

The complete genome sequence of five pre-2013 *Escherichia coli* sequence type (ST)1193 strains reveals insights into an emerging pathogen

Rhys T. White^{1,2,3,*}, Melinda M. Ashcroft^{1,2}, Michelle J. Bauer⁴, Jan Bell⁵, Dominika Butkiewicz⁴, Laura Álvarez-Fraga^{1,6}, Justine S. Gibson⁷, Amanda K. Kidsley⁵, Joanne L. Mollinger⁸, Kate M. Peters^{1,9}, Minh-Duy Phan^{1,9}, Leah W. Roberts^{1,2,4}, Benjamin A. Rogers¹⁰, Mark A. Schembri^{1,9,11}, Darren J. Trott⁵, John Turnidge⁷, Brian M. Forde^{9,11,†} and Scott A. Beatson^{1,2,11,*†}

Abstract

Fluoroquinolone-resistant *Escherichia coli* sequence type (ST)1193 is a profound, emerging lineage associated with systemic, urinary tract and neonatal infections. Humans, companion animals and the environment are reservoirs for ST1193, which has been disseminated globally. Following its detection in 2007, ST1193 has been identified repeatedly amongst fluoroquinolone-resistant clones in Australia. However, despite the growing importance of ST1193, only three complete genomes are published in the literature, none of which are from Australia. Here we expand on the available ST1193 resources with the complete genomes of five ST1193 strains sequenced using Oxford Nanopore Technologies and Illumina. Using *in silico* genotyping, we found that all strains were multi-drug resistant, including resistances to fluoroquinolones and cephalosporins. *In vitro* antibiotic susceptibility testing mostly correlated with individual genotypes. The exception was MS8320, which had additional *in vitro* resistance to piperacillin/tazobactam, ampicillin/sulbactam, cefazolin and doripenem (carbapenem). Further investigation identified seven additional copies of an IS26 transposable unit carrying a *bla*_{TEM-1B} beta-lactamase gene, suggesting this tandem amplification is associated with extended resistance phenotypes. Uropathogenicity factors, including three separate siderophore-encoding loci, were conserved in chromosomal and plasmid regions. Using all complete genomes, we further elucidated the recombination events surrounding the previously described K5/K1 capsular locus switch. Phenotypic confirmation of differing capsules in Australian ST1193 strains, coupled with genetic analysis revealing insertions downstream of the capsular locus, underscored the genetic distinctions between K5 and K1 capsule encoding strains. This study provides five new reference ST1193 genomes from Australia. These include the earliest complete K5-capsule ST1193 genomes on record (collected 2007), alongside our reference genome (MS10858), a clinical isolate obtained early during the ST1193 expansion and representative of the predominant K1-associated clade. These findings lay the foundations for further genomic and molecular analyses that may help understand the underlying reasons for the rapid global expansion of ST1193.

Impact Statement

This study enhances the understanding of fluoroquinolone-resistant *Escherichia coli* sequence type (ST)1193, a growing concern in urinary tract infection management. By understanding the genomic background of fluoroquinolone resistance and predicting potential recombination events contributing to antibiotic resistance acquisition, we deliver insights into the evolutionary trajectory and dissemination of antibiotic resistance in ST1193. Through genomic analyses, including the generation of complete genomes for five representative strains, we broaden the ST1193 genomic dataset and reveal key virulence and antibiotic resistance gene content. Our findings can help guide therapeutic strategies, surveillance initiatives and infection control measures, aiming to alleviate the global health and economic ramifications of antibiotic-resistant *E. coli* infections.

DATA SUMMARY

The study sequences are available in the National Center for Biotechnology Information (NCBI) under BioProject accession number PRJNA493523. The raw sequence read data generated in this study have been deposited in the NCBI sequence read archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) under accession numbers SRR7912362, SRR7912381, SRR7912383, SRR7912392, SRR7912411, and SRR12008669 to SRR12008673. The complete assemblies have been deposited in GenBank under accession numbers CP058897 to CP058900, CP058917 to CP058924, and CP059332 to CP059342. The software used to analyse raw sequence reads for polymorphism discovery and whole-genome sequencing-based phylogenetic reconstruction are available as described in the Methods. The authors confirm that all supporting data protocols have been provided in the article or supplementary data files.

INTRODUCTION

Uropathogenic *Escherichia coli* (UPEC) are significant community- and hospital-acquired pathogens [1, 2]. The emergence of multidrug resistance complicates treatment options and exacerbates the health and economic impacts of UPEC infections. Among the widely prescribed broad-spectrum antibiotics for urinary tract infections (UTIs), fluoroquinolones took precedence, accounting for 70% of UTI treatments by 2001 in some jurisdictions [3, 4]. In Australia, despite restrictions on quinolone use in humans and regulations prohibiting their use in food-producing animals [5], there has been a notable increase in *E. coli* resistance to fluoroquinolones, particularly in the Northern Territory and Western Australia [6, 7]. Globally, the extensive use of fluoroquinolones has driven the emergence of fluoroquinolone-resistant UPEC lineages across all World Health Organization (WHO) regions [8, 9].

The development of fluoroquinolone resistance in UPEC can be rapid [10–15] and may result in prolonged or recurrent infection [16]. The globally disseminated *E. coli* sequence type (ST)131 is the most well-studied UPEC clonal lineage. High rates of resistance to fluoroquinolones (and co-resistance to extended-spectrum beta-lactam antibiotics) among ST131 probably facilitated its global dissemination over the past two decades [11, 15, 17, 18].

Another emerging UPEC lineage, *E. coli* ST1193, has been associated with UTI [19], bloodstream infections [20] and neonatal invasive diseases [21, 22]. Like *E. coli* ST131, ST1193 frequently exhibits multidrug resistance [23] and high rates of fluoroquinolone resistance. Platell *et al.* [19] were among the first to report clinical fluoroquinolone-resistant ST1193 among isolates collected from cases of UTI. Between 2007 and 2009, ST1193 represented 53% of non-ST131 fluoroquinolone-resistant isolates from humans and domesticated canines in Australia [19]. Since then, fluoroquinolone-resistant ST1193 have emerged and disseminated globally [13]. ST1193 has been isolated from humans [13, 20, 22–31], companion animals [29, 30, 32, 33] and the environment [34, 35]. Unlike ST131, evidence for stepwise acquisition of single-nucleotide polymorphisms (SNPs) associated with

Access Microbiology is an open research platform. Pre-prints, peer review reports, and editorial decisions can be found with the online version of this article. Received 01 August 2024; Accepted 17 September 2024; Published 18 October 2024

Author affiliations: ¹School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia; ²Australian Centre for Ecogenomics, The University of Queensland, Brisbane, Queensland 4072, Australia; ³Health Group, Institute of Environmental Science and Research, Porirua 5022, New Zealand; ⁴UQ Centre for Clinical Research (UQCCR), Royal Brisbane & Women's Hospital Campus, The University of Queensland, Herston, Queensland 4029, Australia; ⁵School of Animal and Veterinary Sciences, The University of Adelaide, Roseworthy, South Australia 5371, Australia; ⁶INRAE, Le Laboratoire de Biotechnologie de l'Environnement, University of Montpellier, Narbonne 11100, France; ⁷School of Veterinary Science, The University of Queensland, Gatton, Queensland 4343, Australia; ⁸Biosecurity Queensland, Department of Agriculture and Fisheries, Biosecurity Sciences Laboratory, Coopers Plains, Queensland 4108, Australia; ⁹Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia; ¹⁰School of Clinical Sciences, Monash Medical Centre, Monash University, and Monash Infectious Diseases, Monash Health, Clayton, Victoria 3168, Australia; ¹¹Australian Infectious Disease Research Centre, The University of Queensland, Brisbane, Queensland 4072, Australia.

*Correspondence: Rhys T. White, rhystwhite@gmail.com; Rhys.White@esr.cri.nz; Scott A. Beatson, scott.beatson@uq.edu.au; s.beatson@uq.edu.au

Keywords: antibiotic resistance; capsule; long-read sequencing; mobile genetic elements; phylogenetic analysis.

Abbreviations: AGAR, Australian Group on Antimicrobial Resistance; ARC, Australian Research Council; AST, antibiotic susceptibility testing; BMD, broth microdilution method; CC, clonal complex; CDS, coding DNA sequence; CLSI, Clinical and Laboratory Standards Institute; EUCAST, European Committee on Antimicrobial Susceptibility Testing; ExPEC, extra-intestinal pathogenic *Escherichia coli*; GI, genomic island; GTR, general time reversible; IQR, interquartile range; IS, insertion sequence; LB, lysogeny broth; NCBI, National Center for Biotechnology Information; NCRIS, National Collaborative Research Infrastructure Strategy; NHMRC, National Health and Medical Research Council; DNT, Oxford Nanopore Technologies; QFAB, Queensland Facility for Advanced Bioinformatics; ROD, region of difference; SIR, sensitive, intermediate and resistant; SNP, single nucleotide polymorphism; SRA, sequence read archive; ST, sequence type; UPEC, uropathogenic *Escherichia coli*; UQCCR, University of Queensland Centre for Clinical Research; UTI, urinary tract infection; WGS, whole-genome sequencing; WHO, World Health Organization.

National Center for Biotechnology Information (NCBI) BioProject accession number PRJNA493523. NCBI sequence read archive (SRA) accession numbers SRR7912362, SRR7912381, SRR7912383, SRR7912392, SRR7912411 and SRR12008669 to SRR12008673. GenBank accession numbers CP058897 to CP058900, CP058917 to CP058924 and CP059332 to CP059342. More information about sequence data is provided in the data summary section.

