were cataloged with VFDB [63]. Bacteriocins, ribosomally synthesized and post-translationally modified peptides (RiPPs), were recorded with BAGEL4 (http://bagel4.molgenrug.nl/) [91]. Genomic islands were detected with

IslandViewer4 [79–81]. Mobile Genetic Elements (MGEs) were identified in Proksee [70] based on homology to entries in the mobile orthologous groups database (mobileOG-db) [92]. Transposable elements and IS elements were identified with ISEScan (v.1.7.2.3+Galaxy0) [86] and by BLASTn queries [82] against the TnCentral database, comprised of TnCentral+Integrall+ISFinder entries [83–85].

Antimicrobial susceptibility testing

Susceptibilities were determined as described previously [93]. Broth microdilution was performed using the SENSITITRE™ NARMS Gram Negative panel (Fisher Scientific, Cleveland, OH, USA) then output data were interpreted using CLSI M100 (CLSI, 2022) guidelines fitting the National Antimicrobial Resistance Monitoring System (NARMS) standards [94]. The antimicrobial agents included in the panel, along with their abbreviations and resistance breakpoints, were as follows: amikacin (AMI), ≥64 μg ml^{-1}; amoxicillin – clavulanic acid (AMC), ≥32/16 μg ml^{-1}; ampicillin (AMP), ≥32 μg ml^{-1}; cefoxitin (FOX), ≥32 μg ml^{-1}; ceftiofur (TIO), ≥8 μg ml^{-1}; ceftriaxone (AXO), ≥16 μg ml^{-1}; chloramphenicol (CHL), ≥32 μg ml^{-1}; ciprofloxacin (CIP), ≥4 μg ml^{-1}; gentamicin (GEN), ≥16 μg ml^{-1}; kanamycin (KAN), ≥64 μg ml^{-1}; nalidixic acid (NAL), ≥32 μg ml^{-1}; streptomycin (STR), ≥64 μg ml^{-1}; sulfisoxazole (FIS), ≥512 μg ml^{-1}; tetracycline (TET), ≥16 μg ml^{-1}; and trimethoprim – sulfamethoxazole (COT), ≥4/76 μg ml^{-1}. Resistance to three or more antimicrobial classes was classified as MDR. The antimicrobial classes represented in this panel were: aminoglycosides (AMI, GEN, KAN, STR), β-lactam/β-lactamase inhibitor combination (AMC), cephems (FOX, TIO, AXO), folate pathway inhibitors (FIS, COT), penicillin (AMP), phenicol (CHL), quinolones (CIP, NAL), and tetracycline (TET).

Supplemental text

Plasmid profiles of serogroup O118 E. coli

The 359 analyzed closed and draft genomes identified 36 replicons, which provided a testament to the plasticity of plasmids in this group. Among these were incompatibility group families associated with virulence plasmids and colicins (Table S2). The most prevalent replicon in STEC O118 was IncFIB (AP001918) of H2 strains (97.1%, 235/242) and of H16 strains (94.6%, 88/93). These usually 100 kb or larger low-copy plasmids have been associated with disseminating virulence genes in *Enterobacteriaceae* [95]. Other Inc-types with shared distribution were colicin-type Col156, present in H2 (42.6%, 103/242) and H16 (7.5%, 7/93) strains, as well as present in the pO111 plasmid in H16 (29.0%, 27/93) and H2 (14.5%, 35/242) strains. The presence of Col156 is often associated with plasmids that also carry antibiotic resistance genes [96]. Carriage of virulence plasmid pO111 has been linked to severe gastrointestinal disease [97]. Another noteworthy observation was that most STEC H16- and H2-phylogroups feature IncB/O/K/Z_3 groups (95.7% and 90.7%), and IncFII(pCoo) that is present in all H11-strains. It would be valuable to correlate plasmid replicons with AMR and virulence genes. Unfortunately, the majority of O118 genomes are in draft stage, which limits the ability to unambiguously associate gene inventories with the multiple replicons identified in a given strain. The inclusion of closed genomes in future studies will be essential to resolve plasmid structures and enable the accurate mapping and correlation of AMR and virulence determinants.

Results and discussion

Genomes of closed STEC O118

At the time of data collection, three closed serogroup O118 genomes were publicly available, all of which are STEC: H6-strains 2013C-4538 and 2014C-3050 from human stool [98] and H2-strain EC20017429 isolated in 2017 from a gastroenteritis patient [99]. To support comprehensive analyses of O118 STEC, we sequenced two clinical H16-isolates, strains 12089 and 12867, to closure using long-read sequencing [17,100,101]. This approach yielded high-quality closed chromosomes of 5,804,173 and 5,685,925 bp and recovered carried plasmids (Figures 1, 2). Genome statistics and strain-associated metadata are provided in Table S1. The average nucleotide identity processing closed and draft O118 genomes was 99.34%, indicative of the highly conserved chromosomal *E. coli* backbone [102,103] (Figure S1).

Mobilome of closed STEC O118 genomes

Mobile genome elements are major drivers of pathogenome diversification [32,104,105]. The identified prophages, GIs, and IS elements contributed significantly (24.11 to 28.51%) to the genome content of the closed O118 STEC, comparable to observations in other STEC-lineages [105,106] (Table S2). These STEC strains contained about 1 Mb of additional sequence information compared to nonpathogenic *E. coli* strain K-12 substrain MG1655 [50], included in this comparison. These included pathovar-defining ΦStx prophages [25,29,31–33,107–109] and

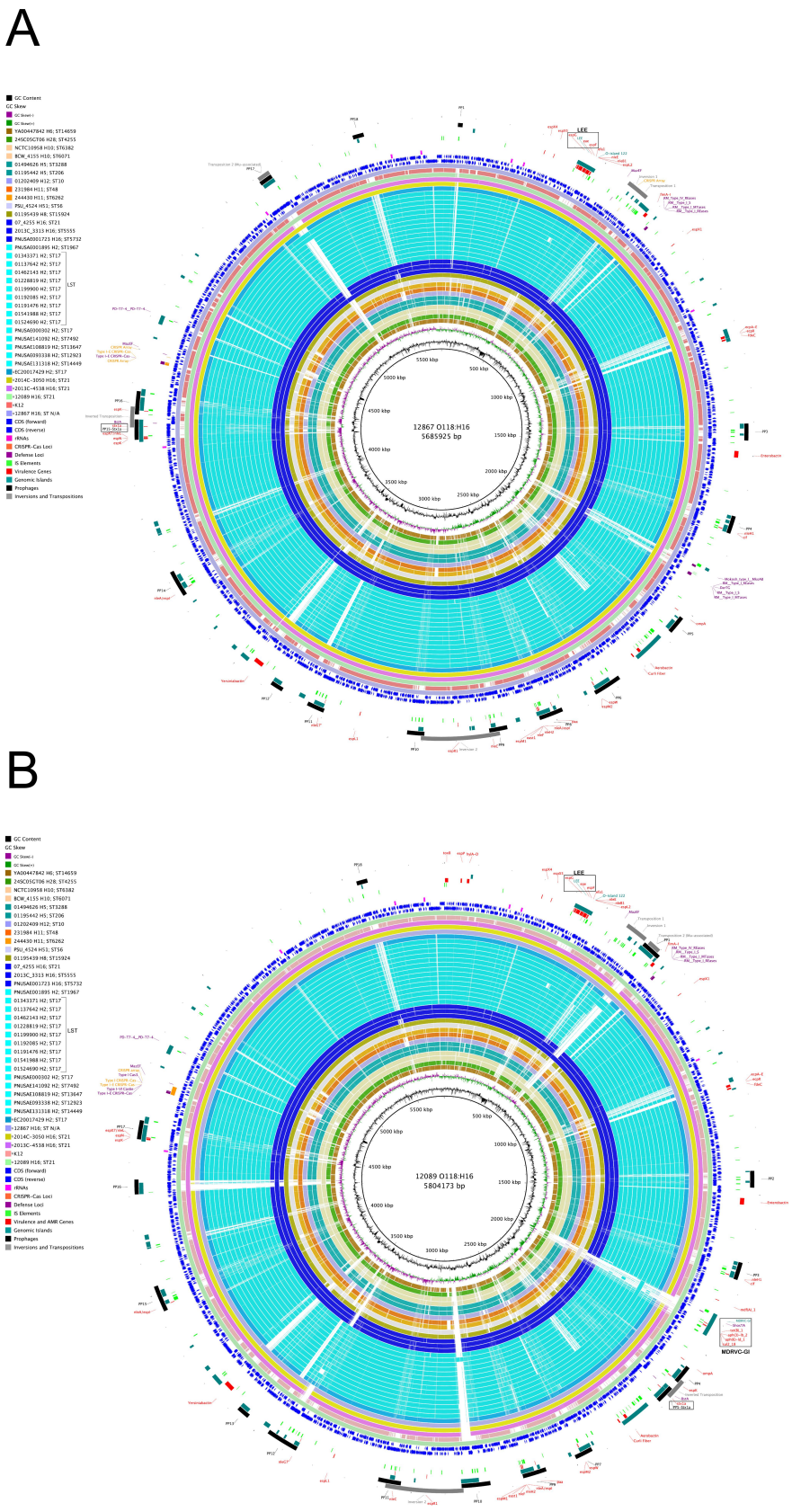

Figure 1. Comparison of O118 STEC chromosomes. BRIG comparison of representative genomes for all identified STEC-phylogroup, considering H-antigen and sequence type, referenced to O118:H16 strains A) 12867 and B) 12089. *E. coli* K-12 is also included to highlight STEC-specific genome content. Query genomes are color-coded, and their order reflects the inferred phylogenomic relationships. Coding sequences are shown as arrows on the + and − strands, and functional annotations for virulence genes and other loci of interest are highlighted.

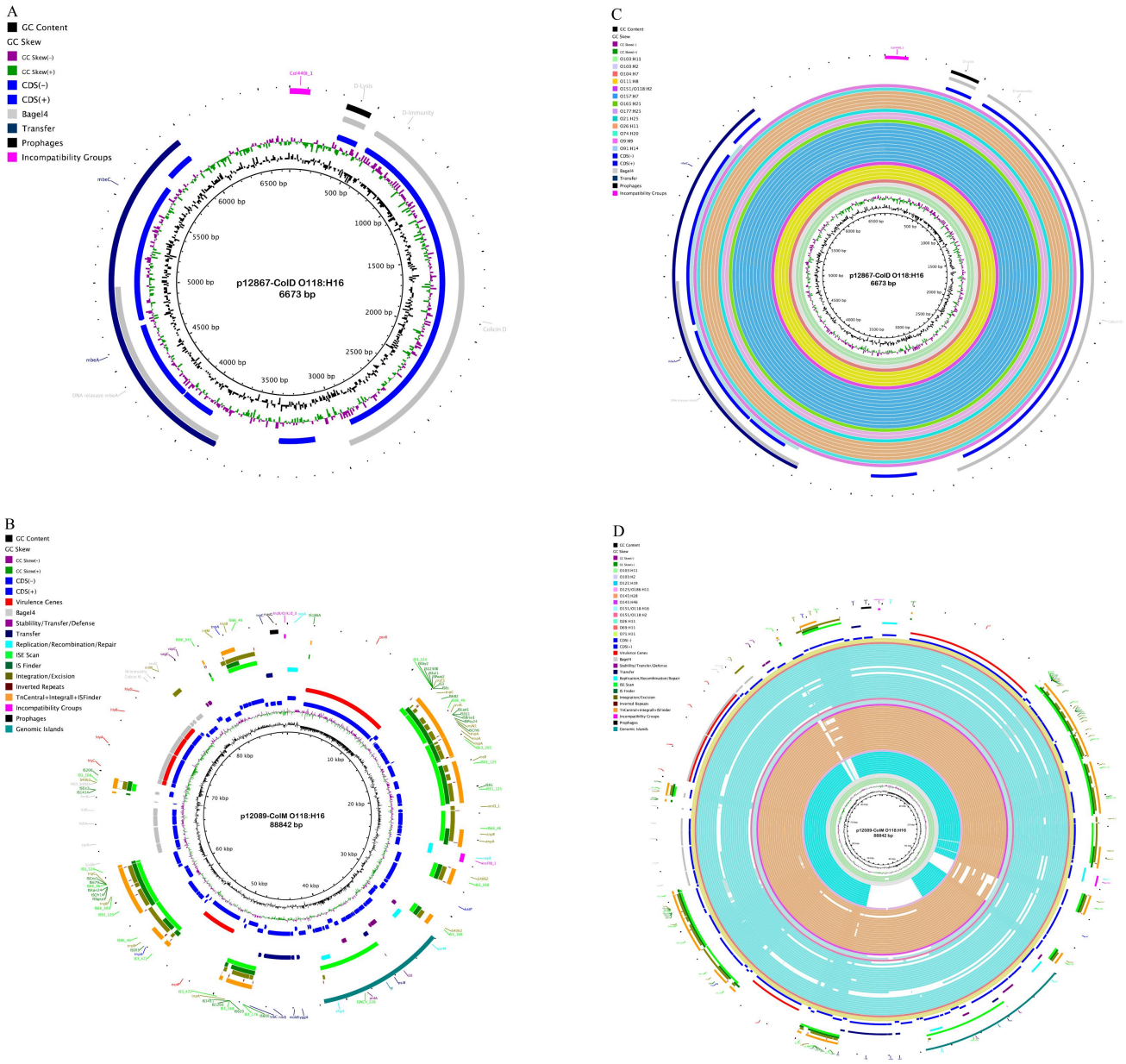

Figure 2. Plasmid inventory of the two sequenced O118 STEC strains. (A) The small 6,673 bp ColD-colicin plasmid in H2-strain 12867, (B) the 88,842 bp multifunctional ColM-colicin and virulence plasmid of H16-strain 12809 features an array of virulence and conjugation loci, (C), and (D) phylogenetically related plasmids were detected in diverse STEC serogroups.

other virulence hallmarks, such as the LEE Pathogenicity Island [25,110,111]. IS elements have been utilized to delineate distinct phylogenetic lineages in STEC [17,112–114]. We recorded IS element prevalence and types in the five closed genomes (Figure S2), showing both shared (IS3_168) but also H-antigen-specific IS elements (e.g. IS3_255, IS4_463).

