Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2026.132

RESEARCH ARTICLE

Isolation and Whole-Genome Analysis of Multidrug-Resistant *Escherichia coli* from Diarrheic Fecal Samples in a Sheep Farm in Ningxia, China

Ziye Zhu[1], Ziyi Xu[1], Xinyue Wu[1], Luying Yang[1], Mingming Liu[2*], Shahid Hussain Farooqi[3] and Kun Li[1*]

[1]College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, 210095, P.R. China; [2]School of Basic Medicine, Hubei University of Arts and Science, Xiangyang, 441000, China; [3]Department of Clinical Sciences, KBCMA College of Veterinary and Animal Sciences, Narowal, Sub-campus UVAS-Lahore, Pakistan.
*Corresponding author: lmm_2010@hotmail.com (ML), lk3005@njau.edu.cn (KL)

ARTICLE HISTORY (26-250)

Received: March 13, 2026
Revised: May 30, 2026
Accepted: June 19, 2026
Published online: June 28, 2026

Key words:
Escherichia coli
Gene
Whole genome sequencing
Resistance
Virulence

ABSTRACT

Multidrug-resistant *Escherichia coli* in livestock is an increasing threat to animals and public health. This study aimed to investigate the genomic characteristics and antimicrobial resistance mechanisms of animal-derived multidrug-resistant *Escherichia coli*. A clinical strain, B19, was isolated from diarrheic fecal samples collected from sheep in Ningxia, China. The strain was subjected to hybrid second- and third-generation whole-genome sequencing followed by comprehensive annotation. The complete genome (total size 4,901,302 bp; 50.65% GC) exhibited high genomic integrity and plasticity. In silico typing revealed that strain B19 belongs to phylogroup B1, sequence type ST101, and serotype Ont:H38. Functional analyses revealed enrichment of genes associated with carbohydrate metabolism and cellular processes, indicating strong metabolic adaptability. Secondary metabolite gene clusters, including NRPS-Metallophore and Terpene pathways, suggested potential ecological roles in iron acquisition and stress defense. CARD analysis identified multiple antimicrobial resistance determinants, predominantly against fluoroquinolones, aminoglycosides, and macrolides, along with efflux pump systems (MdtABC, AcrD). Virulence profiling using VFDB revealed abundant adhesion and iron-transport-related genes, supporting a high colonization and survival potential. Overall, strain B19 carries extensive resistance and virulence determinants, highlighting its adaptive genomic strategies and providing valuable insights into the evolution and dissemination of multidrug-resistant *E. coli* in livestock.

To Cite This Article: Zhu Z, Xu Z, Wu X, Yang L, Liu M, Farooqi S, Li K, 2026. Isolation and whole-genome analysis of multidrug-resistant *Escherichia coli* from diarrheic fecal samples in a sheep farm in Ningxia, China. Pak Vet J, 46(6): 1439-1451. http://dx.doi.org/10.29261/pakvetj/2026.132

INTRODUCTION

Escherichia coli is a common bacterial species that naturally inhabits the intestinal tracts of humans and animals. As an important component of the commensal gut microbiota, it contributes to intestinal homeostasis and microecological balance (Kaper *et al.*, 2004). However, certain strains can acquire specific virulence genes and evolve into opportunistic pathogens, leading to various infectious diseases such as diarrhea, urinary tract infections, septicemia, and meningitis (Pakbin *et al.*, 2021; Ali *et al*., 2022). *E. coli* pathotypes remains a major health issue in livestock for neonatal calf diarrhea (Özavci *et al.,* 2026). Diarrheagenic *E. coli* (DEC) is categorized into six major pathotypes: enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), Shiga toxin-producing *E. coli* (STEC), enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC), and diffusely adherent *E. coli* (DAEC) (Gomes *et al.*, 2016; Rakhalaru *et al.*, 2023). These pathotypes possess specialized virulence factors that play critical roles in adhesion, invasion, toxin secretion, motility, and iron acquisition (Croxen and Finlay, 2010), and can cause clinical manifestations ranging from mild diarrhea to hemorrhagic colitis and hemolytic uremic syndrome (Croxen *et al.*, 2013). The isolation of foodborne pathogens from sheep handlers living in close contact with sheep poses a serious public health concern (Gurjar *et al.,* 2024; Yaseen *et al.,* 2025).

Small ruminants, particularly sheep, are recognized as reservoirs for DEC. Pathogenic *E. coli* strains with zoonotic potential are frequently isolated in sheep feces, posing potential threats to animal health and the

transmission of zoonotic infections (Koleci *et al.*, 2025). In recent years, the growing scale of livestock production and frequent use of antibiotics in animal farming have accelerated the emergence and spread of multidrug-resistant (MDR) *E. coli* in farm environments (Pinheiro *et al.*, 2020; Haulisah *et al.*, 2021). Recent studies have reported that *E. coli* isolates producing extended-spectrum β-lactamases (ESBLs) in sheep and slaughterhouse environments frequently carry CTX-M-type β-lactamase genes, as well as resistance determinants to aminoglycosides, quinolones, and other antibiotic classes, exhibiting a typical multidrug resistance profile (Atlaw *et al.*, 2021; Nawaz *et al*., 2021). Moreover, the intestinal microbiota of livestock can act as a reservoir of antimicrobial resistance genes, facilitating horizontal gene transfer among bacterial species and promoting the spread of resistant traits (Luo *et al.*, 2023). Therefore, understanding the resistance mechanisms in environmental and animal-derived *E. coli* is essential for epidemiological surveillance and public health protection.

Whole-genome sequencing (WGS) has become a powerful approach for elucidating the genetic basis of virulence and antimicrobial resistance in pathogenic microorganisms. Comprehensive genomic annotation enables the identification of resistance genes and virulence factors, as well as systematic evaluation of their functional categories, distribution characteristics, and potential for transmission (Gao *et al.*, 2018). Genome-based analyses help clarify evolutionary relationships, reveal mechanisms of adaptation, and support risk assessment and disease control strategies for livestock-associated pathogens.

Given the increasing occurrence of multidrug-resistant *E. coli* in livestock, understanding the genomic characteristics of strains associated with ovine diarrhea is important for evaluating their pathogenic and resistance potential. In this study, we analyzed a multidrug-resistant *E. coli* strain isolated from diarrheic sheep in Ningxia. A whole-genome approach was applied to characterize its genomic features, including overall genome composition, codon usage patterns, phylogenetic relationships, major functional gene categories, secondary metabolite biosynthetic potential, and key virulence and antimicrobial resistance determinants. These analyses provide a comprehensive genetic overview of the strain and contribute to a deeper understanding of *E. coli* occurring in sheep populations.

MATERIALS AND METHODS

Bacterial Strain Source: *Escherichia coli* quality control strain (ATCC 25922) was procured from the China Institute of Veterinary Drug Control. In September 2024, a total of 19 *E. coli* isolates were obtained from diarrheic sheep fecal samples collected from a sheep farm in Ningxia, China. All 19 isolates were subjected to antimicrobial susceptibility testing for preliminary characterization. Antimicrobial susceptibility testing was performed using the disk diffusion method on Mueller-Hinton agar plates. The results were interpreted according to the Clinical and Laboratory Standards Institute (CLSI) document M100 (most recent edition), and *E. coli* ATCC 25922 was used as the quality control strain for each batch of testing. Subsequently, one multidrug-resistant strain exhibiting resistance to seven antibiotics-cefotaxime, ampicillin, streptomycin, ciprofloxacin, trimethoprim-sulfamethoxazole, ceftazidime, and tetracycline—was selected and designated as strain B19 (Table 1).

Table 1: Summary of antimicrobial resistance patterns among 19 *Escherichia coli* isolates

Isolate ID	AMK	CTX	AMP	GEN	STR	CIP	SXT	CAZ	PB	TET	Number of resistant antibiotics
B1	S	R	R	R	R	R	S	S	S	S	5
B3	S	R	R	R	R	R	R	I	S	S	6
B5	S	R	R	R	S	R	R	I	S	R	6
B6	S	S	R	R	S	R	R	S	S	S	4
B7	R	S	R	R	R	R	R	I	S	S	6
B9	R	S	R	R	R	R	R	S	S	S	6
B10	S	R	R	R	R	S	R	I	S	R	6
B12	S	S	R	R	R	R	R	S	S	R	6
B14	R	S	R	R	R	R	R	S	S	S	6
B15	R	S	R	R	R	R	R	S	S	S	6
B16	R	S	R	R	R	R	R	I	S	S	6
B17	S	S	R	R	R	R	R	I	S	R	6
B19	S	R	R	S	R	R	R	S	R	R	7
B20	S	R	R	S	R	R	R	S	S	R	6
B22	R	S	R	R	R	S	R	S	S	R	6
A7	S	S	R	R	R	R	R	S	S	R	6
A8	S	S	R	R	R	R	R	S	S	R	6
C2	S	S	R	R	R	R	R	S	S	R	6
D1	S	S	R	R	R	R	R	S	S	R	6

AMK, amikacin; CTX, cefotaxime; AMP, ampicillin; GEN, gentamicin; STR, streptomycin; CIP, ciprofloxacin; SXT, trimethoprim-sulfamethoxazole; CAZ, ceftazidime; PB, polymyxin B; TET, tetracycline. R, resistant; I, intermediate; S, susceptible.

DNA Extraction, Genome Sequencing and Assembly: Genomic DNA was extracted from strain B19 using the Qiagen bacterial genomic DNA extraction kit, following the manufacturer's protocol. DNA integrity was assessed by 0.75% agarose gel electrophoresis, and purity was evaluated using OD260/280 and OD260/230 ratios measured on a Nanodrop spectrophotometer. DNA concentration was quantified using the Qubit dsDNA HS assay. High-quality DNA was sent to Wuhan Baiyi Biotechnology Co., Ltd. for whole-genome sequencing on the PromethION platform (Oxford Nanopore Technologies, ONT).

