



Article

Comparative Genomic Analysis of ST131 Subclade C2 of ESBL-Producing *E. coli* Isolates from Patients with Recurrent and Sporadic Urinary Tract Infections

Daniel Jaén-Luchoro ^{1,2,*}, Arezou Kahnamouei ³, Shora Yazdanshenas ⁴, Anna Lindblom ^{1,2,4}, Emma Samuelsson ⁵, Christina Åhrén ^{1,2,6} and Nahid Karami ^{1,2,4}

¹ Institute of Biomedicine, Department of Infectious Diseases, University of Gothenburg, 40530 Gothenburg, Sweden; anna.u.lindblom@vgregion.se (A.L.); christina.ahren@vgregion.se (C.Å.); nahid.karami@microbio.gu.se (N.K.)

² Centre for Antibiotic Resistance Research (CARE), University of Gothenburg, 40530 Gothenburg, Sweden

³ Department of Life Sciences and Systems Biology, University of Turin, 10124 Turin, Italy; arezou.kahnamouei@edu.unito.it

⁴ Sahlgrenska University Hospital, Department of Clinical Microbiology, Region Västra Götaland, 41345 Gothenburg, Sweden

⁵ Sahlgrenska University Hospital, Department of Clinical Genetics and Genomics, Region Västra Götaland, 41345 Gothenburg, Sweden; emma.k.samuelsson@vgregion.se

⁶ Swedish Strategic Program against Antimicrobial Resistance (Strama), Region Västra Götaland, 40544 Gothenburg, Sweden

* Correspondence: daniel.jaen.luchoro@gu.se

Abstract: The global emergence of extended-spectrum beta-lactamase-producing *Escherichia coli* (ESBL-*E. coli*), mainly causing urinary tract infections (UTI), is a major threat to human health. ESBL-*E. coli* sequence type (ST) 131 is the dominating clone worldwide, especially its subclade C2. Patients developing recurrent UTI (RUTI) due to ST131 subclade C2 appear to have an increased risk of recurrent infections. We have thus compared the whole genome of ST131 subclade C2 isolates from 14 patients with RUTI to those from 14 patients with sporadic UTI (SUTI). We aimed to elucidate if isolates causing RUTI can be associated with specific genomic features. Paired isolates from patients with RUTI were identical, presenting 2–18 single nucleotide polymorphism (SNP) differences for all six patients investigated. Comparative genomic analyses, including virulence factors, antibiotic resistance, pangenome and SNP analyses did not find any pattern associated with isolates causing RUTI. Despite extensive whole genome analyses, an increased risk of recurrences seen in patients with UTI due to ST131 subclade C2 isolates could not be explained by bacterial genetic differences in the two groups of isolates. Hence, additional factors that could aid in identifying bacterial properties contributing to the increased risk of RUTI due to ESBL-*E. coli* ST131 subclade C2 remains to be explored.

Keywords: recurrent urinary tract infection; UTI; extended-spectrum beta-lactamase; ESBL; ST131 subclade C2; whole genome sequencing; *E. coli*



Citation: Jaén-Luchoro, D.; Kahnamouei, A.; Yazdanshenas, S.; Lindblom, A.; Samuelsson, E.; Åhrén, C.; Karami, N. Comparative Genomic Analysis of ST131 Subclade C2 of ESBL-Producing *E. coli* Isolates from Patients with Recurrent and Sporadic Urinary Tract Infections. *Microorganisms* **2023**, *11*, 1622. <https://doi.org/10.3390/microorganisms11071622>

Academic Editor: Mark Eppinger

Received: 3 May 2023

Revised: 31 May 2023

Accepted: 17 June 2023

Published: 21 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Extra-intestinal pathogenic *Escherichia coli*, or ExPEC, is a major human pathogen and is the most common cause of urinary tract infections (UTIs) and bloodstream infections [1,2]. Recurrent UTIs (RUTIs) affect approximately one-third of patients with UTI and thus pose an important clinical problem [3,4].

A specific ExPEC clone, sequence type 131 (ST131), is a globally spread multidrug-resistant clone with resistance to fluoroquinolones and producing CTX-M beta-lactamase, (mainly *bla*_{CTX-M-15}, *bla*_{CTX-M-14} and *bla*_{CTX-M-27}) [5]. It is one of the major extended-spectrum beta-lactamase-producing *Escherichia coli* (ESBL-*E. coli*) lineages among human

pathogens worldwide. ST131 is divided into three major clades: A, B and C [6]. Clades A and B are minor subsets and are associated with *fimH41* and *fimH22*, respectively. Clade C, associated with *fimH30*, represents two subclades, C1 (also named *fimH30R*) and C2 (also named *fimH30Rx*). ST131 subclade C2 (hereafter named ST131-C2) mostly harbours *bla*_{CTX-M-15} and represents the dominant population among clade C [7].

Compared to the highly conserved core genome, there is a wide plasticity in the accessory genome of ST131, resulting in the appearance of different virotypes and plasmid contents [8]. Subpopulations of ST131-C2 have been discussed in literature. For example, a subpopulation of ST131-C2 with increased virulence, carriage of *papGII*, increased antimicrobial resistance and chromosomally integrated *bla*_{CTX-M-15} has been described [9,10]. These characteristics of ST131-C2 isolates may lead to a severe and difficult-to-treat infections [9]. Another specific virotype of ST131-C2, with higher rates of aminoglycoside resistance and virulence gene content, shows the shifting dynamics of this pandemic clone in response to antibiotic selection pressure by establishing subsets with high survival potential [11].

We have reported that in the great majority (97%) of RUTIs due to ESBL-*E. coli*, the same strain caused almost all subsequent UTI recurrences for at least six months [12]. We also found that *fimH30Rx*, that is, ST131-C2, isolates are associated with an increased number of UTI recurrences. In a subsequent prospective study of RUTI due to ESBL-*E. coli*, we recently reported that the ST131-C2 clone dominated among ST131 isolates and was more common in patients with RUTI than in those with sporadic UTI (SUTI) (28% vs. 13%) [13]. Additionally, multivariate analysis showed a twofold increased risk for recurrences in patients infected with these isolates as compared to non-ST131 isolates.