Large chromosomal rearrangements in H16-strains 12867 and 12089

Mauve comparison of closed O118 chromosomes revealed largely genome-wide synteny, though we observed large chromosomal rearrangements (LCRs). LCRs can generate rapid phenotypic changes despite identical gene content by modifying chromosome organization [115] and have been associated with environmental fitness, antimicrobial susceptibility, type III secretion, and Stx production [116–118]. As evident in Figure 1 and S3, the inversion and transposition events that differentiate the genomes of the sequenced H16-strains 12867 and 12089 were associated with adjacent or flanking mobility or ribosomal RNA loci [119–122]. Two rearrangements are linked to prophage regions, in similarity to LCRs reported in STEC O157 [118,123], and both inversion 1 and transposition 1 are associated with IS-elements (Figure 1).

Colicinogenic and virulence-associated plasmids in H16-strains 12867 and 12089

The sequenced H16-strain 12867 and H16-strain 12089 are both colicinogenic (Figure 2). Colicins are produced and toxic to *E. coli*, and support the ability of *E. coli* to compete for a shared niche [124]. Plasmid p12867-ColD [124], showed the typical organization into D-lysis, D-immunity, and Colicin D [125], along with a mobilizable DNA-relaxase *mbeA* [126]. Additionally, the plasmid-borne SOS inhibition gene *psiB* was detected [127]. A MASH-based survey identified related plasmids in STEC O157:H7, the "Big Six" serogroups, and other STEC-associated serogroups (Table S2). The larger 88 kb plasmid p12089-ColM encodes Colicin M and its associated M-immunity protein, its other gene content identified it as a virulence plasmid. Notable were hemolysin (*hlyCABD)* [128], serine protease (*espP)* [129], fimbrial regulator *fimB*, and the STEC adherence factor *toxB* [130]. Additionally, modulators of lipopolysaccharide and exopolysaccharide biosynthesis (*lpxM*, *icaB*) [131,132] along with SOS-inhibition gene *psiB* [127], were also present. Approximately 43% of the plasmid sequence are mobility and conjugation loci [119,133]. A MASH-based plasmid survey identified related plasmids among H16 and H2, Big Six and other STEC serogroups [134–136] (Table S2).

Comparison of ΦStx prophages in closed O118 STEC

Comparative genomics provided insights into the shared and specific virulence traits in the identified STEC-phylogroups. Bacteriophages target conserved chromosomal loci and undergo repeated acquisition, loss, and microevolution, thereby collectively shaping bacterial gene content and pathogenicity traits [104,137,138]. A comparison of the complete ΦStx_1 prophages extracted from the closed genomes identified two insertion sites: *torS* and an intergenic insertion upstream of *dinI* with the phage boundaries marked by flanking direct repeats (Figure 3, Table S2). The *torSTR* locus was one of four identified ΦStx_{1a} integration sites in O26:H11 ST21 [139], an intergenic attachment site common in *E. coli* [140]. The ΦStx_{1a} prophages show an overall high-degree of sequence identity and synthetic organization. The *torS* ΦStx_{1a} phage of the H2-strain EC20017429 was distinguished by the presence of an IS3-like element and cluster of three hypothetical genes and a kinase, absent in the other H16 strains. The larger ΦStx_{1a} of H16-strain 12867 carried a DUF550 protein, type III effectors EspK and EspN, and a hypothetical protein.

Locus of enterocyte effacement pathogenicity island in closed O118 STEC

The LEE pathogenicity island is carried by diverse STEC serogroups, forming characteristic attaching and effacing A/E lesions [141–143]. The comparison of LEE islands extracted from the closed O118 STEC genomes (Figure 4) showed they were highly conserved and organized in five polycistronic operons (LEE1 to 5). These encoded T3SS regulators and effectors [144]. All were integrated at *pheU* tRNA, a known LEE integration site in *E. coli* [145–147]. The H2 strain was distinguished by an adjacent O-island 122 within the LEE island boundaries marked by 17 bp direct repeats (Table S2). This island encodes non-LEE-encoded (Nle) effectors and transport and regulatory elements that enhance *E. coli* pathogenicity by promoting intracellular survival, adhesion, immune modulation, and enterotoxin production [148,149]. Its carriage has been associated with highly virulent STEC, causing HUS [149,150]. The H16 phylogroup, represented by strains 2014C-3050, 2013C-4538, 12867, and 12089 features the β-intimin subtype, whereas the ε-intimin subtype is present in the H2 strain EC20017429 (Figure 4).

Pathogenicity associated multi-drug resistance, virulence, and competition island

In strain H16- 12089, a novel 74 kb Resistance, Virulence, and Competition (MDRVC) Island with a complex modular architecture was discovered inserted into tRNA-Ser (Figure 5). This island featured a tyrosine-type integrase [151], with borders marked by direct repeats [152] (Table S2). A core component of this island was the MDR-element AMR-SSuT, which has been previously described in STEC O157, though the genomic context was unclear [153]. This island carries tetracycline, streptomycin, and sulfonamide loci. Embedded within this element was streptomycin resistance transposon Tn5393.2 [154], itself disrupted by an ISVsa5-flanked Tn10 [155], introducing tetracycline [156,157]. The *sul2* locus, however, is located outside this Tn5393.2-Tn10 composite, proximate to IS66-like elements, previously linked to antimicrobial resistance in *E. coli* [158]. Prominently featured were genes that aid interbacterial competition, such as a *cdiA*-mediated contact-dependent inhibition [159] associated with VENN motif pre-toxin [160] and the *cbtA-cbeA* toxin-antitoxin system [161]. A broad array of virulence loci was cataloged; among these was the hemagglutinin adhesin [159] and the hemolysins secretion gene (*fhaC*), which likely interacts with the hemolysin on p12089-ColM (Figure 2). Further diverse

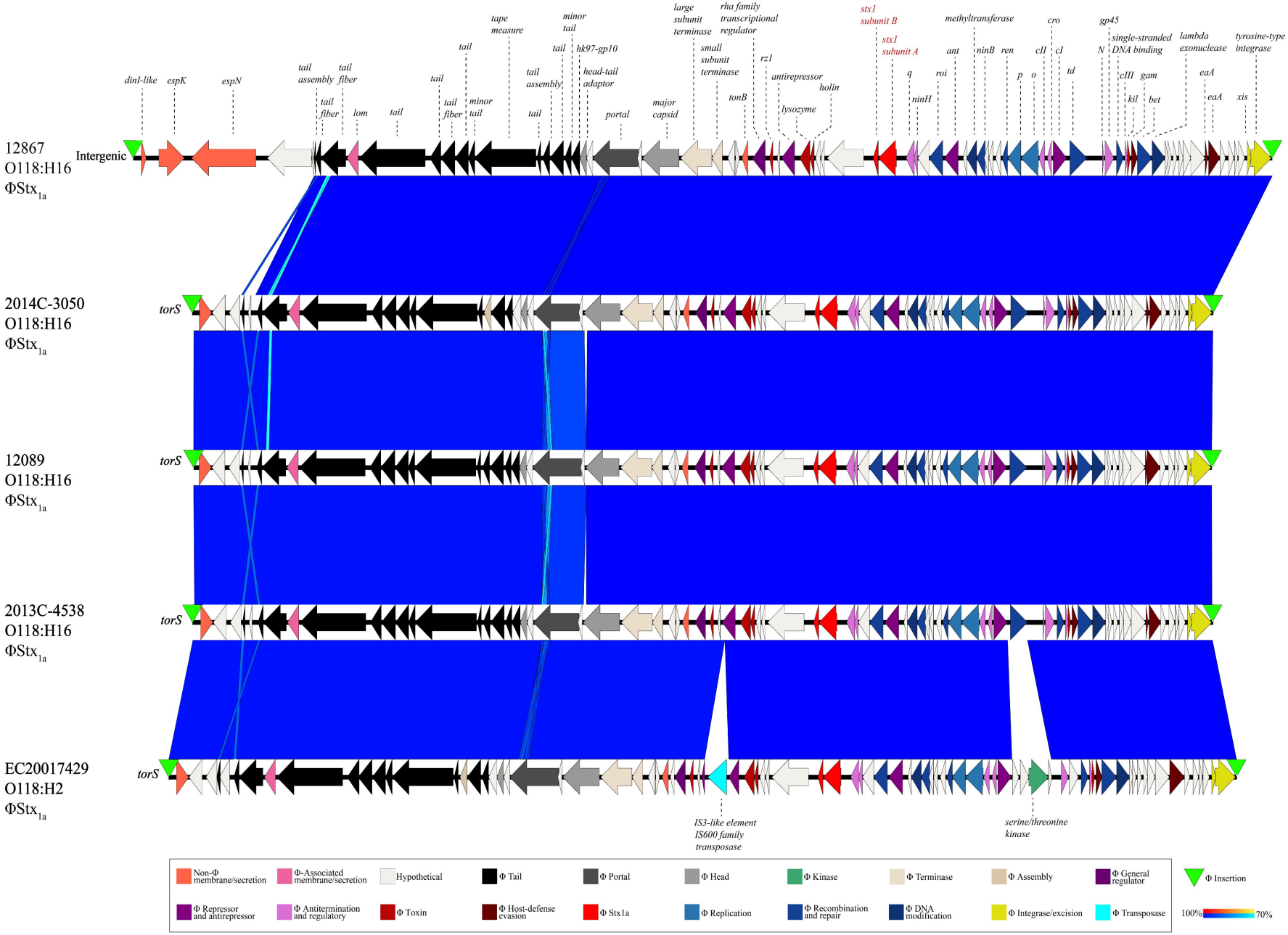

Figure 3. Comparison of ΦStx-prophages in closed O118 STEC. BLASTn-based comparison of the ΦStx_{1a} prophage architectures and gene content. The comparison shows that ΦStx1a phages show considerable genomic plasticity and are inserted at *torS* or intergenically close to the *dinI* locus.

