Research Paper

Stress Resistance and Virulence Gene Profiles Associated with Phylogeny and Phenotypes of *Escherichia coli* from Cattle

Yuan Fang [1], Frances Tran [1], Kim Stanford [2], Xianqin Yang [1],*

[1] *Agriculture and Agri-Food Canada Lacombe Research and Development Centre, 6000 C & E Trail, Lacombe, AB T4L 1W1, Canada*
[2] *University of Lethbridge, Lethbridge, AB T1K 3M4, Canada*

A R T I C L E   I N F O

A B S T R A C T

Seven serogroups of *E. coli* (Top seven *E. coli*) are frequently implicated in foodborne outbreaks in North America, largely due to their carriage of Shiga toxin genes (*stx*). This study aimed to profile resistance genes and virulence factors (VF), and their potential association with phylogeny and phenotypes of Top seven *E. coli* originating from cattle in Canada. 155 Top seven *E. coli* isolates previously characterized for heat and acid resistance and biofilm-forming ability were whole-genome sequenced and analyzed for phylogeny, VF, and stress resistance genes. The 155 *E. coli* strains belonged to six phylogroups: A (n = 32), B1 (n = 93), C (n = 3), D (n = 11), E (n = 15), and G (n = 1). Different phylogroups were clearly separated on the core genome tree, with strains of the same serotype closely clustered. The carriage of *stx* and the transmissible locus of stress tolerance (tLST), the extreme heat resistance marker, was mutually exclusive, in 33 and 15 genomes, respectively. A novel O84:H2 strain carrying stx_{1a} was also identified. In total, 70, 41, and 32 VF, stress resistance genes and antibiotic resistance genes were identified. The stress resistance genes included those for metal (n = 29), biocides/acid (n = 4), and heat (n = 8) resistance. All heat resistance genes and most metal-resistance genes that were differentially distributed among the phylogroups were exclusively in phylogroup A. VF were least and most present in phylogroups A and D, respectively. No specific genes associated with acid resistance or biofilm formation phenotypes were identified. VF were more abundant (P < 0.05) in the non–biofilm-forming population and acid-resistant population.

Escherichia coli is a commensal organism of the intestinal tract of humans and other warm-blooded animals (Gordon & Cowling, 2003), a small fraction of which can cause disease in both humans and animals. Shiga toxin-producing *E. coli* (STEC), also referred to as verotoxigenic *E. coli* (VTEC), is one of the *E. coli* pathotypes (Denamur et al., 2021). STEC is a significant foodborne human pathogen across the globe, causing more than 1 million illnesses, resulting in more than 100 deaths annually (WHO & FAO, 2018). An estimated 6.3 and 3.3 cases of STEC infection per 10,000 population were reported in the USA and Canada in 2019, respectively (Crim et al., 2015; Gill, 2020). The initial recognition of STEC as a significant human pathogen started with STEC O157:H7 from a multistate outbreak in the USA in the1990s (Bell et al., 1994). In recent years, however, there has been increasing incidence of outbreaks associated with non-O157 *E. coli*, in particular O103, O26, O111, O121, O45, and O145. According to FoodNet, the incidence of these six serogroup infections increased from 0.12 per 100,000 population in 2010 to 0.95 per 100,000 population in 2010 in the USA (Gould et al., 2013), in part due to improved

detection methods. Together with serogroup O157, they are often referred to as the Top seven *E. coli*. Beef is the most commonly implicated food commodity in STEC outbreaks in North America (WHO & FAO, 2018). This is hardly surprising as ruminants such as cattle are the primary carrier of STEC and can shed STEC O157 in concentrations of up to 10^6 CFU/g fecal material (Stephens et al., 2009). STEC does not lead to overt symptoms in cattle, due to the lack of the vascular receptor for Shiga toxins (Pruimboom-Brees et al., 2000), and as such, the prevalence of STEC on the hides of animals presented for slaughter can be > 90% (Arthur et al., 2007; Yang et al., 2014). The current control of beef contamination with STEC in commercial practice is largely reliant on postslaughter antimicrobial interventions, including pasteurization of carcasses with steam or hot water and spraying carcasses and cuts with lactic acid and peroxyacetic acid (Yang, 2016).

Previous work has assessed a large number of Top seven *E. coli* (n = 750) from cattle and generic *E. coli* (n = 700) from meat processing environment for their response to heat treatment and their biofilm-forming ability. Top seven *E. coli* from cattle generally lacked strong

* Corresponding author.
  E-mail address: xianqin.yang@agr.gc.ca (X. Yang).

biofilm formers, unlike the generic *E. coli*, i.e., *E. coli* that is not of any particular pathotype, recovered from beef processing environment (Stanford et al., 2021). However, the overall level of heat resistance, as determined by $D_{60°C}$, of these two groups of *E. coli* was similar (Zhang et al., 2020). As well, their carriage of the genetic marker for extreme heat resistance tLST did not differ, as determined by PCR. Despite that, individual strains in the Top seven *E. coli* group varied largely in heat resistance, with $D_{60°C}$ ranging from 0 to 5.60 min (Zhang et al., 2020). Similarly, the reduction in response to lactic acid treatment ranged from 0 to 4.65 log (Castro et al., 2023). The genome of *E. coli* is very plastic, owing to frequently occurring horizontal gene transfer events, and this plasticity allows the organism to survive/ thrive in ecological niches other than its primary host as well as to acquire genes conferring virulence and resistance to stress (Ishii et al., 2006; Zhi et al., 2016). This evolution/diversification of *E. coli* is shaped by its phylogenetic background and habitat to similar extents (Touchon et al., 2020). *E. coli* can be assigned to one of eight phylogroups (Clermont et al., 2013), namely A, B1, B2, and C-G, which may differ in ecological niches and resistance profiles, e.g., antibiotic resistance and heat resistance (Zhang & Yang, 2022). Comparative genomic analysis of generic *E. coli* from meat processing environments showed that all tLST-harboring strains were in phylogroup A (Yang et al., 2021). Also, the tLST-positive strains had fewer genes involved in virulence and epithelial attachment, compared to tLST-negative counterparts from the same environment.

The principal virulence factor of STEC is Shiga toxin, encoded by *stx* located on lambdoid prophages (Hayashi et al., 2001), reflecting their mobile nature. In addition to *stx*, other accessory VF such as *eae* (encoding for intimin which is essential for enterocyte effacement), *tir* (translocated intimin receptor), *espAB* (encoding for components of the needle complex type III secretion system), and *hylA* or *ehxA* (encoding for hemolysin) are often associated with the hemolytic uremic syndrome, which can lead to kidney failure and death in some cases (Fu et al., 2018). These VF are referred to as EHEC (enterohemorrhagic *E. coli*) genes. EHEC is a subgroup of STEC that causes more severe clinical outcomes. Coexistence of tLST and *stx* as determined by PCR has been reported (Ma & Chui, 2017). On the other hand, the presence of *stx* and tLST has been found incompatible by a large-scale genomic analysis (Zhang & Yang, 2022). Ever since the emergence of STEC O157 as an important foodborne pathogen, serotype has been used as a determinant for identifying STEC strains that have the potential to cause severe human diseases, including testing beef trimmings for the presence of the organism (USDA-FSIS, 2013). Even so, serotypes are not virulence factors and they do not predict the presence of *stx* or accessory virulence genes. Thus, a better understanding of virulence and resistance profiles would aid the downstream control of the Top seven *E. coli* in meat processing. The present study then aimed to detect the genes mediating resistance to environmental stress and involved in virulence by AMRFinderPlus and characterize their association with phenotypic characteristics and phylogeny using the Top seven *E. coli* recovered from cattle.

Material and Methods

Bacterial strains. In previous studies, 750 Top seven *E. coli* isolates were assessed for their resistance to heat (Zhang et al., 2020) and lactic acid (Castro et al., 2023), and biofilm-forming ability (Stanford et al., 2021). These *E. coli* isolates were recovered from feedlot cattle or their environment in Alberta (Stanford et al., 2005, 2014, 2013, 2016; Stephens et al., 2009) and were initially confirmed to be one of the Top seven serogroups, O26, O45, O103, O111, O121, O145, and O157, by PCR (Stanford et al., 2016). Based on the phenotypical characteristics of these 750 isolates in relation to biofilm-forming ability, heat and acid resistance, 155 isolates were selected for whole-genome sequencing in this study (Tables 1 and S1). Generally, all

the resistant phenotypes among the 750 isolates were selected. Selection of other phenotypes aimed to have a roughly equal distribution among different serotypes.