To further evaluate if the risk of recurrences can be associated with a certain subpopulation of ST131-C2 or isolates with specific genomic traits or properties among those causing RUTI, we aimed to compare the whole genome of the ST131-C2 isolate from patients with recurrent and sporadic UTI identified in our previous study [13]. Identifying a particular genetic trait may aid in the prediction of RUTI and improve diagnostics and care of patients with a risk of developing recurrences when infected with this emerging pathogen.

2. Materials and Methods

2.1. Patients and Isolates

The 28 patients included in this study were part of a previous study investigating the bacterial features of 297 isolates from patients with recurrent as compared to sporadic UTI due to ESBL-*E. coli* in the Västra Götaland region in western Sweden [13]. In the previous study, between 1 October 2017 and 1 October 2018 all urinary ESBL-*E. coli* isolates from patients ≥ 15 years were collected from all clinical microbiology laboratories in the region. The laboratory databases were searched for urine samples positive for ESBL-*E. coli* in in- and outpatient settings. Epidemiological and clinical data linked to the index ESBL-*E. coli* isolates were extracted from referral data as well for all subsequent UTI episodes with ESBL-*E. coli* for one year following the index episode. Only patients with no previous recorded history of ESBL-producing bacteria in any type of clinical or screen culture were included in the study. Only voided samples with a significant number of ESBL-*E. coli* ($\geq 10^5$ CFU/mL) in monoculture were included. Cultures with isolates of more than one bacterial species were excluded, as were cultures with referral data indicating control purposes or presence of a urinary catheter or urinary abnormalities. Patients who died or moved from the region during the year following their first infection were excluded.

Recurrent UTI was defined according to the clinical international definition (<https://uroweb.org/guidelines/uropathological-infections> (accessed on 1 June 2023)), that is, ≥ 2 infections within six months or at least three within one year with at least 30 days between each episode. If a urine culture with other uropathogens than ESBL-*E. coli*, including *E. coli* not producing ESBL, was detected between ESBL-*E. coli* RUTI episodes, the patient was not included. Sporadic UTI was defined as only one UTI ESBL-*E. coli* episode during one year of follow up.

E. coli isolates were identified according to routine clinical microbiology practice and phenotypic antibiotic resistance was determined using the disc diffusion method and breakpoints according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) (<https://www.eucast.org> (accessed on 1 June 2023)) at the time. Cephalosporin-resistant isolates were screened for the ESBL phenotype using the double-disc diffusion assay and later genetically confirmed. For further methodological details, we refer the reader to our previous manuscript [13].

Out of the 50 ST131-C2 isolates detected in the previous study, almost all those causing RUTI (14/19) and half of those causing SUTI (14/31) were included in this study, aiming to match included patients with RUTI with those with SUTI by age and sex as increasing age was found to also increase the risk of RUTI in a previous study [13]. In addition, both the index and the subsequent isolate in six arbitrarily selected patients with RUTI were compared.

2.2. Whole Genome Sequencing and Assembly

DNA was extracted from fresh biomass, following a previously described protocol [14]. DNA samples were quantified with the Qubit[®] 2.0 fluorimeter and the Qubit[™] dsDNA BR kit (ThermoFisher Scientific, Waltham, MA, USA). Quality was determined by analysis of ratios 260/230 and 260/280 on a NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). Estimation of the distribution of DNA fragment sizes was performed, using a TapeStation 2200 (Agilent Technologies, Santa Clara, CA, USA). Whole genome sequencing was performed externally, using an Illumina NovaSeq 6000 S4 (read mode 2x 150 bp) (Eurofins Genomics, Ebersberg, Germany). All strains were additionally sequenced in-house, using a MinION Mk101B sequencer (Oxford Nanopore Technologies, Oxford, UK), for the generation of longreads. Libraries were prepared, using the rapid barcoding sequencing kit (SQK-RBK004) (Oxford Nanopore Technologies, Oxford, UK) and the sequencing was performed using a FLOW-MIN106 vR9.4 (Oxford Nanopore Technologies, Oxford, UK). The sequencing run was performed for 72 h on MinKNOW software v4.3.12 (Oxford Nanopore Technologies, Oxford, UK). Reads were basecalled and demultiplexed afterwards using Guppy v6.0.1 (Oxford Nanopore Technologies, Oxford, UK). Hybrid assemblies were performed using Unicycler v0.4.8 [15], combining short- and longreads generated by the Illumina and nanopore sequencing platforms, respectively. Quality of the genomes was assessed using the online Quality Assessment Tool (QUAST) [16]. Genomes were annotated using Prokka v1.14.6 [17] for downstream analysis and using the Prokaryotic Genome Annotation Pipeline (PGAP) for submission to GenBank [18].

2.3. Genomic Characterization of the Isolates

Genome sequences generated characterized by in silico multi-locus sequence typing (MLST), serotyping and phylogroup typing. MLST was determined using the software mlst 2.19.0 (<https://github.com/tseemann/mlst>). Serogroups were analysed using ectyper v1.0.0 [19], and phylogroups were inferred using EZclermont v0.6.3 [20]. Determination of *fimH* type was performed using FimTyper v1.1 [21]. Determination of antimicrobial resistance genes and virulence factors was performed using the Comprehensive Antibiotic Resistance Database (CARD) [22], and the Virulence Factors Database (VFDB) [23], respectively, with the software abricate v1.0.1 (<https://github.com/tseemann/abricate>).

2.4. Pan and Core Genome Analysis

The pangenome was calculated as previously described [24], using the protein sequence file obtained in the previous annotation with Prokka. Clusters of homologous proteins identified were used to build a matrix of the presence/absence of proteins. This matrix was used for clustering analysis, generating a heatmap of the presence/absence of proteins with the ComplexHeatmap package in Rstudio v1.2.5033 (<http://www.rstudio.com>) [25].

Additionally, core genome single-nucleotide polymorphism (SNP) analysis and SNP-core tree were performed using snippy v4.6.0 (<https://github.com/tseemann/snippy>).