High-molecular-weight DNA was subjected to damage repair, end repair, and barcoding using ONT NBD104 and NBD114 kits. Sequencing adapters were ligated to the barcoded fragments to construct the PromethION library. Library concentration was quantified using Qubit prior to loading onto the PromethION flow cell for long-read sequencing.

For Illumina short-read sequencing, a paired-end library with an insert size of 350 bp was constructed and sequenced on the NovaSeq platform (Illumina, San Diego, CA, USA) using the PE150 strategy, following the manufacturer's protocols.

Raw ONT reads were quality-filtered using NanoFilt with minimum read length ≥1,000 bp and Q-score ≥7. This cutoff was selected as a commonly used criterion for Nanopore long-read assembly to balance sequencing depth and read quality. After filtering, 258,518 reads (96.7% of the raw reads) were retained, yielding a total of 1,168,498,218 bp, with a mean read length of 4,520 bp and an N50 read length of 7,490 bp. High-quality reads were assembled using Unicycler v0.5.1, which integrates

long reads for scaffold extension. The draft assembly was polished using Pilon (Illumina short-read polishing) and NextPolish to correct base-level errors. Circularization of chromosomal and plasmid contigs was performed using Circlator. Sequencing depth and genome coverage were evaluated using Minimap2 and SAMtools, and plasmid sequences were identified through BLASTn searches against the NCBI plasmid database. Assembly quality was assessed using QUAST, and genome completeness was evaluated using BUSCO with the Enterobacterales lineage dataset.

Genome Annotation and Functional Prediction: Protein-coding sequences (CDSs) were predicted using Prodigal v2.6.3**.** Transfer RNA (tRNA) and ribosomal RNA (rRNA) genes were identified using tRNAscan-SE and RNAmmer**,** respectively. Additional non-coding RNAs were detected using Infernal with the Rfam database. CRISPR arrays were predicted using MinCED**,** whereas genomic islands and prophage regions were detected using IslandPath-DIMOB and PhiSpy**,** respectively.

Functional annotation was performed by aligning predicted proteins against multiple databases, including CAZy, KEGG, COG, GO, RefSeq, Pfam, SwissProt, and TIGRFAMs, using BLASTp (e-value < 1e−5) and InterProScan. KEGG pathway assignment was conducted using KEGG Automatic Annotation Server (KAAS) with the bidirectional best-hit (BBH) method. COG functional categories and GO terms (molecular function, cellular component, biological process) were assigned using eggNOG-mapper against the eggNOG 5.0 database. Secondary metabolite biosynthetic gene clusters were predicted using antiSMASH (default parameters). Signal peptides and transmembrane helices were identified using SignalP and TMHMM, respectively. Genomic visualization, including genome structure and functional distribution, was generated using Circos.

Genome-Based Typing and Serotyping: MLST analysis was performed using the MLST Finder available at the Center for Genomic Epidemiology (CGE). Serotype prediction was conducted using SerotypeFinder, and phylogenetic grouping was determined using ClermonTyping.

Phylogenetic Analysis: Whole-genome sequences of *Escherichia coli* B19 together with publicly available reference genomes downloaded from NCBI were analyzed using Parsnp under a Linux operating system to obtain a core-genome alignment. Based on this alignment, a phylogenetic tree was inferred with the maximum-likelihood method implemented in IQ-TREE v2. The optimal nucleotide substitution model (GTR+G) was determined by ModelFinder. Statistical support for tree topology was evaluated using 1,000 ultrafast bootstrap iterations. *Salmonella enterica* subsp. *enterica* LT2 (AE006468.2) served as the outgroup for tree rooting, and the resulting phylogeny was displayed and annotated using iTOL.

Codon Usage and RSCU Analysis: All coding sequences (CDSs) of *E. coli* B19 were extracted from the annotated genome. Codon frequency and relative synonymous codon usage (RSCU) values were calculated using a custom R script. Before analysis, thymine (T) in CDSs was converted to uracil (U), and all codons in this study are reported in mRNA notation (U instead of T).

RSCU values were determined as the ratio of the observed codon frequency to the expected frequency under equal usage of synonymous codons. Codons with RSCU > 1.0 were considered preferentially used. All codon usage visualizations were generated using the R package ggplot2.

Virulence and Antimicrobial Resistance Gene Identification: Antimicrobial resistance (AMR) genes were identified using Abricate with the Comprehensive Antibiotic Resistance Database (CARD) version 20231121. Hits with identity ≥80% and coverage ≥70% were retained. Virulence genes were detected using VFDB (core dataset) version 20240322 through Abricate with the same cutoff (identity ≥70%, coverage ≥70%). VFDB results were classified into major functional categories, including adhesion, invasion, toxin production, and iron acquisition. These analyses provided an overall assessment of the virulence and resistance potential of strain B19.

RESULTS

The genome of *Escherichia coli* strain B19 was sequenced using a combination of second- and third-generation sequencing technologies. The circular genome map of strain B19 is shown in Fig. 1a. The complete genome of strain B19 consisted of one circular chromosome of 4,671,443 bp and three circular plasmids of 108,446 bp, 81,824 bp, and 39,589 bp, with sequencing depths of 210.55× for the chromosome and 259.48×, 257.96×, and 542.59× for the three plasmids, respectively. The total genome size was 4,901,302 bp with a GC content of 50.65%. A total of 4,598 protein-coding genes were predicted, with a cumulative length of 4,281,141 bp and an average gene length of 931 bp, accounting for 87.35% of the entire genome sequence. Additionally, the annotation also identified seven genomic islands, one prophage region, and two CRISPR structures. A total of 218 repeat sequences were detected, with an average fragment length of 95 bp. Among the non-coding RNAs, 88 tRNA genes and 22 rRNA genes were predicted, the latter consisting of seven 16S, seven 23S, and eight 5S rRNAs. Furthermore, three putative plasmid sequences were identified, all exhibiting circular topologies, with GC contents ranging from 44.21% to 51.79%. The assembly quality was high, with uniform sequencing depth and excellent genome coverage.

Genomic typing and serotype identification: Multilocus sequence typing (MLST) classified strain B19 as sequence type ST101. In silico serotyping assigned the strain to serotype Ont:H38, where Ont denotes an O-non-typeable serotype. Phylogenetic grouping further placed strain B19 into phylogroup B1.

Codon Usage Analysis: Codon usage analysis of the *E. coli* B19 genome revealed a clear and non-random bias

toward specific synonymous codons (Table 2). Overall, B19 showed a strong preference for codons ending with C or G, a pattern that aligns with the typical codon usage tendencies reported for *E. coli*. These trends are visually illustrated in Fig. 1b. The RSCU analysis further confirmed these patterns, demonstrating that several amino acids were predominantly encoded by GC-ending codons, including alanine (GCC, GCG), glycine (GGC), valine (GUG), proline (CCG), and arginine (CGC).

For amino acids with multiple synonymous codons, one or two codons exhibited dominant usage. Glycine was most frequently encoded by GGC, followed by GGU, whereas GGA and GGG were used at substantially lower frequencies. Alanine displayed a strong preference for GCG, the most frequently used codon for this amino acid. Aspartic acid favored GAU and GAC, and asparagine preferentially used AAC. Arginine showed an evident bias toward CGC, with other synonymous codons occurring at much lower frequencies. These patterns reflect the overall translational preference of B19 for codons that enhance decoding efficiency. For stop codons, UAA was the most frequently used, exceeding UAG and UGA, suggesting a preferred translational termination signal in B19.

Phylogenetic Analysis: A phylogenetic tree was constructed to investigate the evolutionary relationship of *Escherichia coli* strain B19 with representative *Escherichia* species and closely related bacteria. A total of 14 reference genomes were included in the analysis, with *Salmonella enterica* subsp. *enterica* LT2 (AE006468.2) as the outgroup. All accession numbers are indicated in Fig. 2. Maximum likelihood (ML) analysis was performed with 1,000 bootstrap replicates to assess the robustness of the inferred tree.