Whole-genome sequencing, assembly, and annotation. The selected *E. coli* isolates were each streaked onto MacConkey agar (Oxoid) and incubated at 35°C for 24 ± 2 h. A single colony was then subcultured in 5 mL Luria-Bertani (LB) broth (BD Difco, Fisher Scientific, Canada) and incubated for 18 h at 35°C, shaking at 80–100 rpm. The overnight cultures were subjected to DNA extraction using a QIA-GEN DNeasy Blood & Tissue kit (QIAGEN) according to the manufacturer's instructions. DNA samples were prepared into shotgun libraries and subsequently sequenced using an Illumina NovaSeq 6000 or a HiSeq platform (150 bp, PE) with a targeted sequencing depth >300 X (Genome Quebec).

The quality of reads was assessed using FastQC v0.11.7, and the adaptor sequences and low-quality reads were removed by Trimmomatic v0.39, using default parameters (Bolger et al., 2014). SPAdes v3.14.0 was used to assemble the genomes with the most optimal k-mer value for each genome (Bankevich et al., 2012). The quality of the assemblies was assessed by Quast v5.0.2 (Gurevich et al., 2013). Finally, the genomes were annotated using Prokka v1.14.6 (https:// github.com/tseemann/prokka). Genome data are deposited under BioProject PRJNA846735.

In silico **sero- and phylo-typing.** The species identity and serotype of each genome were predicted using ECTyper v1.0 (https://github.-com/phac-nml/ecoli_serotyping) with default settings. Phylogroups were determined by Clermont Typing 20.03 (https://github.com/A-BN/ClermonTyping), an *in silico* PCR method (Clermont et al., 2013).

Core genome phylogenetic analysis. The pan-genome of the isolates was parsed using Roary v3.13.0 (Page et al., 2015). Genes shared by all the genomes with ≥95% amino acid identity were regarded as core genes (*n* = 3087). The core genes were concatenated and used to construct a maximum likelihood phylogenetic tree using RAxML with the general time reversible gamma nucleotide model (GTRGAMMA) and bootstrapping for 1,000 iterations (Stamatakis, 2014). *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 (GenBank accession: GCA_000006945.2) was included as an outgroup (McClelland et al., 2001).

Detection of resistance genes and virulence factors (VF) All genomes were screened for tLST using ABRicate v1.0.1 (https:// github.com/tseemann/abricate) with ≥80% identity and ≥80% coverage using the following as reference sequences for two tLST variants of tLST1 and tLST2. The tLST1 was retrieved from *E. coli* AW1.7 and CX16 which were recovered from beef processing plants (GenBank accession: LDYJ01000141; CP081748.1) (Mercer et al., 2015; Yang et al., 2021). The tLST2 was retrieved from a clinical isolate *E. coli* FAM21805 on a plasmid, and a beef processing plant isolate *E. coli* 0H15 (GenBank accession: KY416992.1; CP081678.1) (Boll et al.,

Table 1

Phenotypic characteristics of the Top seven *E. coli* isolates included in this study

Phenotypes and categories	Category	Criteria	No. of isolates
Heat	Resistant	$D_{60°C} \geq 2.0$ min	22
	Sensitive	$D_{60°C} < 2.0$ min	133
Acid	Resistant	$Log(N_0/N) \leq 0.5$	28
	Mediocre	$0.5 < Log(N_0/N) < 4.0$	117
	Sensitive	$Log(N_0/N) \geq 4.0$	10
Biofilm	Forming	$OD > 2 \times ODc$[a]	22
	Nonforming	$OD < ODc$	120

[a] Biofilm formation was assessed by the crystal violet staining method. ODc, cut-off optical density (OD) equals three times the standard deviation of the OD of the negative control plus the average OD of the negative control. The weak biofilm former with OD between ODc and 2xODc (*n* = 13) were not included in biofilm-related comparisons.

2017; Yang et al., 2021). Each *E. coli* genome was screened for virulence, antibiotic resistance (AMR), and stress resistance genes using the AMRFinderPlus 3.10.23 (Feldgarden et al., 2021) with default settings. Gene names and resistance classification were retrieved from AMRFinderPlus. Stx subtypes in the AMRFinderPlus were further verified by examining individual sequences. Stress resistance genes included those that are involved/predicted to be involved in heat, biocide/acid, and metal resistance (Feldgarden et al., 2021). AMR included genes required for a specific or multiple drug resistance.

Detection of genes discriminating phylogroups A and B1. To determine the accessory genes that discriminate *E. coli* phylogroups A and B1, *E. coli* genomes were searched against sequences in the Roary pan-genome using ABRricate v1.0.1 with >80% identity and >80% coverage. The ABRicate output was further analyzed using Scoary (Brynildsrud et al., 2016). Phylogroup A or B1-associated genes were determined as those presented in >80% in a phylogroup but in <20% in the other phylogroup with a significant difference ($P < 0.05$).

Statistical analysis. Fisher's exact test was used to assess the distribution of genes among different phylogroups and phenotypes, including heat, acid phenotypes, and biofilm formation of *E. coli*. Significant differences were determined by a P value < 0.05.

Results

Prevalence of tLST. The tLST1 was found in 15 genomes (9.7%); however, no tLST2 were found. All but one of the 16 strains positive for tLST determined by PCR (Zhang et al., 2020) harbored tLST in their genomes. The tLST1 from *E. coli* AW1.7 and CX16 differed by the *ftsH* sequence, which is truncated in the former but in full length in the latter strain (Yang et al., 2021). Of the 15 tLST1-positive *E. coli*, only one *E. coli* (ID: E001) encoded the partial component of *ftsH*, and the rest had the full length of *ftsH*. None of the tLST-positive strains harbored any *stx* genes. Eight heat resistance genes (*hsp20, shsP, yfdX1, yfdX2, hdeD.GI, trxLHR, kefB.GI, psi. GI*) were found from the AMRFinderPlus search. These genes were 100% identical with the eight ORFs of tLST1 and were used as heat resistance markers in further analysis in this work.

Population structure. The 155 *E. coli* belonged to six phylogroups: A ($n = 32$), B1 ($n = 93$), C ($n = 3$), D ($n = 11$), E ($n = 15$), and G ($n = 1$), accounting for 20.6, 60.0, 1.9, 7.1, 9.7, and 0.6% of the total population, respectively (Table 2). Interestingly, the population structures of the biofilm-forming group and heat-resistant group differed from that of the overall population, reflected by the much higher fractions of phylogroup A relative to B1. On the contrary, the slightly lower fraction of phylogroup B1 in the acid-resistant population, compared to the overall population structure, was compensated by the increase in phylogroups D and E. Different phylogroups were clearly separated on the core genome tree (Fig. 1A). Of the 155 genomes, 15 were predicted to be non-Top seven O-types (Table 3) which differs from the serotyping by PCR. For the other 140 genomes, the serotypes O103:H2, O121:H7, O145:H-, O157:H12, O157:H7, and O26:H11

each accounted for a sizable fraction of the respective O-group (>40%). *E. coli* of the same serotype clustered in closely related lineages (Fig. 1A). For example, *E. coli* O157:H12 and O157:H7 were in phylogroups A and E, respectively, and all *E. coli* O145: H- were in phylogroup D.

Phylogroups A and B1 differed by more than 200 genes (Fig. 1B), with 99 and 128 genes overrepresented in the respective groups (Table S2). Phylogroup A and B1 had 55 and nine genes exclusively. The genes only found in phylogroup A were *chiA,* the *gsp* cluster, two toxin-antitoxin systems (*mqsAR, yafNO*), as well as genes encoding for a fosmidomycin efflux transporter and multidrug efflux transporter *MdfA*, among others. The *ydj* operon, coding for a novel carbohydrate pathway, was highly abundant in group A (81-97%). Around half of the genes (47.6%) unique to phylogroup B1 were involved in putative functions.