Core genome MLST (cgMLST) analysis and cgMLST tree were performed using the on-line tool cgMLSTFinder v1.2 from the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org/>), using the *Escherichia coli* (Enterobase) database [26,27].

Gene–trait association analysis through a microbial pan-GWAS was performed, using Scoary v1.6.16 [28]. The traits were “recurrent infection” (R) and “sporadic infection” (S). Input files for Scoary were generated by an additional pangenome calculation performed with Roary v3.13.0 [29], using the default parameters. The same analysis but based on SNPs was performed, using as input the SNP presence absence matrix previously generated with snippy.

3. Results and Discussion

A total of 28 ST131 subclade C2 ESBL-*E. coli* isolates from 14 patients with RUTI and 14 patients with SUTI were compared in the present study (Table 1). The subsequent recurrent isolate from six patients with RUTI were also included. All isolates were whole-genome sequenced by Illumina (short-read) and nanopore (long-read) technologies. The objective was to obtain the complete and closed genome of all isolates in order to have the entire genetic information for a complete comparison. The recurrent isolates from six patients with RUTI were sequenced with Illumina only and the data obtained were used to compare the first and second isolate of the respective patient. Accession numbers for Illumina and nanopore data of each isolate are included in Table S1.

Table 1. Characteristics for the 28 patients with recurrent or sporadic UTI and their studied ESBL-*E. coli* isolates.

Patient and Isolates ID	Sex	Age (Year)	HC/PC ^b	Date of Isolation	Phenotypic Resistance to			Days to First Recurrence	Virotype ^d
					CIP ^c	TOB ^c	TMP ^c		
Recurrent UTI									
R1	M	77	PC	Oct. 2017	R	S	R	40	C
R2	M	80	HC	Oct. 2017	R	R	R	68	C
R3 ^a	M	51	HC	Oct. 2017	R	R	S	42	C
R4	M	59	PC	Nov. 2017	R	R	S	158	C
R5	F	88	HC	Nov. 2017	R	R	R	58	C
R6 ^a	F	21	HC	Dec. 2017	R	R	S	78	C
R7	M	86	HC	Feb. 2018	R	R	S	106	A
R8	F	65	HC	Mar. 2018	R	R	S	41	C
R9 ^a	F	59	HC	May 2018	R	S	S	33	A
R10	M	78	HC	Jun. 2018	R	R	S	31	C
R11 ^a	F	89	HC	Jun. 2018	R	R	S	62	A
R12	F	34	HC	Aug. 2018	R	R	R	73	C
R13 ^a	F	80	HC	Sep. 2018	R	R	R	146	C
R14 ^a	F	74	PC	Sep. 2018	R	R	R	54	C
Sporadic UTI									
S1	M	73	HC	Nov. 2017	R	R	R		C
S2	F	78	PC	Dec. 2017	R	R	S		C
S3	M	74	PC	Jan. 2018	R	R	S		None
S4	M	74	HC	Jan. 2018	R	R	R		C
S5	M	51	HC	Jan. 2018	R	R	R		C
S6	F	68	PC	Feb. 2018	R	R	S		C
S7	F	85	PC	Mar. 2018	R	R	R		C
S8	F	25	HC	May 2018	R	R	S		C
S9	F	69	PC	May 2018	R	R	R		C
S10	F	47	HC	May 2018	R	R	R		C
S11	F	18	HC	Jul. 2018	R	R	R		C
S12	M	84	HC	Aug. 2018	R	R	R		C
S13	F	86	HC	Sep. 2018	R	R	S		C
S14	M	63	PC	Sep. 2018	R	R	R		C

a—The second isolate from these patients was also analysed by SNP analyses. b—Index isolate detected during hospital care (HC) or in primary care (PC). c—ciprofloxacin (CIP), tobramycin (TOB), trimetoprim (TMP): R = resistant, S = sensitive. d—virotype determination according to [30].

3.1. Genome-Based Typing

Genome assembly generated 28 whole genome sequences, from which 23 were completely closed and 5 were high-quality draft genomes. Those five isolates (R1, R5, S1, S10 and S11) for which only a draft genome was obtained were re-sequenced to obtain additional long-read information for completing the genome. However, the additional reads obtained were not enough to fulfil the purpose. Basic quality indicators, annotation results and accession numbers are included in Tables S2 and S3.

With the genomes assembled, checked, and characterized, they were compared genomically to identify the relatedness among strains and specific traits that could be linked to the recurrency of the infection, i.e., a particular trait that can identify and distinguish between R and S isolates. Other studies have been performed to identify traits associated with infection recurrency, with no conclusive results [31–33].

Firstly, genome sequences were analysed through several typing tools for sequence typing, serotyping, phylogroup typing and *fimH*-typing, to confirm the previous experimental determinations of these parameters. There was a 100% correlation; all isolates belonged to the ST131-C2, phylogroup B2 and serotype O25:H4, and were of *fimH*30 type.

Secondly, we investigated their identity by SNP analyses between the index and subsequent recurrent isolate in six patients with RUTI. The differences in the coupled isolates varied from 2 to 18 SNPs. (Table S4). We could thus confirm our previous findings [12] and findings from others that the same strain reoccurs, causing the subsequent UTI episode [33,34], most likely by reintroduction from the faecal or vaginal flora or from persistent intrabladder colonisation after the first infection [13,34,35]. Accordingly, it will be sufficient to analyse the patients' index isolates when comparing R and S isolates. It also suggests that a change in strain properties from the first to the subsequent UTI does not explain the occurrence of the recurrence event.

Thirdly, the 28 R and S isolates were compared for various genetic properties as outlined below.