†These authors contributed equally to this work

Eight supplementary figures and seven supplementary tables are available with the online version of this article

000894.v3 © 2024 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

fluoroquinolone resistance has not been established for ST1193. Instead, a contemporary study suggested that ST1193 acquired fluoroquinolone resistance as recently as 2005 through multiple recombination events with other antimicrobial-resistant *E. coli* clones [36].

Molecular genotyping of the O-antigen polysaccharide shows ST1193 is an O75-associated lineage belonging to clonal complex (CC)14 [23] within the *E. coli* phylogroup B2 [29, 33]. A deletion event triggering a frameshift mutation in *lacY* (encoding lactose permease) has induced a lactose non-fermentation phenotype in ST1193 [33] and other CC14 strains [13]. This particular phenotype might have contributed to the under-reporting of ST1193 before 2010, possibly due to selective and differential use of clinical diagnostic tests relying on MacConkey agar [37].

The genetic landscape of ST1193 comprises two major clades: ‘K5-associated’ or ‘K1-associated’ ST1193 [13], which is related to differences in the capsular antigen. This capsule-switching phenomenon, from a K5 to a K1 antigen, probably underpins the rapid global expansion of ST1193 [13]. However, at the time of writing, only three complete ST1193 genomes (all ‘K1-associated’) were published in the literature, with two strains from humans [21, 38], and one from river water [39]. To expand the genomic dataset and unravel the genomic content of ST1193, the genomes of three human-derived ST1193 strains with a K1 capsule and two strains with a K5 capsule underwent whole-genome sequencing (WGS) using both long-read and short-read sequencing. We generated complete genomes for all five ST1193 strains, including manual curation of mobile genetic elements, and defined the full complement and genomic context of virulence and antibiotic resistance genes. The addition of five complete ST1193 genomes across both capsule types will facilitate future comparative genomics studies of this important UPEC clone.

METHODS

Bacterial strains

The five *E. coli* ST1193 strains analysed in this study were obtained from three different sources. One isolate (MS10711, Alternative ID: GNB 2889) originated from the Australian Group on Antimicrobial Resistance (AGAR), a collaboration between clinicians and microbiologists (www.agargroup.org, accessed 4 April 2024). The remaining four isolates were sourced from previously published studies [19, 28]. Detailed information on the five ST1193 strains sequenced in this study are listed in Table S1 (available in the online version of this article).

Phenotypic assays

To confirm the capsule type phenotypically, the cross-brush method was performed as previously described [40] using the K1 and K5 lytic phage sourced from the Statens Serum Institut (Denmark). After crossing the phage suspension line, growth inhibition was considered a positive reaction. For antibiotic susceptibility testing (AST), isolates were sent to The University of Queensland Centre for Clinical Research (UQCCR) for AST using a broth microdilution method (BMD) with *E. coli* strain ATCC 25922 as quality control. Cultures were inoculated from -80°C storage onto Columbia Horse Blood Agar and incubated at 37°C overnight. A 0.5 McFarland suspension was prepared in sterile water (ISO20776.1-2017). Eleven microlitres of suspension was added to 11 ml cation-adjusted Mueller Hinton Broth with TES buffer, and $50\ \mu\text{l}$ of broth suspension was added to each well of a GN6F Sensititre Plate (Thermo Fisher). Plates were sealed and incubated at 35°C for 19 h before determining the minimum inhibitory concentration (MIC) using the Thermo Scientific Sensititre Vizion system instrument with SWIN software v3.4. MICs and sensitive, intermediate and resistant (SIR) interpretations were reported according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) and Clinical and Laboratory Standards Institute (CLSI) breakpoints (User Manual SWIN Software System, v3.4).

Genomic DNA extraction, library preparation and sequencing

Bacteria were grown in lysogeny broth (LB), and genomic DNA was extracted using the MoBio UltraClean Microbial DNA isolation kit (as per the manufacturer’s instructions). Genome sequencing was performed using Illumina short-read and nanopore long-read sequencing as previously described [41–43]. We performed Illumina sequencing using the NextSeq 500 platform at the Australian Centre for Ecogenomics (University of Queensland, Brisbane, Australia). DNA libraries were prepared with the Nextera XT Library Preparation Kit and indexed with the Nextera XT Index Kit (both from Illumina). Sequencing generated 150 bp paired-end reads. The Oxford Nanopore Technologies (ONT) sequencing libraries were prepared using 400–600 ng of genomic DNA, according to the rapid barcoding sequencing kit (catalogue number SQK-RBK004, 12 barcodes mixed in equal volume) as per the manufacturer’s instructions. The entire library was loaded onto an R9.4 flow cell and run on a MinION device for approximately 48 h (using MinKNOW v1.10.16). The reads generated from the MinION sequencing run were basecalled and binned using the ONT Guppy basecalling suite v3.1.5 – Guppy basecaller and Guppy barcoder, respectively. Further details

of methods including nanopore read quality control, Illumina read quality control and genome assembly are available in the Supplementary Materials.

Genome annotations

The assemblies were annotated using Prokka v1.14.0 [44], with the complete genome of *E. coli* ST131 strain EC958 (GenBank: HG941718) as a reference for trusted proteins. Prophage regions were identified using PHASTER [45, 46]. Mobile genetic elements were identified using IslandViewer 4 [47] and ISSaga v2.0 [48] (ISfinder platform [49]), followed by manual curation using Artemis v18.1.0 [50]. Methods for *in silico* multilocus sequence typing, genotyping the O, H and K-antigens (lipopolysaccharide, flagellar and capsule, respectively), along with virulence and antibiotic resistance gene genotyping (including chromosomal point mutations) have been described previously [15].

Investigating regions of recombination

The complete chromosome sequences from the five ST1193 genomes sequenced in this study underwent a reference-free whole-chromosome alignment with three publicly available, complete ST1193 chromosomes using progressiveMauve [51] in Mauve (version snapshot_2015-02-13) [52]. SNPs were extracted from the whole-chromosome alignment using HarvestTools v1.2 [53]. To identify regions of recombination, the whole-chromosome alignment was input into Gubbins v2.4.1 [54] [default settings, 'raxml mode' with the General Time Reversible (GTR) GAMMA correction]. The predicted regions of recombination were visualized using the ggplot2 v3.3.5 library [55] in R v4.1.0 [56].

Compiling a high-quality ST1193 global dataset and identifying genetic variants

We sequenced five Australian ST1193 genomes to compare them with other ST1193 genomes sampled worldwide. To identify additional ST1193 genomes, we first screened the Enterobase database v1.1.2 (<https://enterobase.warwick.ac.uk/>, accessed 21 October 2021) for strains belonging to ST1193. This screening identified 616 ST1193 genomes. We then downloaded the corresponding epidemiological data (e.g. collection date, source and country of origin) from Enterobase. Subsequently, we retrieved the sequence read data for these 616 ST1193 genomes from the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) using the 'prefetch' and 'fastq-dump' tools within the SRA Toolkit v2.9.0-mac64 (<https://github.com/ncbi/sra-tools>, accessed 21 October 2021).

We integrated multiple complete chromosomes by simulating error-free reads using ART (version ART-MountRainier-2016-06-05) [57] to 60× coverage with an insert size of 340±40 bp. These included the three publicly available and published chromosomes [09-02E (GenBank: AP022650), AVS0096 (GenBank: CP076344) and MCJCHV-1 (GenBank: CP030111)]. A read-mapping approach for identifying genetic variants was performed using the SPANDx v3.2 pipeline, as previously described [58]. In brief, trimmed Illumina reads were mapped to the complete chromosome MS10858, originally isolated from a female with a urine infection in Australia in October 2007 [19]. Our final dataset consisted of 619 genomes representing previously published datasets (including three complete genomes) plus our Australian collection ($n=5$).

High-resolution phylogeny of ST1193

The quality-trimmed paired-end Illumina reads from the 623 ST1193 genomes were mapped onto the chromosome of MS10858 as described above. SNPs within regions of high-density clusters (≥ 3 SNPs found within a 10 bp window) and predicted recombination sites (identified using Gubbins) were removed from the core-genome alignment. Core-genome SNP alignments were independently run through jModelTest v2.1.10 [59, 60], to identify the best-fit evolutionary model using 12 candidate models from three substitution schemes with base tree for likelihood calculations optimized for maximum likelihood. jModelTest included models with equal/unequal base frequencies (+F), with/without a proportion of invariable sites (+I) and with/without rate variation among sites (+G) (four rate categories). Maximum likelihood phylogenetic trees were generated from the orthologous core-genome SNP alignments using RAxML v8.2.10 [61] (GTR correction) through optimization of ten distinct, randomized maximum parsimony trees, before adding 1000 bootstrap replicates. The resulting phylogenetic trees were visualized using FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>, accessed 8 April 2024).