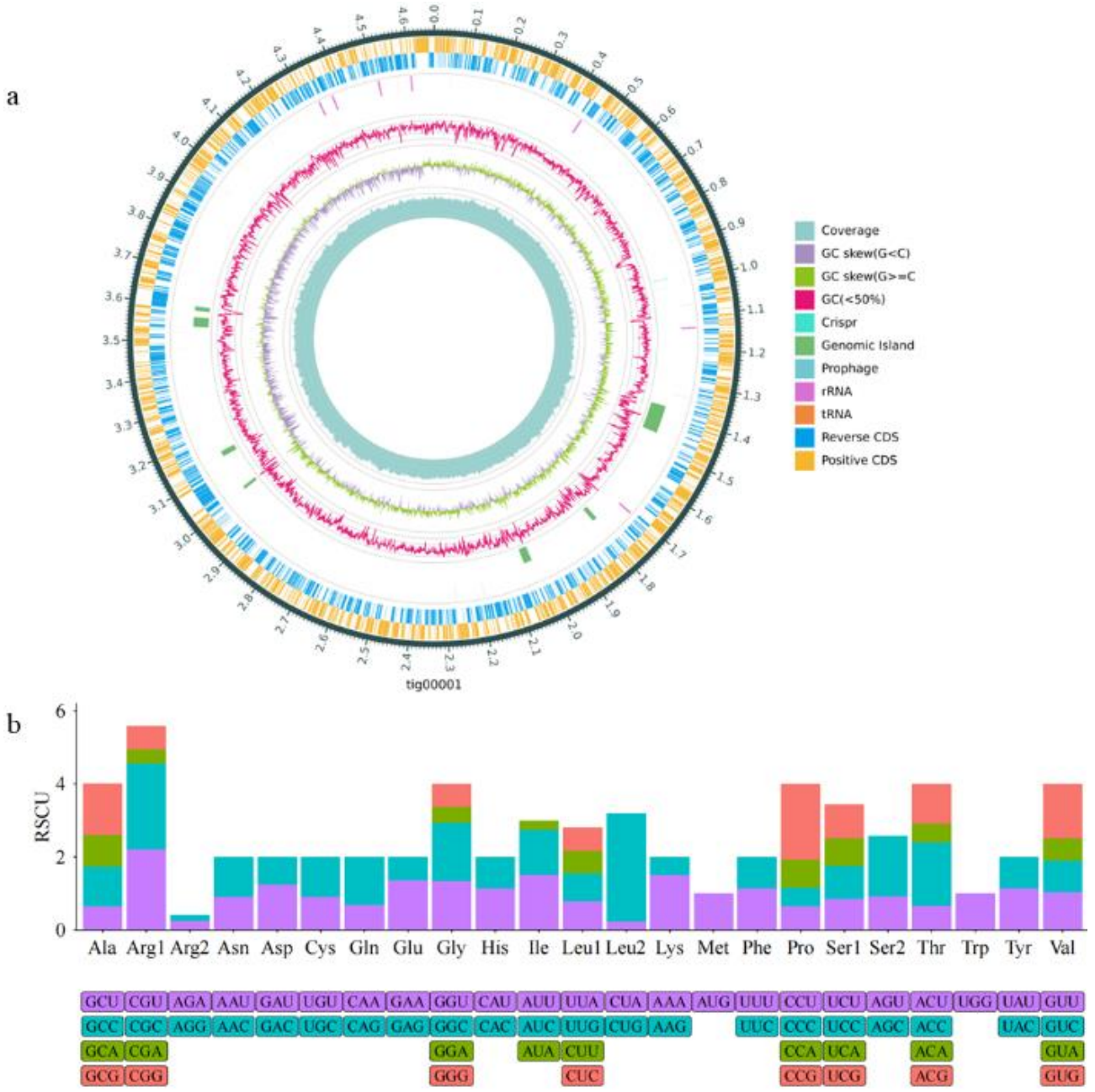

Fig. 1: Genome overview and codon usage pattern of *Escherichia coli* strain B19. a: Circular genomic map of *E. coli* B19. b: Relative synonymous codon usage (RSCU) values of the B19 genome.

Table 2: Codon usage frequency statistics of *Escherichia coli* strain B19. Stop codons are represented by an asterisk (*) and correspond to UAA, UAG, and UGA.

#AA	Codon	Number	/1000
Gly	GGG	15,997	11.21
Gly	GGA	11,704	8.2
Gly	GGU	35,075	24.58
Gly	GGC	41,666	29.2
Glu	GAG	25,691	18
Glu	GAA	56,637	39.69
Asp	GAU	46,036	32.26
Asp	GAC	27,382	19.19
Val	GUG	37,295	26.13
Val	GUA	15,568	10.91
Val	GUU	26,123	18.31
Val	GUC	21,677	15.19
Ala	GCG	47,545	33.32
Ala	GCA	28,870	20.23
Ala	GCU	22,026	15.43
Ala	GCC	36,356	25.48
Arg	AGG	1,880	1.32
Arg	AGA	3,171	2.22
Ser	AGU	12,887	9.03
Ser	AGC	23,041	16.15
Lys	AAG	14,983	10.5
Lys	AAA	48,139	33.73
Asn	AAU	25,723	18.03
Asn	AAC	30,626	21.46
Met	AUG	39,551	27.72
Ile	AUA	6,679	4.68
Ile	AUU	43,082	30.19
Ile	AUC	35,457	24.85
Thr	ACG	20,599	14.43
Thr	ACA	10,203	7.15
Thr	ACU	12,785	8.96
Thr	ACC	32,883	23.04
Trp	UGG	21,636	15.16
*	UGA	1,363	0.96
Cys	UGU	7,317	5.13
Cys	UGC	9,008	6.31
*	UAG	373	0.26
*	UAA	2,862	2.01
Tyr	UAU	23,423	16.41
Tyr	UAC	17,629	12.35
Leu	UUG	19,162	13.43
Leu	UUA	19,770	13.85
Phe	UUU	31,890	22.35
Phe	UUC	23,469	16.45
Ser	UCG	12,733	8.92
Ser	UCA	10,414	7.3
Ser	UCU	12,072	8.46
Ser	UCC	12,457	8.73
Arg	CGG	8,201	5.75
Arg	CGA	5,008	3.51
Arg	CGU	29,491	20.67
Arg	CGC	31,092	21.79
Gln	CAG	41,528	29.1
Gln	CAA	21,616	15.15
His	CAU	18,301	12.82
His	CAC	13,662	9.57
Leu	CUG	75,051	52.59
Leu	CUA	5,572	3.9
Leu	CUU	16,101	11.28
Leu	CUC	15,854	11.11
Pro	CCG	32,775	22.97
Pro	CCA	11,892	8.33
Pro	CCU	10,177	7.13
Pro	CCC	7,811	5.47

The resulting phylogeny (Fig. 2) shows that strain B19 falls within the *Escherichia coli* clade, which includes *E. coli* K-12, the pathogenic strain O157 644-PT8, and other *E. coli* reference strains, with high bootstrap support. B19 is more distantly related to other *Escherichia* species, including *E. fergusonii*, *E. ruysiae*, *E. whittamii*, *E. marmotae*, and *E. albertii*, which form separate branches with strong bootstrap support. The tree also includes *Atlantibacter* and *Citrobacter* species as closely related outgroups, with *Salmonella enterica* subsp. *enterica* LT2 serving as the distant outgroup. Overall, the tree confirms that strain B19 belongs to the *E. coli* lineage.

CAZy Functional Annotation: The CAZy (Carbohydrate-Active Enzymes) database was used to analyze and compare enzyme families in the *E. coli* strain B19 involved in the synthesis, degradation, and modification of complex carbohydrates. A total of 807 CAZyme-related domains were identified (Fig. 3), encompassing six major functional categories. Among these, 380 genes encoded glycoside hydrolases (GHs), accounting for 47.1% of all annotated CAZymes; 308 genes encoded glycosyl transferases (GTs) (38.2%); 56 genes encoded carbohydrate-binding modules (CBMs) (6.9%); 36 genes encoded carbohydrate esterases (CEs) (4.5%); and 13 genes each encoded auxiliary activity enzymes (AAs) and polysaccharide lyases (PLs). These results suggest that strain B19 possesses diverse carbohydrate metabolism capabilities, primarily mediated by glycoside hydrolases and glycosyl transferases.

KEGG Functional Annotation: The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was employed to functionally characterize the gene products of *E. coli* strain B19, focusing on their associated metabolic pathways and cellular functions. As shown in Fig. 4, a total of 3,356 coding genes were assigned to KEGG pathway categories. Among the Level 1 categories, the largest number of genes (2,579 genes) were clustered under the "Human Diseases" primary category. Functionally, these genes are predominantly involved in conserved pathways governing essential cellular and genetic mechanisms, such as protein families: signaling and cellular processes (1,016 genes), protein families: genetic information processing (716 genes), and protein families: metabolism (304 genes). Additionally, 1,062 genes were involved in metabolism pathways, predominantly enriched in carbohydrate metabolism (320 genes), amino acid metabolism (224 genes), and cofactor and vitamin metabolism (210 genes). Other genes were associated with environmental information processing (452 genes), cellular processes (251 genes), genetic information processing (202 genes), and organismal systems (55 genes). Notably, the metabolic pathway analysis revealed 38 genes involved in terpenoid and polyketide metabolism and 142 genes related to glycan biosynthesis and metabolism, suggesting that strain B19 may possess biosynthetic potential for antimicrobial compounds and participate in glycoprotein formation processes.

COG Functional Annotation: The predicted protein sequences of *E. coli* strain B19 were compared against the Clusters of Orthologous Groups (COG) database to classify and predict their potential biological functions. As shown in Fig. 5a, a total of 4,310 genes were assigned to COG functional categories. Among these, the largest group was involved in carbohydrate transport and metabolism (G), accounting for 403 genes (9.4%),

followed by amino acid transport and metabolism (E) with 377 genes (8.7%), and transcription (K) with 320 genes (7.4%). Additionally, 304 genes (7.1%) were associated with cell wall/membrane/envelope biogenesis (M). Other functional categories included energy production and conversion (C, 7.0%), translation, ribosomal structure and biogenesis (J, 6.8%), signal transduction mechanisms (T, 4.9%), coenzyme transport and metabolism (H, 4.9%), and inorganic ion transport and metabolism (P, 5.8%). These results suggest that strain B19 has a wide range of metabolic and regulatory capabilities, particularly in nutrient utilization, energy metabolism, and cell envelope synthesis, which may contribute to its environmental adaptability and potential pathogenicity.

GO Enrichment Analysis: Gene Ontology (GO) annotation of the *E. coli* B19 genome identified a total of 8,936 genes assigned to various GO terms Fig. 5b. Among these, 5,214 genes (58.3%) were classified under biological processes, mainly associated with metabolic processes, cellular processes, and localization. In the molecular function category, 2,545 genes (28.5%) were annotated, primarily involved in catalytic activity and binding functions. The cellular component category included 1,177 genes (13.2%), mainly related to cell and cellular structures. Notably, genes associated with membrane structure were significantly enriched, suggesting that strain B19 may possess a strong potential for biofilm formation, which could enhance its resistance to adverse environmental conditions. Furthermore, enrichment of GO terms related to localization and establishment of localization indicates that this strain carries genes involved in bacterial adhesion and colonization on host cell surfaces, providing a genetic basis for its persistence and potential activity within the host intestinal environment.