EHEC genes and other VF. A total of 33 *E. coli* genomes (21.3%) carried *stx* genes, with subtype stx_{1a} being the most prevalent ($n = 16$; Fig. 2). Other subtypes included stx_{1d} ($n = 7$), and three stx_2 subtypes 2a ($n = 11$), 2c ($n = 3$), and 2e ($n = 1$). Of the 33 genomes, five harbored both stx_1 and stx_2 with strains 177 and E117 having stx_{1a} and stx_{2a} and strains E015, E030, and 94 having stx_{1a} and stx_{2c}. These five strains were all of serotype O157:H7. The remaining 28 genomes only had one *stx*. All but E122 (O84:H2) were among the Top seven serogroups and this potentially emerging STEC carried stx_{1a}. All *stx* subtypes cooccurred with the additional five EHEC genes (*eae, tir, espA, espB,* and *ehxA*) except for stx_{1d} and stx_{2e} which were found in genomes harboring only one of the additional EHEC genes, *hylA* (Table S3).

In total, 70 VF were found in 146 of the 155 (94.2%) genomes (Table S3) and the number of virulence genes in individual genomes ranged from 1 to 30. Interestingly, the genomes carrying stx_{1d} and stx_{2e}, which were all O121, had the least number of VF, 7-10, while the other *stx*-positive genomes had 19-30. VF present in more than 50% of the genomes included *ssIE* (65.8%), *IfpA* (64.5%), *fdeC* (87.1%), *Iss* (51.0%), and *espX1* (66.5%).

Stress resistance genes and ARGs. In total, 41 and 32 genes related to stress and antibiotic resistance were identified (Table S3). The stress resistance genes included those for metal ($n = 29$), biocides/acid ($n = 4$), and heat resistance ($n = 8$). Of the 155 genomes, 76 (49%) harbored metal resistance genes, including those for tellurium (*terDWZ*), arsenate (*arsADR*), mercury (*merABCDEPTR*), copper (*pocABCDERS*), and silver (*silABCEFPRS*) resistance. The most prevalent genes were *terDWZ*, at 40%, and all other genes were at <6%. All O157:H7 harbored *terDWZ*. Six genomes (E001, E054, E055, E065, E081, and E082) harbored between 12 and 17 metal resistance genes, for resistance to copper, silver, and tellurium. The remaining 70 genomes had ≤7 metal resistance genes primarily with a complete operon for resistance to one type of metal, with two genomes for mercury resistance, two genomes for arsenate and most of the remaining genomes for tellurium.

For biocide/acid resistance genes, *ymgB,* involved in acid resistance and regulating biofilm formation through repression of biofilm formation (Lee et al., 2007) and *emrE* (encoding for a multidrug resistance

Table 2
Population structure of Top seven *E. coli* isolates[a]

Phenotypes	Phylogroups (%)					
	A	B1	C	D	E	G
Biofilm former	59.1	31.8	4.5	–	4.5	–
Heat resistant	54.5	45.5	–	–	–	–
Acid resistant	25.0	50.0	–	10.7	14.3	–
Overall population	20.6	60.0	1.9	7.1	9.7	0.6

-, that phylogroup was not found in a particular population.
[a] The population structure is defined as the composition of phylogroups in each phenotype group or the overall population. Biofilm former, acid-resistant, heat-resistant phenotype group, and overall population included 22, 28, 22, and 155 isolates.
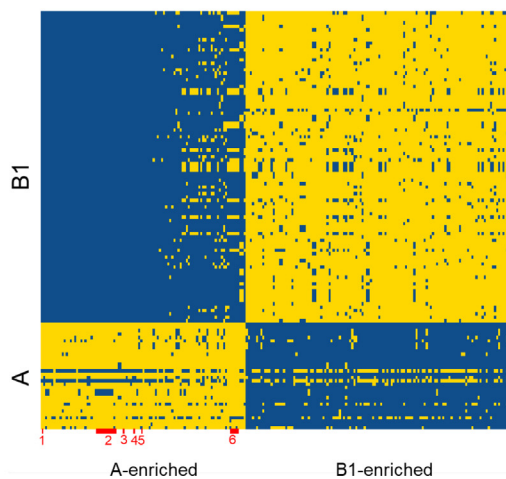
A
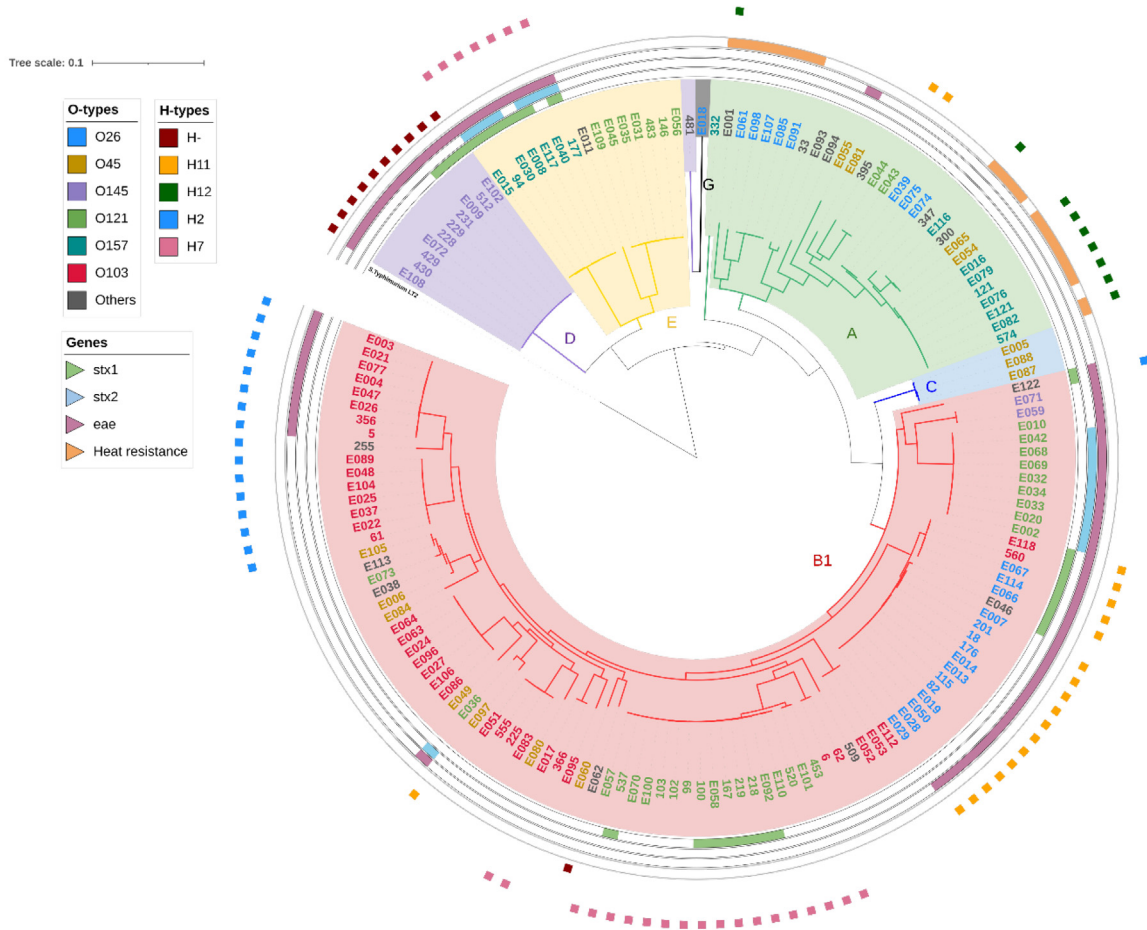
B

Figure 1. Phylogeny of the *E. coli* genomes. A core genome phylogenetic tree was constructed using 155 *E. coli* genomes (**A**). Branches were color-coded to represent different phylogroups: A (green), B1 (red), C (blue), D (purple), E (yellow), and G (black). Isolates were color-labeled to represent different O-types shown in the legend. Color strips in the rings, from outer to inner, indicate the presence of genes for heat resistance (*yfdX2, yfdX1, trxLHR, shsP, psi.GI, kefB.GI, hsp20, hdeD.GI*), and *stx₂, stx₁,* and *eae*. Next to the rings, squares with different colors represent the major H-types. Heat map (**B**) shows the genes distinguishing phylogroups A and B1 *E. coli*: 1. *chiA,* endochitinase; 2. *gsp cluster,* type II secretion system, 3. *mqsAR,* toxin-antitoxin system, 4. *yafNO,* toxin-antitoxin system, 5 and 6, *ydjL* and *ydjEFGHJ,* novel long-chain sugar-acid pathway. The grid is colored coded: blue (absence) and yellow (presence).

efflux transporter) were found in 92.3% (*n* = 143) and 54.8% (*n* = 85) of the genomes. The other two genes *qacE*Δ1 and *qacL* were found in two and one genomes in total, respectively.