RESULTS

Genomic descriptions of complete ST1193 representative genomes

The complete genomes of five Australian ST1193 and three published completed ST1193 genomes are described here. *In silico* genotyping confirmed all eight strains as ST1193, with an O75:H5 serotype and *fim*H64 type 1 fimbrial adhesin, consistent with previous phenotypic characterizations of ST1193 [23, 33]. Additionally, general features such as the number of predicted coding DNA sequences (CDS) and GC content were consistent across each strain and are listed in Table 1. The ST1193 strain MS10858 was selected as the reference genome for comparative analyses, as it represents the earliest reported clinical UTI isolate to date, collected in October 2007 (as of 21 October 2021). To characterize chromosomal features and identify

Table 1. Genomic descriptions of complete *Escherichia coli* sequence type (ST)1193 representative genomes

<i>Escherichia coli</i> ST1193 strain									
Characteristics	MS10711	MS10858	MS10860	MS8320	MS8324	MCJCHV-1*	09-02E*	AVS0096*	
O- and H-genotype	O75:H5	O75:H5	O75:H5	O75:H5	O75:H5	O75:H5	O75:H5	O75:H5	
Type 1 fimbriae	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	<i>fimH64</i>	
Capsule	K5	K1	K5	K1	K1	K1	K1	K1	
Collection year	2010	2007	2007	2012	2011	2015	2018	2020	
Location	Australia	Australia	Australia	Australia	Australia	USA	Vietnam	Switzerland	
Host	Human	Human	Human	Human	Human	Human	Human	River water	
Clinical presentation	Urinary tract infection	Urinary tract infection	Urinary tract infection	Urinary tract infection	Urinary tract infection	Neonatal meningitis	Healthy	-	
Anatomical site	Urine	Urine	Urine	Urine	Urine	Cerebrospinal fluid	Faeces	-	
Alternative ID	GNB 2889	QU054	QU059	CS12-NT018	CS12-WA103	NMEC-O75	-	-	
BioSample	SAMN10141433	SAMN10141460	SAMN10141462	SAMN10141488	SAMN10141490	SAMN09462202	SAMD000198301	SAMN19493560	
GenBank accession	CP058922 to CP058924	CP058897 to CP058900	CP059332 to CP059338	CP059339 to CP059342	CP058917 to CP058921	CP030111 to CP030116	AP022650 to AP022659	CP076344 to CP076347	
SRA accession	SRR12008673	SRR12008672	SRR12008671	SRR12008670	SRR12008669	-	-	SRS9116829	
Genome coverage	184x	276x	177x	97x	196x	320x	143x	125x	
Chromosome length (bp)	4976233	4955580	5023252	5082313	4939743	4939457	5075911	4944762	
Predicted CDS	4642	4604	4709	4758	4605	4848	4831	4645	
G+C content (%)	50.59	50.63	50.63	50.51	50.61	50.60	50.62	50.62	
No. of tRNA genes	92	91	90	92	91	90	94	85	
No. of rRNA operons	21	22	21	22	22	22	22	22	
No. of plasmids	2	3	6	3	4	5	9	3	
F-type plasmid	F-:A1:B20	F-:A1:B10	F-:A1:B20	F-:A1:B10	F-:A1:B10 FIIK:A--B-	F-:A1:B10	F-:A1:B10	F-:A1:B10	
IncX1	Not detected	Not detected	Not detected	Not detected	Not detected	Not detected	Present	Present	
Colicimogenic plasmid	Col(BS512)	Col(BS512)	ColpVC Col(BS512)	Col(BS512)	Col(BS512)	ColpVC Col(BS512)	Col(MP18) Col(BS512)	Col(BS512)	
Untypeable cryptic plasmids	0	1	3	1	1	2	5	0	

*Not sequenced as part of this study.

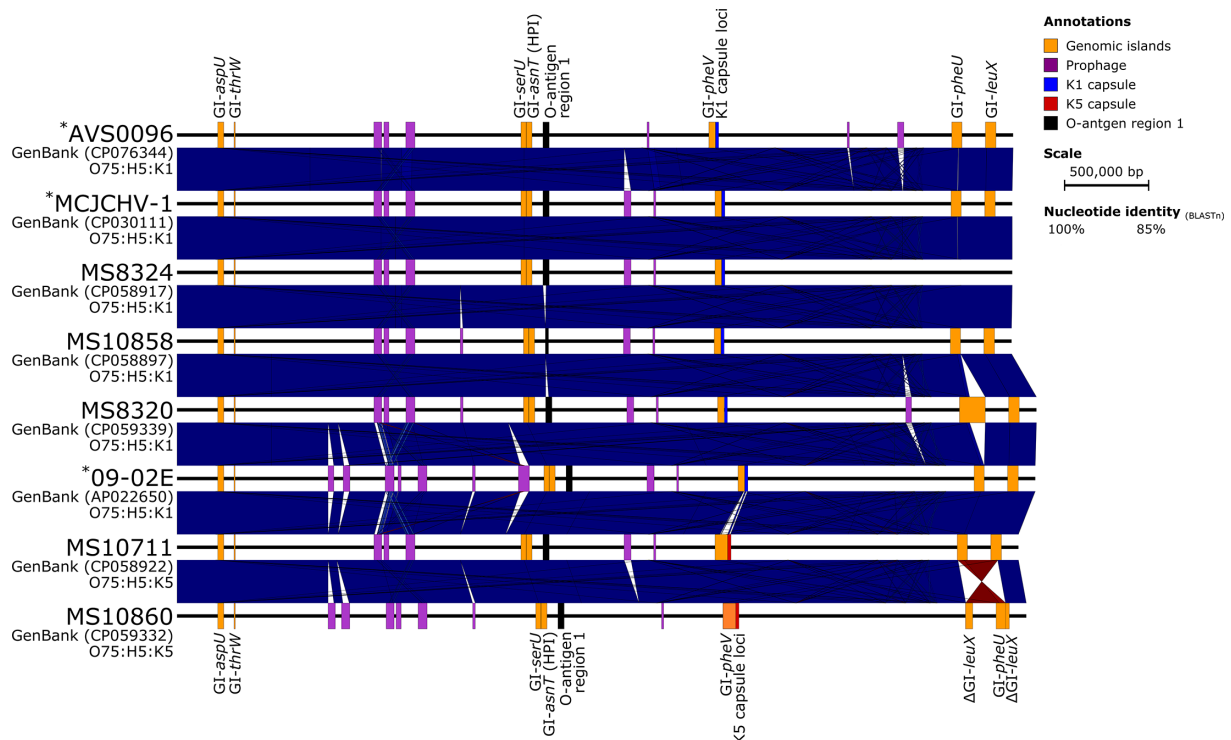


Fig. 1. Whole-genome nucleotide pairwise comparisons between *Escherichia coli* sequence type (ST)1193 strains. Linear nucleotide alignment of the whole genome, highlighting chromosomal mobile genetic elements shared between genomes. Shared prophages and genomic islands (GIs) are shown in purple and orange, respectively. Blue and red (inversions) shading indicates nucleotide identity between sequences according to BLASTn (85–100%). Strains not sequenced in this study are indicated (*). Image created using Easyfig [80].

differences, the chromosome of MS10858 underwent a BLASTn comparison against the remaining seven ST1193 complete genomes (Fig. 1), and nine other complete *E. coli* genomes chosen to represent different phylogroups and pathotypes (Fig. S1). A total of 33 regions of difference (RODs), between 1.9 and 62.6 kb in length, were defined in the chromosome of MS10858 (Fig. S1; Tables S2 and S3). These include seven genomic islands (GIs), four intact prophages and two partial prophages. The 20 remaining RODs primarily comprised genes associated with UPEC virulence. They included loci encoding flagella, O-antigen, the capsule, several putative fimbrial-like adhesins and a 1947 bp GI-*argU* remnant. GIs were conserved across all eight ST1193 genomes (Fig. 1). However, some large insertions and translocations were present, including a 34598 bp insertion in GI-*pheV* in MS10711. In MS10860, recombination between GI-*pheU* and GI-*leuX* resulted in a large (>100 kb) inversion of the intervening chromosome. Assembly error was ruled out by examining the read mapping in this region (Fig. S2). There was also a 92421 bp insertion into GI-*pheU* in MS8320. This region starts after the frameshifted integrase gene (locus tag: MS8320_4488) and carried numerous hypothetical proteins ($n=63$) and a putative fimbriae locus (locus tags: MS8320_4505 and MS8320_4506). In addition, BLASTn showed that this ROD was present in multiple other species, including *Proteus mirabilis*, *Morganella morganii* and *Providencia stuartii* (Table S4).

All eight ST1193 strains display multidrug resistance

The five Australian ST1193 strains were subjected to AST and exhibited resistance (based on CLSI clinical breakpoints) to ampicillin (beta-lactam) and ciprofloxacin (fluoroquinolone). They also showed variable resistance to antibiotics in the aminoglycoside, folate inhibitor, beta-lactam/beta-lactamase inhibitor and tetracycline classes (Table 2). Using long-read sequencing we were able to fully characterize the repertoire and genomic context of all antibiotic resistance genes in all eight ST1193 genomes, and correlate resistance genotype with observed phenotypic susceptibilities (Table 2). Antibiotic resistance genes were exclusively carried on F-type plasmids among our isolates (Fig. S3). However, no antibiotic resistance gene was conserved entirely across all eight strains, resulting in some variation in both plasmid structure and antibiotic resistance gene content (Figs S4 and S5). All eight strains also carried known chromosomal point mutations in genes encoding DNA gyrase (*gyrA* S83L and D87N), topoisomerase IV subunit A (*parC* S80I) and topoisomerase IV subunit B (*parE* L416F), conferring resistance to fluoroquinolones (MICs >2 mg l⁻¹). Additional details and findings regarding the genomic context of these antibiotic resistance genes and plasmids are summarized in the Supplementary Materials.