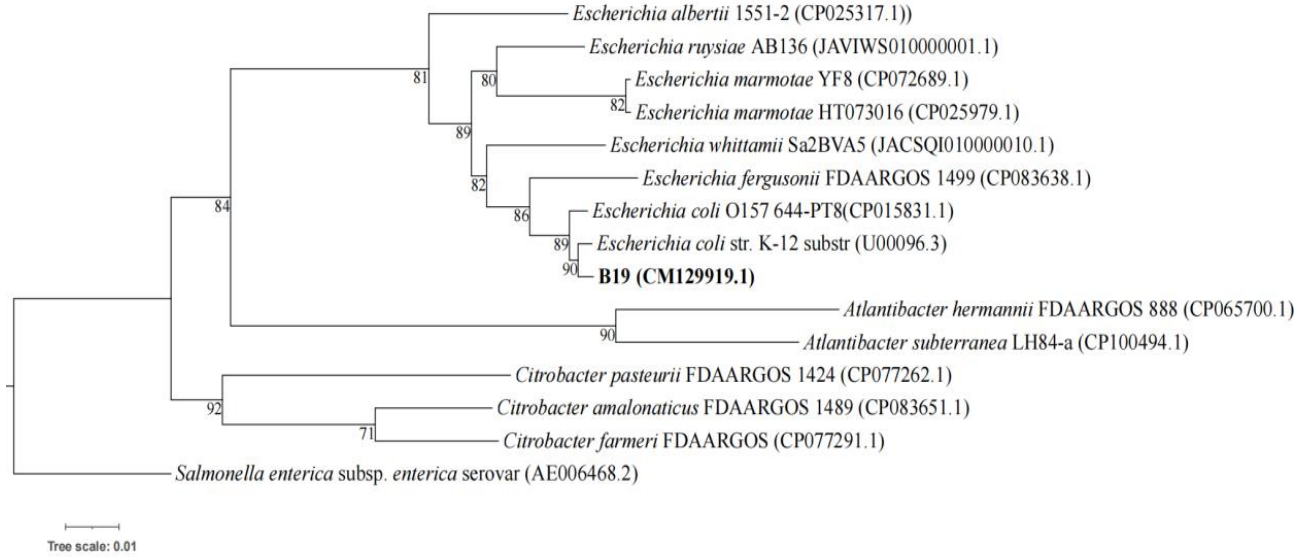

Fig. 2: Phylogenetic tree of *Escherichia coli* strain B19 based on whole-genome sequences.

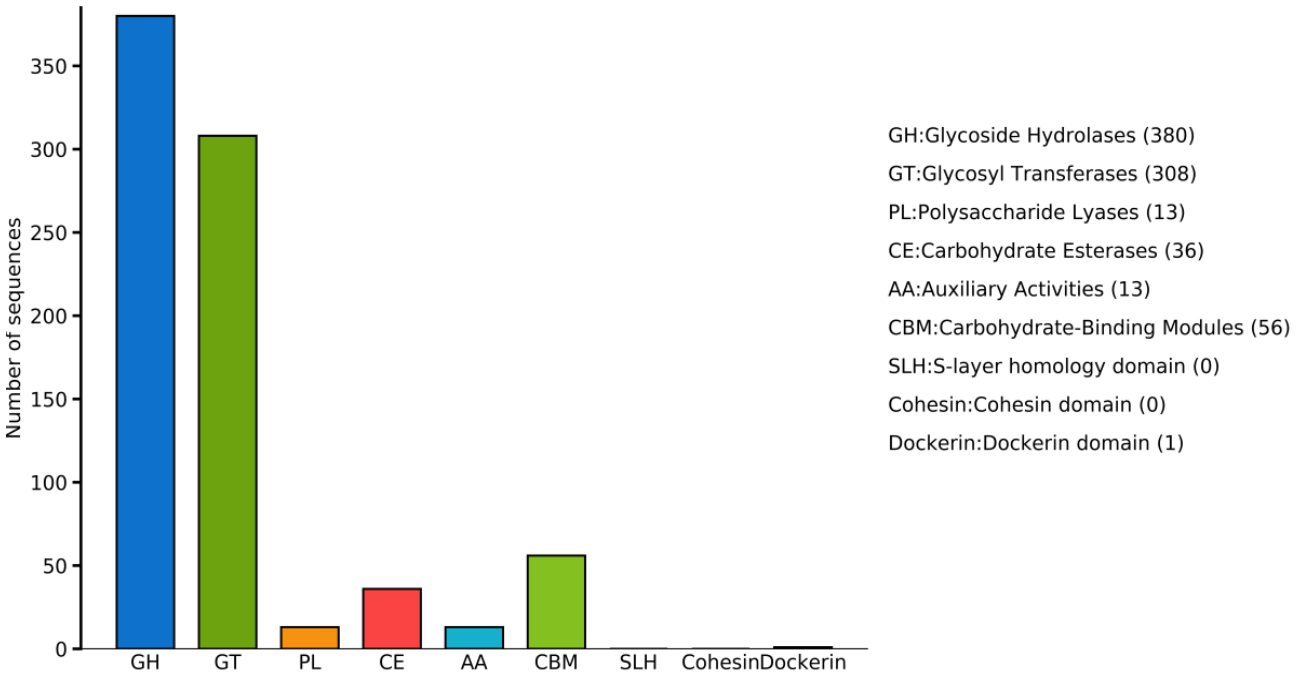

Fig. 3: CAZy functional classification of *Escherichia coli* B19 genome.

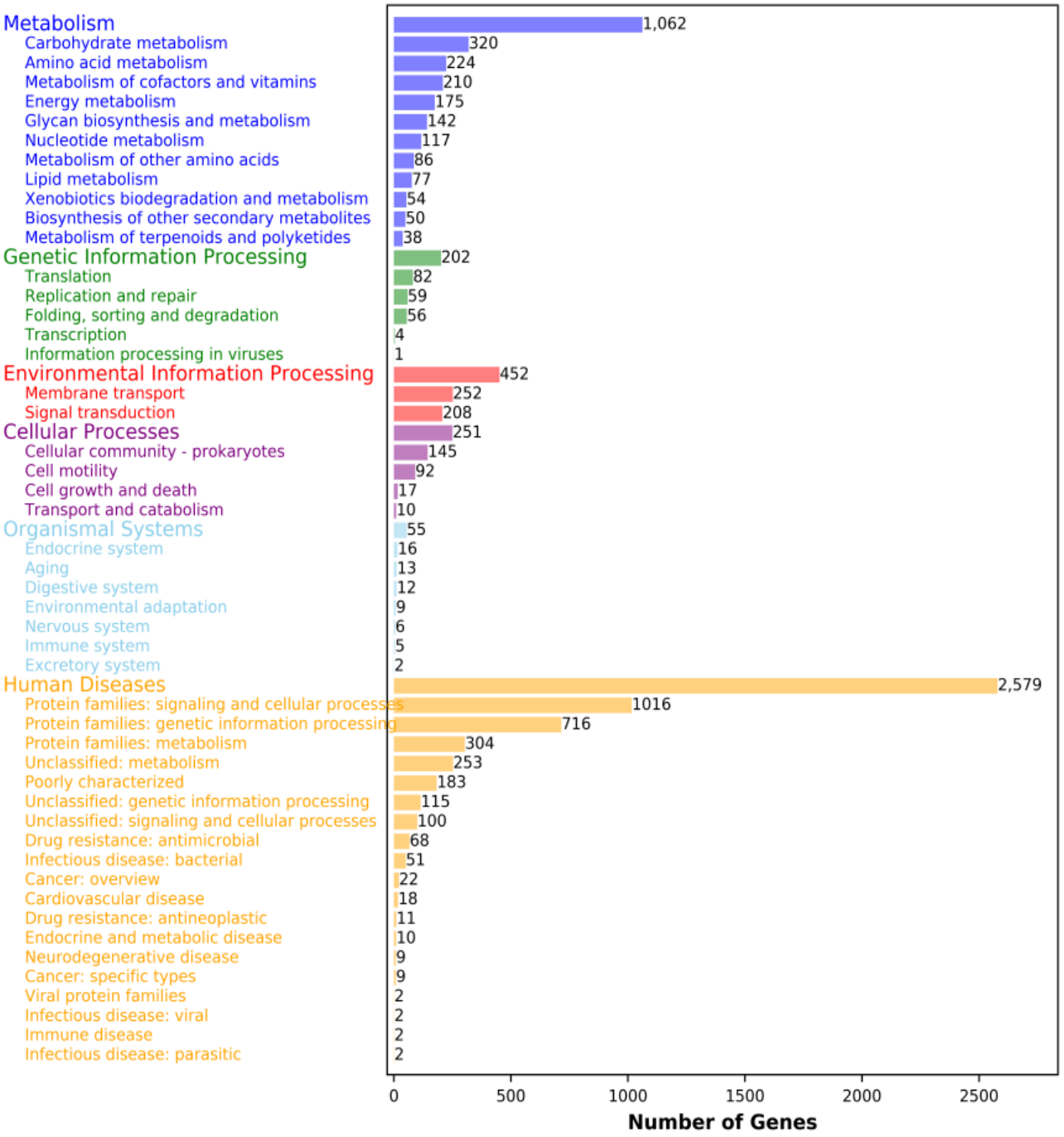

Fig. 4: KEGG functional classification of *Escherichia coli* B19 genome.

Secondary Metabolite Gene Cluster Analysis: The secondary metabolite biosynthetic gene clusters (BGCs) of *E. coli* strain B19 were predicted using the antiSMASH program. As shown in Fig. 6, a total of four secondary metabolite gene clusters were identified, corresponding to Tripeptide, Terpene, NRPS-Metallophore, and Azole RiPP, each represented by a single cluster. Further analysis revealed that the Tripeptide cluster contained 23 genes, the Terpene cluster contained 24 genes, the NRPS-Metallophore cluster was the largest with 48 genes, and the Azole RiPP cluster contained 20 genes. These results suggest that strain B19 has the potential to synthesize a variety of secondary metabolites, some of which, such as NRPS- and RiPP-related products, may play roles in antimicrobial activity, metal ion acquisition, or bacterial competition.