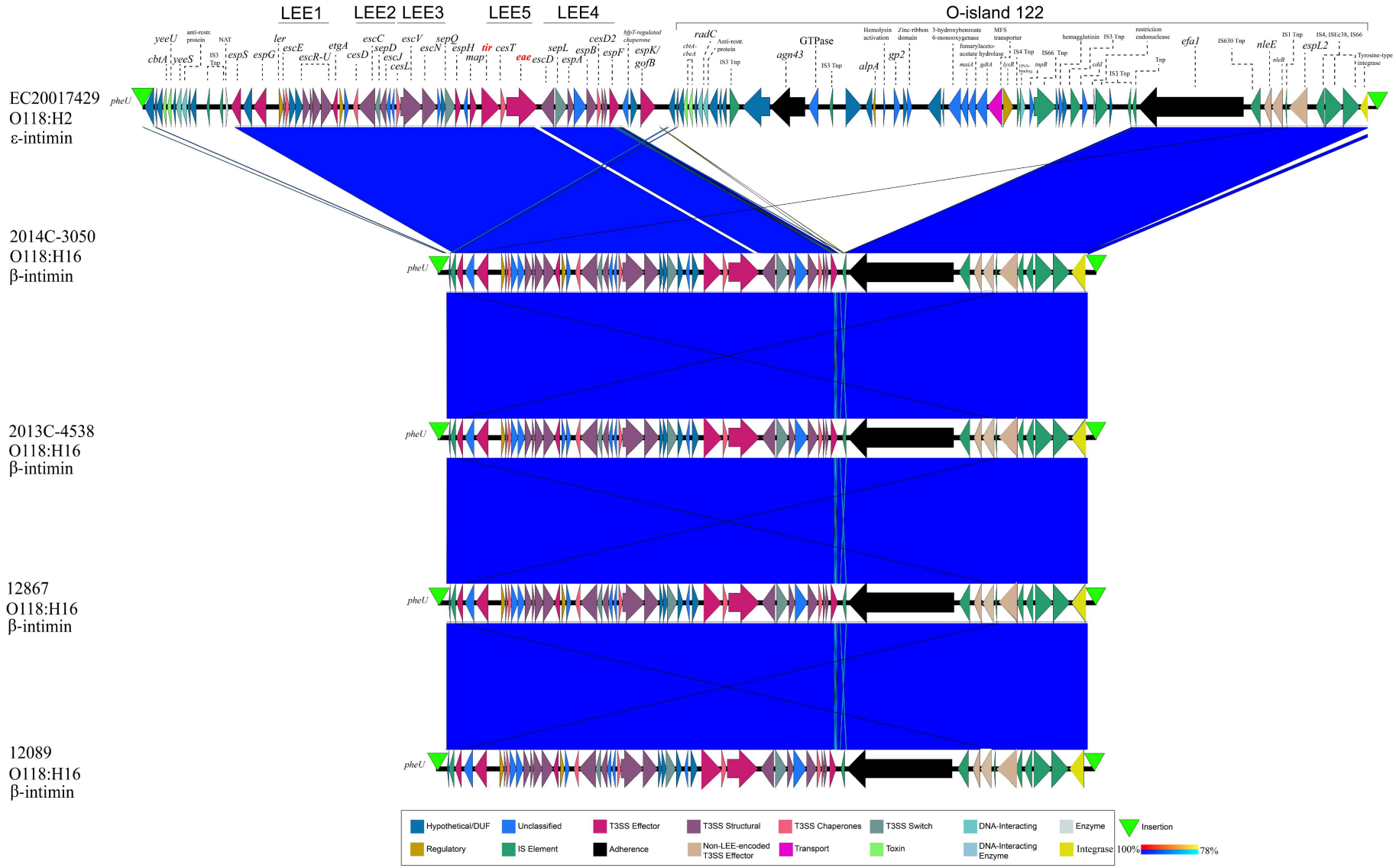

Figure 4. LEE islands of closed O118 STEC. BLASTn-based comparison of LEE islands inserted into the *pheU*. The organization of the LEE1 to LEE5 operons and O-island 122, disrupting the LEE island in H2-strain EC20017429, is indicated. The order of strains reflects their inferred phylogenomic position.

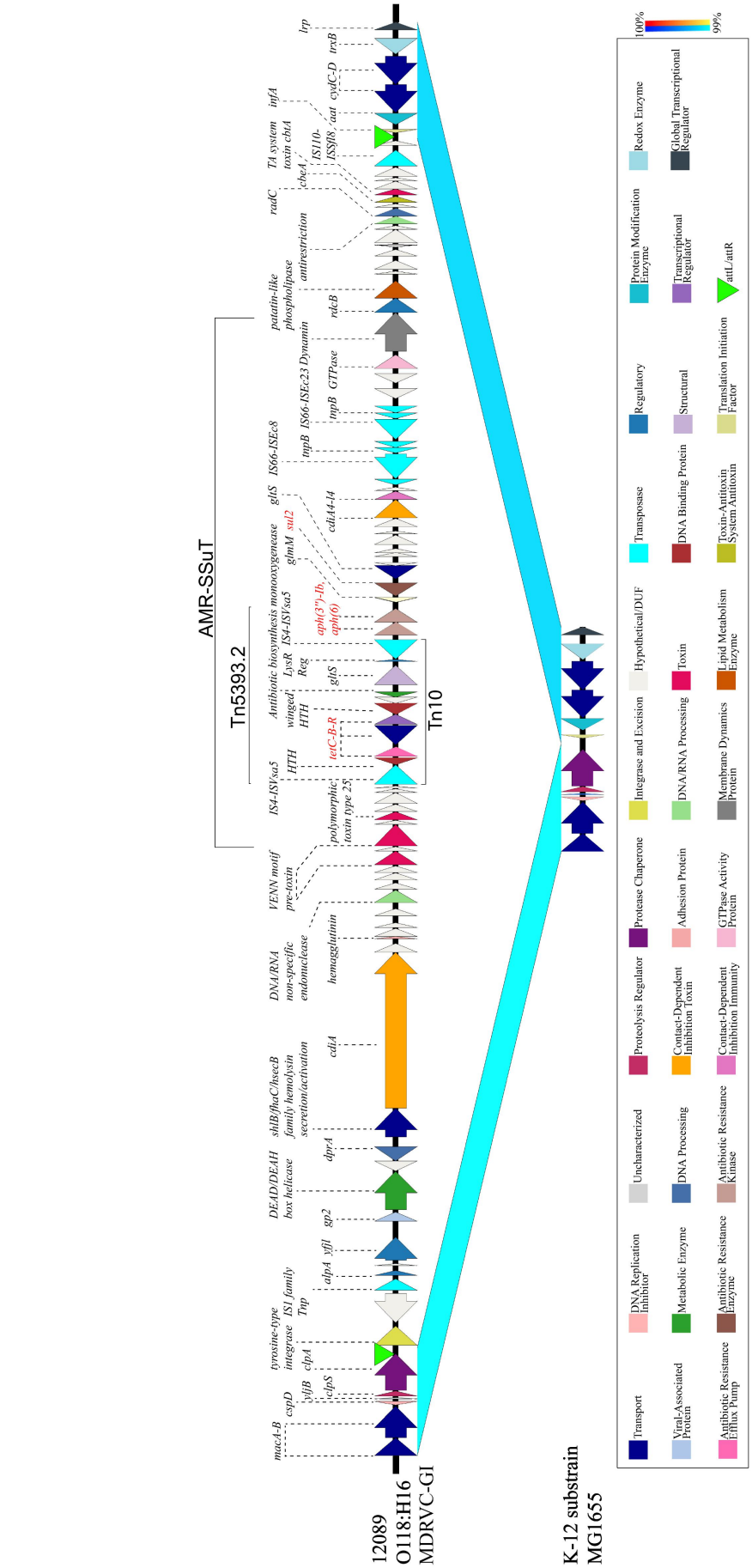

Figure 5. Modular architecture of the MDRVC island. This newly discovered 74kb multifunctional island, inserted into the *tRNA-ser* of H16-strain 12089, features a complex modular architecture. It integrates an antimicrobial multi-drug resistance cassette with gene loci, promoting virulence, interbacterial competition, stress resilience, and overall fitness.

regulator genes were associated with stress-response. This included *lysR* [162], a DEAD/DEAH-box helicase [163], *alpA* [164], and a bacteriophage defense gene *gp2* [165], along with membrane modeling genes, patatin-like phospholipases and dynamin-like protein *rdcB* [166,167]. DNA-modification enzymes support genomic integrity through repair (*radC*) or endonuclease-mediated defense [168,169]. Altogether, the carriage of this multifunctional GI by strain 12089 suggests increased pathogenic potential and fitness, promoting virulence, antimicrobial resistance, stress adaptation, and also genomic stability.

Phylogenomic framework and pathovar boundaries of serogroup O118/O151 STEC

To support broad-scale analysis, we included all publicly available serogroup O118, O151, or O118/O151 genomes, as identified in the NCBI Pathogen Detection Isolate Browser, for which a common ancestry has been proposed [170]. O118 and O151 strains are differentiated by subtle antigenic changes in their O-antigen gene clusters (OGCs). Their OGCs share > 99% nucleotide sequence identity and differ by only two nonsynonymous substitutions (O118 → O151): a V292A substitution in *wzy* and a P88S substitution in *fnlB*. To comprehensively capture this diversity, we included all genomes annotated as O118/O151, O118, or O151 in our analyses. Notably, 358 of 359 analyzed genomes possessed the O151-*wzy* variant, while a single strain carried O118-*wzy* allele (Table S1) [170–172]. Strain-associated metadata and genome statistics of this expanded strain panel can be found in Table S1. This panel of 359 O118 genomes features ten H-antigens, as determined by *in silico* flagellin (*fliC*) subtyping [173,174]. The shared gene inventory was determined at 4,160 genes, comprised of 2,437 core and 1,723 accessory loci, indicative of the genome plasticity found in heterogenous O118 serogroup *E. coli* (Figure 7, Table S4). The ANI of 96.32% was computed through an all-vs-all comparison of the 359 O118 genomes. The UPGMA-clustering of ANI values recovered the different H-phylogroups (Figure S1). The majority belonged to H2 ($n = 242$), followed by H16 ($n = 93$) and H11 ($n = 11$), with relatively few available representative genomes of H10 and H12 ($n = 3$), H5 and H28 ($n = 2$), and a single strain for H6, H8, and H51 (Table S1). To investigate the phylogenomic and virulence boundaries of emerging O118 STEC among the heterogenous serogroup O118/O151 strains, we established a phylogenomic framework informed by targeted and core genome MLST (Figure 6, 7). The tree topology partitions the isolates by H-antigen and ST. Pathovars were inferred *in silico* from their VFDB virulence gene profiles. The STEC phylogroups among serogroup O118 were comprised of H2, H6, H12, and H16 strains. Non-STEC pathovars belonged to EPEC (H5), ETEC (H10, H51), UPEC (H28, H10, H12, H11), unclassifiable (H11), and a single strain containing virulence genes associated with both ETEC and EPEC (Figure 7, Table S1, Table S3). Notably, H10 antigen belonged to EPEC and UPEC, which formed distinct clusters, also evident in the ANI heatmap (Fig. S1).

Correlation between stx status, H-antigens, and intimin subtypes

For STEC, the following *stx/eae*-genotype profiles were recorded: H6 (stx_{2f}, *eae-ι*), H12 (stx_{2b}, *eae*(-)), H16 (stx_{1a}, *eae-β*) and H2 (stx_{1a}, *eae-ε*; stx_{2a}, *eae-ε*; stx_{1a}, stx_{2c}, *eae-ε*; and stx_{1a}, stx_{2a}, *eae-ε*) [77]. The best-represented STEC groups in the analyzed strain panel were H16 ($n = 93$) and H2 ($n = 242$). H16 STEC was genetically uniform (stx_{1a}, *eae-β*) compared to H2 STECs with significant toxicity profile heterogeneity. Among H2 STEC, the predominant *stx* genotype was stx_{1a} only ($n = 215$), followed by $stx_{1a} + stx_{2c}$ ($n = 11$), $stx_{1a} + stx_{2a}$ ($n = 5$), and stx_{2a} only ($n = 2$). Within the dominant H2 STEC, we further identified a distinct *stx* (-) sublineage. Further analyses of the H2 ΦStx_{1a} phage insertion site at *torS* (Figure 3) detected phage remnants (Figure S4). We thus speculate that this lineage evolved from ST-17, stx_{1a}+ ancestors and represents Lost Shiga Toxin (LST) isolates [100]. Their closest relatives were stx_{2a}+ strains that similarly lack stx_{1a}+. However, the draft status of these genomes hindered further investigation into the ΦStx insertion locus at *torS* to evaluate the ancestral or evolved ΦStx-status (Figure S4). We thus speculate that this STEC lineage may have secondarily acquired stx_{2a}, essentially transitioning between *stx*(+/−) states [100,175,176]. Similarly, the LEE-H12-group contained both stx_{2b}+ and *stx*- isolates. However, the draft status of the genomes hindered further investigation of the ΦStx status.