Table 2. Resistant MIC breakpoints ($\mu\text{g ml}^{-1}$) and associated acquired antibiotic resistance genes and/or chromosomal mutations

Escherichia coli sequence type (ST1193) strain							
Characteristic	MS10711	MS10860	MS10858	MS8320	MS8324	MCJCHV-1*	AVS0096*†
Capsule (chromosome)	K5	K5	K1	K1	K1	K1	K1
Plasmid	pMS10711A	pMS10860A	pMS10858A	pMS8320A	pMS8324A	pNMEC-O75A	pAVS0096-a
GenBank accession	CP058923	CP059333	CP058898	CP059340	CP058918	CP030112	CP076345
Plasmid length (bp)	98451	104838	93277	181199	89795	88,421	101018
Plasmid multi-locus sequence type	F-:A1:B20	F-:A1:B20	F-:A1:B10	F-:A1:B10	F-:A1:B10	F-:A1:B10	F-:A1:B10
Beta-lactams							
Amoxicillin/clavulanic acid	ND	ND	ND	ND	ND	ND	21 mm (S)
Ampicillin	>16 (R)	>16 (R)	>16 (R)	>16 (R)	> 6 (R)	>16 (R)	6 mm (R)
Piperacillin/tazobactam	≤8 (S)	≤8 (S)	≤8 (S)	=128 (R)	≤8 (S)	ND	ND
Ampicillin/sulbactam	=16 (I)	>16 (R)	=16 (I)	>16 (R)	=16 (I)	ND	ND
Cefazolin	FAIL	=2 (S)	=4 (I)	>16 (R)	=8 (R)	=4 (I)	6 mm (R)
Cefepime	≤2 (S)	≤2 (S)	=4 (SDD)	=8 (SDD)	≤2 (S)	≤4 (SDD)	21 mm (I)
Cefotaxime	ND	ND	ND	ND	ND	ND	13 mm (R)
Ceftazidime	≤1 (S)	=2 (S)	≤1 (S)	=2 (S)	≤1 (S)	≤1 (S)	ND
Ceftazidime/avibactam	≤2 (No Intp)	≤2 (No Intp)	≤2 (No Intp)	≤2 (No Intp)	≤2 (No Intp)	ND	ND
Ceftriaxone	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	ND
Ceftolozane/tazobactam	≤1 (S)	=2 (S)	≤1 (S)	=2 (S)	≤1 (S)	ND	ND
Aztreonam	≤1 (S)	≤1 (S)	=4 (S)	≤1 (S)	≤1 (S)	ND	ND
Doripenem	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	=4(R)	=1(S)	ND	ND
Ertapenem	≤0.25 (S)	≤0.25 (S)	≤0.25 (S)	≤0.25 (S)	≤0.25 (S)	ND	ND
Imipenem	=2 (I)	≤1 (S)	≤1 (S)	≤1 (S)	≤1 (S)	≤0.5 (S)	ND
Meropenem	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	≤0.5 (S)	=2 (I)	ND	ND
Antibiotic resistance genes	<i>bla</i> _{TEM-1B}	<i>bla</i> _{TEM-1B}	<i>bla</i> _{TEM-1B}	<i>bla</i> _{TEM-1B}	<i>bla</i> _{TEM-1B}	<i>bla</i> _{TEM-1B}	<i>bla</i> _{CTX-M-27}
Fluoroquinolones							
Ciprofloxacin	>2 (R)	>2 (R)	>2 (R)	>2 (R)	>2 (R)	>2 (R)	8 mm (R)
Levofloxacin	=8 (R)	=8 (R)	>8 (R)	=8 (R)	=8 (R)	ND	ND
Chromosomal point mutations	<i>gyrA</i> (S83L and D87N), <i>parC</i> (S80I) & <i>parE</i> (L416F)						
Aminoglycosides							
Tobramycin	>8 (R)	=8 (I)	=8 (I)	≤2 (S)	≤2 (S)	=8 (I)	ND
Amikacin	≤8 (S)	≤8 (S)	=16 (S)	≤8 (S)	≤8 (S)	≤8 (S)	ND
Gentamicin	>8 (R)	>8 (R)	>8 (R)	≤2 (S)	≤2 (S)	>8 (R)	20 mm (S)
Antibiotic resistance genes	<i>aac(3)-IId</i>	<i>aac(3)-IId</i>	<i>aac(3)-IId</i>	Undetected	Undetected	<i>aac(3)-IId</i>	Undetected
	<i>aadA5</i>	<i>aadA5</i>	<i>aadA5</i>	Undetected	<i>aadA5</i>	Undetected	<i>aadA5</i>

Continued

Table 2. Continued

Escherichia coli sequence type (ST1193) strain							
Characteristic	MS10711	MS10860	MS10858	MS8320	MS8324	MCJCHV-1*	AVS0096*†
Macrolides							
Azithromycin	ND	ND	ND	ND	ND	ND	9 mm (R)
Antibiotic resistance genes	<i>mphA</i>	<i>mphA</i>	<i>mphA</i>	<i>mphA</i>	<i>mphA</i>	Undetected	<i>mphA</i>
Tetracyclines							
Tetracycline	≤4 (S)	>8 (R)	>8 (R)	≤4 (S)	≤4 (S)	≤4 (S)	ND
Tigecycline	≤1 (No Intp)	≤1 (No Intp)	≤1 (No Intp)	≤1 (No Intp)	=4 (No Intp)	≤1 (No Intp)	7 mm (R)
Antibiotic resistance genes	Undetected	<i>tetA</i> & <i>tetR</i>	<i>tetA</i> & <i>tetR</i>	Undetected	Undetected	Undetected	<i>tetA</i> & <i>tetR</i>
Trimethoprim							
Trimethoprim/sulfamethoxazole	>4 (R)	>4 (R)	>4 (R)	≤2 (S)	>4 (R)	≤2 (S)	6 mm (R)
Antibiotic resistance genes	<i>sul1</i> & <i>sul2</i>	<i>sul1</i> & <i>sul2</i>	<i>sul1</i> & <i>sul2</i>	Undetected	<i>sul1</i>	Undetected	<i>sul1</i> & <i>sul2</i>
	<i>drfA17</i>	<i>drfA17</i>	<i>drfA17</i>	Undetected	<i>drfA17</i>	Undetected	<i>drfA17</i>
Other							
Streptomycin	ND	ND	ND	ND	ND	ND	6 mm (R)
Antibiotic resistance genes	Undetected	<i>strA</i> & <i>strB</i>	<i>strA</i> & <i>strB</i>	Undetected	Undetected	Undetected	<i>strA</i> & <i>strB</i>
Chloramphenicol	ND	ND	ND	ND	ND	ND	23 mm (S)
Nitrofurantoin	≤32 (No Intp)	≤32 (No Intp)	≤32 (No Intp)	≤32 (No Intp)	≥64 (No Intp)	ND	20 mm (S)

*Complete ST1193 genomes not sequenced in this study. Antibiotic resistance genes, Resfinder (updated 19 April 2020); (R) resistant; (I) intermediate; (SDD) susceptible dose dependent; (S) susceptible; (No Intp) no interpretation; (ND) not determined.

†Disc zone interpretive criteria.

Key virulence-associated regions of difference are adjacent to major regions of recombination distribution

The capsular locus was the most notable chromosomal difference between the eight ST1193 strains, aside from mobile genetic elements (Fig. 2). Specifically, changes in the polysialic acid biosynthesis (*kps*) cluster have probably resulted from a recombination-mediated switch from a K5 (MS10711 and MS10860) to a K1 (MS10858, MS8320 and MS8324) capsule [13]. To confirm this, the eight complete ST1193 chromosomes underwent a reference-free global alignment (5399477 bp), which was then used to predict regions of recombination using Gubbins. When using MS10858 as a reference, this identified several major regions of recombination with high SNP densities in and surrounding key regions associated with virulence, including GIs, prophages and the capsular locus (Fig. S6). Four regions of recombination between 9151 and 19547 bp were found to encompass the *GI-pheV* and capsular locus, confirming that the capsular switch is probably the result of recombination.

Furthermore, iron acquisition is a critical feature of ST1193, evidenced by the presence of siderophores, along with various virulence genes associated with extra-intestinal pathogenic *E. coli* (ExPEC) virulence, as well as haem receptor and ferric citrate transporter systems found in Australian ST1193 chromosomes. All eight ST1193 genomes analysed in this study carried essential siderophores, such as enterobactin, yersiniabactin and aerobactin. These siderophores, along with the *iucABCD/iutA*, *sat* and *iha* loci, shared high nucleotide identity with virulence plasmids in other *E. coli* strains, further underlining their role in ExPEC virulence. Additionally, the Australian ST1193 chromosomes exhibited haem receptors encoded by the *chu* operon and the ferric citrate transporter system encoded by *fecIR* and *fecABCDE*, along with conserved principal iron/manganese transporters encoded by the *sitABCD* locus. Further details of the iron acquisition mechanisms can be found in the Supplementary Materials. Additional virulence genes, including those associated with iron acquisition, are summarized in Table S5.