All genomes carried ARGs, with the number per genome ranging from 2 to 15 (Table S3). One or more genes encoded for resistance to β-lactam (93.5%), fosfomycin (88.4%), tetracycline (18.1%), streptomycin (10.3%), sulphonamide (5.1%), trimethoprim (2.6%), florfenicol (1.3%), bleomycin (1.3%), streptothricin (0.6%), chloramphenicol (0.6%), and multidrug transport systems (100%). The most prevalent genes were *blaEC* (93.5%), *acrF* (94.2%), *mdtM*

Table 3
Summary of O- and H-types[a] of the *E. coli* strains as determined by *in silico* serotyping

O-types	H-types (% of the O-type)	No. of genomes
O103	H11, H14, H16, **H2 (41.6)**, H21, H38, H7, H8	36
O121	H10, H11, H19, H21, H46, **H7 (45.9)**	37
O145	H25, **H- (83.3)**	12
O157	**H12 (50), H7 (43.8)**, H29	16
O26	**H11 (62.5)**, H18, H32, H46, H9	24
O45	H11, H14, H19, H30, H34, H4, H45, H51, H8, H9	14
O111	H8	1
Others: -O, O128	-:H34, -:-, O128:H12, O51:H14,	15
O51, O53, O76, O8, O84,	O53:H32, O76:H34, O84:H2, O88:H4,	
O88, O9	O8:H25, H2 or H10, O9:H4, H9 or H30	

[a] H-types in bold represent the H-type of *E. coli* accounting for >40% of the total of the corresponding O-group indicated in the same row.
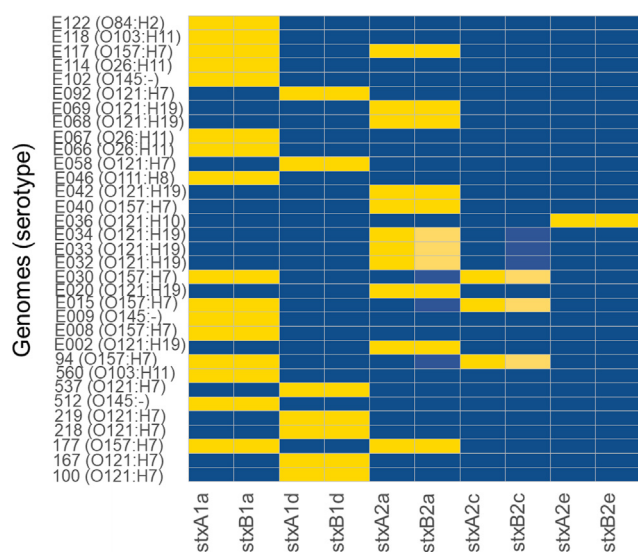

Figure 2. Subtypes of *stx* genes in *E. coli* predicted by *in silico* genotyping. The grid is colored coded: blue (absence) and yellow (presence). The subtypes included 1a (*n* = 16), 1d (*n* = 7), 2c (*n* = 3), 2e (*n* = 1), and 2a (*n* = 11). In all cases, *stx* contained both subunits: A1 and B1 in the case of *stx₁* or A2 and B2 in the case of *stx₂*.

(100%), and *glpT_E448K* (88.4%). The other 28 genes were mostly found in ≤5% of the genomes.

Differential distribution of stress resistance genes, ARGs and VF in phylogroups. Strains in phylogroups A, B1, D, and E were used to compare gene distribution as phylogroups C and G each included ≤3 genomes. Of the 41 stress resistance genes and 32 ARGs, 29 and 6 were differentially ($P < 0.05$) distributed among the phylogroups (Fig. 3A; Table S4). All heat resistance genes ($n = 8$) were differentially distributed and were exclusively in phylogroup A. All differentially distributed metal resistance genes ($n = 18$) except for the *ter* cluster ($n = 3$) were also exclusively in phylogroup A, while the differentially distributed genes involved in resistance to biocides ($n = 2$), antibiotics ($n = 6$), or tellurium ($n = 3$) were more spread out among the different phylogroups. The gene *emrE*, contributing to the resistance to multiple drugs/antiseptics (Kovacevic et al., 2016; Ma & Chang, 2004), was present in more than 50% of all phylogroups except for phylogroup D where it was absent. In contrary to the higher prevalence of heat and most metal resistance genes in phylogroup A, virulence genes were least present in phylogroup A, followed by B1, E, and D, with this trend most marked for genes encoding adhesins and secretion systems (Fig. 3B).

Differential distribution of stress resistance genes, ARGs, and VF in different pheno-groups Stress-resistant populations, i.e., resistant to heat and acid, and biofilm formers as determined by

phenotyping in previous studies, each accounted for ≤16% of the total population (Table S1). Genes related to heat resistance were significantly more prevalent in the heat-resistant group, 50% vs 3% (Fig. 4A; Table S5). The biofilm-forming population had a significantly higher prevalence of genes involved in heat resistance (~fourfold; 8/8 genes), metal resistance (~sevenfold; 12/15 genes), biocide resistance (*emrE*), and antibiotic resistance (7/8 genes). However, there was no difference in *ymgB* between biofilm and non–biofilm-forming populations ($P > 0.05$). The prevalence of *ymgB* and *emrE* decreased with increasing acid resistance.

The heat-resistant and heat-sensitive populations had equal numbers, but different set of significantly more abundant VF, with *efa*, *toxB*, and *espF* being more abundant in the resistant population and *ipfA1*, *hlyA*, *sslE*, and *espX1* being more abundant in the sensitive group (Fig. 4B). In contrast to the higher abundance of stress genes in the biofilm-forming population, VFs were mostly more abundant in the non–biofilm-forming population when differences in carriage were found. For adhesions and toxins, the distribution of VF was mixed, with *eae*, *iha*, *katP*, and *toxB* exclusively in the acid-resistant population. All significantly different secretion system genes except for *ssIE* were only found in the acid-resistant group. Overall, heat-resistant and biofilm-forming *E. coli* were positively correlated with stress resistance genes, whereas the acid-resistant population was positively correlated with VF, in particular the EHEC genes, genes in the type III secretion system, and non-LEE effectors (*nle*).