Virulence and Pathogenicity Analysis: The amino acid sequences of *E. coli* strain B19 were compared against the Pathogen–Host Interaction (PHI) database using BLAST, and the identified genes were integrated with corresponding functional annotations to determine their potential roles in pathogenicity. As shown in Fig. 7, the annotated results revealed that the majority of genes were associated with the "reduced virulence" mutation phenotype, followed by those categorized as "unaffected pathogenicity." This distribution suggests that while strain B19 carries multiple genes linked to virulence-related pathways, some mutations may have attenuated their pathogenic potential, whereas others may maintain or contribute to the strain's adaptability and survival during host–pathogen interactions.

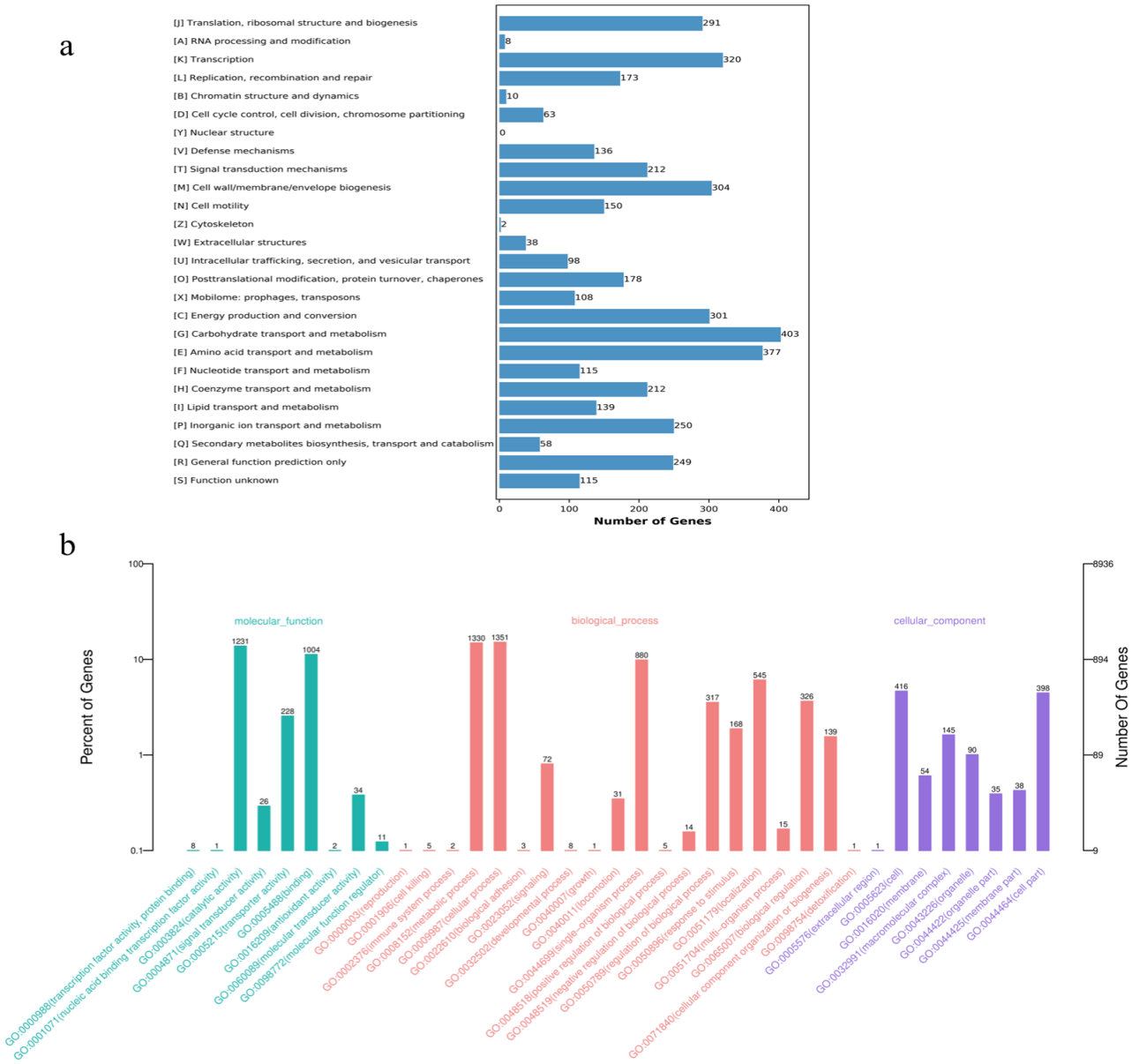

Fig. 5: Functional annotation of predicted genes in the B19 genome. a: COG functional classification. b: Gene Ontology (GO) functional classification.

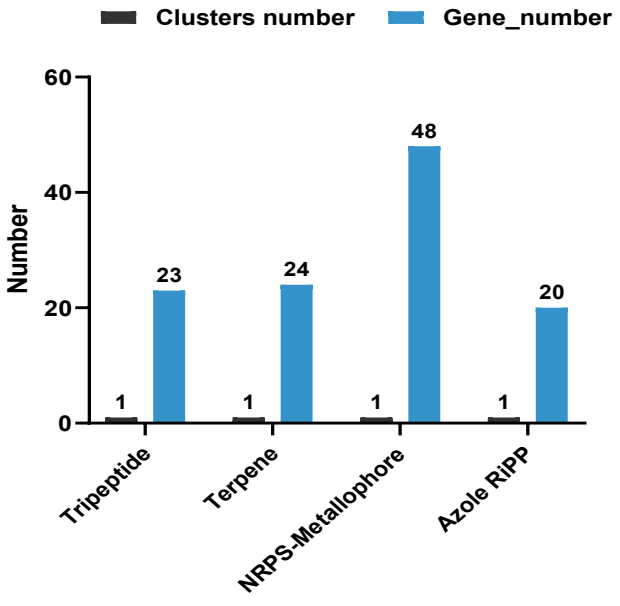

Fig. 6: Statistics of secondary metabolic gene clusters and number of genes.

CARD Database Annotation Analysis: Resistance genes of strain B19 were annotated using the Comprehensive Antibiotic Resistance Database (CARD) via the Abricate tool. As shown in Fig. 8a, multiple antibiotic resistance genes were identified, covering a broad spectrum of antimicrobial classes. Summary analysis revealed that the most abundant resistance determinants were associated with fluoroquinolones (25%), followed by aminoglycosides (18%) and macrolides (18%), which collectively accounted for over 60% of all annotated resistance genes. This distribution suggests that strain B19 may exhibit high resistance levels to these commonly used antibiotic classes. In addition, resistance genes related to tetracyclines (9%), and peptide antibiotics (9%) were also detected, indicating a potential broad-spectrum resistance profile. Other resistance determinants were distributed among multiple antibiotic categories, including nucleosides (6%), sulfonamides (3%), cephalosporins (3%), diaminopyrimidines (2%), glycopeptides (2%), as well as genes conferring tolerance to disinfectants and

preservatives (2%). The presence of resistance genes against both conventional antibiotics and newer antimicrobial agents, such as next-generation quinolones and chemical disinfectants, reflects the strain's adaptation to diverse antimicrobial selection pressures. Moreover, analysis of specific resistance genes revealed frequent occurrence of efflux pump-associated genes, such as *mdtA*, *mdtB*, *mdtC*, and *acrD*, which may enhance substrate extrusion and contribute to cross-resistance against multiple antibiotic classes. Additionally, resistance determinants specific to drugs such as novobiocin, kanamycin A, nalidixic acid, and ciprofloxacin were detected, further confirming the multidrug resistance (MDR) phenotype of strain B19.

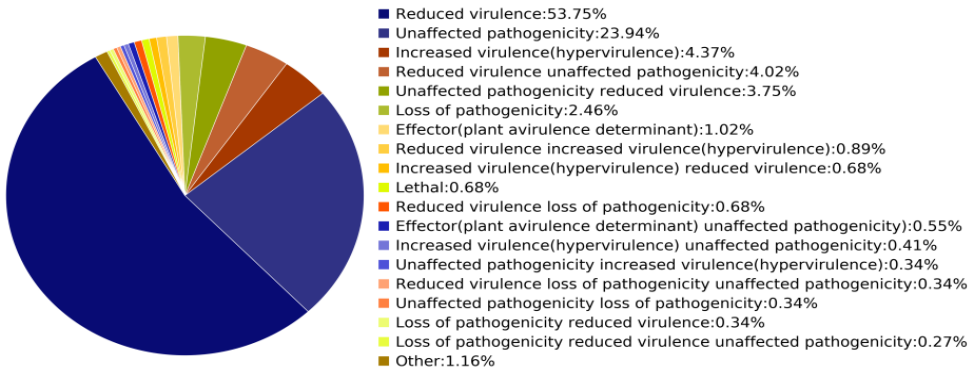

Fig. 7: Statistics on the number of genes with different PHI phenotype mutation types in genome B19.