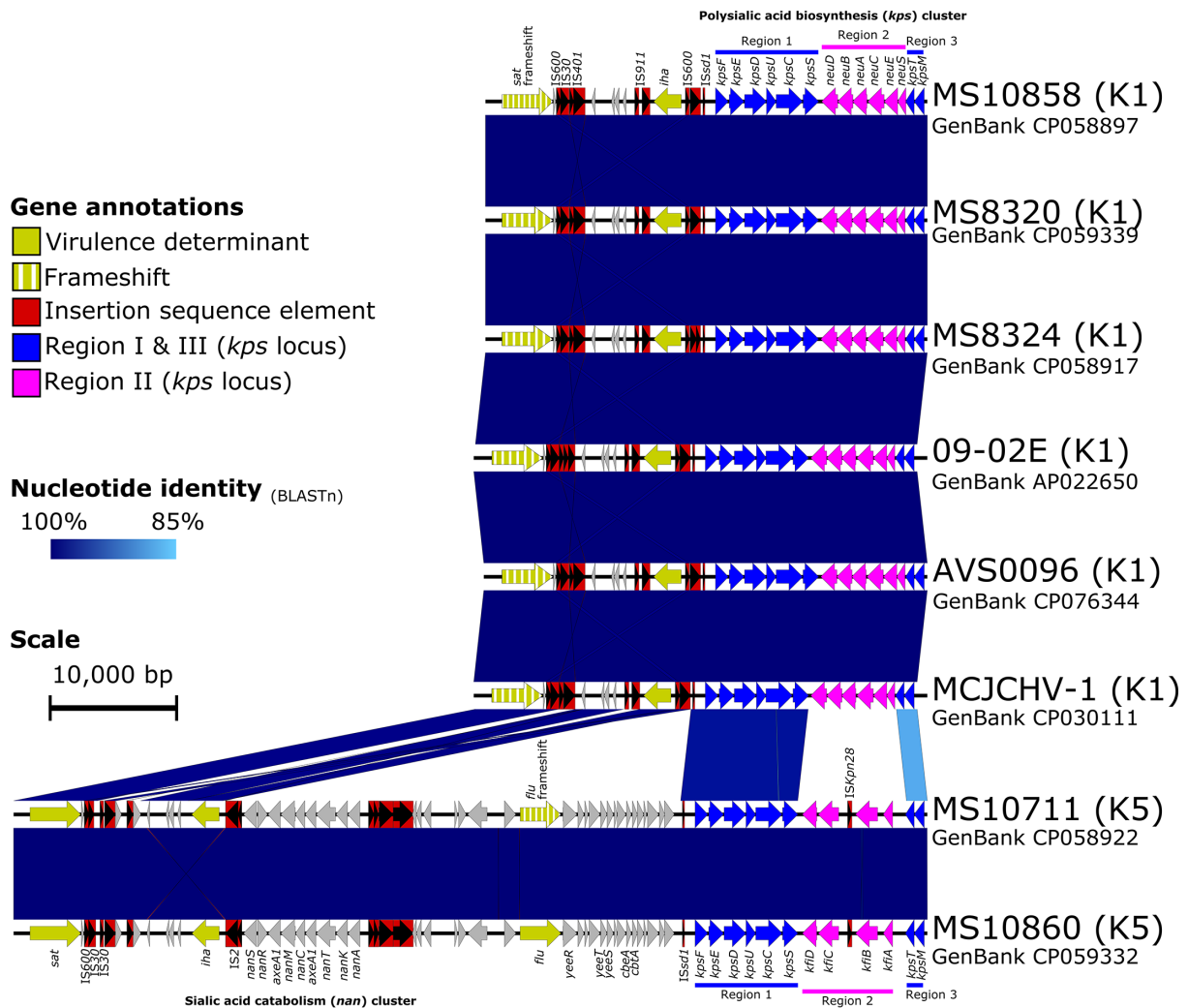


Fig. 2. Linear nucleotide comparisons of the group 2 capsules in *Escherichia coli* sequence type (ST)1193 strains. Blue shading indicates nucleotide identity between sequences according to BLASTn. Key genomic regions are indicated: insertion sequences (IS): red; conserved capsular regions I and III: blue; differing capsular regions (region II): pink; virulence determinants: yellow; other CDS: grey. Image created using Easyfig [80].

We confirmed phenotypically that the Australian ST1193 strains have differing capsules (Fig. S7). K5 capsule encoding strains had an ~34.6 kb insertion downstream of *GI-pheV* that was absent in the K1 capsule encoding strains. This ~34.6 kb region carried the *flu* gene (encodes antigen 43) and a *nan*-operon (sialic acid catabolism), and is downstream of the *iha* adhesin (locus tags MS10711_3080 and MS10860_3153). Notably, the *flu* gene was intact in MS10860 (locus tag MS10860_3176) but was disrupted by frameshift mutations (possible pseudogene) in MS10711 (locus tags MS10711_3103 to MS10711_3106).

Phylogenetic relationship of the complete ST1193 genomes

We revisited the population structure of ST1193 by incorporating the five Australian ST1193 genomes with a global ST1193 collection ($n=619$). By using the chromosome of MS10858 as a reference, the core-genome alignment between the 624 genomes comprised approximately 2461300 bp (based on regions estimated to the nearest 100 bp with $\geq 95\%$ coverage across all genomes), which represents 49.9% of the MS10858 chromosome. A total of 14327 orthologous, biallelic core-genome SNPs called against the reference chromosome MS10858 were identified. After removing 4640 SNPs within identified regions of recombination, an alignment of 9687 non-recombinant, orthologous, biallelic core-genome SNPs remained and was used to construct a maximum likelihood phylogenetic tree (Fig. 3).

This analysis of 624 ST1193 strains showed the lineage tightly clustered, with a median pairwise SNP distance of 55 SNPs [interquartile range (IQR): 44–69]. Notably, the maximum pairwise SNP distance was 314 SNPs between Australian strains

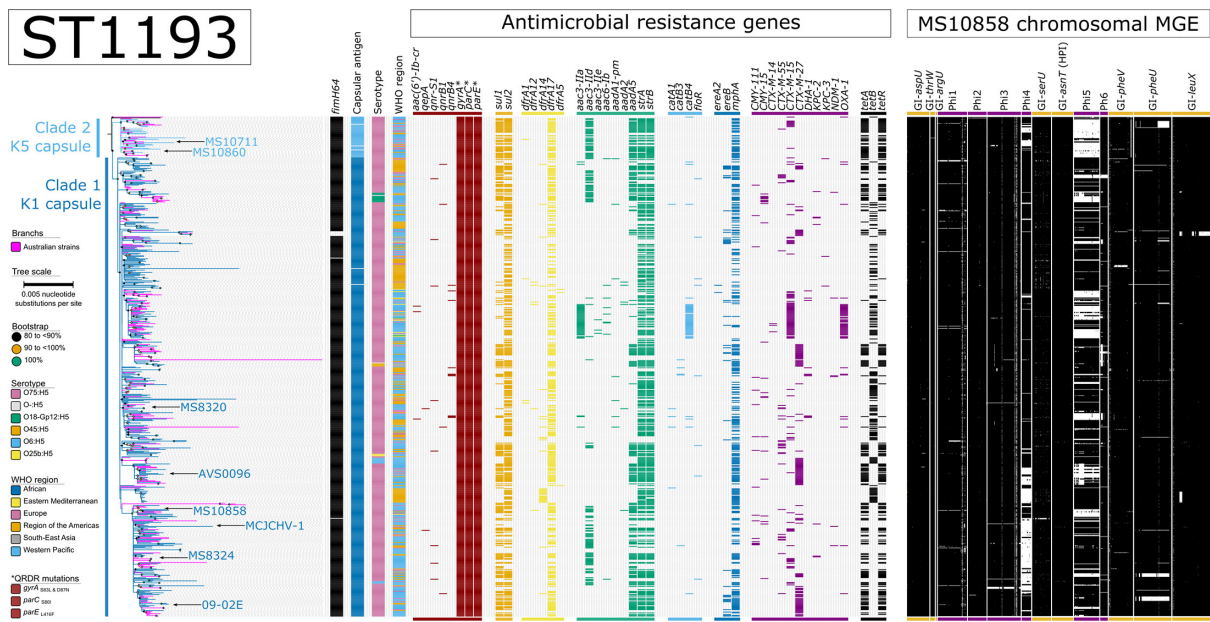


Fig. 3. Maximum likelihood phylogeny of *Escherichia coli* sequence type (ST)1193 isolates, alongside the antimicrobial resistance genotype and mobile genetic element (MGE) complement. The phylogeny was inferred from 9687 non-recombinant orthologous biallelic core-genome SNPs from 624 strains. SNPs were derived from a core-genome alignment of 2461300bp and are called against the reference chromosome MS10858 (GenBank: CP058897). The phylogenetic tree is rooted according to the strain HICF686 (SRA: SRS1221162) outgroup, which has been omitted for visual purposes. Bootstrapping using 1000 replicates demonstrates the robustness of the branches. The presence/absence analysis of loci is based on the uniform coverage at each 100bp window size in SPANDx. Coverage is shown as a heat map where $\geq 50\%$ identity is highlighted in yellow and $\geq 80\%$ identity is highlighted in black. White plots indicate regions with $< 50\%$ coverage and may contain regions that are absent.

AUSMDU00035640 (collected in 2019; SRA: SRS9466797) and AUSMDU00018204 (collected from human sputum in 2019; SRA: SRS8335820).