3.2. Virulence and Antibiotic Resistance Determinants

Genomes were analysed to identify all possible encoded virulence factors (VF). The objective was to determine the possible differential distribution of VF, which could be related to the capacity to develop a recurrent infection, or the lack thereof. A total of 98 VF or virulence determinants were identified using the VFDB, from which 67 VFs were present in almost all isolates (or missing only in one of the isolates) (Table S5). The remaining 31 VFs investigated were distributed differentially among the different isolates (Table 2). Even though some virulence factors have been indicated to be over-represented in RUTIs in previous studies [36], other studies have indicated that R and S isolates are similar in terms of VF and that no virulence traits can be associated with recurrency [33,37]. For example, Ejrnæs et al. indicated that 12 VF had a higher prevalence in isolates causing relapse or persistent infections [38]. The analysis with the VFDB showed that, from the 12 VFs indicated by the mentioned study, seven were not found in any of the isolates studied in the present work (*sfa/focDE*, *iroN*, *KpsM II*, *KpsM II K2*, *malX*, *usp*, *agn43* and *PapA*), *PapH* was present in 20 of them, and *ChuA* and *FyuA* were present in all. Additionally, from all the virulence factors identified in our study, some predominance of certain VFs in one of the two subsets of isolates was noted. For example, *hlyABCD* is mostly, but not exclusively, present in S isolates (9/14) rather than R isolates (5/14), while *cnf1* is predominant in R isolates (9/14) rather than S isolates (4/14). None the less, we found no distribution pattern particularly linked to the R or S isolates, but rather a random distribution of the VF between both types of isolates, which has also been reported previously [38]. Nor did we find any association according to the virotypes described by Blanco et al. [30], based on presence or absence of the operon *afa*, the siderophore *iroN*, the brain endothelium invasion factor *ibeA*, and the secreted autotransporter toxin *sat*; 24 of the isolates were classified as virotype C (*afa* negative, *iroN* negative, *ibeA* negative, and *sat* positive), which has been indicated as the most predominant type [30]. For the remaining four isolates, three recurrent isolates

(R7, R9, R11) were virotype A (*afa* positive, *iroN* negative, *ibeA* negative, and *sat* positive); and one sporadic isolate (S3) did not match any of the four virotypes A-D. Nor did we find associations to RUTI when analysing the 55 VFs and virulence scoring for extraintestinal *E. coli* described by Johansson et al. [39]. Twelve of the VFs listed in the aforementioned study were already found by the previous BLAST analysis with the VFDB. primers provided by Johansson et al. [39] for the 55 VFs used for in silico PCR analysis, identifying five extra VFs (*papG* allele II, *yfcV*, *hra*, *kpsM II*, and *kfiC (K5)*) that showed no distribution correlated to S or R isolates. Altogether, these results reinforce the statement that recurrency may not be linked to a particular VF trait, but instead as suggested by others the success of ST131, the clade C may be especially associated with enhanced persistence within its host, evading host defence and antibiotic clearance [7,33,40–42].

Table 2. Virulence factors (VF) identified using the Virulence Factor DataBase (VFDB). Only VFs distributed differentially among the 28 ESBL-*E. coli* isolates are shown out of 98 VFs explored.

Isolate	Adhesins/Fimbria				Toxins			T2SS	Hemolysins		Others	
	<i>afaA,B-I,C-I,D</i>	<i>afaE-I</i>	<i>papCDEFGHIJK</i>	<i>papE</i>	<i>fimB</i>	<i>senB</i>	<i>cnf1</i>	<i>gspCDEFGHIJK</i>	<i>hlyABD</i>	<i>hlyC</i>	<i>draP</i>	<i>astA</i>
R1	.	.	Y	Y
R2	Y
R3	.	.	Y	.	.	Y	Y	.	Y	Y	.	.
R4	.	.	Y	.	.	.	Y	.	Y	Y	.	.
R5	.	.	Y	.	.	Y
R6	.	.	Y	.	.	Y	Y	.	Y	Y	.	.
R7	Y	Y	.	.	.	Y	Y	.
R8	Y	Y1	.	Y
R9	Y	Y	.	.	Y	.
R10	.	.	Y
R11	Y	Y	.	.	Y	.
R12	.	.	Y	.	.	.	Y	.	Y	Y	.	.
R13	.	.	Y	.	.	Y
R14	.	.	Y	.	.	.	Y	.	Y	Y	.	.
S1	.	.	Y	.	.	Y	Y	.	Y	Y	.	.
S2	Y	Y	.	.	.	Y	Y	.
S3	.	.	Y	.	.	.	Y	.	Y	Y	.	.
S4	.	.	Y	.	.	.	Y	.	Y	Y	.	.
S5	.	.	Y	.	.	Y	Y	.	Y	Y	.	.
S6	.	.	Y	.	.	.	Y	.	Y	Y	.	.
S7	Y	Y	.
S8	.	.	Y	.	.	Y	Y	.	Y	Y	.	.
S9	.	.	Y	.	.	Y
S10	.	.	Y	.	.	Y	Y	.	Y	Y	.	Y
S11	Y	Y	.	.	Y	.
S12	.	.	Y	.	.	.	Y	.	Y	Y	.	.
S13	.	.	Y	.	.	.	Y	.	Y	.	.	.
S14	.	.	Y	Y	Y	Y	.	Y	Y	Y	.	.

Y = yes, presence of the gene.

Since recurrency of infection has been previously linked to antibiotic-resistance traits [43], the same approach applied for VF in the present work was undertaken focused on antibiotic-resistance genes. In this case, 64 antibiotic-resistance-related genes in the CARD were detected, from which 50 were present in all isolates or absent in almost isolates (Table S6). The remaining 14 resistance genes were differentially distributed among the isolates but mixed between R and S strains (Table 3), with no clear differentiation between the two types of isolates. Similarly, to VF, some antibiotic resistance genes seem to be more present in one type of isolate. In this sense, *mphA*, *sul1* and *dfrA17* are present in 9/14 S isolates, while is only present in 4/14, 5/14 and 2/14 R isolates, respectively. In any case, as stated previously, the distribution of the genes does not point to a strong link to R or S isolates in the group of strains studied.

Table 3. Identified genes related to antibiotic resistance using CARD database. Only genes distributed differentially among the 28 isolates are shown out of 64 resistance genes explored.