Except for the H12 STEC, the LEE island is present in all other identified STEC phylogroups (H2, H6, H12, and H16) (Figure 7). The LEE island is also present in EPEC (H5) and a strain containing genes associated with both ETEC and EPEC (H8) (Figure 7). Over 20 intimins mediating the attachment to intestinal cells are currently described [17]. Previous studies indicated a correlation between the H-antigen and the intimin subtype in LEE-positive *E. coli* [177]. Indeed, we found a correlation between H-antigens and the five *eae*-subalelles present in serogroup O118. The intimin ε-subtype is associated

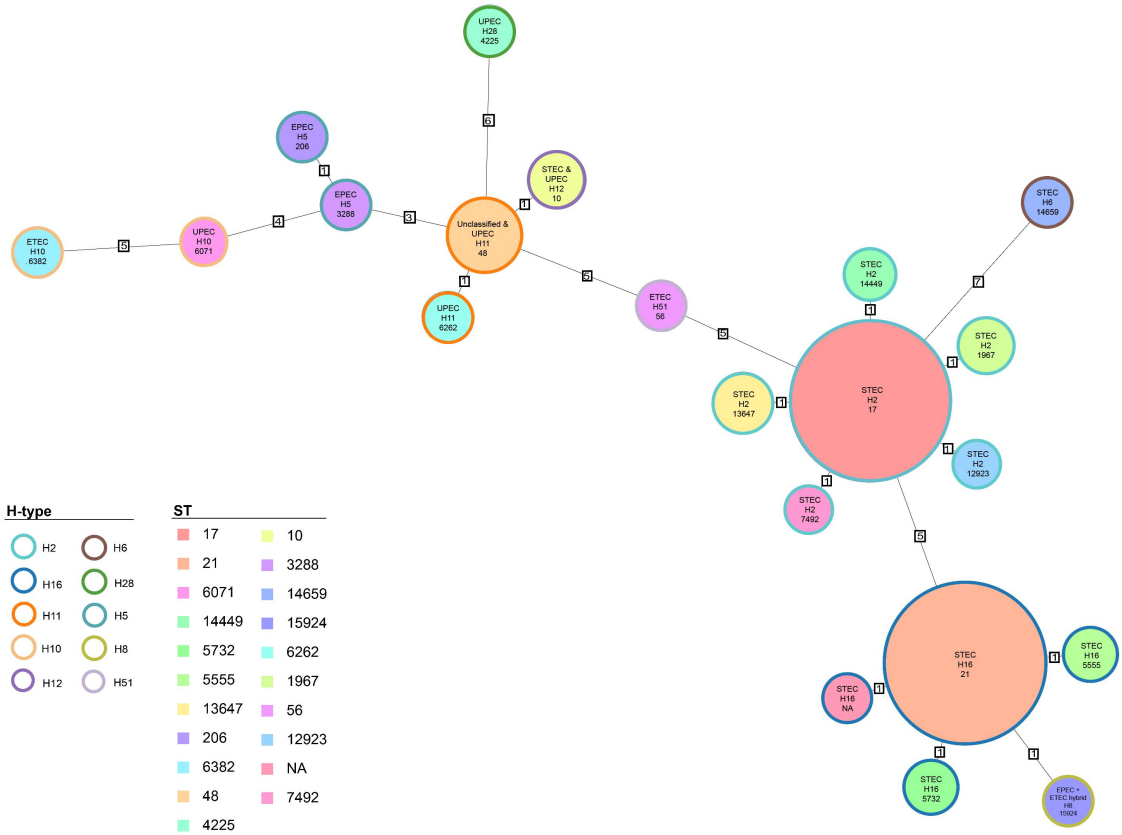

Figure 6. MLST-inferred phylogeny of serogroup O118. The relatedness of O118 strains was determined using MLST in ridom SeqSphere+ through both targeted seven-gene MLST accessed in the EnteroBase website. Numbers on connecting branches indicate the number of genes with differing allele status. The shared gene inventory of 4,160 genes comprises 2,437 core and 1,723 accessory loci according to the inclusion/exclusion criteria of the SeqSphere+ target Definer. Circle size corresponds to the number of isolates with identical st.

with H2 STEC, exemplified by strain EC20017429, and LST strains of this group (Figure 7). This group also contains *eae*-negative strain PNUSAE161722, missing integral LEE components and effectors (Figure 7). Its position in the tree and intimate relationship to ST-17 LEE+ isolates may suggest secondary loss rather than an ancestral LEE-status. The H16-phylogroup features the β-intimin, represented by strains 2014C-3050, 2013C-4538, 12867, and 12089 (Figure 4). We further observed the following correlations H6-*eae*-ι, H5-*eae*-κ, and H8-*eae*-θ, in the other LEE+ phylogroups, though these were represented by only a few strains.

Antimicrobial- and multidrug-resistance loci in serogroup O118

AR and MDR *E. coli* O118 STEC strains have been previously reported [41,99]. Consistent with these observations, AR/MDR strains were detected within STEC phylogroups H16 and H2 (Figure 8), and not in other STEC lineages. Among the resistant O118 STEC, we identified genes associated with resistance to ten antibiotic classes: β-Lactam, Chloramphenicol (CHL), Quinolone (QNL), Tetracycline (Tet), Polymyxin (PMX), Macrolide – Lincosamide – Streptogramin (MLS), Aminoglycoside, Sulfonamide, Trimethoprim, and Rifampin (RIF). Numerous MDR isolates, as defined by resistance to three or more antibiotic classes, were detected. A prominent three-class profile – tetracyclines, aminoglycosides, and sulfonamides – was observed frequently in H16 and H2 STEC, similar to the one observed in the MDRVC-Island of strain 12089 (H16, ST21) (Figure 5). Antimicrobial susceptibility testing showed that the 12089 isolate was resistant to streptomycin (≥32 µg/mL), sulfisoxazole (≥512 µg/mL), and tetracycline (≥16 µg/mL). We also noted that resistance gene presence and combinations followed phylogroup-specific patterns, as evident from the cgMLST-based phylogeny (Figure 8). However, due to the draft status of most genomes, we were limited in our ability to determine the genomic context, whether these AR loci are chromosomal (Figure 5), plasmid-borne [99], or phage-associated [12] (Figure 8).

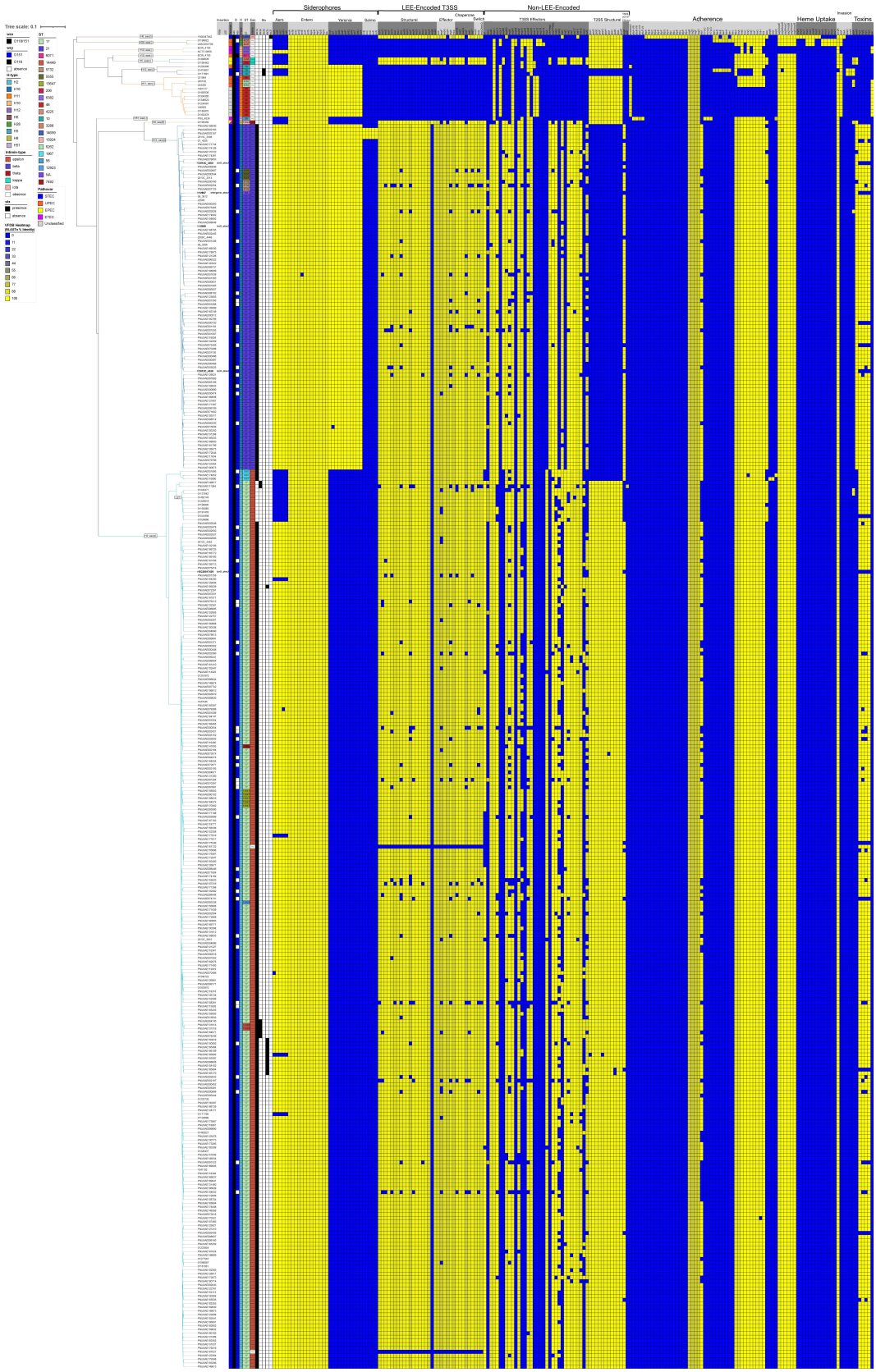

Figure 7. Phylogeny and distribution of virulence genes. The relatedness of analyzed O118/O151 strains was determined using cgMLST in ridom SeqSphere+ using the closed chromosome of *E. coli* strain K12 substrain MG1655 as seed, complemented by the 357 publicly available strains and two closed genomes strains 12089 and 12867. The shared gene inventory was determined at 4,160 genes. These comprise 2,437 core and 1,723 accessory loci according to the inclusion/exclusion criteria of the SeqSphere+ target Definer. The topology and clusters indicate a strong association between H-antigen, *eae*-subtype, and st in LEE+ isolates. A total of 199 distinct virulence genes were cataloged, of which 16 are shared. The STEC-pathogroups are comprised of H-antigens 6, 12, 16, and 2 and carry the $stx_{1a,2a,2b,2c,2f}$ suballeles.

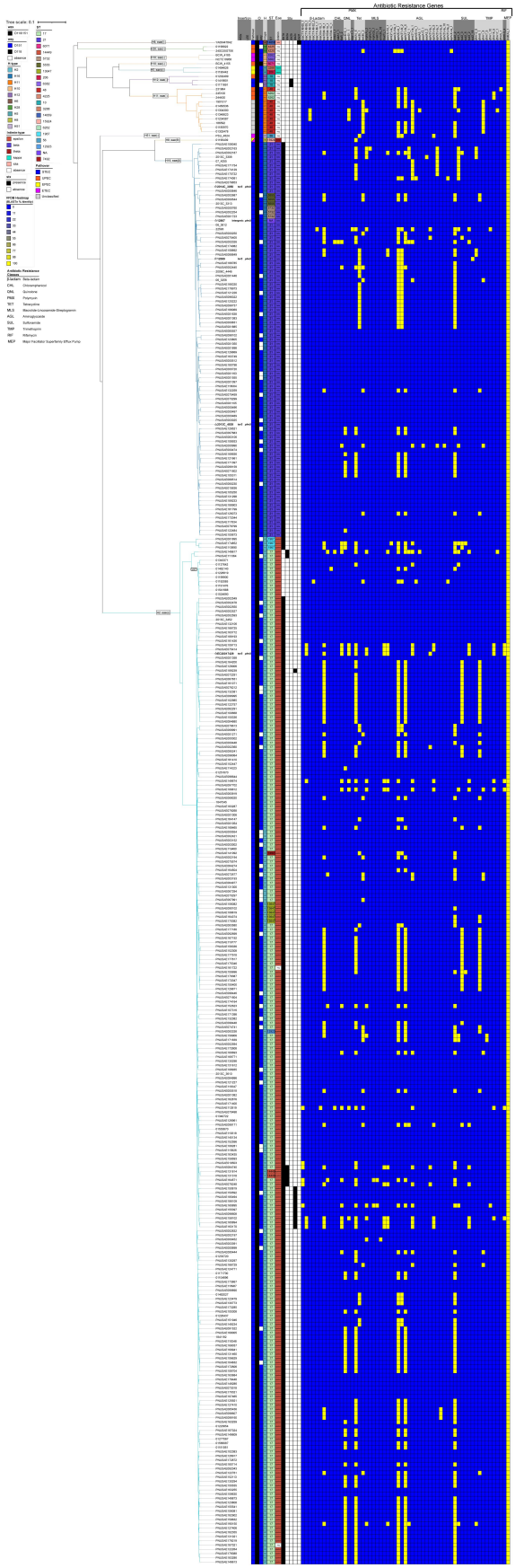

Figure 8. Phylogeny and distribution of antibiotic-resistance genes. Percentage identities for each antibiotic-resistance gene identified in ResFinder are visualized in a heatmap. The strains feature a total of 58 antibiotic-resistance genes that were classified into 10 antibiotic classes.