The ST1193 phylogeny resolved into two well-supported and diverse clades which have been named Clade 1 and Clade 2 (Fig. 3). Clade 1 was predominant ($n=573/624$ strains, 91.8%) amongst the global ST1193 population, whereas Clade 2 was less common (49/624 strains, 7.9%). The 573 strains in Clade 1 were collected across 30 countries spanning six WHO regions. Over 75% of the strains in Clade 1 were collected from Australia ($n=200$, 34.9%), the USA ($n=127$, 22.2%), Canada ($n=37$, 6.5%), Singapore ($n=36$, 6.3%) and China ($n=31$, 5.4%). The 49 strains within Clade 2 were collected across ten countries spanning four WHO regions. More than half of the strains within Clade 2 were collected in Australia ($n=29/49$, 59.2%). There was no discernible difference in host specificity between Clades 1 and 2. Clade 1 primarily comprised strains collected from humans ($n=533/573$, 93.0%) and domesticated dogs ($n=18/573$, 3.1%), with fewer strains from the environment ($n=10/573$ 1.7%), silver gulls ($n=7,573$ 10.2%) and food ($n=4/573$, 0.7%). Similarly, Clade 2 consisted primarily of strains from humans ($n=42/49$, 89.4%) and domesticated dogs ($n=7/49$, 14.9%). Of the 533 Clade 1 strains collected from humans, 314 have descriptive clinical sample metadata; 145 (46.2%) were from urine, 88 from blood (28.0%), 62 (19.7%) from faecal samples and 8 (2.5%) from sputum. Of the 42 Clade 2 strains collected from humans, 26 (53.1%) were collected from urine, three (6.1%) from faecal samples and a single strain from blood. The other 12 genomes were from the public domain and are missing sample type metadata.

Relative to MS10858, only five and eight SNPs defined the terminal branches to Clade 1 and Clade 2, respectively. Global strains within Clade 1 were separated by a maximum pairwise SNP distance of 314 SNPs (median: 54, IQR: 43–66). Meanwhile, global strains within Clade 2 were separated by a maximum pairwise SNP distance of 113 SNPs (median: 39, IQR: 27–50).

DISCUSSION

The fluoroquinolone-resistant ST1193 lineage has emerged worldwide as a significant extraintestinal pathogen [13]. The five Australian ST1193 strains sequenced in this study using long-read nanopore sequencing provide a broader representation of the known ST1193 major clades [13] and include the earliest known reported cases of clinical ST1193 involved in UTIs [19], thereby expanding on previous descriptions from incomplete draft assemblies [35]. Long-read sequencing highlighted (i) the low

pairwise genomic diversity of strains from both Australia and globally, and (ii) the high conservation of the core and accessory genome (both chromosome and plasmid). Additionally, this study confirmed that the major differences in the capsular loci (K1 and K5) resulted from recombination, similar to previous reports [27, 36]. These genomes also contained a similar F-type plasmid backbone, with six of the eight strains carrying a similar multidrug resistance region defined by either short or long variants of this resistance region (Fig. S5).

The inclusion of the reference genome for strain MS10858 provides crucial insight, representing one of the initial clinical isolates sequenced within the prevalent 'K1-associated' clade of ST1193 [19]. Contrasting the remaining seven complete ST1193 genomes with MS10858, we observed a remarkable level of genomic preservation, characterized by minimal pairwise SNP distances, except for localized regions of intense SNP clustering attributed to potential recombination events (Fig. S8). Analysis of the accessory genome also indicated a close evolutionary relationship, with four prophages and all GIs largely conserved across all strains (albeit with some large indels and translocations). Notably, while the prophages in MS10858 were specific primarily to our ST1193 collection, most GIs that encode virulence determinants have been identified in other ExPEC complete genomes. For example, the content and arrangement of modules in the MS10858 GI-*pheV* and GI-*leuX* were homologous with the ST131 reference strain EC958, suggesting that these regions are important to ST1193 fitness, as shown in ST131 [62]. A recent study has demonstrated that ST1193 emerged via simultaneous homologous recombinations in 11 gene loci from other *E. coli* [36]. This may be responsible for the high number of putative GIs in the chromosome of MS10858 seen in our results.

The collection of eight complete genomes now represents fully resolved capsular loci from both K5- and K1-associated ST1193. Our analysis showed that four regions of recombination between 9151 and 19547 bp encompassed the GI-*pheV* and capsular locus, confirming that the capsular switch is probably the result of recombination. However, whether this was a single recombination event or multiple events remains to be elucidated. K1 and K5 capsule loci share conserved regions (regions I and III) that encode the transmembrane complex involved in the export and assembly of *E. coli* Group 2 capsular polysaccharides [40, 63, 64]. However, region II is serotype-specific and encodes enzymes responsible for synthesizing the capsular polysaccharide. The K5 antigen is a heteropolymer of glucuronic acid and *N*-acetyl glucosamine [65], whereas the K1 antigen is a homopolymer of α -2-8 linked *N*-acetylneuraminic acid (NeuNAc) [66, 67], which is immunochemically identical to the group B *Neisseria meningitidis* capsular polysaccharide [68]. The K5 capsule, like many other K antigens, only provides protection to serum killing [69], whereas the K1 capsule also confers resistance to phagocytic killing [70]. Together, these observations indicate that *E. coli* presenting a K1 antigen can evade host immunogenic responses and may reach a threshold level of bacteraemia necessary for meningeal invasion [71], as was seen in the case of lethal neonatal meningoencephalitis, associated with O75:H5:K1-ST1193 strain MCJCHV-1 [21].

Outside of the ST1193 lineage, exchange of the *kps* locus has also been reported in the globally predominant *E. coli* ST131 lineage [72]. The switching of region II in the *kps* locus is described as the predominant mechanism of K-antigen-type sharing between genome pairs belonging to different clonal lineages [73]. Our findings concerning the role of recombination imply that the capsule locus in ST1193 is under selective pressure, where the acquisition of a capsular locus with a different *kps* region II leading to the switch between K5 and K1 antigen (thus increasing in capacity to evade host immunogenic responses) may have primed the ST1193 lineage for its recent population expansion [13]. A similar capsule locus switching was associated with the emergence of vaccine escape variants and subsequent population expansions in *Streptococcus pneumoniae* [74]. However, no evidence of a 'straightforward relationship' between the *kps* locus switching and different lineages was detected when investigating population frequency and locus gene content [73].

The availability of eight complete ST1193 genomes allowed us to extend previous ST1193 studies that targeted either a specific repertoire of genes, or were unable to fully resolve the surrounding genomic context of key virulence and resistance genes [26, 33]. For example, we showed that the yersiniabactin and aerobactin siderophore loci were carried on mobile genetic elements in ST1193. ST1193 also contained the enterobactin siderophore, like that present in the ST131 strain EC958. The conservation of this siderophore in other UPEC lineages indicates its importance in the fitness of ST1193. This has been exemplified in the work by Pi *et al.* [75], demonstrating the essential function of enterobactin secretion and transportation in colonizing the healthy murine gastrointestinal tract by *E. coli*. When compared to the importance of enterobactin secretion and transportation within the host [75], an increased siderophore repertoire (such as the yersiniabactin and aerobactin loci in ST1193) carried on mobile genetic elements may have been more important in the evolution and dissemination of ST1193, as these strains are more likely to outcompete other bacteria in sequestering iron in the densely populated gastrointestinal tract [76, 77]. Notably, an earlier study that screened ST1193 isolates from a Chinese hospital between 2014 and 2015 reported the genes *iutA* and *fyuA* (carried on the aerobactin and yersiniabactin loci, respectively) at a frequency of over 90% [26].

AST of these eight ST1193 strains showed multidrug resistance, including resistance to critically important antibiotics like fluoroquinolones and cephalosporins. In most cases, the phenotype matched the genotype, but strain MS8320 also showed *in vitro* resistance to piperacillin/tazobactam, ampicillin/sulbactam, cefazolin and doripenem (carbapenem), suggesting that eight copies of the *bla*_{TEM-1B} gene carried on IS26-mediated transposable units in a tandem array confer additional

resistance phenotypes. A similar mechanism of piperacillin/tazobactam resistance due to IS26-mediated amplification of *bla*_{TEM-1B} in *E. coli* has previously been reported [78]. Similarly, tandem arrays which increase the copy number of genes encoding resistance to beta-lactams have been shown to confer resistance to carbapenems [79]. Fluoroquinolone resistance was attributed to vertically transmitted point mutations within the quinolone resistance-determining regions of *gyrA*, *parC* and *parE*, like other UPEC lineages. However, unlike the stepwise acquisition of mutations conferring fluoroquinolone resistance in the *E. coli* ST131 lineage, resistance to fluoroquinolones in ST1193 results from multiple recombination events [36]. Six of the eight ST1193 strains (excluding MCJCHV-1 and MS8320) also featured a similar resistance region on a homologous F-type plasmid, differing primarily on whether the short or long variant was present and the arrangement of gene modules. Although we cannot be sure of the events leading to the evolution of these plasmids, it is notable that variants of this resistance region featured in strains across both major clades of the ST1193 lineage.

The mutations in DNA gyrase and topoisomerase genes that encode resistance to fluoroquinolones were not present in non-ST1193 isolates within the basal CC14 lineages [13], except for the outgroup strain HICF686 (ST6460; K5; SRA: SRS1221162), which contains a single mutation in *gyrA* (S83L). Notably, the type 1 fimbrial adhesin was defined as type *fimH*64, which has been previously reported as a single-locus variant of *fimH*27 [C314T (P125L)] relative to other CC14 strains such as HICF686 [23]. Of most interest and supported by Johnson *et al.* [13], the capsular locus in ST14 sub-lineages is K5. This indicates that the ancestral capsule type in the ST1193 lineage is K5 and that a switch to K1 has occurred prior to the expansion of the 'K1-associated' ST1193 clade. A defining characteristic of Clade 1 strains is carriage of the K1 capsular antigen (referred to as the 'K1-associated' lineage by Johnson *et al.* [13]), in contrast to carriage of the K5 capsular antigen in Clade 2 strains (referred to as the 'K5-associated' lineage [13]).