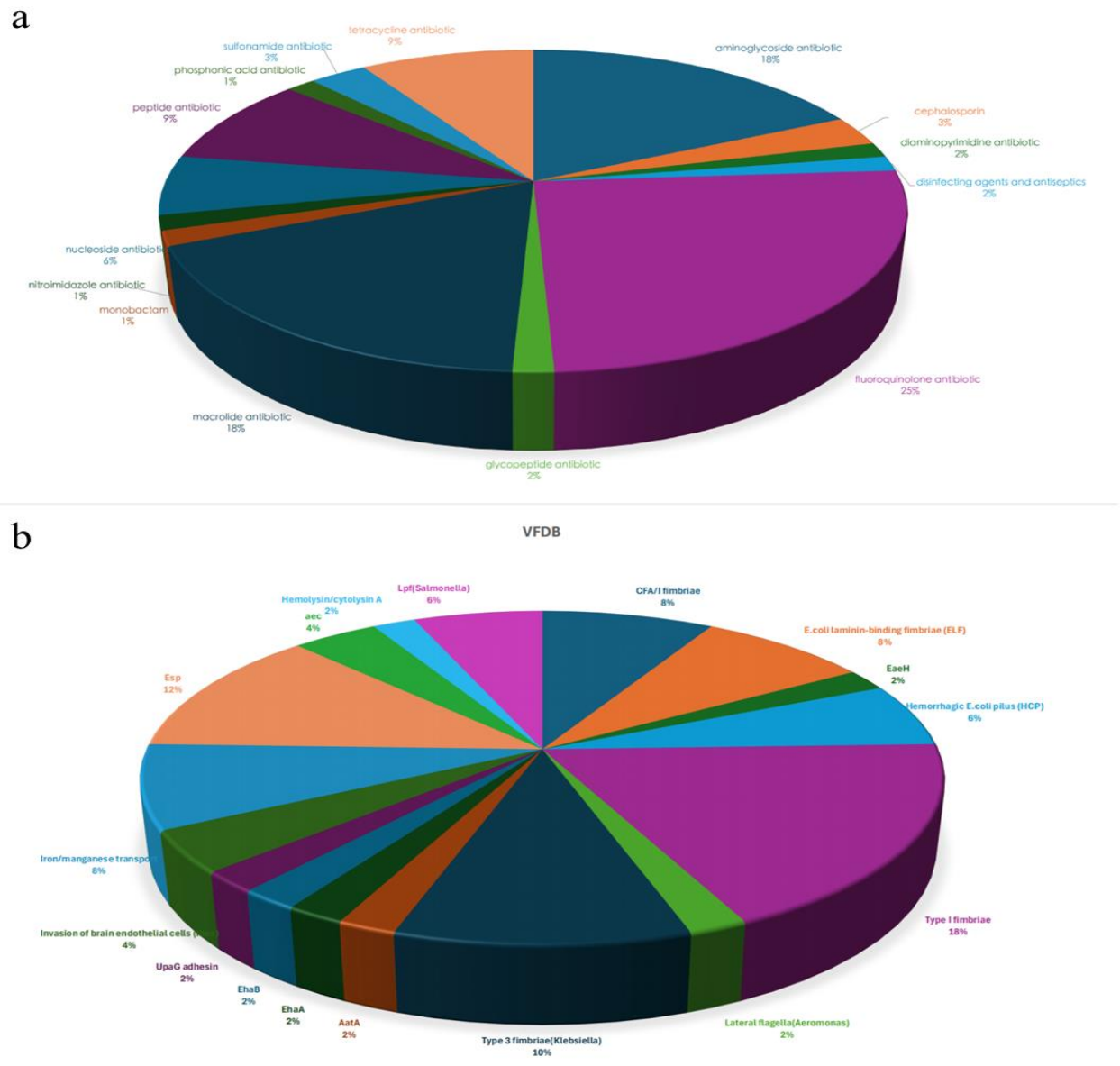

Fig. 8: Distribution of virulence factors and antibiotic resistance genes in *Escherichia coli* strain B19. a: Distribution of antibiotic resistance gene classes in *E. coli* strain B19 based on CARD annotation. b: Virulence factor distribution in *E. coli* strain B19 as annotated by the VFDB database.

VFDB Database Annotation Analysis: Virulence factors of strain B19 were annotated using the Virulence Factors of Pathogenic Bacteria (VFDB) database through the Abricate tool. The annotation results are presented in Fig. 8b. Multiple virulence-associated genes were identified, which were categorized into several functional classes, including adhesion, invasion, toxin production, and iron acquisition systems. Overall, adhesion-related genes accounted for the largest proportion of annotated virulence factors. Among them, Type I fimbriae were the most prevalent category (18%), followed by Type 3 fimbriae of *Klebsiella* origin (10%), suggesting that strain B19 possesses strong adhesive capabilities that may facilitate colonization within the host. Additional adhesion-associated determinants included CFA/I fimbriae (8%), *E. coli* laminin-binding fimbriae (ELF, 8%), Hemorrhagic *E. coli* pilus (HCP, 6%), Lpf fimbriae of *Salmonella* origin (6%), and various single adhesion-related proteins such as *AatA, EhaA, EhaB,* and *UpaG* (each 2%). Furthermore, the presence of *Esp* (12%), a structural component linked to Type III secretion systems, implies that the strain may harbor a secretion-associated pathogenic mechanism. Virulence factors related to iron acquisition, such as the iron/manganese transport system, accounted for 8%, indicating that the strain may enhance its survival and competitiveness within the host environment. Invasion-associated genes, exemplified by the Ibe (Invasion of brain endothelial cells) family, comprised 4%, suggesting a potential capability for crossing cellular barriers. Additionally, several toxin-related genes were detected, including hemolysin/cytolysin A (2%) and *aec* (4%), further supporting the pathogenic potential of strain B19.

DISCUSSION

In this study, whole-genome sequencing of *Escherichia coli* strain B19 was carried out by integrating short-read and long-read sequencing approaches, and the generated genome data were archived in the NCBI repository with the accession number PRJNA1346393. The total genome size of strain B19 was 4,901,302 bp, comprising a circular chromosome of 4,671,443 bp and three plasmids of 108,446 bp, 81,824 bp, and 39,589 bp. A total of 218 dispersed and tandem repeat sequences were identified within the genome, indicating a certain level of genomic redundancy. Previous studies have demonstrated that dispersed and tandem repeats are widely present in bacterial genomes, where they contribute to genomic plasticity by promoting events such as gene recombination, deletion, and duplication (Aras *et al.*, 2003, Treangen *et al.*, 2009). These repetitive structures can induce reversible changes in gene expression regulation under environmental stress, thereby enhancing the bacterium's adaptability and reproductive capacity in complex microbial environments (Delihas, 2011; Zhou *et al.*, 2014). In addition, tandem repeats can generate phenotypic diversity by altering the copy number of promoter regions or regulatory sequences (Subirana and Messeguer, 2020). This genomic plasticity may provide a genetic basis for the environmental adaptability of *E. coli* B19.

The codon usage pattern of strain B19 also provides further insight into its evolutionary adaptation and translational optimization strategies. The strain shows a clear preference for codons ending in C or T, with dominant usage of specific synonymous codons across multiple amino acids, such as GGC, GCG, and GAT. These findings are consistent with the codon usage biases observed in other *E. coli* strains (Parvathy *et al.*, 2022). Previous studies have demonstrated that codon usage bias is often associated with efficient and accurate protein translation, particularly among highly expressed housekeeping genes (Johnson *et al.*, 2023). Such bias may reflect selective pressures that drive the optimization of protein synthesis under specific ecological or host-associated conditions, thereby enhancing bacterial growth rate and competitiveness within microbial communities (Plotkin and Kudla, 2011). In addition, the preferred use of TAA as the major stop codon suggests that B19 may adopt a conserved and efficient translational termination strategy, which can reduce readthrough errors and ensure proper protein folding and maturation (Korkmaz *et al.*, 2014).

Phylogenetic analysis placed strain B19 within the *E. coli* clade with high bootstrap support. Specifically, B19 formed a subclade with *E. coli* K-12 (90% bootstrap support), and together with *E. coli* O157 644-PT8 they formed a well-supported *E. coli* lineage (89% bootstrap support). The distinct separation of B19 from other *Escherichia* species, including *E. marmotae*, *E. ruysiae*, and *E. fergusonii*, further confirms its phylogenetic placement within the *E. coli* lineage.

It should be noted that core-genome phylogeny primarily reflects the evolutionary relationships of conserved housekeeping genes. In contrast, virulence-associated genes are often located on mobile genetic elements, including plasmids, prophages and genomic islands, which can be acquired or lost through horizontal gene transfer independent of the core genome (Dobrindt *et al.*, 2004; Croxen and Finlay, 2010). Accordingly, strains with closely related core genomes, such as B19 and *E. coli* K-12, may exhibit markedly different virulence gene profiles. Similar observations have been widely reported in pathogenic *E. coli* and other bacterial species.

Comparative genomic analysis with previously reported sheep-derived MDR *E. coli* isolates from other regions revealed both commonalities and distinctions. A study by Furlan *et al*. (2019) characterized 14 MDR *E. coli* isolates from sheep feces in Brazil and identified 13 distinct sequence types (including ST48, ST155, ST162, and ST642), with phylogenetic groups predominantly belonging to B1 and A (Furlan *et al.*, 2019). Zhao *et al*. (2024) reported that ESBL-producing MDR *E. coli* isolates from sheep in China were assigned to ST10, ST162, and ST457, also dominated by phylogroups B1 and A (Zhao *et al.*, 2024). In the present study, strain B19 belongs to ST101 and phylogroup B1. This matches the general prevalence of phylogroup B1 among sheep-derived *E. coli*, while ST101 is a rarely documented sequence type in this host. In terms of antimicrobial resistance genes, the profile of B19, including *bla*CTX-M-55, *tet*(A), *sul*2 and *qnr*S1, is largely consistent with findings from prior studies on sheep isolates. Overall, these results demonstrate that B19 shares typical phylogenetic and resistance characteristics with other

sheep-origin MDR *E. coli*, whereas its ST101 genotype represents an uncommon clonal lineage within this ecological niche.