Conclusion

Whole-genome sequencing (WGS) typing strategies have significantly advanced our understanding of the genomic makeup and evolution of *E. coli* pathovars [11]. The high-resolution framework established for the *E. coli* serogroup O118/O151 complex allowed us to define the phylogenomic boundaries and virulence traits of emerging O118 STEC. For partitioning O118 serogroup isolates, their H-antigen was a robust phylogenetic marker [178], although genetically unrelated *E. coli* strains may share H-types due to lateral acquisition [179]. Among the delineated STEC O118 phylogroups, we observed substantial plasticity in virulence and resistance inventories, characterized by diverse *stx* and *eae* suballeles. The presence or absence of *stx* was shaped by dynamic ΦStx prophage acquisitions but did not necessarily reflect evolutionary relationships [176]. Notably, we identified a secondary Stx prophage loss event in an *stx*-negative H2 subgroup branching from ST-17 STEC, suggesting that such losses, followed by potential reacquisition, may lead to transitional shifts in pathogenic potential [100]. Our findings showed a strict correlation between the H-antigen and LEE-borne intimin subtype in STEC (H6/*eae-ι*, H16/*eae-β*, H2/*eae-ε*) and non-STEC (H5/*eae-κ*, H8/*eae-θ*) isolates, reinforcing the hypothesis of evolutionary lineage sorting [177,180]. We identified horizontally acquired pathogenicity-associated islands (PAI) that enhance virulence potential. Notable was the O-island 122 inserted into the LEE-island of clinical H2-strain EC20017429, a locus previously associated with severe disease outcomes [148–150]. The discovered multifunctional island in clinical H16-strain 12089 enhances interbacterial competitiveness, virulence, stress adaptation, and may increase the strain's fitness and pathogenic potential. However, disease outcome is multifactorial, resulting from the complex interplay between the infective agent, the host microbiota [181,182], and the infected individual [183,184], and thus cannot be reliably predicted *in silico*. The emergence of clinically relevant O118 STEC in both human and animal reservoirs, with the potential to cause severe foodborne disease, highlights the need for ongoing genomic surveillance. Insights into pathogenome plasticity and evolutionary trajectories, gained through comparative genomic approaches, are critical for informing public health strategies, risk assessment, and infection control efforts.

Implications for public health and risk assessment

By linking phylogroup structure to virulence and AMR inventories, this work aids in the phylogenetically informed surveillance, risk assessment, and outbreak investigations of emerging O118 lineages. The established phylogroup-linked and gene – content – based framework for O118 allows in-depth cluster interpretations, particularly for lineages with high-risk combinations of virulence and resistance determinants. Risk assessment can be clearly improved by integrating phylogroup membership with key virulence determinants. Across O118 STEC, we observed phylogroups spanning multiple H-types with diverse Stx profiles and frequent carriage of LEE, highlighting that the O118 serogroup contains multiple lineages with distinct virulence potential (Figure 7, Table S1, Table S3). The consistent association between H-type and *eae*-subtype (e.g. H6/*eae-ι*, H16/*eae-β*, H2/*eae-ε*) provides an additional discriminator that can be used alongside *stx*-subtyping. As reported, a subset of H2-STEC lacked *stx*, consistent with potential secondary ΦStx phage loss [100]. This observation is important for surveillance and laboratory interpretation because *stx*-negative derivatives can complicate routine screening and may be misclassified without phylogenomic context. The described AMR/MDR inventory, including a newly resolved pathogenicity-associated MDR island that harbors loci related to colonization and interbacterial competition, underscores that subsets of O118 STEC infections may exhibit enhanced pathogenic potential and limited therapeutic options.

Acknowledgements

This work received computational support from the High-Performance Computing Cluster (HPCC) operated by Tech Solutions at UTSA. The use of product and company names is necessary to accurately report the methods and results; however, the United States Department of Agriculture (USDA) neither guarantees nor warrants the standard of the products, and the use of names by the USDA implies no approval of the product to the exclusion of others that may also be suitable. The USDA is an equal opportunity provider and employer. We wish to thank Drs. Peter Iwen and Paul Fey of the University of Nebraska Medical Center College of Medicine and the Nebraska State Health Laboratory for providing *E. coli* strains 12089 and 12867. We would like to acknowledge Mariana Sainz Garcia, Felix Borrego, and Jacob Alford for assistance with data visualization.

Conceived and designed the experiments: JMB and ME. Analyzed the data: IR, SSK, JMB, and ME. Drafted the manuscript: ME. Contributed to and edited the manuscript: IR, SKK, JMB, and ME. All authors have read and agreed to the published version of this manuscript.

Author contributions

CRediT: **Irvin Rivera:** Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – review

& editing; **Sara S.K. Koenig:** Formal analysis, Investigation, Project administration, Visualization, Writing – review & editing; **Armando L. Rodriguez:** Methodology, Resources, Software; **Joseph M. Bosilevac:** Data curation, Formal analysis, Investigation, Methodology, Resources, Writing – review & editing; **Mark Eppinger:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

Research reported in this publication was supported by the National Institutes of Health under Award Number [SC1GM135110] to ME and the South Texas Center for Emerging Infectious Diseases (STCEID).

Disclosure statement

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.

Data availability statement

The sequence data sets generated and analyzed in this study have been deposited in the Sequence Read Archive (SRA) and GenBank at NCBI under BioProject PRJNA1103183 (https://www.ncbi.nlm.nih.gov/search/all/?term=PRJNA1103183). Accessions for genomic reads, assembled annotated chromosomes, and plasmids, along with strain-associated metadata, are provided in Table S1. Supplementary tables and figures are available from figshare: (https://doi.org/10.6084/m9.figshare.30194302) [185]. A preprint is available from bioRxiv (https://doi.org/10.1101/2025.04.29.651274) [186].

ORCID

Irvin Rivera http://orcid.org/0009-0005-0647-3559

References

[1] Stordeur P, China B, Charlier G, et al. Clinical signs, reproduction of attaching/effacing lesions, and enterocyte invasion after oral inoculation of an O118 enterohaemorrhagic Escherichia coli in neonatal calves. Microbes Infect. 2000;2(1):17–24. doi: 10.1016/S1286-4579(00)00290-2

[2] Karmali MA, Petric M, Lim C, et al. Escherichia coli cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis. Lancet. 1983;322(8362):1299–1300. doi: 10.1016/S0140-6736(83)91167-4

[3] Majowicz SE, Scallan E, Jones-Bitton A, et al. Global incidence of human Shiga toxin-producing Escherichia coli infections and deaths: a systematic review and knowledge synthesis. Foodborne Pathog Dis. 2014;11(6):447–455. doi: 10.1089/fpd.2013.1704

[4] Beutin L, Bulte M, Weber A, et al. Investigation of human infections with verocytotoxin-producing strains of Escherichia coli (VTEC) belonging to serogroup O118 with evidence for zoonotic transmission. Epidemiol Infect. 2000;125(1):47–54. doi: 10.1017/S0950268899004094

[5] Abu-Ali GS, Lacher DW, Wick LM, et al. Genomic diversity of pathogenic Escherichia coli of the EHEC 2 clonal complex. BMC Genomics. 2009;10(1):296. doi: 10.1186/1471-2164-10-296

[6] Maria Ferreira Cavalcanti A, Tavanelli Hernandes R, Harummyy Takagi E, et al. Virulence profiling and molecular typing of Shiga toxin-producing E. coli (STEC) from human sources in Brazil. Microorganisms. 2020;8(2):171. doi: 10.3390/microorganisms8020171

[7] Pearce MC, Evans J, McKendrick IJ, et al. Prevalence and virulence factors of Escherichia coli serogroups O26, O103, O111, and O145 shed by cattle in Scotland. Appl Environ Microbiol. 2006;72(1):653–659. doi: 10.1128/AEM.72.1.653-659.2006

[8] Nataro JP, Kaper JB. Diarrheagenic Escherichia coli. Clin Microbiol Rev. 1998;11(1):142–201. doi: 10.1128/CMR.11.1.142

[9] Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol. 2004;2(2):123–140. doi: 10.1038/nrmicro818

[10] Basavaraju M, Gunashree BS. Escherichia coli: an overview of main characteristics. In: Starčič Erjavec M, editor. Escherichia coli - Old and new insights. London: IntechOpen; 2022. p. 21. doi: 10.5772/intechopen.105508

[11] Rusconi B, Sanjar F, Koenig SS, et al. Whole genome sequencing for genomics-guided investigations of Escherichia coli O157: H7 outbreaks. Front Microbiol. 2016;7:985.

[12] Eppinger M, Mammel MK, Leclerc JE, et al. Genomic anatomy of Escherichia coli O157: H7 outbreaks. Proc Natl Acad Sci U S A. 2011;108(50):20142–20147. doi: 10.1073/pnas.1107176108

[13] Glassman H, Ferrato C, Chui L. Epidemiology of non-O157 Shiga toxin-producing Escherichia coli in the province of Alberta, Canada, from 2018 to 2021. Microorganisms. 2022;10(4):814. doi: 10.3390/microorganisms10040814

[14] Vishram B, Jenkins C, Greig DR, et al. The emerging importance of Shiga toxin-producing Escherichia coli other than serogroup O157 in England. J Med Microbiol. 2021;70(7). doi: 10.1099/jmm.0.001375

[15] Gould LH, Mody RK, Ong KL, et al. Increased recognition of non-O157 Shiga toxin-producing Escherichia coli infections in the United States during 2000–2010: epidemiologic features and comparison with E. coli O157 infections. Foodborne Pathog Dis. 2013;10(5):453–460. doi: 10.1089/fpd.2012.1401

[16] Tarr GAM, Rounds J, Vachon MS, et al. Differences in risk factors for transmission among Shiga toxin-producing Escherichia coli serogroups and stx profiles. J Infect. 2023;87(6):498–505. doi: 10.1016/j.jinf.2023.10.017