This high-resolution study in the genomics of ST1193 had two main strengths. The first was the contribution of three additional complete reference genomes representing strains from the major 'K1-associated' lineage, and two genomes from the 'K5-associated' lineage. The addition of these five strains with the three already available complete reference genomes of 'K1-associated' strains will help improve our knowledge of the pathobiology of ST1193. The second strength was that long-read sequencing enabled in-depth analysis of the core and accessory genome. This allowed for confirmation of previous assertions regarding capsule switching due to recombination. Long-read sequencing also allowed high-resolution examination of both antibiotic-resistance gene context and copy number, which was important for predicting antibiotic resistance phenotypes from the genotype. However, sequencing only five genomes to completion was a limitation in this study. As such, we acknowledge that phylogenetic inference on this small sample size alone would inappropriately underestimate the overall mutations across the ST1193 population. However, the availability of eight publicly available and published genomes (as of 2021) that are representative of the major 'K1-associated' and 'K5-associated' ST1193 lineages now offers opportunities for further large-scale phylogenetic analyses using more appropriate, closely related reference genomes. A larger sample size is critical for further understanding of transmission, adaptation and the evolution of ST1193.

CONCLUSION

We used ONT and Illumina sequencing to complete the first five representative ST1193 genomes from patients with UTIs in Australia. These genomes represent high-quality, circularized, complete reference genomes for future comparative genomics studies of ST1193. Here, we expanded on previous studies of ST1193 by providing the genomic context of virulence and antibiotic resistance gene content and highlighted the important role of mobile genetic elements in shaping this lineage. While ST1193 represents a new threat to global human health, its emergence and spread as a fluoroquinolone-resistant clone appears to be driven by different mechanisms than ST131, including the acquisition of fluoroquinolone resistance due to recombination prior to a switch from the ancestral K5 (CC14 lineage) to K1 capsule type, and the acquisition of numerous mobile genetic element-borne virulence-associated siderophore loci. Detailed ongoing analyses of ST1193, and the incorporation of detailed epidemiological data, may help us to pinpoint triggers that further define the successful global expansion of this lineage and allow us to predict the emergence of new resistant clones in the future.

Funding information

This work was supported by the Australian Research Council (ARC) Linkage grant (LP130100736) and a National Health and Medical Research Council (NHMRC) grant (GNT1067455). Scott Beatson and Mark Schembri are supported by NHMRC fellowships.

Acknowledgements

We are grateful to the Australian Group on Antimicrobial Resistance for contributing isolates from their surveillance programme. We acknowledge the facilities, and the scientific and technical assistance of staff at the Australian Centre for Ecogenomics (Brisbane, Australia) and Queensland Health (Brisbane, Australia). We acknowledge Jenny Robson from Sullivan Nicolaides Pathology for providing strain metadata. We thank Roger Stephan (Institute for Food Safety and Hygiene Vetsuisse Faculty at the University of Zurich) and Catherine Logue (Poultry Diagnostic and Research Center at the University of Georgia) for providing MIC data for antibiotics against *E. coli* ST1193 strains AVS0096 and MCJCHV-1, respectively. This research was supported by QRIScloud and by use of the Nectar Research Cloud. The Nectar Research Cloud is a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS). Finally, we thank Derek Sarovich and Erin Price (Centre for Bioinnovation at the

University of the Sunshine Coast, and the Sunshine Coast Health Institute) and Thomas Cuddihy (QFAB Bioinformatics and Research Computing Centre, The University of Queensland) for high-performance computing support and helpful discussions about software functionality.

Author contributions

Conceptualization: R.T.W.; investigation: R.T.W.; funding acquisition: M.A.S., S.A.B.; formal analysis: R.T.W.; wet-lab experiments: M.J.B., D.B., L.A.F., K.M.P., M-D.P. and L.W.R.; data analysis: R.T.W. and M.M.A.; data curation: R.T.W., J.B., J.S.G., A.K.K., J.L.M., B.A.R., D.J.T., J.T. and M.A.S.; writing (original draft preparation): R.T.W. and M.M.A.; writing (review and editing): R.T.W., M.M.A., M.J.B., J.B., D.B., L.A.F., J.S.G., A.K.K., J.L.M., K.M.P., M-D.P., L.W.R., B.A.R., M.A.S., D.J.T., J.T., B.M.F. and S.A.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare that there are no conflicts of financial, general or institutional competing interests.

Ethical statement

This work was undertaken on stored bacterial cultures and no samples were collected from any persons and/or animals during this study.

References

1. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Am J Med* 2002;113:5–13.
2. Stamm WE. Scientific and clinical challenges in the management of urinary tract infections. *Am J Med* 2002;113:1–4.
3. Acar JF, Goldstein FW. Trends in bacterial resistance to fluoroquinolones. *Clin Infect Dis* 1997;24:S67–S73.
4. Hensley S. Cipro loses share in traditional market, as doctors seek other cures in shortage. *Wall Street J* 2001.
5. Cheng AC, Turnidge J, Collignon P, Looke D, Barton M, et al. Control of fluoroquinolone resistance through successful regulation, Australia. *Emerg Infect Dis* 2012;18:1453–1460.
6. White RT. *Escherichia coli*: placing resistance to third-generation cephalosporins and fluoroquinolones in Australia and New Zealand into perspective. *Microbiol Aust* 2021;42:104–110.
7. Australian Commission on Safety and Quality in Health Care. AURA 2023: fifth Australian report on antimicrobial use and resistance in human health. Sydney, Australia: Australian Commission on Safety and Quality in Health Care; 2023. <https://www.safetyandquality.gov.au/publications-and-resources/resource-library/aura-2023-fifth-australian-report-antimicrobial-use-and-resistance-human-health-report> [accessed 9 July 2021].
8. World Health Organization. Critically important antimicrobials for human medicine. 5 ed. Geneva: Switzerland; 2017. <https://www.who.int/publications/i/item/9789241512220> [accessed 9 July 2021].
9. World Health Organization. Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2017–2018. Geneva, Switzerland. World Health Organization; 2017. <https://www.who.int/publications/i/item/9789241513449> [accessed 30 June 2022].
10. Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, Khanh Nhu NT, Roberts LW, et al. Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. *mBio* 2016;7:e00347–16.
11. Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, et al. Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *mBio* 2016;7:e02162.
12. Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, et al. Systematic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res* 2017;27:1437–1449.
13. Johnson TJ, Elnekave E, Miller EA, Munoz-Aguayo J, Flores Figueroa C, et al. Phylogenomic analysis of extraintestinal pathogenic *Escherichia coli* sequence type 1193, an emerging multidrug-resistant clonal group. *Antimicrob Agents Chemother* 2019;63.
14. Ludden C, Decano AG, Jamrozy D, Pickard D, Morris D, et al. Genomic surveillance of *Escherichia coli* ST131 identifies local expansion and serial replacement of subclones. *Microb Genom* 2020;6:e000352.
15. White RT, Bull MJ, Barker CR, Arnott JM, Wootton M, et al. Genomic epidemiology reveals geographical clustering of multidrug-resistant *Escherichia coli* ST131 associated with bacteraemia in Wales. *Nat Commun* 2024;15:1371.
16. Forde BM, Roberts LW, Phan M-D, Peters KM, Fleming BA, et al. Population dynamics of an *Escherichia coli* ST131 lineage during recurrent urinary tract infection. *Nat Commun* 2019;10:3643.
17. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc Natl Acad Sci USA* 2014;111:5694–5699.
18. Price LB, Johnson JR, Aziz M, Clabots C, Johnston B, et al. The epidemic of extended-spectrum- β -lactamase-producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, H30-Rx. *mBio* 2013;4:e00377–13.
19. Platell JL, Cobbold RN, Johnson JR, Trott DJ. Clonal group distribution of fluoroquinolone-resistant *Escherichia coli* among humans and companion animals in Australia. *J Antimicrob Chemother* 2010;65:1936–1938.
20. Peirano G, Matsumara Y, Nobrega D, DeVinney R, Pitout J. Population-based epidemiology of *Escherichia coli* ST1193 causing blood stream infections in a centralized Canadian region. *Eur J Clin Microbiol Infect Dis* 2021.
21. Iqbal J, Dufendach KR, Wellons JC, Kuba MG, Nickols HH, et al. Lethal neonatal meningoencephalitis caused by multi-drug resistant, highly virulent *Escherichia coli*. *Infect Dis* 2016;48:461–466.
22. Ding Y, Zhang J, Yao K, Gao W, Wang Y. Molecular characteristics of the new emerging global clone ST1193 among clinical isolates of *Escherichia coli* from neonatal invasive infections in China. *Eur J Clin Microbiol Infect Dis* 2021;40:833–840.
23. Tchesnokova VL, Rechkina E, Larson L, Ferrier K, Weaver JL, et al. Rapid and extensive expansion in the United States of a new multidrug-resistant *Escherichia coli* clonal group, sequence type 1193. *Clin Infect Dis* 2019;68:334–337.
24. Harris PN, Peleg AY, Iredell J, Ingram PR, Miyakis S, et al. Meropenem versus piperacillin-tazobactam for definitive treatment of bloodstream infections due to ceftriaxone non-susceptible *Escherichia coli* and *Klebsiella* spp (the MERINO trial): study protocol for a randomised controlled trial. *Trials* 2015;16:24.
25. Xia L, Liu Y, Xia S, Kudinha T, Xiao S-N, et al. Prevalence of ST1193 clone and Inc1/ST16 plasmid in *E-coli* isolates carrying bla_{CTX-M-55} gene from urinary tract infections patients in China. *Sci Rep* 2017;7:44866.
26. Wu J, Lan F, Lu Y, He Q, Li B. Molecular characteristics of ST1193 clone among phylogenetic group B2 non-ST131 fluoroquinolone-resistant *Escherichia coli* *Front Microbiol* 2017;8:2294.
27. Johnson JR, Johnston BD, Porter SB, Clabots C, Bender TL, et al. Rapid emergence, subsidence, and molecular detection of *Escherichia coli* sequence type 1193-*fimH64*, a new disseminated multidrug-resistant commensal and extraintestinal pathogen. *J Clin Microbiol* 2019;57:e01664–18.
28. Rogers BA, Ingram PR, Runnegar N, Pitman MC, Freeman JT, et al. Community-onset *Escherichia coli* infection