Isolate	Aminoglycosides					Macrolides		β-lactams		Diaminopyrimidines		Sulfonamides		Tetracyclins
	AAC(3)-Ile	AAC(6′)-Ib-cr	APH(3′′)-Ib	APH(6)-Id	aadA2	aadA5	mphA	OXA-1	TEM-1	dfrA12	dfrA17	sul1	sul2	tet(A)
R1
R2	.	Y	.	.	Y	.	Y	Y	.	Y	.	Y	.	Y
R3	Y	Y	Y
R4	Y	Y	Y
R5	.	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
R6	.	Y	Y	Y
R7	.	Y	Y	Y	.	.	.	Y	Y	Y
R8	Y	Y	Y	.	.	.	Y	.	.
R9
R10	.	Y	Y	Y
R11	.	Y	Y	Y
R12	Y	Y	Y
R13	.	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
R14	Y	Y	.	.	Y	.	Y	Y	.	Y	.	Y	.	Y
S1	Y	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
S2	.	Y	Y	Y	.	.	.	Y	Y	Y
S3	Y	Y	Y
S4	Y	Y	Y	Y	.	.	Y	Y	.	Y
S5	Y	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
S6	.	Y	Y
S7	Y	Y	Y	Y	.	Y	Y	Y	Y	.	Y	Y	Y	Y
S8	Y	Y	Y	Y
S9	.	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
S10	Y	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	.
S11	.	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	.
S12	Y	Y	.	.	.	Y	Y	Y	.	.	Y	Y	.	Y
S13	Y	Y	Y
S14	.	.	Y	Y	.	Y	Y	.	Y	.	Y	Y	Y	Y

Y = yes, presence of the gene.

McNully et al. also identified an enrichment of variants in anaerobic metabolism genes in *E. coli* ST131 clade C [42]. We could not obtain publicly available accession numbers or locus tags for each gene highlighted in this study. Because of that, we were only able to correlate four genes by name, which presented at least three alleles in the mentioned study, and analyse them in the present work (*aroD*, *eutB*, *nemA*, *nirB*) under the hypothesis that different alleles could be present only in R or S isolates. The result showed 100% identity between sequences of both kind of strains. Even though this result, a more in-depth analysis should be performed with more genes involved in anaerobic metabolism to obtain a more comprehensive result in this aspect.

3.3. Pangenome Analysis

A pangenome was calculated, using all 28 genome sequences for the index isolates included in this study. The pangenome revealed a total of 6,432 homologous protein clusters (Figure S1). The pangenome matrix, indicating the presence or absence of a particular cluster in a particular isolate, was used for clustering analysis. The hypothesis was that differential distribution of genes conforming to the pangenome could separate the two kinds of isolates, and that the functions developed for these specific proteins could give clues about the capacity of an isolate to cause recurrent UTI. The heatmap generated showed no differential distribution of R and S isolates, indicating that there was not a particular pattern of presence/absence of genes associated with R or S isolates (Figure 1).

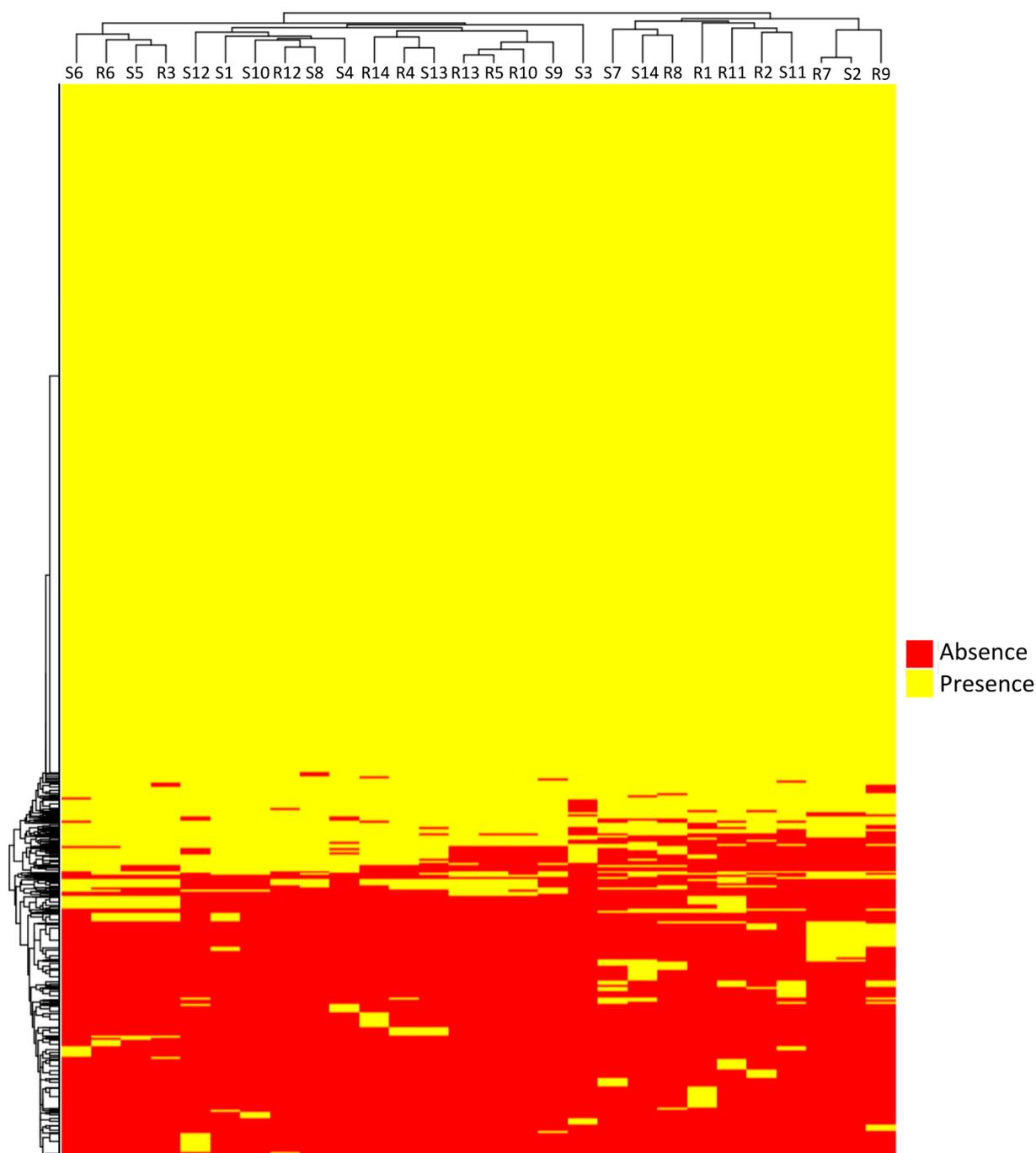


Figure 1. Clustering presence/absence of proteins based on the pangenome matrix in 28 ESBL-*E. coli* isolates from patients with recurrent (R) or sporadic (S) urinary tract infections.