[17] Kalalah AA, Koenig SSK, Bono JL, et al. Pathogenomes and virulence profiles of representative big six non-O157 serogroup Shiga toxin-producing Escherichia coli. Front Microbiol. 2024;15:1364026. doi: 10.3389/fmicb.2024.1364026
[18] (CDC) CfDCaP. National shiga toxin-producing Escherichia coli (STEC) surveillance annual report, 2017. Atlanta GUDoHaHS; 2021.
[19] Tseng M, Sha Q, Rudrik JT, et al. Increasing incidence of non-O157 Shiga toxin-producing Escherichia coli (STEC) in Michigan and association with clinical illness. Epidemiol Infect. 2016;144(7):1394–1405. doi: 10.1017/S0950268815002836
[20] Blankenship HM, Mosci RE, Phan Q, et al. Genetic diversity of non-O157 Shiga toxin-producing Escherichia coli recovered from patients in Michigan and Connecticut. Front Microbiol. 2020;11:529. doi: 10.3389/fmicb.2020.00529
[21] Luna-Gierke RE, Griffin PM, Gould LH, et al. Outbreaks of non-O157 Shiga toxin-producing Escherichia coli infection: USA. Epidemiol Infect. 2014;142(11):2270–2280. doi: 10.1017/S0950268813003233
[22] Blankenship HM, Mosci RE, Dietrich S, et al. Population structure and genetic diversity of non-O157 Shiga toxin-producing Escherichia coli (STEC) clinical isolates from Michigan. Sci Rep. 2021;11(1):4461. doi: 10.1038/s41598-021-83775-z
[23] Alharbi MG, Al-Hindi RR, Esmael A, et al. The "Big Six": hidden emerging foodborne bacterial pathogens. Trop Med Infect Disease. 2022;7(11):356. doi: 10.3390/tropicalmed7110356
[24] Hashimoto H, Mizukoshi K, Nishi M, et al. Epidemic of gastrointestinal tract infection including hemorrhagic colitis attributable to Shiga toxin 1-producing Escherichia coli O118: H2 at a junior high school in Japan. Pediatrics. 1999;103(1):E2. doi: 10.1542/peds.103.1.e2
[25] Wieler LH, Busse B, Steinruck H, et al. Enterohemorrhagic Escherichia coli (EHEC) strains of serogroup O118 display three distinctive clonal groups of EHEC pathogens. J Clin Microbiol. 2000;38(6):2162–2169. doi: 10.1128/JCM.38.6.2162-2169.2000
[26] Orskov F, Orskov I, Furowicz AJ. Four new Escherichia coli O antigens, O148, O151, O152, O153, and one new H antigen, H50, found in strains isolated from enteric diseases in man with a discussion on the future numbering of K antigens. Acta Pathol Microbiol Scand B Microbiol Immunol. 1972;80(3):435–440. doi: 10.1111/j.1699-0463.1972.tb00057.x
[27] Steinruck H, Mochmann H, Kontny I, et al. Escherichia coli O151: k-: h10 ("krim") as etiologic agent of diarrhoea (author's transl). Zentralbl Bakteriol A. 1980;248(3):335–344. doi: 10.1016/S0174-3031(80)80007-2
[28] Wieler LH, Schwanitz A, Vieler E, et al. Virulence properties of Shiga toxin-producing Escherichia coli (STEC) strains of serogroup O118, a major group of STEC pathogens in calves. J Clin Microbiol. 1998;36(6):1604–1607. doi: 10.1128/JCM.36.6.1604-1607.1998
[29] Kruger A, Lucchesi PM. Shiga toxins and Stx phages: highly diverse entities. Microbiol (Read). 2015;161(3):451–462. doi: 10.1099/mic.0.000003
[30] Skinner C, Patfield S, Stanker LH, et al. New high-affinity monoclonal antibodies against Shiga toxin 1 facilitate the detection of hybrid Stx1/Stx2 in vivo. PLOS ONE. 2014;9(6):e99854. doi: 10.1371/journal.pone.0099854
[31] Zuppi M, Tozzoli R, Chiani P, et al. Investigation on the evolution of Shiga toxin-converting phages based on whole genome sequencing. Front Microbiol. 2020;11:1472. doi: 10.3389/fmicb.2020.01472
[32] Asadulghani M, Ogura Y, Ooka T, et al. The defective prophage pool of Escherichia coli O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. PLOS Pathog. 2009;5(5):e1000408. doi: 10.1371/journal.ppat.1000408
[33] Huang A, Friesen J, Brunton JL. Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in Escherichia coli. J Bacteriol. 1987;169(9):4308–4312. doi: 10.1128/jb.169.9.4308-4312.1987
[34] Fuller CA, Pellino CA, Flagler MJ, et al. Shiga toxin subtypes display dramatic differences in potency. Infect Immun. 2011;79(3):1329–1337. doi: 10.1128/IAI.01182-10
[35] McNichol BA, Bova RA, Torres K, et al. Switching Shiga toxin (stx) type from Stx2d to Stx2a but not Stx2c alters virulence of stx-producing Escherichia coli (STEC) strain B2F1 in streptomycin (str)-treated mice. Toxins (Basel). 2021;13(1):64. doi: 10.3390/toxins13010064
[36] Usein CR, Ciontea AS, Militaru CM, et al. Molecular characterisation of human Shiga toxin-producing Escherichia coli O26 strains: results of an outbreak investigation, Romania, February to August 2016. Euro Surveill. 2017;22(47). doi: 10.2807/1560-7917.ES.2017.22.47.17-00148
[37] Yamasaki E, Watahiki M, Isobe J, et al. Quantitative detection of Shiga toxins directly from stool specimens of patients associated with an outbreak of enterohemorrhagic Escherichia coli in Japan—quantitative Shiga toxin detection from stool during EHEC outbreak. Toxins (Basel). 2015;7(10):4381–4389. doi: 10.3390/toxins7104381
[38] Petro CD, Trojnar E, Sinclair J, et al. Shiga toxin type 1a (Stx1a) reduces the toxicity of the more potent Stx2a in vivo and in vitro. Infect Immun. 2019;87(4). doi: 10.1128/IAI.00787-18
[39] Dutta S, Pazhani GP, Nataro JP, et al. Heterogenic virulence in a diarrheagenic Escherichia coli: evidence for an EPEC expressing heat-labile toxin of ETEC. Int J Med Microbiol. 2015;305(1):47–54. doi: 10.1016/j.ijmm.2014.10.006
[40] Carrilero L, Dunn SJ, Moran RA, et al. Evolutionary responses to acquiring a multidrug resistance plasmid are dominated by metabolic functions across diverse Escherichia coli lineages. mSystems. 2023;8(1):e0071322. doi: 10.1128/msystems.00713-22
[41] Maidhof H, Guerra B, Abbas S, et al. A multiresistant clone of Shiga toxin-producing Escherichia coli O118:[H16] is spread in cattle and humans over different European countries. Appl Environ Microbiol. 2002;68(12):5834–5842. doi: 10.1128/AEM.68.12.5834-5842.2002

[42] Ogura Y, Ooka T, Iguchi A, et al. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic Escherichia coli. Proc Natl Acad Sci U S A. 2009;106 (42):17939–17944. doi: 10.1073/pnas.0903585106
[43] Caprioli A, Morabito S, Brugère H, et al. Enterohaemorrhagic Escherichia coli: emerging issues on virulence and modes of transmission. Vet Res. 2005;36(3):289–311. doi: 10.1051/vetres:2005002
[44] Liu Y, Gilchrist A, Zhang J, et al. Detection of viable but nonculturable Escherichia coli O157: H7 bacteria in drinking water and river water. Appl Environ Microbiol. 2008;74(5):1502–1507. doi: 10.1128/AEM.02125-07
[45] Community TG, Nekrutenko A, Grüning BA. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 2022;50(W1):W345–W351. doi: 10.1093/nar/gkac247
[46] Kolmogorov M, Yuan J, Lin Y, et al. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–546. doi: 10.1038/s41587-019-0072-8
[47] Tatusova T, DiCuccio M, Badretdin A, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016;44(14):6614–6624. doi: 10.1093/nar/gkw569
[48] Shen W, Sipos B, Zhao L. SeqKit2: a Swiss army knife for sequence and alignment processing. Imeta. 2024;3 (3):e191. doi: 10.1002/imt2.191
[49] Shen W, Le S, Li Y, et al. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLOS ONE. 2016;11(10):e0163962. doi: 10.1371/journal.pone.0163962
[50] Blattner FR, Plunkett G, Bloch CA, et al. The complete genome sequence of Escherichia coli K-12. Science. 1997;277(5331):1453–1462. doi: 10.1126/science.277.5331.1453
[51] Riley M, Abe T, Arnaud MB, et al. Escherichia coli K-12: a cooperatively developed annotation snapshot–2005. Nucleic Acids Res. 2006;34(1):1–9. doi: 10.1093/nar/gkj405
[52] Jünemann S, Sedlazeck FJ, Prior K, et al. Updating benchtop sequencing performance comparison. Nat Biotechnol. 2013;31(4):294–296. doi: 10.1038/nbt.2522
[53] Foley SL, Lynne AM, Nayak R. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. Infect Genet Evol. 2009;9 (4):430–440. doi: 10.1016/j.meegid.2009.03.004
[54] Zhou Z, Alikhan NF, Mohamed K, et al. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res. 2020;30(1):138–152. doi: 10.1101/gr.251678.119
[55] Díaz L, Gutierrez S, Moreno-Switt AI, et al. Diversity of non-O157 Shiga toxin-producing Escherichia coli isolated from cattle from central and southern Chile. Anim (Basel). 2021;11(8):11. doi: 10.3390/ani11082388
[56] Kruskal JB. On the shortest spanning subtree of a graph and the traveling salesman problem. Proc Amer Math Soc. 1956;7(1):48–50. doi: 10.1090/S0002-9939-1956-0078686-7
[57] Francisco AP, Bugalho M, Ramirez M, et al. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. BMC Bioinf. 2009;10(1):152. doi: 10.1186/1471-2105-10-152
[58] Darling AE, Mau B, Perna NT. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLOS ONE. 2010;5(6):e11147. doi: 10.1371/journal.pone.0011147
[59] Shimoyama Y. Pygenomeviz: a genome visualization Python package for comparative genomics. 2024.
[60] Waskom ML. Seaborn: statistical data visualization. J Open Source Softw. 2021;6(60):3021. doi: 10.21105/joss.03021
[61] Alikhan NF, Petty NK, Ben Zakour NL, et al. Blast ring image generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 2011;12(1):402. doi: 10.1186/1471-2164-12-402
[62] Shaw J, Yu YW. Fast and robust metagenomic sequence comparison through sparse chaining with skani. Nat Methods. 2023;20(11):1661–1665. doi: 10.1038/s41592-023-02018-3
[63] Liu B, Zheng D, Zhou S, et al. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022;50(D1):D912–D917. doi: 10.1093/nar/gkab1107
[64] Florensa AF, Kaas RS, Clausen P, et al. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom. 2022;8(1). doi: 10.1099/mgen.0.000748
[65] Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–410. doi: 10.1016/S0022-2836(05)80360-2
[66] Baral SK. Characterization of virulence factors in multidrug-resistant Escherichia coli isolated from intestinal and extra-intestinal clinical samples. J Manmohan Meml Inst Health Sci. 2024;9(2):13–18. doi: 10.3126/jmmihs.v9i2.71802
[67] Boudjerda D, Lahouel M. Virulence and antimicrobial resistance of Escherichia coli isolated from chicken meat, beef, and raw milk. Austral J Vet Sci. 2022;54(3):115–125. doi: 10.4067/S0719-81322022000300115
[68] Seemann T. Barrnap 0.7: rapid ribosomal RNA prediction. 2013.
[69] Couvin D, Bernheim A, Toffano-Nioche C, et al. Crisprcasfinder, an update of Crisrfinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 2018;46(W1):W246–W251. doi: 10.1093/nar/gky425
[70] Grant JR, Enns E, Marinier E, et al. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 2023;51(W1):W484–W492. doi: 10.1093/nar/gkad326
[71] Tesson F, Herve A, Mordret E, et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat Commun. 2022;13(1):2561. doi: 10.1038/s41467-022-30269-9
[72] Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27 (7):1009–1010. doi: 10.1093/bioinformatics/btr039