- resistant to expanded-spectrum cephalosporins in low-prevalence countries. *Antimicrob Agents Chemother* 2014;58:2126–2134.
29. Chang J, Yu J, Lee H, Ryu H, Park K, et al. Prevalence and characteristics of lactose non-fermenting *Escherichia coli* in urinary isolates. *J Infect Chemother* 2014;20:738–740.
 30. Matsumura Y, Noguchi T, Tanaka M, Kanahashi T, Yamamoto M, et al. Population structure of Japanese extraintestinal pathogenic *Escherichia coli* and its relationship with antimicrobial resistance. *J Antimicrob Chemother* 2017;72:1040–1049.
 31. Rogers BA, Ingram PR, Runnegar N, Pitman MC, Freeman JT, et al. Sequence type 131 fimH30 and fimH41 subclones amongst *Escherichia coli* isolates in Australia and New Zealand. *Int J Antimicrob Agents* 2015;45:351–358.
 32. Kidsley AK, White RT, Beatson SA, Saputra S, Schembri MA, et al. Companion animals are spillover hosts of the multidrug-resistant human extraintestinal *Escherichia coli* pandemic clones ST1131 and ST1193. *Front Microbiol* 2020;11:1968.
 33. Platell JL, Trott DJ, Johnson JR, Heisig P, Heisig A, et al. Prominence of an O75 clonal group (clonal complex 14) among non-ST131 fluoroquinolone-resistant *Escherichia coli* causing extraintestinal infections in humans and dogs in Australia. *Antimicrob Agents Chemother* 2012;56:3898–3904.
 34. Gomi R, Matsuda T, Fujimori Y, Harada H, Matsui Y, et al. Characterization of pathogenic *Escherichia coli* in river water by simultaneous detection and sequencing of 14 virulence genes. *Environ Sci Technol* 2015;49:6800–6807.
 35. Cabal A, Peischl N, Rab G, Stöger A, Springer B, et al. Draft genome sequence of a multidrug-resistant *Escherichia coli* sequence type 1193 pandemic clone isolated from wastewater in Austria. *Microbiol Resour Announc* 2021;10:e0076221.
 36. Tchesnokova V, Radey M, Chattopadhyay S, Larson L, Weaver JL, et al. Pandemic fluoroquinolone resistant *Escherichia coli* clone ST1193 emerged via simultaneous homologous recombinations in 11 gene loci. *Proc Natl Acad Sci USA* 2019;116:14740–14748.
 37. Elazhary MA, Saheb SA, Roy RS, Lagacé A. A simple procedure for the preliminary identification of aerobic gram negative intestinal bacteria with special reference to the Enterobacteriaceae. *Can J Comp Med* 1973;37:43–46.
 38. Mohsin M, Tanaka K, Kawahara R, Kondo S, Noguchi H, et al. Whole-genome sequencing and comparative analysis of the genomes of *Bacteroides thetaiotaomicron* and *Escherichia coli* isolated from a healthy resident in Vietnam. *J Glob Antimicrob Resist* 2020;21:65–67.
 39. Biggel M, Hoehn S, Schmitt K, Frei A, Jans C, et al. Complete genome sequence of *Escherichia coli* sequence type 1193 isolate AVS0096, recovered from river water in Switzerland. *Microbiol Resour Announc* 2021;10:e0060721.
 40. Goh KGK, Phan M-D, Forde BM, Chong TM, Yin W-F, et al. Genome-wide discovery of genes required for capsule production by uropathogenic *Escherichia coli* mBio 2017;8:e01558-17.
 41. Nhu NTK, Phan M-D, Peters KM, Lo AW, Forde BM, et al. Discovery of new genes involved in curli production by a uropathogenic *Escherichia coli* strain from the highly virulent O45:K1:H7 lineage. *mBio* 2018;9:e01462-18.
 42. Nhu NTK, Phan M-D, Forde BM, Murthy AMV, Peters KM, et al. Complex multilevel control of hemolysin production by uropathogenic *Escherichia coli*. *mBio* 2019;10:e02248-19.
 43. Hancock SJ, Lo AW, Ve T, Day CJ, Tan L, et al. Ucl fimbriae regulation and glycan receptor specificity contribute to gut colonisation by extra-intestinal pathogenic *Escherichia coli*. *PLoS Pathog* 2022;18:e1010582.
 44. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
 45. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. *Nucleic Acids Res* 2011;39:W347–W352.
 46. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016;44:W16–W21.
 47. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 2017;45:W30–W35.
 48. Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* 2011;12:R30.
 49. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 2006;34:D32–D36.
 50. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000;16:944–945.
 51. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:e11147.
 52. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–1403.
 53. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;15:524.
 54. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
 55. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. New York, United States of America: Springer; 2016.
 56. R Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019. <https://www.r-project.org/> [accessed 3 November 2021].
 57. Huang W, Li L, Myers JR, Marth GT. ART: a next-generation sequencing read simulator. *Bioinformatics* 2012;28:593–594.
 58. Sarovich DS, Price EP. SPANdx: a genomics pipeline for comparative analysis of large haploid whole genome re-sequencing datasets. *BMC Res Notes* 2014;7:618.
 59. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 2012;9:772.
 60. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
 61. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
 62. Pitout JDD, DeVinney R. *Escherichia coli* ST131: a multidrug-resistant clone primed for global domination. *F1000Res* 2017;6.
 63. Whitfield C. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* 2006;75:39–68.
 64. Whitfield C, Roberts IS. Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* 1999;31:1307–1319.
 65. Vann WF, Schmidt MA, Jann B, Jann K. The structure of the capsular polysaccharide (K5 antigen) of urinary-tract-infective *Escherichia coli* O10:K5:H4. A polymer similar to desulfo-heparin. *Eur J Biochem* 1981;116:359–364.
 66. Dewitt CW, Rowe JA. Sialic acids (N,7-O-diacetylneuraminic acid and N-acetylneuraminic acid) in *Escherichia coli*. I. Isolation and identification. *J Bacteriol* 1961;82:838–848.
 67. McGuire EJ, Binkley SB. The structure and chemistry of colominic acid. *Biochemistry* 1964;3:247–251.
 68. Kasper DL, Winkelhake JL, Zollinger WD, Brandt BL, Artenstein MS. Immunochemical similarity between polysaccharide antigens of *Escherichia coli* O7:K1(L):NM and group B *Neisseria meningitidis*. *J Immunol* 1973;110:262–268.

69. Glynn AA, Howard CJ. The sensitivity to complement of strains of *Escherichia coli* related to their K antigens. *Immunology* 1970;18:331–346.
70. Cross AS, Zollinger W, Mandrell R, Gemski P, Sadoff J. Evaluation of immunotherapeutic approaches for the potential treatment of infections caused by K1-positive *Escherichia coli*. *J Infect Dis* 1983;147:68–76.
71. Kim KS, Itabashi H, Gemski P, Sadoff J, Warren RL, et al. The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat. *J Clin Invest* 1992;90:897–905.
72. Alqasim A, Scheutz F, Zong Z, McNally A. Comparative genome analysis identifies few traits unique to the *Escherichia coli* ST131 H30Rx clade and extensive mosaicism at the capsule locus. *BMC Genom* 2014;15:830.
73. Holt KE, Lassalle F, Wyres KL, Wick R, Mostowy RJ. Diversity and evolution of surface polysaccharide synthesis loci in Enterobacteriales. *ISME J* 2020;14:1713–1730.
74. Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat Genet* 2013;45:656–663.
75. Pi H, Jones SA, Mercer LE, Meador JP, Caughron JE, et al. Role of catecholate siderophores in gram-negative bacterial colonization of the mouse gut. *PLoS One* 2012;7:e50020.
76. Hagan EC, Lloyd AL, Rasko DA, Faerber GJ, Mobley HLT. *Escherichia coli* global gene expression in urine from women with urinary tract infection. *PLoS Pathog* 2010;6:e1001187.
77. Watts RE, Totsika M, Challinor VL, Mabbett AN, Ulett GC, et al. Contribution of siderophore systems to growth and urinary tract colonization of asymptomatic bacteriuria *Escherichia coli*. *Infect Immun* 2012;80:333–344.
78. Hubbard ATM, Mason J, Roberts P, Parry CM, Corless C, et al. Piperacillin/tazobactam resistance in a clinical isolate of *Escherichia coli* due to IS26-mediated amplification of bla_{TEM-1B}. *Nat Commun* 2020;11:4915.
79. Forde BM, Henderson A, Playford EG, Looke D, Henderson BC, et al. Fatal respiratory diphtheria caused by β -lactam-resistant *Corynebacterium diphtheriae*. *Clin Infect Dis* 2021;73:e4531–e4538.
80. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011;27:1009–1010.

The Microbiology Society is a membership charity and not-for-profit publisher.

Your submissions to our titles support the community – ensuring that we continue to provide events, grants and professional development for microbiologists at all career stages.

Find out more and submit your article at microbiologyresearch.org