Accessory genes were also investigated, determining exclusive genes for each strain, that is, genes present in a particular strain but absent in the rest. In the case of R isolates, the number of exclusive genes per strain ranged from 1 to 119; while for S isolates the number ranged from 0 to 112 (Table S7). The mean of exclusive genes per strain was 30.8 for R isolates, and 27.6 in S isolates. The variation in the number of exclusive genes as well as the similar mean of exclusive genes per strain between R and S isolates may indicate that there is no correlation between the capacity to accumulate differential genes and the ability to cause recurrent infections. Genes present only in all R isolates but not in S isolates, and vice versa, were also studied. No exclusive genes present neither only in R nor S isolates were found.

McNally et al. established a strong correlation between the distribution of intermediate frequency genes (present in >5% and <95% of the genomes studied) and sequence clusters [42]. We speculated about the possibility of a correlation of the distribution of these genes and the ability of an isolate to cause recurrent infections. A total of 920 intermediate frequency genes were identified, and clustering analysis based on these genes among the 28 isolates studied did not show a clear separation between S and R isolates (Figure 2).

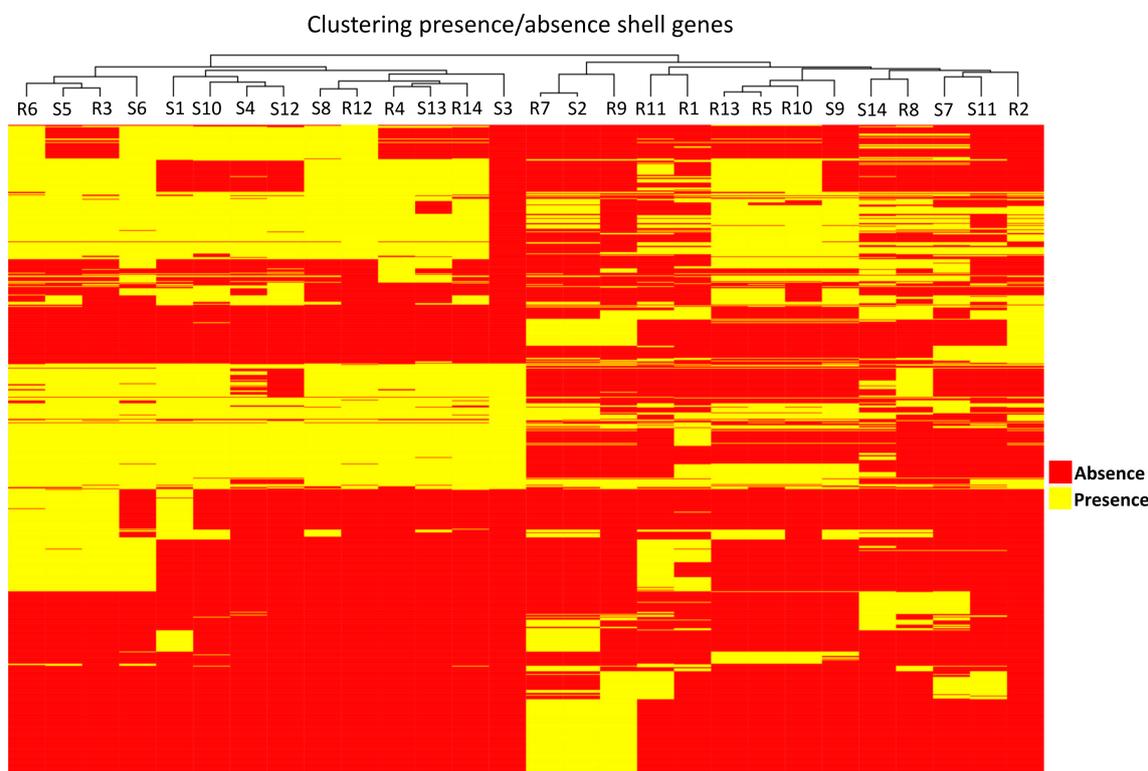


Figure 2. Clustering analysis based on presence/absence of genes which have been annotated in <5% and >95% of the genomes studied, i.e., intermediate genes present in 28 ESBL-*E. coli* isolates from patients with recurrent (R) or sporadic (S) urinary tract infections.

A more in-depth analysis from the pangenome point of view was also performed, using a microbial pan-GWAS approach, trying to establish a gene–trait association, being in this case defined as a trait representing R or S isolates. The objective was to find a gene, or group of genes, strongly associated with a particular trait, i.e., a gene or group of genes predominantly, or completely, associated with R or S isolates. When R was considered the reference trait, the two top-ranked proteins identified were dihydrofolate reductase type 1 and streptomycin 3'-adenylyltransferase, for which a sensitivity of 14.29% and a specificity of 35.71% (odds ratio: 0.09; p -value: 0.02) were calculated. The same calculations using S as a reference trait were done, and indicated, for these same top-ranked proteins, a sensitivity and specificity of 64.29% and 85.71%, respectively (odds ratio: 10.80 and p -value: 0.02). The obtained percentages do not allow for an interpretation of the importance these two proteins could have in terms of recurrency or sporadic infections. This means that the results showed no clear gene–trait association that could determine if, when that gene was found, a particular isolate would represent a potential R or S isolate, even though percentages are higher when S is taken as a reference trait. The same analysis was performed, but focusing on SNP, i.e., if a particular SNP (or group of SNPs) could be linked to S or R isolates, with no results obtained in terms of SNPs strongly linked to a particular trait.