[73] Wishart DS, Han S, Saha S, et al. PHASTEST: faster than PHASTER, better than PHAST. Nucleic Acids Res. 2023;51(W1):W443–W450. doi: 10.1093/nar/gkad382
[74] Ogura Y, Mondal SI, Islam MR, et al. The Shiga toxin 2 production level in enterohemorrhagic Escherichia coli O157: H7 is correlated with the subtypes of toxin-encoding phage. Sci Rep. 2015;5(1):16663. doi: 10.1038/srep16663
[75] Llarena AK, Aspholm M, O'Sullivan K, et al. Replication region analysis reveals non-lambdoid Shiga toxin converting bacteriophages. Front Microbiol. 2021;12:640945. doi: 10.3389/fmicb.2021.640945
[76] Fagerlund A, Aspholm M, Węgrzyn G, et al. High diversity in the regulatory region of Shiga toxin encoding bacteriophages. BMC Genomics. 2022;23(1):230. doi: 10.1186/s12864-022-08428-5
[77] Scheutz F, Teel LD, Beutin L, et al. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing stx nomenclature. J Clin Microbiol. 2012;50(9):2951–2963. doi: 10.1128/JCM.00860-12
[78] Carrillo CD, Koziol AG, Mathews A, et al. Comparative evaluation of genomic and laboratory approaches for determination of Shiga toxin subtypes in Escherichia coli. J Food Prot. 2016;79(12):2078–2085. doi: 10.4315/0362-028X.JFP-16-228
[79] Bertelli C, Laird MR, Williams KP, et al. Islandviewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 2017;45(W1):W30–W35. doi: 10.1093/nar/gkx343
[80] Bertelli C, Brinkman FSL. Improved genomic island predictions with IslandPath-DIMOB. Bioinformatics. 2018;34(13):2161–2167. doi: 10.1093/bioinformatics/bty095
[81] Bertelli C, Tilley KE, Brinkman FSL. Microbial genomic island discovery, visualization and analysis. Brief Bioinform. 2018;20(5):1685–1698. doi: 10.1093/bib/bby042
[82] Camacho C, Coulouris G, Avagyan V, et al. Blast+: architecture and applications. BMC Bioinf. 2009;10(1):421. doi: 10.1186/1471-2105-10-421
[83] Ross K, Varani AM, Snesrud E, et al. TnCentral: a prokaryotic transposable element database and web portal for transposon analysis. mbio. 2021;12(5):e0206021. doi: 10.1128/mBio.02060-21
[84] Siguier P, Perochon J, Lestrade L, et al. Isfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 2006;34(90001):D32–6. doi: 10.1093/nar/gkj014
[85] Liu M, Li X, Xie Y, et al. Iceberg 2.0: an updated database of bacterial integrative and conjugative elements. Nucleic Acids Res. 2019;47(D1):D660–D665. doi: 10.1093/nar/gky1123
[86] Xie Z, Tang H. Isescan: automated identification of insertion sequence elements in prokaryotic genomes. Bioinformatics. 2017;33(21):3340–3347. doi: 10.1093/bioinformatics/btx433
[87] Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4(8). doi: 10.1099/mgen.0.000206
[88] Carattoli A, Zankari E, Garcia-Fernandez A, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895–3903. doi: 10.1128/AAC.02412-14
[89] Galata V, Fehlmann T, Backes C, et al. PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res. 2019;47(D1):D195–D202. doi: 10.1093/nar/gky1050
[90] Ondov BD, Treangen TJ, Melsted P, et al. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 2016;17(1):132. doi: 10.1186/s13059-016-0997-x
[91] van Heel AJ, de Jong A, Song C, et al. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res. 2018;46(W1):W278–W281. doi: 10.1093/nar/gky383
[92] Brown CL, Mullet J, Hindi F, et al. MobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. Appl Environ Microbiol. 2022;88(18):e0099122. doi: 10.1128/aem.00991-22
[93] Guragain M, Schmidt JW, Bagi LK, et al. Antibiotic resistance and disinfectant resistance among Escherichia coli isolated during red meat production. J Food Prot. 2024;87(6):100288. doi: 10.1016/j.jfp.2024.100288
[94] Administration USFaD. NARMS integrated laboratory manual for antimicrobial susceptibility testing. National antimicrobial resistance monitoring system (NARMS). 2020.
[95] Johnson TJ, Nolan LK. Pathogenomics of the virulence plasmids of Escherichia coli. Microbiol Mol Biol Rev. 2009;73(4):750–774. doi: 10.1128/MMBR.00015-09
[96] Messele YE, Trott DJ, Hasoon MF, et al. Phylogenetic analysis of Escherichia coli isolated from Australian feedlot cattle in comparison to pig faecal and poultry/human extraintestinal isolates. Antibiot (Basel). 2023;12(5):12. doi: 10.3390/antibiotics12050895
[97] Wu R, Lv L, Wang C, et al. Is26-mediated formation of a hybrid plasmid carrying mcr-1.1. Infect Drug Resist. 2022;15:7227–7234. doi: 10.2147/IDR.S390765
[98] Patel PN, Lindsey RL, Garcia-Toledo L, et al. High-quality whole-genome sequences for 77 Shiga toxin-producing Escherichia coli strains generated with PacBio sequencing. Genome Announc. 2018;6(19):00391–18. doi: 10.1128/genomeA.00391-18
[99] Abdullah S, Almusallam A, Li M, et al. Whole genome-based genetic insights of bla NDM producing clinical E. Coli Isolates Hosp Settings Pakistant. 2023. doi: 10.1128/spectrum.00584-23
[100] Nyong EC, Zaia SR, Allue-Guardia A, et al. Pathogenomes of atypical non-shigatoxigenic Escherichia coli NSF/SF O157: H7/NM: comprehensive phylogenomic analysis using closed genomes. Front Microbiol. 2020;11:619. doi: 10.3389/fmicb.2020.00619
[101] Allue-Guardia A, Koenig SSK, Martinez RA, et al. Pathogenomes and variations in Shiga toxin production among geographically distinct clones of Escherichia coli O113: H21. Microb Genom. 2022;8(4). doi: 10.1099/mgen.0.000796
[102] Rasko DA, Rosovitz MJ, Myers GS, et al. The pangenome structure of Escherichia coli: comparative

genomic analysis of E. coli commensal and pathogenic isolates. J Bacteriol. 2008;190(20):6881–6893. doi: 10.1128/JB.00619-08

[103] Jain C, Rodriguez-R LM, Phillippy AM, et al. High throughput ANI analysis of 90k prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9(1):5114. doi: 10.1038/s41467-018-07641-9

[104] Tozzoli R, Grande L, Michelacci V, et al. Shiga toxin-converting phages and the emergence of new pathogenic Escherichia coli: a world in motion. Front Cell Infect Microbiol. 2014;4:80. doi: 10.3389/fcimb.2014.00080

[105] Delannoy S, Mariani-Kurkdjian P, Webb HE, et al. The mobilome; a major contributor to Escherichia coli stx2-positive O26: h11 strains intra-serotype diversity. Front Microbiol. 2017;8:1625. doi: 10.3389/fmicb.2017.01625

[106] Perna NT, Plunkett G, Burland V, et al. Genome sequence of enterohaemorrhagic Escherichia coli O157: h7. Nature. 2001;409(6819):529–533. doi: 10.1038/35054089

[107] Alizade H, Hosseini Teshnizi S, Azad M, et al. An overview of diarrheagenic Escherichia coli in Iran: a systematic review and meta-analysis. J Res Med Sci. 2019;24(1):23. doi: 10.4103/jrms.JRMS_256_18

[108] Bielaszewska M, Mellmann A, Zhang W, et al. Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect Dis. 2011;11(9):671–676. doi: 10.1016/S1473-3099(11)70165-7

[109] Sperandio V, Hovde CJ. Enterohemorrhagic Escherichia coli and other Shiga toxin-producing E. Washington (DC): coli ASM Press; 2015.

[110] Franzin FM, Sircili MP. Locus of enterocyte effacement: a pathogenicity island involved in the virulence of enteropathogenic and enterohemorragic Escherichia coli subjected to a complex network of gene regulation. Biomed Res Int. 2015;2015:534738.

[111] Kaper JB. Pathogenic Escherichia coli. Int J Med Microbiol. 2005;295(6–7):355–356. doi: 10.1016/j.ijmm.2005.06.008

[112] Toro M, Rump LV, Cao G, et al. Simultaneous presence of insertion sequence-excision enhancer (IEE) and insertion sequence IS629 correlates with increased diversity and virulence in Shiga-toxin producing Escherichia coli (STEC). J Clin Microbiol. 2015;53(11):3466–3473. doi: 10.1128/JCM.01349-15

[113] Stanton E, Park D, Dopfer D, et al. Phylogenetic characterization of Escherichia coli O157: h7 based on IS629 distribution and Shiga toxin genotype. Microbiology. 2014;160(3):502–513. doi: 10.1099/mic.0.073437-0

[114] Yokoyama E, Hashimoto R, Etoh Y, et al. Biased distribution of IS629 among strains in different lineages of enterohemorrhagic Escherichia coli serovar O157. Infect Genet Evol. 2011;11(1):78–82. doi: 10.1016/j.meegid.2010.10.007

[115] Bryant JA, Sellars LE, Busby SJ, et al. Chromosome position effects on gene expression in Escherichia coli K-12. Nucleic Acids Res. 2014;42(18):11383–11392. doi: 10.1093/nar/gku828

[116] Raeside C, Gaffe J, Deatherage DE, et al. Large chromosomal rearrangements during a long-term evolution experiment with Escherichia coli. Mbio. 2014;5(5):e01377–14. doi: 10.1128/mBio.01377-14

[117] Le VVH, Leon-Quezada RI, Biggs PJ, et al. A large chromosomal inversion affects antimicrobial sensitivity of Escherichia coli to sodium deoxycholate. Microbiol (Read). 2022;168(8):168. doi: 10.1099/mic.0.001232

[118] Fitzgerald SF, Lupolova N, Shaaban S, et al. Genome structural variation in Escherichia coli O157: h7. Microb Genom. 2021;7(11). doi: 10.1099/mgen.0.000682

[119] Frost LS, Leplae R, Summers AO, et al. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005;3(9):722–732. doi: 10.1038/nrmicro1235

[120] Hacker J, Blum-Oehler G, Hochhut B, et al. The molecular basis of infectious diseases: pathogenicity islands and other mobile genetic elements. A review. Acta Microbiol Immunol Hung. 2003;50(4):321–330. doi: 10.1556/AMicr.50.2003.4.1

[121] Dobrindt U, Blum-Oehler G, Nagy G, et al. Genetic structure and distribution of four pathogenicity islands (PAI I(536) to PAI IV(536)) of uropathogenic Escherichia coli strain 536. Infect Immun. 2002;70(11):6365–6372. doi: 10.1128/IAI.70.11.6365-6372.2002

[122] Hashimoto JG, Stevenson BS, Schmidt TM. Rates and consequences of recombination between rRNA operons. J Bacteriol. 2003;185(3):966–972. doi: 10.1128/JB.185.3.966-972.2003

[123] Iguchi A, Iyoda S, Terajima J, et al. Spontaneous recombination between homologous prophage regions causes large-scale inversions within the Escherichia coli O157: h7 chromosome. Gene. 2006;372:199–207. doi: 10.1016/j.gene.2006.01.005

[124] Cascales E, Buchanan SK, Duche D, et al. Colicin biology. Microbiol Mol Biol Rev. 2007;71(1):158–229. doi: 10.1128/MMBR.00036-06

[125] Masaki H, Ogawa T. The modes of action of colicins E5 and D, and related cytotoxic tRNases. Biochimie. 2002;84(5–6):433–438. doi: 10.1016/S0300-9084(02)01425-6

[126] Budič M, Rijavec M, Petkovšek Ž, et al. Escherichia coli bacteriocins: antimicrobial efficacy and prevalence among isolates from patients with bacteraemia. PLOS ONE. 2011;6(12):e28769. doi: 10.1371/journal.pone.0028769

[127] Petrova V, Chitteni-Pattu S, Drees JC, et al. An SOS inhibitor that binds to free RecA protein: the PsiB protein. Mol Cell. 2009;36(1):121–130. doi: 10.1016/j.molcel.2009.07.026

[128] Gigliucci F, van Hoek A, Chiani P, et al. Genomic characterization of hlyF-positive Shiga toxin-producing Escherichia coli, Italy and the Netherlands, 2000–2019. Emerg Infect Dis. 2021;27(3):853–861. doi: 10.3201/eid2703.203110

[129] Tontanahal A, Sperandio V, Kovbasnjuk O, et al. IgG binds Escherichia coli serine protease EspP and protects mice from E. coli O157: h7 infection. Front Immunol. 2022;13:807959. doi: 10.3389/fimmu.2022.807959