Discussion

On the basis of amino acid sequences, Stx (Stx1 and Stx2) were initially grouped into 10 subtypes by Scheutz et al. in 2012 (Scheutz et al., 2012), and since then, eight additional subtypes have been reported (Gill et al., 2022). The Stx subtype may vary with animal host and region of isolation. In the present work, five subtypes (1a, 1d, 2a, 2c, and 2e) were identified, with Stx₁ₐ being the most prevalent and most STEC genomes (28/33) harboring only one Stx. All genomes with both Stx1 and Stx2 were O157:H7. This is largely in agreement with reports on Top seven STEC from cattle in North America (Chui et al., 2015; Shridhar et al., 2017; Zhang et al., 2021). All *stx* genes were accompanied with the five additional EHEC genes except for Stx₂ₑ (O121:H10) and Stx₁d (O121:H7). These two groups not only lacked most of the additional EHEC genes but also missed most of the accessory virulence genes found in other Stx subtype genomes in the present work. A study assessing the pathogenicity of STEC of different virulence profiles showed a very low incidence of severe outcomes associated with the subtype *stx₁d* or *stx₂ₑ* STEC (EFSA Biohaz Panel et al., 2020). Another study showed that most strains of O121 which harbored *stx₁d* as well as other subtypes of Stx lacked both *eae* and *ehxA*, in contrast to those O121 strains that did not harbor *stx₁d* (Carter et al., 2019).
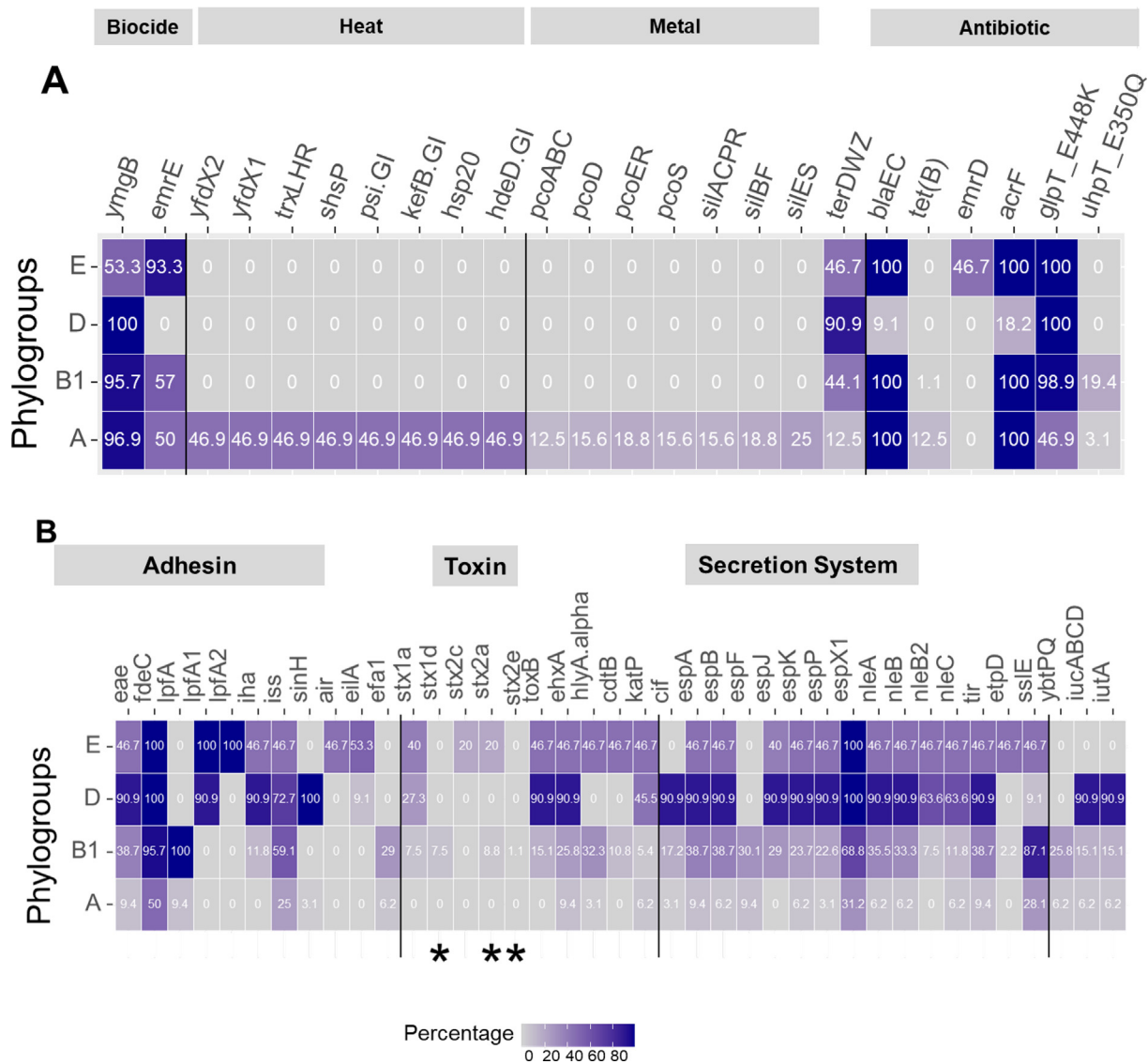
Figure 3. Differential distribution of antibiotic and stress resistance genes (A) and virulence factors (B) in different *E. coli* phylogroups. Genes that differed in their distribution among the groups and all *stx* subtypes were included. The two *stx* genes did not differ significantly in their distribution and were indicated by a single asterisk (*P* > 0.05). Phylogroups A, B1, D, and E included 32, 93, 11, and 15 isolates, respectively. Gene prevalence for each phylogroup represented as percentage (%).

The proportion of Top seven carrying *stx* was 22.9%, with 100% for O157:H7 and 18.8% for the other six serogroups. All 155 isolates used in this study had been previously tested positive for *stx* and were of the seven serogroups as determined by PCR (Zhang et al., 2020). This inconsistency likely resulted from the difference in methodology for detection, PCR vs whole-genome sequencing (Castro et al., 2021). Regardless, the carriage of *stx* in different serogroups varies with cattle origin and serogroups. A Canadian study found that *stx* was present in >90% of O157, and 65.3–100% in the six non-O157 serogroup isolates from western-Canadian slaughter cattle (Stanford et al., 2016). Pearce *et al.* found the prevalence of *stx* in O145 and O103 was rare (<5%) and was at 28.9% in O26 and 49.0% in O145 shed by cattle in Scotland (Pearce et al., 2006). In a US study on Top 7 *E. coli* in feedlot cattle, almost all O157 harbored *stx*, while its presence in the other six serogroups was rare (Dewsbury et al., 2015). These findings suggest that virulence profile rather than serogroup-focused testing for STEC in food commodities would better inform risk assessment, especially as one of the STEC in the present study was O84:H2, which does not fall under the Top seven *E. coli*. Similarly, focusing on serogroup

would also exclude the *E. coli* 104 strain which caused a very large outbreak in Germany in 2011 (Grad et al., 2012)

In the present study, sizable fractions (10-93.5%) of the 155 genomes harbored one or more genes conferring resistance to β-lactam, fosfomycin, tetracycline, and streptomycin. A large-scale analysis (n = 18,959) of *E. coli* genomes found genes associated with resistance to β-lactam, aminoglycoside, sulfonamide, and tetracycline were present in more than 40% of the genomes, while <10% of the genomes had resistance genes to fosfomycin (Zhang & Yang, 2022). The high prevalence of fosfomycin resistance in the present work was driven by a point mutation in *glpT_E448K*. Information on fosfomycin resistance in *E. coli* from cattle is scarce. This point mutation *glpT_E448K* was found in 90% of *E. coli* from clinical and environmental samples (Altayb et al., 2022). Among all detected ARGs and point mutations, the *blaEC* gene (93.5%), which encodes class C beta-lactamase, the multidrug efflux-related transporter protein-encoding genes *acrF* (94.2%) and *mdtM* (100%) and *emrE* (92.3%) were most prevalent in the Top seven *E. coli* genomes in the present study. A comparable prevalence for all these four genes has been reported for non-Top
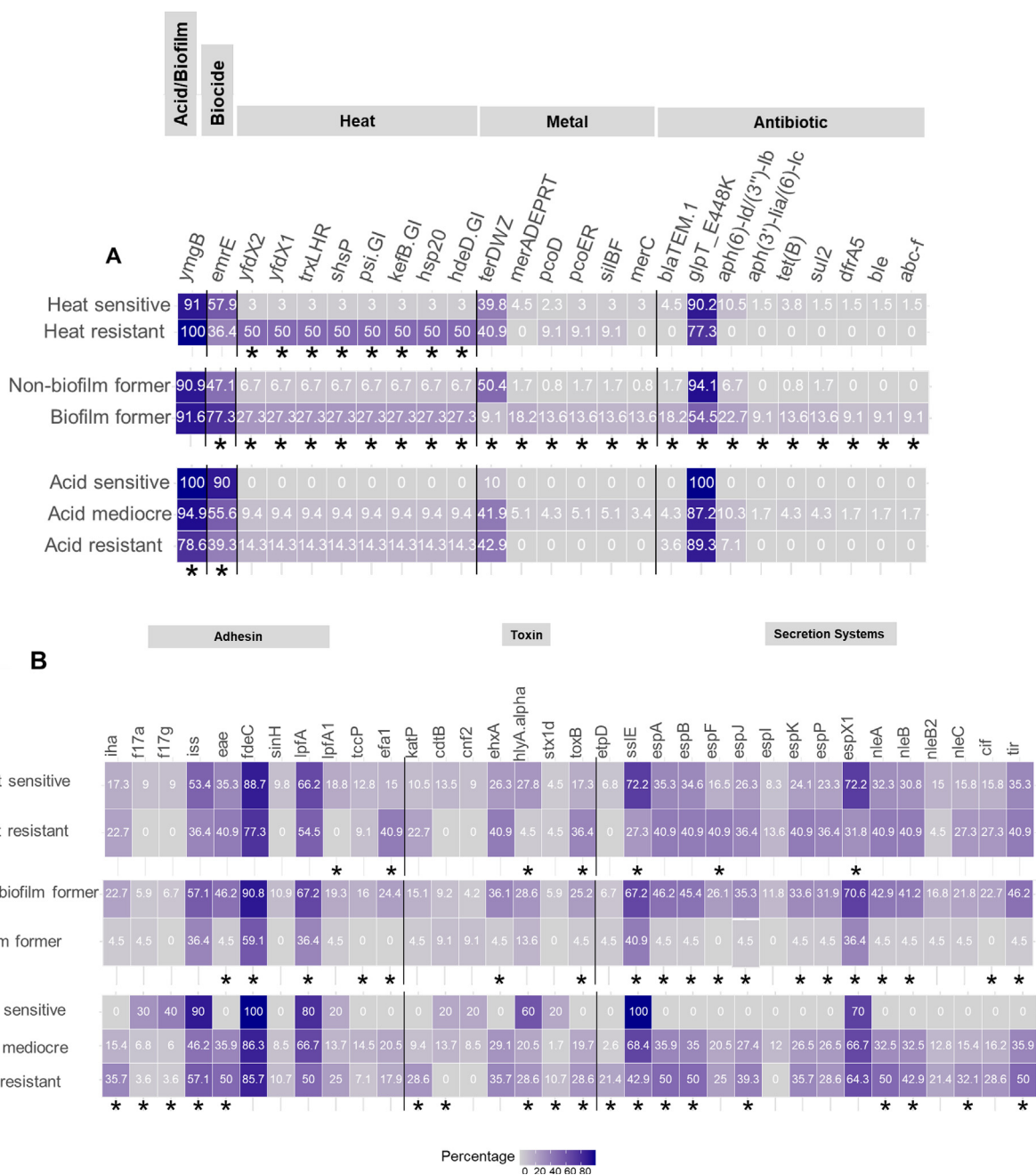
Figure 4. Distribution of genes encoding antibiotic resistance and stress resistance (A), and virulence factors (B) in *E. coli* of different phenotypes. Panels show the gene prevalence (%) in *E. coli* grouped by their heat, acid, and biofilm-forming phenotypes. The composition of the population was: heat sensitive ($n = 133$) and resistant ($n = 22$); acid sensitive (10), mediocre ($n = 117$), and resistant ($n = 28$); biofilm-forming ($n = 22$) and nonforming ($n = 120$). Genes significantly associated with different categories of phenotypes were indicated by a single asterisk ($P < 0.05$).