3.4. SNP Analysis

A SNP-core tree was also determined, assessing the identify differences between R and S isolates based on SNP-phylogeny (Figure 3). Differences ranging from 1 to more than 2000 SNPs were found among the isolates. This difference is larger than those previously reported for ST131-C2 UTI isolates, where the diversity was strikingly low [43,44]. However, the SNP analyses revealed that some isolates seem more closely related to each other than others, but this could not be explained by an epidemiological relationship between cases as a particular ward or hospital, long-term facility or a certain part of a region, contrary to what have been described by others [44,45]. The distribution of R and S isolates was mixed, showing no differences traceable to a particular kind of isolate. The difference for R isolates (1-1,780 SNPs) and for S isolates (36 -2,133) was of a similar magnitude, further indicating the lack of a particular trait in R isolates. Taken together, from the core-genome perspective, the R and S isolates included in this study cannot be differentiated by variations in the core genome. This agrees with results previously described for RUTI due to a polyclonal set of sensitive *E. coli* without ESBL, lacking however ST131 isolates [37].

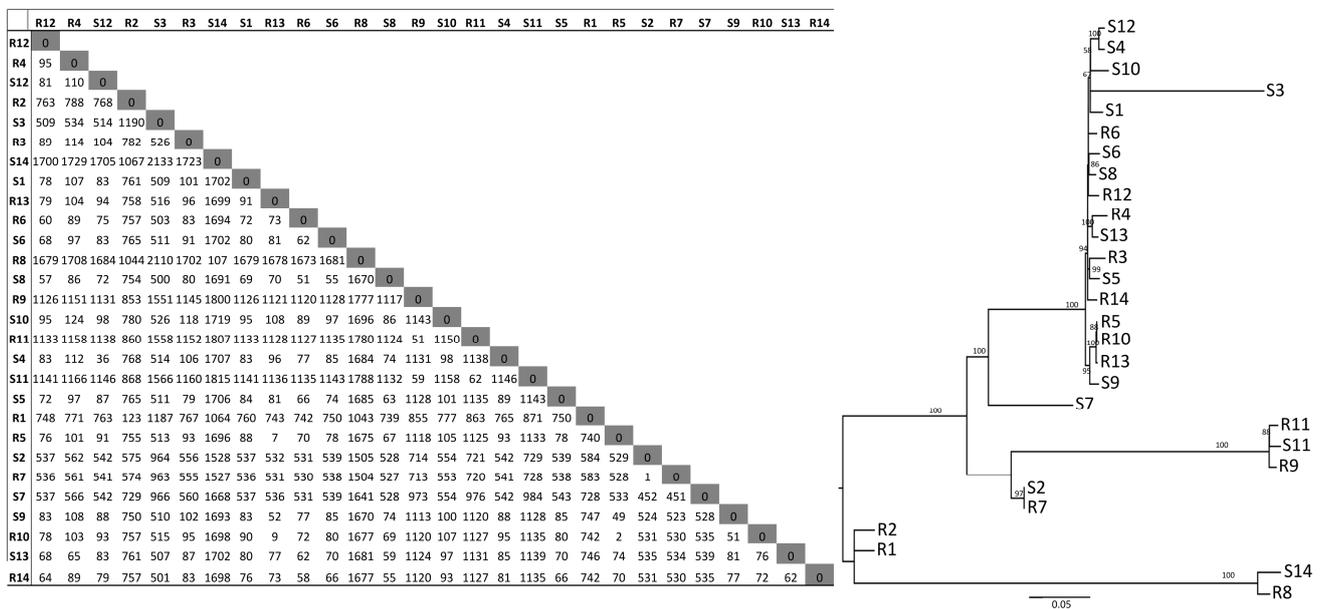


Figure 3. SNP core analysis performed on 28 ESBL-*E. coli* isolates from patients with recurrent (R) or sporadic (S) urinary tract infections. Number of SNPs between the different isolates, as well as the SNP-core tree, are shown.

Finally, a cgMLST approach was also assessed. Of a total of 2513 loci present in the database, more than 94% were called and matched in all the isolates (Figure 4). The tree based on the cgMLST was built, showing, once more, no differential distribution of R isolates and S isolates, indicating that they cannot be distinguished by cgMLST determinations