CAZy annotation revealed that a total of 380 glycoside hydrolases (47.1%), 308 glycosyltransferases (38.2%), 56 carbohydrate-binding modules (6.9%), 36 carbohydrate esterases (4.5%), 13 auxiliary activities, and 13 polysaccharide lyases were identified. The high proportions of glycoside hydrolases and glycosyltransferases suggest that B19 possesses a strong capacity for carbohydrate degradation and modification, which may facilitate its interaction with host glycan structures or other microorganisms. Similar observations have been reported in the CAZy annotations of various bacterial genomes, where the abundance of glycoside hydrolases is positively correlated with enhanced metabolic versatility and ecological functional potential (Geng *et al.*, 2021). KEGG pathway annotation revealed that the genes were primarily enriched in 1,016 signal transduction and cellular process pathways, 716 genetic information processing pathways, 452 environmental information processing pathways, 320 carbohydrate metabolism pathways, and 304 protein family metabolism pathways, suggesting metabolic network complexity and strong adaptive potential of the strain (Busi *et al.*, 2022; Fang *et al.*, 2022). Additionally, a substantial portion of the annotated genes was cataloged under the "Human Diseases" level 1 category. This phenomenon does not imply that the ovine-derived strain B19 possesses specific traits for human-specific diseases; rather, it is an artifact of the inherent classification layout of the KEGG database. In KEGG, numerous evolutionarily highly conserved pathways that regulate fundamental cellular homeostasis, signal transduction, and basic macromolecular metabolism are systemically archived under this primary disease-related label. Therefore, the high gene enrichment in this category further underscores the robust cellular viability and environmental responsiveness of strain B19, rather than zoonotic disease potential per se. COG-based functional categorization showed that genes related to carbohydrate utilization and transport, together with those associated with amino acid utilization and transport, were highly represented. Additionally, approximately 7.1% of genes were associated with cell wall/membrane/envelope biogenesis, indicating the potential capability of B19 to remodel the cell envelope or form biofilms in response to environmental stress (Silhavy *et al.*, 2010; Agius *et al.*, 2021). GO functional annotation further showed that the genes were mainly enriched in metabolic processes, molecular binding, cellular processes, and cellular components, reflecting extensive basal metabolic capacity and cellular maintenance functions (Gene Ontology, 2021). Secondary metabolite gene cluster analysis revealed that the major biosynthetic gene clusters (BGCs) in B19 included NRPS-Metallophore, terpene, tripeptide, and azole RiPP. These secondary metabolites are known to play crucial roles in bacterial ecology and host–microbe interactions. NRPS-Metallophore products typically participate in metal ion, particularly iron, chelation and acquisition, facilitating bacterial survival and competitiveness in micronutrient-limited environments (Awaya and Dubois, 2008). Terpenes may act as signaling molecules or antimicrobial compounds, mediating microbial competition and communication (Guimarães *et al.*, 2019). Therefore, the enrichment of these secondary metabolite gene clusters in B19 not only provides chemical defense and resource acquisition advantages but may also enhance its competitiveness in complex microbial communities or within the host gut niche.

Based on the CARD database annotation, the resistance genes of strain B19 were predominantly associated with fluoroquinolones (25%), aminoglycosides (18%), and macrolides (18%), accounting for over 60% collectively. This suggests that the strain may possess a high level of tolerance to multiple commonly used antibiotics. Fluoroquinolone resistance is typically associated with mutations in *gyrA* and *parC*, the acquisition of transferable qnr genes, or the upregulation of efflux pump systems (Morgan-Linnell *et al.*, 2009; Shigemura *et al.*, 2012). In *E. coli*, double gyrA mutations (e.g., S83L and D87N), in combination with S80I substitution in parC, have frequently been reported to confer high-level fluoroquinolone resistance (Tewawong *et al.*, 2025). Resistance to aminoglycosides is commonly attributed to the presence of aminoglycoside-modifying enzymes (AMEs) or 16S rRNA methyltransferases, both of which are capable of conferring high-level resistance and are frequently spread via plasmid-mediated transfer (Doi and Arakawa, 2007; Garneau-Tsodikova and Labby, 2016; Wachino *et al.*, 2020). Macrolide resistance is primarily mediated by erm or mph genes, which inhibit drug action through ribosomal methylation or antibiotic phosphorylation, respectively (Svetlov *et al.*, 2021; Chen *et al.*, 2023). In *E. coli*, *mphA* and *mphB*, which encode macrolide phosphotransferases, have been shown to confer high-level resistance to erythromycin and clarithromycin (Noguchi *et al.*, 1998). In addition, efflux pump-related genes such as *mdtA*, *mdtB*, *mdtC*, and *acrD* were detected. Efflux systems such as MdtABC and *AcrD* in Gram-negative bacteria play a major role in expelling various antibiotics and disinfectants, thereby contributing to cross-resistance. In *E. coli*, the MdtABC-TolC efflux system forms a heterotrimeric $MdtB_2C$ complex that uses a proton motive force to extrude diverse toxic substrates, significantly enhancing multidrug resistance (Kim *et al.*, 2010). Furthermore, the *AcrD* efflux pump in *E. coli* not only efficiently expels aminoglycosides but also synergizes with the major AcrAB-TolC pump, thereby broadening the resistance spectrum and enhancing bacterial survival under diverse environmental conditions (Anes *et al.*, 2015). Additionally, resistance determinants associated with specific antibiotics such as novobiocin, kanamycin A, nalidixic acid, and ciprofloxacin were also identified, indicating a diverse resistance profile. Notably, genes associated with tolerance to disinfectants and preservatives were present, suggesting that exposure to chemical disinfectants in the environment may contribute to the co-selection, persistence, and dissemination of antibiotic resistance genes. In *E. coli*, adaptive exposure to triclosan has been shown to induce upregulation of efflux pump genes such as acrAB, thereby enhancing resistance not only to triclosan but also to multiple antibiotics (Zeng *et al.*, 2020).

In the VFDB database annotation analysis, strain B19 was found to harbor multiple virulence-associated

factors, including those involved in adhesion, invasion, toxin production, and iron acquisition systems. Notably, Type I fimbriae accounted for the highest proportion (18%), followed by Type 3 fimbriae of *Klebsiella* origin (10%). Fimbriae are key virulence structures in *Escherichia coli*, mediating host adhesion, colonization and biofilm formation (Klemm *et al.*, 2010). These observations suggest that strain B19 possesses strong adhesive capabilities, which may enhance its colonization potential within the host. Iron acquisition system-related factors, such as components of the iron/manganese transport system, accounted for 8% of the annotated virulence genes. Iron is an essential element for bacterial growth and pathogenicity, while hosts commonly employ iron-restriction strategies to limit pathogen proliferation (Wooldridge and Williams, 1993). To overcome this limitation, bacteria synthesize high-affinity iron-chelating molecules known as siderophores, which play critical roles in iron acquisition and virulence. Studies have demonstrated that disruption of siderophore-mediated iron uptake systems can significantly attenuate bacterial pathogenicity (Wells *et al.*, 2013; Schalk, 2025).

.

Limitations of the Study: Although this study provides a comprehensive genomic blueprint of the multidrug-resistant *Escherichia coli* strain B19 isolated from diarrheic sheep in Ningxia, several limitations should be acknowledged. First, the phenotypic antimicrobial susceptibility profiling of the 19 isolates in this study relied entirely on the qualitative disk diffusion method. This approach was primarily deployed as a preliminary screening tool to identify and select the most resistant isolate for subsequent whole-genome sequencing. Second, this study characterizes the complete genome architecture and functional repertoire of a single representative multidrug-resistant strain. While the hybrid long-read and short-read sequencing approach ensured high genomic integrity and excellent plasmid resolution, a single isolate cannot fully capture or reflect the wider clonal diversity, plasmid heterogeneity, or complex mobile genetic element dynamics of *E. coli* circulating within the broader livestock population. Lastly, the sampling scope was restricted to a single sheep farm in Ningxia within a specific temporal window (September 2024). Therefore, the findings may have localized constraints and cannot be generalized to track long-term evolutionary trends or wide-scale transmission pathways of animal-origin antimicrobial resistance. To overcome these limitations, future multi-center epidemiological surveillance and large-scale longitudinal investigations incorporating quantitative MIC determinations are highly warranted to systematically monitor the dissemination and public health risks of livestock-derived multidrug-resistant *E. coli*.

Conclusions: In this study, the complete genome of multidrug-resistant *Escherichia coli* B19 from sheep was resolved using a combination of second- and third generation sequencing technologies, revealing its integrated adaptive features in genome structure, metabolism, and antimicrobial resistance mechanisms. The abundance of repetitive sequences and the diversity of CAZy enzyme-encoding genes in the B19 genome indicate high genomic plasticity and carbohydrate metabolic potential, providing a genetic basis for survival in the host gut and environmental niches. The presence of a unique NRPS-Metallophore secondary metabolite cluster suggests a potential advantage in iron acquisition and ecological competition. The coexistence of multiple antibiotic resistance genes and efflux pump systems (e.g., MdtABC, AcrD), along with the detection of genes associated with disinfectant tolerance, implies that environmental selective pressures may facilitate the development of cross-resistance. Overall, this study provides novel insights into the genomic features, ecological adaptation mechanisms, and environmental co-selection of sheep-derived multidrug-resistant *E. coli*, establishing a genomic foundation for understanding the evolution and dissemination of animal-origin antimicrobial-resistant bacteria.

Ethics approval statement: All the experiment procedures in this study were under the ethics committee of Nanjing Agricultural University (NJAU.No 20240910161).

Data availability: The genome sequence raw data was deposited in NCBI database under accession number: PRJNA1346393.

Competing interests: The authors declare no competing interests.

Funding: None.

Author contributions: Conceptualization, KL; methodology, ZZ, ZX, XW, and LY; software, ZZ; validation, KL; formal analysis, ZZ; investigation, ZZ, XW, and LY; resources, KL; data curation, ZZ, writing-original draft preparation, ZZ, KL, Farooqi S, and ML; writing-review and editing, KL and ML; visualization, KL; supervision, KL; project administration. All authors have read and agreed to the published version of the manuscript.

Acknowledgements: None.

REFERENCES

Agius JE, Phalen DN, Rose K, *et al.*, 2021. Genomic insights into the pathogenicity of a novel biofilm-forming *Enterococcus* sp. bacteria (*Enterococcus lacertideformus*) identified in reptiles. Frontiers in Microbiology 12: 635208.

Ali S and Alsayeqh AF, 2022. Review of major meat-borne zoonotic bacterial pathogens. Frontiers in Public Health 10:1045599.

Anes J, McCusker MP, Fanning S, *et al.*, 2015. The ins and outs of RND efflux pumps in *Escherichia coli*. Frontiers in Microbiology 6: 587.

Aras RA, Kang J, Tschumi AI, *et al.*, 2003. Extensive repetitive DNA facilitates prokaryotic genome plasticity. Proceedings of the National Academy of Sciences of the United States of America 100(23): 13579-84.