seven STEC ($n = 284$) based on genomic analysis (Huang et al., 2021). Heavy metal resistance in *E. coli* originating from livestock is not often examined. Genes encoding resistance to copper, silver, and mercury were found in 8.8, 23.8, and 17.5% of *E. coli* isolated from pig slaughterhouse in the UK (Yang et al., 2020). The prevalence of these genes found in the present study was <6%, slightly lower than that in the previous report. The findings that all *E. coli* O157:H7 harbored *ter* genes and the other serogroups had an overall prevalence of 40% but a varying distribution in individual serogroups were in agreement with a previous study reporting variability in tellurite resistance and the *ter* gene cluster among STEC from different sources (Orth et al., 2007).

The 155 *E. coli* were primarily of phylogroup B1 (60%) and A (20.6%), and the predominance of these two phylogroups reflects the population structure of *E. coli* of bovine origin and of beef processing environments (Bok et al., 2015; Fang et al., 2022; Yang et al., 2021). However, the *E. coli* populations from animals had a higher fraction of phylogroups D and E, compared to those from meat processing environments. The finding that all tLST (or the eight heat resistance marker genes) were found in phylogroup A is also in agreement with previous reports (Zhang & Yang, 2022; Zhi et al., 2016). It is then not surprising that the heat-resistant population had an increased proportion of phylogroup A (54.5%). However, not all heat-resistant phenotype harbored tLST and not all tLST-harboring

E. coli were phenotypically heat resistant, and this inconsistency would suggest mechanisms other than tLST mediating heat resistance in *E. coli*. In contrast with the heat resistance and metal resistance genes being most concentrated in phylogroup A, virulence genes particularly those for adhesins and secretion systems were mostly found in D and E, suggesting differences in stress relevant to intestinal environments of humans and animals and the environments shape the *E. coli* populations surviving in these ecological niches (Touchon et al., 2020). The higher prevalence of stress-resistance genes (biocide, heat, metal, and antibiotics) in the biofilm-forming population may further increase the survival of these bacteria in adverse environments. In the present study, only one of the biofilm formers harbored *stx* genes (E117) and belonged to phylogroup E. STEC are in general poor biofilm formers, compared to generic *E. coli* (Stanford et al., 2021), and this has been attributed to the frequent inactivation of the biofilm regulatory gene *mlrA* by the insertion of the Stx prophage (Uhlich et al., 2013).

All five additional EHEC genes (*eae, tir, espA, espB,* and *ehxA*) were exclusively present in the acid-resistant population along with VF in the secretion system. Four systems have been recognized for acid resistance in *E. coli* (Kanjee & Houry, 2013), none of which involves these EHEC genes. However, it has been reported that STEC are more resistant to lactic acid treatment than generic *E. coli* (Gill et al., 2019) and the *E. coli* population harboring *eae* was more acid-resistant than other *E. coli* in samples from cattle hides (Wang et al., 2014). These findings suggest that even though these virulence genes do not contribute to acid resistance, the genetic background for acquiring these virulence genes may be beneficial for their survival in acid environments such as the gastric system. Despite the difference in VF and stress resistance genes between phylogroups, phylogroups A and B1 also differed significantly in gene content, including toxin-antitoxin systems, which may be associated with ecological benefits. It was hypothesized that toxin-antitoxin systems such as *yafNO* and *mqsAR* are part of stress responses and involved in adaptation to nutrient starvation by arresting cell growth (Christensen-Dalsgaard et al., 2010). The *mqsR* (**m**otility **q**uorum-**s**ensing **r**egulators) was found positively affecting flagella and motility gene expression, and negatively affecting biofilm formation (Sun et al., 2017).

In conclusion, the findings of this study demonstrate that only a fraction of Top seven *E. coli* harbored the principal virulence factor *stx* for STEC and Stx subtypes 1d and 2e were not accompanied by genes commonly recognized for severe human infection outcomes. Thus, virulence profile based rather than serotype-based testing is more appropriate for low infectious dose STEC in food. The Top seven *E. coli* population diverged into different groups for better host survival (adhesions and secretion system) or adverse secondary environmental survival (metal resistance, heat resistance, and biofilm formation). Even so, the association of virulence genes with the acid-resistant population would suggest pathogenic strains should be included when validating acid-based antimicrobial interventions for food. The consistency between genotype to phenotype of antibiotic resistance was higher than >90% using AMRFinder (Feldgarden et al., 2019). Comparing to antibiotic resistance, the consistency between the two methods was much lower for acid and heat resistance, and biofilm formation. Thus, further development of a database of stress resistance would greatly benefit the application of sequence-based tools for accurate phenotype prediction.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jfp.2023.100122.

References

Arthur, T. M., Bosilevac, J. M., Brichta-Harhay, D. M., Guerini, M. N., Kalchayanand, N., Shackelford, S. D., Wheeler, T. L., & Koohmaraie, M. (2007). Transportation and lairage environment effects on prevalence, numbers, and diversity of *Escherichia coli* O157:H7 on hides and carcasses of beef cattle at processing. *Journal of Food Protection, 70*(2), 280–286.

Altayb, H. N., Elbadawi, H. S., Alzahrani, F. A., Baothman, O., Kazmi, I., Nadeem, M. S., Hosawi, S., & Chaieb, K. (2022). Co-occurrence of β-lactam and aminoglycoside resistance determinants among clinical and environmental isolates of *Klebsiella pneumoniae* and *Escherichia coli*: A genomic approach. *Pharmaceuticals, 15*(1011), 1–21.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology, 19*(5), 455–477.

Bell, B. P., Goldoft, M., Griffin, P. M., Davis, M. A., Gordon, D. C., Tarr, P. I., Bartleson, C. A., Lewis, J. H., Barrett, T. J., & Wells, J. G. (1994). A multistate outbreak of *Escherichia coli* O157: H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers: The Washington experience. *Journal of the American Medical Association, 272*(17), 1349–1353.

Bok, E., Mazurek, J., Stosik, M., Wojciech, M., & Baldy-Chudzik, K. (2015). Prevalence of virulence determinants and antimicrobial resistance among commensal *Escherichia coli* derived from dairy and beef cattle. *International Journal of Environmental Research and Public Health, 12*(1), 970–985.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114–2120.