[130] Yin X, Zhu J, Feng Y, et al. Differential gene expression and adherence of Escherichia coli O157: h7 in vitro and in ligated pig intestines. PLOS ONE. 2011;6(2): e17424. doi: 10.1371/journal.pone.0017424
[131] Abdel-Shafi S, El-Serwy H, El-Zawahry Y, et al. The association between icaA and icaB genes, antibiotic resistance and biofilm formation in clinical isolates of Staphylococci spp. Antibiot (Basel). 2022;11(3):11. doi: 10.3390/antibiotics11030389
[132] Anisimov AP, Shaikhutdinova RZ, Pan'kina LN, et al. Effect of deletion of the lpxM gene on virulence and vaccine potential of Yersinia pestis in mice. J Med Microbiol. 2007;56(4):443–453. doi: 10.1099/jmm.0.46880-0
[133] Smillie C, Garcillan-Barcia MP, Francia MV, et al. Mobility of plasmids. Microbiol Mol Biol Rev. 2010;74(3):434–452. doi: 10.1128/MMBR.00020-10
[134] Hu B, Perepelov AV, Liu B, et al. Structural and genetic evidence for the close relationship between Escherichia coli O71 and Salmonella enterica O28 O-antigens. FEMS Immunol Med Microbiol. 2010;59(2):161–169. doi: 10.1111/j.1574-695X.2010.00676.x
[135] Capps KM, Ludwig JB, Shridhar PB, et al. Identification, Shiga toxin subtypes and prevalence of minor serogroups of Shiga toxin-producing Escherichia coli in feedlot cattle feces. Sci Rep. 2021;11(1):8601. doi: 10.1038/s41598-021-87544-w
[136] King LA, Filliol-Toutain I, Mariani-Kurkidjian P, et al. Family outbreak of Shiga toxin-producing Escherichia coli O123: h-, France, 2009. Emerg Infect Dis. 2010;16 (9):1491–1493. doi: 10.3201/eid1609.100472
[137] Canchaya C, Fournous G, Brussow H. The impact of prophages on bacterial chromosomes. Mol Microbiol. 2004;53(1):9–18. doi: 10.1111/j.1365-2958.2004.04113.x
[138] Fortier LC, Sekulovic O. Importance of prophages to evolution and virulence of bacterial pathogens. Virulence. 2013;4(5):354–365. doi: 10.4161/viru.24498
[139] Yano B, Taniguchi I, Gotoh Y, et al. Dynamic changes in Shiga toxin (stx) 1 transducing phage throughout the evolution of O26: h11 stx-producing Escherichia coli. Sci Rep. 2023;13(1):4935. doi: 10.1038/s41598-023-32111-8
[140] Ramisetty BCM, Sudhakari PA. Bacterial 'grounded' prophages: hotspots for genetic renovation and innovation. Front Genet. 2019;10:65. doi: 10.3389/fgene.2019.00065
[141] Jerse AE, Yu J, Tall BD, et al. A genetic locus of enteropathogenic Escherichia coli necessary for the production of attaching and effacing lesions on tissue culture cells. Proc Natl Acad Sci U S A. 1990;87 (20):7839–7843. doi: 10.1073/pnas.87.20.7839
[142] McDaniel TK, Jarvis KG, Donnenberg MS, et al. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc Natl Acad Sci U S A. 1995;92(5):1664–1668. doi: 10.1073/pnas.92.5.1664
[143] Sperandio V, Kaper JB, Bortolini MR, et al. Characterization of the locus of enterocyte effacement (LEE) in different enteropathogenic Escherichia coli (EPEC) and Shiga-toxin producing Escherichia coli (STEC) serotypes. FEMS Microbiol Lett. 1998;164 (1):133–139. doi: 10.1111/j.1574-6968.1998.tb13078.x
[144] Furniss RCD, Clements A. Regulation of the locus of enterocyte effacement in attaching and effacing pathogens. J Bacteriol. 2018;200(2). doi: 10.1128/JB.00336-17
[145] Jores J, Rumer L, Wieler LH. Impact of the locus of enterocyte effacement pathogenicity island on the evolution of pathogenic Escherichia coli. Int J Med Microbiol. 2004;294(2–3):103–113. doi: 10.1016/j.ijmm.2004.06.024
[146] Sváb D, Falgenhauer L, Mag T, et al. Genomic diversity, virulence gene, and prophage arrays of bovine and human Shiga toxigenic and enteropathogenic Escherichia coli strains isolated in Hungary. Front Microbiol. 2022;13:896296. doi: 10.3389/fmicb.2022.896296
[147] Saile N, Schuh E, Semmler T, et al. Determination of virulence and fitness genes associated with the pheU, pheV and selC integration sites of LEE-negative food-borne Shiga toxin-producing Escherichia coli strains. Gut Pathog. 2018;10(1):43. doi: 10.1186/s13099-018-0271-8
[148] Memariani M, Najar-Peerayeh S, Zahraei Salehi T. Multi locus VNTR (MLVA) typing and detection of the OI-122 pathogenicity island in typical and atypical enteropathogenic Escherichia coli isolated from children with acute diarrhea. Arch Clin Infect Dis. 2019;14 (3):e65855. doi: 10.5812/archcid.65855
[149] Ju W, Shen J, Toro M, et al. Distribution of pathogenicity islands OI-122, OI-43/48, and OI-57 and a high-pathogenicity island in Shiga toxin-producing Escherichia coli. Appl Environ Microbiol. 2013;79 (11):3406–3412. doi: 10.1128/AEM.03661-12
[150] Konczy P, Ziebell K, Mascarenhas M, et al. Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing Escherichia coli. J Bacteriol. 2008;190(17):5832–5840. doi: 10.1128/JB.00480-08
[151] Desvaux M, Dalmasso G, Beyrouthy R, et al. Pathogenicity factors of genomic islands in intestinal and extraintestinal Escherichia coli. Front Microbiol. 2020;11:2065. doi: 10.3389/fmicb.2020.02065
[152] Azpiroz MF, Bascuas T, Lavina M. Microcin H47 system: an Escherichia coli small genomic island with novel features. PLOS ONE. 2011;6(10):e26179. doi: 10.1371/journal.pone.0026179
[153] Ziebell K, Johnson RP, Kropinski AM, et al. Gene cluster conferring streptomycin, sulfonamide, and tetracycline resistance in Escherichia coli O157: h7 phage types 23, 45, and 67. Appl Environ Microbiol. 2011;77 (5):1900–1903. doi: 10.1128/AEM.01934-10
[154] Partridge SR, Kwong SM, Firth N, et al. Mobile genetic elements associated with antimicrobial resistance. Clin Microbiol Rev. 2018;31(4):31. doi: 10.1128/CMR.00088-17
[155] Haniford DB, Ellis MJ. Transposons Tn10 and Tn5. Microbiol Spectr. 2015;3(1):MDNA3–0002–2014. doi: 10.1128/microbiolspec.MDNA3-0002-2014
[156] Chopra I. Plasmid-determined tetracycline resistance in Escherichia coli K12: lack of evidence that resistance is related to changes in lipid metabolism [proceedings]. Biochem Soc Trans. 1978;6(2):431–433. doi: 10.1042/bst0060431

[157] Aldema ML, McMurry LM, Walmsley AR, et al. Purification of the Tn10-specified tetracycline efflux antiporter TetA in a native state as a polyhistidine fusion protein. Mol Microbiol. 1996;19(1):187–195. doi: 10.1046/j.1365-2958.1996.359886.x

[158] Awosile B, Fritzler J, Levent G, et al. Genomic characterization of fecal Escherichia coli isolates with reduced susceptibility to beta-lactam antimicrobials from wild hogs and coyotes. Pathogens. 2023;12(7):929. doi: 10.3390/pathogens12070929

[159] Beck CM, Willett JL, Cunningham DA, et al. CdIA effectors from uropathogenic Escherichia coli use heterotrimeric osmoporins as receptors to recognize target bacteria. PLoS Pathog. 2016;12(10):e1005925. doi: 10.1371/journal.ppat.1005925

[160] Bartelli NL, Passanisi VJ, Michalska K, et al. Proteolytic processing induces a conformational switch required for antibacterial toxin delivery. Nat Commun. 2022;13(1):5078. doi: 10.1038/s41467-022-32795-y

[161] Heller DM, Tavag M, Hochschild A. Cbta toxin of Escherichia coli inhibits cell division and cell elongation via direct and independent interactions with FtsZ and MreB. PLoS Genet. 2017;13(9):e1007007. doi: 10.1371/journal.pgen.1007007

[162] Wang Q, Wei Y, Huang Y, et al. Z3495, a LysR-type transcriptional regulator encoded in O island 97, regulates virulence gene expression in enterohemorrhagic Escherichia coli O157: h7. Microorganisms. 2024;12(1):140. doi: 10.3390/microorganisms12010140

[163] Taschuk F, Cherry S. Dead-box helicases: sensors, regulators, and effectors for antiviral defense. Viruses. 2020;12(2):181. doi: 10.3390/v12020181

[164] Wang X, Kim Y, Ma Q, et al. Cryptic prophages help bacteria cope with adverse environments. Nat Commun. 2010;1(1):147. doi: 10.1038/ncomms1146

[165] Camara B, Liu M, Reynolds J, et al. T7 phage protein Gp2 inhibits the Escherichia coli RNA polymerase by antagonizing stable DNA strand separation near the transcription start site. Proc Natl Acad Sci U S A. 2010;107(5):2247–2252. doi: 10.1073/pnas.0907908107

[166] Li F, Cao L, Bahre H, et al. Patatin-like phospholipase CapV in Escherichia coli - morphological and physiological effects of one amino acid substitution. NPJ Biofilms Microbiomes. 2022;8(1):39. doi: 10.1038/s41522-022-00294-z

[167] Junkermeier EH, Hengge R. A novel locally c-di-GMP-controlled exopolysaccharide synthase required for bacteriophage N4 infection of Escherichia coli. Mbio. 2021;12(6):e0324921. doi: 10.1128/mbio.03249-21

[168] Tsutakawa SE, Lafrance-Vanasse J, Tainer JA. The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. DNA Repair (Amst). 2014;19:95–107. doi: 10.1016/j.dnarep.2014.03.022

[169] Katsiou E, Nickel CM, Garcia AF, et al. Molecular analysis and identification of the radC gene from the phototrophic bacterium Rhodobacter capsulatus B10. Microbiol Res. 1999;154(3):233–239. doi: 10.1016/S0944-5013(99)80020-2

[170] Liu Y, Fratamico P, Debroy C, et al. DNA sequencing and identification of serogroup-specific genes in the Escherichia coli O118 O antigen gene cluster and demonstration of antigenic diversity but only minor variation in DNA sequence of the O antigen clusters of E. coli O118 and O151. Foodborne Pathog Dis. 2008;5(4):449–457. doi: 10.1089/fpd.2008.0096

[171] Liu B, Furevi A, Perepelov AV, et al. Structure and genetics of Escherichia coli O antigens. FEMS Microbiol Rev. 2020;44(6):655–683. doi: 10.1093/femsre/fuz028

[172] Liu B, Perepelov AV, Guo D, et al. Structural and genetic relationships between the O-antigens of Escherichia coli O118 and O151. FEMS Immunol Med Microbiol. 2010;60(3):199–207. doi: 10.1111/j.1574-695X.2010.00738.x

[173] Macnab R. Flagella and motility. In: Curtiss R, editor. Escherichia coli and Salmonella: cellular and molecular biology. ASM Press; 1996. p. 123–145.

[174] Reid SD, Selander RK, Whittam TS. Sequence diversity of flagellin (fliC) alleles in pathogenic Escherichia coli. J Bacteriol. 1999;181(1):153–160. doi: 10.1128/JB.181.1.153-160.1999

[175] Kalalah AA, Koenig SSK, Feng P, et al. Pathogenomes of Shiga toxin positive and negative Escherichia coli O157: h7 strains TT12A and TT12B: comprehensive phylogenomic analysis using closed genomes. Microorganisms. 2024;12(4):699. doi: 10.3390/microorganisms12040699

[176] Ferdous M, Zhou K, Mellmann A, et al. Is Shiga toxin-negative Escherichia coli O157: h7 enteropathogenic or enterohemorrhagic Escherichia coli? Comprehensive molecular analysis using whole-genome sequencing. J Clin Microbiol. 2015;53(11):3530–3538. doi: 10.1128/JCM.01899-15

[177] Oswald E, Schmidt H, Morabito S, et al. Typing of intimin genes in human and animal enterohemorrhagic and enteropathogenic Escherichia coli: characterization of a new intimin variant. Infect Immun. 2000;68(1):64–71. doi: 10.1128/IAI.68.1.64-71.2000

[178] Schutz K, Cowley LA, Shaaban S, et al. Evolutionary context of non-sorbitol-fermenting Shiga toxin-producing Escherichia coli O55: h7. Emerg Infect Dis. 2017;23(12):1966–1973. doi: 10.3201/eid2312.170628

[179] Wang L, Rothemund D, Curd H, et al. Species-wide variation in the Escherichia coli flagellin (H-antigen) gene. J Bacteriol. 2003;185(9):2936–2943. doi: 10.1128/JB.185.9.2936-2943.2003

[180] Strachan NJ, Rotariu O, Lopes B, et al. Whole genome sequencing demonstrates that geographic variation of Escherichia coli O157 genotypes dominates host association. Sci Rep. 2015;5(1):14145. doi: 10.1038/srep14145

[181] Gamage SD, Patton AK, Strasser JE, et al. Commensal bacteria influence Escherichia coli O157: h7 persistence and Shiga toxin production in the mouse intestine.

Infect Immun. 2006;74(3):1977–1983. doi: 10.1128/IAI.74.3.1977-1983.2006

[182] Figler HM, Dudley EG. The interplay of Escherichia coli O157: h7 and commensal E. coli: the importance of strain-level identification. Expert Rev Gastroenterol Hepatol. 2016;10(4):415–417. doi: 10.1586/17474124.2016.1155449

[183] Wong CS, Jelacic S, Habeeb RL, et al. The risk of the hemolytic-uremic syndrome after antibiotic treatment of Escherichia coli O157: h7 infections. N Engl J Med. 2000;342(26):1930–1936. doi: 10.1056/NEJM200006293422601

[184] Dundas S, Todd WT, Stewart AI, et al. The central Scotland Escherichia coli O157: h7 outbreak: risk factors for the hemolytic uremic syndrome and death among hospitalized patients. Clin Infect Dis. 2001;33(7):923–931. doi: 10.1086/322598

[185] Rivera I, Koenig S, Rodriguez A, et al. Supplementary material - phylogenomic framework and virulence gene boundaries of emerging Shiga toxin-producing *Escherichia coli* O118 informed by the comprehensive profiling of 359 O118 genomes. Figshare. 2026. doi: 10.6084/m9.figshare.30194302.v6.figshare

[186] Rivera I, Konig SSK, Rodriguez AL, et al. Phylogenomic framework and virulence gene boundaries of emerging Shiga toxin-producing Escherichia coli O118 informed by the comprehensive profiling of 359 O118 genomes. bioRxiv. 2025. doi: 10.1101/2025.04.29.651274