Atlaw NA, Keelara S, Correa M, *et al.*, 2021. Identification of CTX-M Type ESBL *E. coli* from sheep and their abattoir environment using whole-genome sequencing. Pathogens 10(11): 1480.

Awaya JD and Dubois JL, 2008. Identification, isolation, and analysis of a gene cluster involved in iron acquisition by *Pseudomonas mendocina* ymp. Biometals 21(3): 353-366.

Busi SB, Bourquin M, Fodelianakis S, *et al.*, 2022. Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. Nature Communications 13(1): 2168.

Chen CW, Leimer N, Syroegin EA, *et al*., 2023. Structural insights into the mechanism of overcoming Erm-mediated resistance by macrolides acting together with hygromycin-A. Nature Communications 14(1): 4196.

Croxen MA and Finlay BB, 2010. Molecular mechanisms of *Escherichia coli* pathogenicity. Nature Reviews Microbiology 8(1): 26-38.

Croxen MA, Law RJ, Scholz R, *et al*., 2013. Recent advances in understanding enteric pathogenic *Escherichia coli*. Clinical Microbiology Reviews 26(4): 822-880.

Delihas N, 2011. Impact of small repeat sequences on bacterial genome evolution. Genome Biology and Evolution 3: 959-973.

Dobrindt U, Hochhut B, Hentschel U, *et al*., 2004. Genomic islands in pathogenic and environmental microorganisms. Nature Reviews Microbiology 2(5): 414-24.

Doi Y and Arakawa Y, 2007. 16S Ribosomal RNA Methylation: Emerging resistance mechanism against aminoglycosides. Clinical Infectious Diseases 45(1): 88-94.

Fang GY, Chai LJ, Zhong XZ, *et al*., 2022. Comparative genomics unveils the habitat adaptation and metabolic profiles of *Clostridium* in an artificial ecosystem for liquor production. mSystems 7(3): e0029722.

Furlan JPR, Gallo IFL, de Campos A, *et al*., 2019. Molecular characterization of multidrug-resistant Shiga toxin-producing *Escherichia coli* harboring antimicrobial resistance genes obtained from a farmhouse. Pathogens and Global Health 113(6): 268-74.

Gao X, Yang X, Noll L, *et al*., 2018. Whole genome shotgun sequencing revealed highly polymorphic genome regions and genes in *Escherichia coli* O157:H7 isolates collected from a single feedlot. PloS One 13(8): e0202775.

Garneau-Tsodikova S and Labby KJ, 2016. Mechanisms of resistance to aminoglycoside antibiotics: Overview and perspectives. Medchemcomm 7(1): 11-27.

Gene Ontology C, 2021. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research 49(D1): D325-D34.

Geng A, Jin M, Li N, *et al*., 2021. New insights into the co-occurrences of glycoside hydrolase genes among prokaryotic genomes through network analysis. Microorganisms 9(2): 427.

Gomes TA, Elias WP, Scaletsky IC, *et al*., 2016. Diarrheagenic *Escherichia coli*. Brazilian Journal of Microbiology 47 (Suppl 1): 3-30.

Guimarães AC, Meireles LM, Lemos MF, *et al*., 2019. Antibacterial activity of terpenes and terpenoids present in essential oils. Molecules 24(13): 2471.

Gurjar T, Singathia R, Sharma DK, *et al.* 2024. *Escherichia coli* in diarrhoeic lambs: Prevalence, virulence and antibiotic resistance. Polish Journal of Veterinary Sciences 27(3): 451-457.

Haulisah NA, Hassan L, Bejo SK, *et al*., 2021. High levels of antibiotic resistance in isolates from diseased livestock. Frontiers in Veterinary Science 8: 652351.

Johnson MM, Hockenberry AJ, McGuffie MJ, *et al*., 2023. Growth-dependent gene expression variation influences the strength of codon usage biases. Molecular Biology and Evolution 40(9): msad189.

Kaper JB, Nataro JP and Mobley HL, 2004. Pathogenic *Escherichia coli*. Nature Reviews Microbiology 2(2): 123-40.

Kim HS, Nagore D and Nikaido H, 2010. Multidrug efflux pump MdtBC of *Escherichia coli* is active only as a B2C heterotrimer. Journal of Bacteriology 192(5): 1377-86.

Klemm P, Hancock V and Schembri MA, 2010. Fimbrial adhesins from extraintestinal *Escherichia coli*. Environmental Microbiology Reports 2(5): 628-40.

Koleci X, Zalla P, Sulce M, *et al*., 2025. Isolation of shiga toxin-producing *Escherichia coli* from sheep faecal samples: bacteriological findings. Access Microbiology 7(7): 001004-v3.

Korkmaz G, Holm M, Wiens T, *et al*., 2014. Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. Journal of Biological Chemistry 289(44): 30334-42.

Luo S, Liao C, Peng J, *et al*., 2023. Resistance and virulence gene analysis and molecular typing of *Escherichia coli* from duck farms in Zhanjiang, China. Frontiers in Cellular and Infection Microbiology 13: 1202013.

Morgan-Linnell SK, Boyd LB, Steffen D, *et al*., 2009. Mechanisms accounting for fluoroquinolone resistance in *Escherichia coli* clinical isolates. Antimicrobial Agents and Chemotherapy 53(1): 235-41.

Nawaz Z, Aslam B, Zahoor MA, *et al*., 2021. Frequency of extended spectrum beta lactamase producing *Escherichia coli* in fresh and frozen meat. Pakistan Veterinary Journal 41(1): 102-106.

Noguchi N, Tamura Y, Katayama J, *et al*., 1998. Expression of the mphB gene for macrolide 2'-phosphotransferase II from *Escherichia coli* in *Staphylococcus aureus*. FEMS Microbiology Letters 159(2): 337-42.

Özavci V, Kirkan Ş, Yüksel Dolgun HT, *et al*., 2026. Prevalence of virulent *Escherichia coli* pathotypes in calf diarrhoea: molecular characterisation and antibiotic resistance profiles in Türkiye. Polish journal of Veterinary Sciences 29(1): 111-120.

Pakbin B, Bruck WM and Rossen JWA, 2021. Virulence Factors of Enteric Pathogenic *Escherichia coli*: A Review. International Journal of Molecular Sciences 22(18): 9922.

Parvathy ST, Udayasuriyan V and Bhadana V, 2022. Codon usage bias. Molecular Biology Reports 49(1): 539-65.

Pinheiro REE, Chaves TP, Melo ES, *et al*., 2020. Modulatory-antibiotic activity of the essential oil from *Eucalyptus citriodora* against MDR bacterial strains. Cellular and Molecular Biology 66(4): 60-64.

Plotkin JB and Kudla G, 2011. Synonymous but not the same: the causes and consequences of codon bias. Nature Reviews Genetics 12(1): 32-42.

Rakhalaru P, Munzhedzi L, Abia ALK, *et al*., 2023. Prevalence and antimicrobial resistance profile of diarrheagenic *Escherichia coli* from fomites in rural households in South Africa. Antibiotics (Basel) 12(8).

Schalk IJ, 2025. Bacterial siderophores: diversity, uptake pathways and applications. Nature Reviews Microbiology 23(1): 24-40.

Shigemura K, Tanaka K, Yamamichi F, *et al*., 2012. Does mutation in gyrA and/or parC or efflux pump expression play the main role in fluoroquinolone resistance in *Escherichia coli* urinary tract infections?: A statistical analysis study. International Journal of Antimicrobial Agents 40(6): 516-20.

Silhavy TJ, Kahne D and Walker S, 2010. The bacterial cell envelope. Cold Spring Harbor Perspectives in Biology 2(5): a000414.

Subirana JA and Messeguer X, 2020. Unique features of tandem repeats in bacteria. Journal of Bacteriology 202(21): 10.1128/jb.00229-20.

Svetlov MS, Syroegin EA, Aleksandrova EV, *et al*., 2021. Structure of Erm-modified 70S ribosome reveals the mechanism of macrolide resistance. Nature Chemical Biology 17(4): 412-20.

Tewawong N, Kowaboot S, Lektrakul W, *et al*., 2025. Mechanisms of fluoroquinolone resistance among *Escherichia coli* isolates from urinary tract infections in Thailand. PloS One 20(5): e0325175.

Treangen TJ, Abraham AL, Touchon M, *et al*., 2009. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiology Reviews 33(3): 539-71.

Wachino JI, Doi Y and Arakawa Y, 2020. Aminoglycoside Resistance: Updates with a focus on Acquired 16S Ribosomal RNA Methyltransferases. Infectious Disease Clinics of North America 34(4): 887-902.

Wells RM, Jones CM, Xi Z, *et al*., 2013. Discovery of a siderophore export system essential for virulence of Mycobacterium tuberculosis. PloS Pathogens 9(1): e1003120.

Wooldridge KG and Williams PH, 1993. Iron uptake mechanisms of pathogenic bacteria. FEMS Microbiology Reviews 12(4): 325-48.

Yaseen K, Ali S, Rahman SU and Sajid MS, 2025. Comparative molecular virulence typing and antibiotic resistance of *Campylobacter* species at the human–animal–environment interface. Foodborne Pathogens and Disease 22(2): 109-117.

Zeng W, Xu W, Xu Y, *et al*., 2020. The prevalence and mechanism of triclosan resistance in *Escherichia coli* isolated from urine samples in Wenzhou, China. Antimicrobial Resistance and Infection Control 9(1): 161.

Zhao X, Chen H, Bi W, *et al*., 2024. Coexistence and genomics characterization of mcr-1 and extended-spectrum-β-lactamase-producing *Escherichia coli*, an emerging extensively drug-resistant bacteria from sheep in China. Science of the Total Environment 955: 177016.

Zhou K, Aertsen A and Michiels CW, 2014. The role of variable DNA tandem repeats in bacterial adaptation. FEMS Microbiology Reviews 38(1): 119-41.