Boll, E. J., Marti, R., Hasman, H., Overballe-Petersen, S., Stegger, M., Ng, K., Knøchel, S., Krogfelt, K. A., Hummerjohann, J., & Struveet, C. (2017). Turn up the heat—Food and clinical *Escherichia coli* isolates feature two transferrable loci of heat resistance. *Frontier Microbiology, 8*.

Brynildsrud, O., Bohlin, J., Scheffer, L., & Eldholm, V. (2016). Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biology, 17*(1), 238.

Castro, V. S., Fang, Y., Yang, X., & Stanford, K. (2023). Association of resistance to quaternary ammonium compounds and organic acids with genetic markers and their relationship to *Escherichia coli* serogroup. *Food Microbiology, 113* 104267.

Castro, V. S., Ortega Polo, R., Figueiredo, E. E. d. S., Bumunange, E. W., McAllister, T., King, R., Conte-Junior, C. A., & Stanford, K. (2021). Inconsistent PCR detection of Shiga toxin-producing *Escherichia coli*: Insights from whole genome sequence analyses. *PLoS One, 16*(9), e0257168.

Christensen-Dalsgaard, M., Jørgensen, M. G., & Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Molecular Microbiology, 75*(2), 333–348.

Chui, L., Li, V., Fach, P., Delannoy, S., Malejczyk, K., Patterson-Fortin, L., Poon, A., King, R., Simmonds, K., Scott, A. N., & Lee, M. C. (2015). Molecular profiling of *Escherichia coli* O157:H7 and non-O157 strains isolated from humans and cattle in Alberta, Canada. *Journal of Clinical Microbiology, 53*(3), 986–990.

Clermont, O., Christenson, J. K., Denamur, E., & Gordon, D. M. (2013). The Clermont *Escherichia coli* phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups. *Applied Microbiology International, 5*(1), 58–65.

Crim, S. M., Griffin, P. M., Tauxe, R., Marder, E. P., Gilliss, D., Cronquist, A. B., Cartter, M., Tobin-D'Angelo, M., Blythe, D., Smith, K., Lathrop, S., Zansky, S., Cieslak, P. R., Dunn, J., Holt, K. G., Wolpert, B., & Henao, O. L. (2015). Preliminary incidence and trends of infection with pathogens transmitted commonly through food - Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2006–2014. *MMWR, 64*(18), 495–499.

Denamur, E., Clermont, O., Bonacorsi, S., & Gordon, D. (2021). The population genetics of pathogenic *Escherichia coli. Nature Reviews Microbiology, 19*(1), 37–54.

Dewsbury, D. M. A., Renter, D. G., Shridhar, P. B., Noll, L. W., Shi, X., Nagaraja, T. G., & Cernicchiaro, N. (2015). Summer and winter prevalence of Shiga toxin–producing *Escherichia coli* (STEC) O26, O45, O103, O111, O121, O145, and O157 in feces of feedlot cattle. *Foodborne Pathogens and Disease, 12*(8), 726–732.

EFSA Biohaz Panel Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L., Hilbert, F., Lindqvist, R., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Jenkins, C., Monteiro Pires, S., Morabito, S., Niskanen, T., Scheutz, F., da Silva Felício, M. T., Messens, W., & Bolton, D. (2020). Pathogenicity assessment of Shiga toxin-

producing *Escherichia coli* (STEC) and the public health risk posed by contamination of food with STEC. *EFSA Journal, 18*(1), e05967.

Fang, Y., Stanford, K., & Yang, X. (2022). Lactic acid resistance and population structure of *Escherichia coli* from meat processing environment. *Microbiology Spectrum, 10*(5) e0135222.

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., Hoffmann, M., Pettengill, J. B., Prasad, A. B., Tillman, G. E., Tyson, G. H., & Klimke, W. (2021). AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific Reports, 11*(1), 12728.

Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., Tyson, G. H., Zhao, S., Hsu, C. H., McDermott, P. F., Tadesse, D. A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J. P., & Klimke, W. (2019). Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Scientific Reports, 63*(11) e00483-00419.

Fu, S., Bai, X., Fan, R., Sun, H., Xu, Y., & Xiong, Y. (2018). Genetic diversity of the enterohaemolysin gene (*ehxA*) in non-O157 Shiga toxin-producing *Escherichia coli* strains in China. *Scientific Reports, 8*(1), 4233.

Gill, A. (2020). *Review of the state of knowledge on verotoxigenic Escherichia coli and the role of whole genome sequencing as an emerging technology supporting regulatory food safety in*. Canada: Government of Canada.

Gill, A., Dussault, F., McMahon, T., Petronella, N., Wang, X., Cebelinski, E., Scheutz, F., Weedmark, K., Blais, B., & Carrillo, C. (2022). Characterization of atypical shiga toxin gene sequences and description of stx2j, a new subtype. *Journal of Clinical Microbiology, 60*(3) e0222921.

Gill, A., Tamber, S., & Yang, X. (2019). Relative response of populations of *Escherichia coli* and *Salmonella enterica* to exposure to thermal, alkaline and acidic treatments. *International Journal of Food Microbiology, 293*, 94–101.

Gordon, D. M., & Cowling, A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: Host and geographic effects. *Microbiology, 149*(12), 3575–3586.

Gould, L. H., Mody, R. K., Ong, K. L., Clogher, P., Cronquist, A. B., Garman, K. N., Lathrop, S., Medus, C., Spina, N. L., Webb, T. H, White, P. L., Wymore, K., Gierke, R. E., Mahon, B. E., & Griffin, P. M. (2013). 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 Pathogens and Disease, 10*(5), 453–460.

Grad, Y. H., Lipsitch, M., Feldgarden, M., Arachchi, H. M., Cerqueira, G. C., FitzGerald, M., Godfrey, P., Haas, B. J., Murphy, C. I., Russ, C., Sykes, S., Walker, B. J., Wortman, J. R., Young, S. Zeng, Q., Abouelleil, A., Bochicchio, J., Chauvin, S., DeSmet, T., Gujja, S., McCowan, C., Montmayeur, A., Steelman, S., Frimodt-Møller, J., Petersen, A. M., Struve, C., Krogfelt, K. A., Bingen, E., Weill, F. X., Lander, E. S., Nusbaum, C., Birren, B. W., Hung, D. T., & Hanage, W. P. (2012). Genomic epidemiology of the *Escherichia coli* O104: H4 outbreaks in Europe, 2011. *Proceedings of the National Academy of Sciences, 109*(8), 3065–3070.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics, 29*(8), 1072–1075.

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C. G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M., & Shinagawa, H. (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Research, 8*(1), 11–22.

Huang, X., Yang, X., Shi, X., Erickson, D. L., Nagaraja, T. G., & Meng, J. (2021). Whole-genome sequencing analysis of uncommon Shiga toxin-producing *Escherichia coli* from cattle: Virulence gene profiles, antimicrobial resistance predictions, and identification of novel O-serogroups. *Food Microbiology, 99* 103821.

Ishii, S., Ksoll, W. B., Hicks, R. E., & Sadowsky, M. J. (2006). Presence and growth of naturalized *Escherichia coli* in temperate soils from lake superior watersheds. *Applied and Environmental Microbiology, 72*(1), 612–621.

Kanjee, U., & Houry, W. A. (2013). Mechanisms of Acid Resistance in *Escherichia coli*. *Annual Review of Microbiology, 67*(1), 65–81.

Stanford, K., Hannon, S., Booker, C. W., & Jim, G. K. (2014). Variable efficacy of a vaccine and direct-fed microbial for controlling *Escherichia coli* O157:H7 in feces and on hides of feedlot cattle. *Foodborne Pathogens and Disease, 11*(5), 379–387.

Kovacevic, J., Ziegler, J., Wałecka-Zacharska, E., Reimer, A., Kitts, D. D., Gilmour, M. W., & Drakeet, H. L. (2016). Tolerance of *Listeria monocytogenes* to quaternary ammonium sanitizers is mediated by a novel efflux pump encoded by *emrE*. *Applied and Environmental Microbiology, 82*(3), 939–953.

Lee, J., Page, R., García-Contreras, R., Palermino, J.-M., Zhang, X.-S., Doshi, O., Wood, T. K., & Peti, W. (2007). Structure and function of the *Escherichia coli* protein YmgB: A protein critical for biofilm formation and acid-resistance. *Journal of Molecular Biology, 373*(1), 11–26.

Ma, A., & Chui, L. (2017). Identification of heat resistant *Escherichia coli* by qPCR for the locus of heat resistance. *Journal of Microbiological Methods, 133*, 87–89.

Ma, C., & Chang, G. (2004). Structure of the multidrug resistance efflux transporter *EmrE* from *Escherichia coli*. *Biophysics and Computational Biology, 101*(9), 2852–2857.

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, zM., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., & Wilson, R. K. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature, 413* (6858), 852–856.

Mercer, R., Zheng, J., Garcia-Hernandez, R., Ruan, L., Gänzle, M., & McMullen, L. (2015). Genetic determinants of heat resistance in *Escherichia coli*. *Frontiers in Microbiology, 6*(932).

Carter, M. Q., Tan, Z. F., Pham, A., Carychao, D. K., & Coole, a. M. B. (2019). A clonal Shiga toxin–producing *Escherichia coli* O121:H19 population exhibits diverse carbon utilization patterns. *Foodborne Pathogens and Disease, 16*(6), 384–393.

Orth, D., Grif, K., Dierich, M. P., & Würzner, R. (2007). Variability in tellurite resistance and the *ter* gene cluster among Shiga toxin-producing *Escherichia coli* isolated from humans, animals and food. *Research in Microbiology, 158*(2), 105–111.

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., Fookes, M., Falush, D., Keane, J. A., & Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics, 31*(22), 3691–3693.

Pearce, M. C., Evans, J., McKendrick, I. J., Smith, A. W., Knight, H. I., Mellor, D. J., Woolhouse, M. E. J., Gunn, G. J., & Low, J. C. (2006). Prevalence and virulence factors of *Escherichia coli* serogroups O26, O103, O111, and O145 shed by cattle in Scotland. *Applied and Environment Microbiology, 72*(1), 653–659.

Pruimboom-Brees, I. M., Morgan, T. W., Ackermann, M. R., Nystrom, E. D., Samuel, J. E., Cornick, N. A., & Moon, H. W. (2000). Cattle lack vascular receptors for *Escherichia coli* O157:H7 Shiga toxins. *Proceedings of the National Academy of Sciences of the United States of America, 97*(19), 10325–10329.

Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R., Morabito, S., Strockbine, N. A., Melton-Celsa, A. R., Sanchez, M., Persson, S., & O'Brien, A. D. (2012). Multicenter evaluation of a sequence-based protocol for subtyping shiga toxins and standardizing stx nomenclature. *Journal of Clinical Microbiology, 50*(9), 2951–2963.

Shridhar, P. B., Siepker, C., Noll, L. W., Shi, X., Nagaraja, T. G., & Bai, J. (2017). Shiga toxin subtypes of non-O157 *Escherichia coli* serogroups isolated from cattle feces. *Frontiers in Cellular and Infection Microbiology, 7*, 121.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312–1313.

Stanford, K., Bach, S. J., Marx, T. H., Jones, S., Hansen, J. R., Wallins, G. L., Zahiroddini, H., & Mcallister, T. A. (2005). Monitoring *Escherichia coli* O157:H7 in inoculated and naturally colonized feedlot cattle and their Environment. *Journal of Food Protection, 68*(1), 26–33.

Stanford, K., Gibb, D., & McAllister, T. A. (2013). Evaluation of a shelf-stable direct-fed microbial for control of *Escherichia coli* O157 in commercial feedlot cattle. *Canadian Journal of Animal Science, 93*(4), 535–542.

Stanford, K., Johnson, R. P., Alexander, T. W., McAllister, T. A., & Reuter, T. (2016). Influence of season and feedlot location on prevalence and virulence factors of seven serogroups of *Escherichia coli* in feces of western-Canadian slaughter cattle. *PLoS One, 11*(8) e0159866.

Stanford, K., Tran, F., Zhang, P., & Yang, X. (2021). Biofilm-forming capacity of *Escherichia coli* isolated from cattle and beef packing plants: Relation to virulence attributes, stage of processing, antimicrobial interventions, and heat tolerance. *Applied and Environmental Microbiology, 87*(23) e01126-01121.

Stephens, T. P., McAllister, T. A., & Stanford, K. (2009). Perineal swabs reveal effect of super shedders on the transmission of *Escherichia coli* O157:H7 in commercial feedlots. *Journal of Animal Science, 87*(12), 4151–4160.

Sun, C., Guo, Y., Tang, K., Wen, Z., Li, B., Zeng, Z., & Wang, X. (2017). MqsR/MqsA toxin/antitoxin system regulates persistence and biofilm formation in *Pseudomonas putida* KT2440. *Frontiers in Microbiology, 8*, 840.

Touchon, M., Perrin, A., de Sousa, J. A. M., Vangchhia, B., Burn, S., O'Brien, C. L., Denamur, E., Gordon, D., & Rocha, E. P. (2020). Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*. *PLoS Genetics, 16*(6) e1008866-e1008866.

Uhlich, G. A., Chen, C. Y., Cottrell, B. J., Hofmann, C. S., Dudley, E. G., Strobaugh, T. P., & Nguyen, L. H. (2013). Phage insertion in *mlrA* and variations in *rpoS* limit curli expression and biofilm formation in *Escherichia coli* serotype O157: H7. *Microbiology, 159*(Pt 8), 1586–1596.

USDA-FSIS. (2013). *About the microbiological testing program for E. coli O157:H7 and non-O157 Shiga toxin-producing E. coli (STEC)*. Retrieved September 22, 2016, from https://www.fsis.usda.gov/science-data/data-sets-visualizations/microbiology/microbiological-testing-program-escherichia-coli

Wang, H., Gill, C. O., & Yang, X. (2014). Development of a real-time PCR procedure for quantification of viable *Escherichia coli* in populations of *E. coli* exposed to lactic acid, and the acid tolerance of verotoxigenic *E. coli* (VTEC) from cattle hides. *Food Control, 43*, 104–109.

WHO & FAO (2018). *Shiga toxin-producing Escherichia coli (STEC) and food: Attribution, characterization, and monitoring: Report*. Rome: World Health Organization.

Yang, H., Wei, S. H., Hobman, J. L., & Dodd, C. E. R. (2020). Antibiotic and metal resistance in *Escherichia coli* isolated from pig slaughterhouses in the United Kingdom. *Antibiotics, 9*(11).

Yang, X. (2016). Microbial ecology of beef carcasses and beef products. In A. d. S. Sant'Ana (Ed.), *Quantitative Microbiology in Food Processing: Modeling the Microbial Ecology*. Chichester, UK: John Wiley and Sons.

Yang, X., Badoni, M., Wang, H., & Gill, C. O. (2014). Effects of mild and pasteurizing heat treatments on survival of generic and verotoxigenic *Escherichia coli* from beef enrichment cultures. *Food Control, 39*, 100–104.

Yang, X., Tran, F., Zhang, P., & Wang, H. (2021). Genomic and phenotypic analysis of heat and sanitizer resistance in *Escherichia coli* from beef in relation to the locus of heat resistance. *Applied and Environment Microbiology, 87*(23) e01574-01521.

Zhang, P., Essendoubi, S., Keenliside, J., Reuter, T., Stanford, K., King, R., Lu, P., & Yang, X. (2021). Genomic analysis of Shiga toxin-producing *Escherichia coli* O157:H7 from cattle and pork-production related environments. *NPJ Science of Food, 5*(1), 15.

Zhang, P., Tran, F., Stanford, K., & Yang, X. (2020). Are antimicrobial interventions associated with heat-resistant *Escherichia coli* on meat? *Applied and Environmental Microbiology, 86*(13) e00512-00520.

Zhang, P., & Yang, X. (2022). Genetic characteristics of the transmissible locus of stress tolerance (tLST) and tlLST harboring *Escherichia coli* as revealed by large-scale genomic analysis. *Applied and Environment Microbiology, 88*(7) e02185-02121.

Zhi, S., Banting, G., Li, Q., Edge, T. A., Topp, E., Sokurenko, M., Scott, C., Braithwaite, S., Ruecker, N. J., Yasui, Y., McAllister, T. A., Chui, L., Neumann, N. F., & Stams, A. J. M. (2016). Evidence of naturalized stress-tolerant strains of *Escherichia coli* in municipal wastewater treatment plants. *Applied and Environmental Microbiology, 82*(18), 5505–5518.