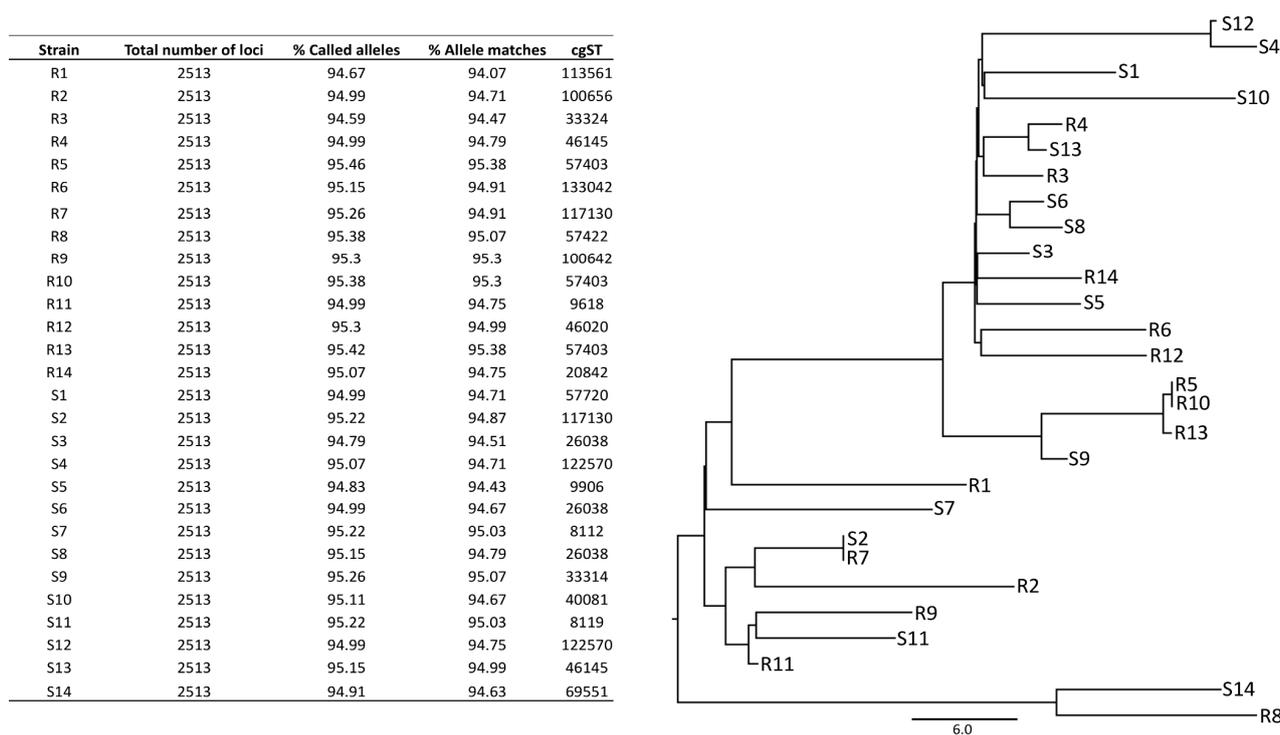


Figure 4. cgMLST results after analysis of the 28 isolates from patients with recurrent (R) and sporadic (S) infections. The percentages of alleles called and matched, and cgMLST tree, are indicated.

3.5. Limitations

These isolates were part of a large cohort of well-characterized isolates from patients with RUTI and SUTI due to ESBL-*E. coli* from an entire region, minimising the risk of clonal selection and epidemiologically linked isolates being included. Nonetheless, the low number of isolates in this study is a limitation. Despite this limitation, results correlate with previous studies performed with larger populations of more diverse strains, in which no genetic features were found to conclusively identify recurrency [37]. Additionally, there were no trends in dissimilarities detected in the various types of comparisons, suggesting that adding additional isolates would not alter the conclusions.

Secondly, we cannot exclude that despite presenting the same genes, the expression of a particular gene or combination of genes may vary between R and S isolates, and that this may have an impact on virulence or recurrency. From one of our previous studies, it can be seen that the same genes in different context, such as strains with different genetic backgrounds, can have different levels of expression [46]. Similarly, differences in phenotypic traits expressed in various environments and in importance in the ESBL-*E. coli* UTI pathogenesis may not have been revealed by the genetic analyses. Differences in early biofilm formation and type 1 fimbriae production between clade B and C isolates of ST131 have been described, but no differences were noted between subclade C1 and C2 ST131 ESBL-*E. coli* isolates [40]. In any case, finding a sequence or genetic trait that can be easily detected and linked to recurrent isolates of the emerging ST131 C2 subclade would be of great aid in foreseeing the risk of RUTI in patients infected with these isolates. It is also important to distinguish them from isolates that do not have that capacity, and hence, can be less problematic for the patient in the long term. In this sense, the present study tries to find genomic markers linked to recurrent or sporadic ST131-C2 *E. coli* isolates using different genomic approaches (pangenome, cgMLST, SNPs, pan-GWAS, VF, antibiotic resistance). Despite extensive genomic analyses, the results have shown no particular genomic traits underlying that distinction.

4. Conclusions

The genomic analysis and comparisons from different points of view presented here did not find any genetic traits or SNP particularly associated with recurrency, nor any gene-based clustering that correlates with recurrent isolates and differentiates them from isolates causing sporadic urinary tract infections. The increased risk of recurrences seen in patients with UTI due to ST131-C2 isolates can thus not solely be explained by bacterial genetic differences in the two groups of isolates. Other explanations need to be searched, including the expression of genes, phenotypic traits and patient and bacterial interactions that could favour the development of recurrent UTI.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms11071622/s1>, Table S1. Sequence Read Archive (SRA) accession numbers for Illumina and Nanopore Read datasets generated for each isolate; Table S2. Basic characteristics of the genome sequences obtained for all recurrent isolates; Table S3. Basic characteristics of the genome sequences obtained for all sporadic isolates; Table S4. Number of SNPs identified between the pairs of isolates from the same patient with recurrent infection (Index and subsequent isolate).

Author Contributions: Project conceptualization: D.J.-L., C.Å., A.L. and N.K. Investigation: D.J.-L., A.K. and S.Y. Data curation: D.J.-L. and E.S. Writing—Original draft: D.J.-L., C.Å. and N.K. Writing—Review and editing: D.J.-L., C.Å., A.L., E.S. and N.K. Supervision: D.J.-L., C.Å., E.S. and N.K. Project administration: D.J.-L., C.Å., E.S. and N.K. Funding acquisition C.Å., E.S. and N.K. All authors have read and agreed to the published version of the manuscript.

Funding: CÅ obtained funding from Region Västra Götaland (ALFGBG-965000) and (VGFOUREG-929979) and the Centre for Antibiotic Resistance Research (CARE) at University of Gothenburg and NK from Sahlgrenska University Hospital, C4A (no. 2020-02518).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Regional Ethical Review Board in Gothenburg, Sweden (recordal:170-17).

Informed Consent Statement: Not applicable.

Data Availability Statement: The sequencing data, including raw reads and genome sequences, are available in GenBank and the Sequence Read Archive (SRA) at NCBI. Accession numbers are included in Tables S1–S3.

Acknowledgments: The authors thank the staff at the Clinical Microbiological Laboratories for their help with collecting isolates.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Hooton, T.M. Clinical practice. Uncomplicated urinary tract infection. *N. Engl. J. Med.* **2012**, *366*, 1028–1037. [[CrossRef](#)] [[PubMed](#)]
2. Russo, T.A.; Johnson, J.R. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: Focus on an increasingly important endemic problem. *Microbes Infect.* **2003**, *5*, 449–456. [[CrossRef](#)] [[PubMed](#)]
3. Foxman, B.; Gillespie, B.; Koopman, J.; Zhang, L.; Palin, K.; Tallman, P.; Marsh, J.V.; Spear, S.; Sobel, J.D.; Marty, M.J.; et al. Risk factors for second urinary tract infection among college women. *Am. J. Epidemiol.* **2000**, *151*, 1194–1205. [[CrossRef](#)] [[PubMed](#)]
4. Lindblom, A.; Karami, N.; Magnusson, T.; Ahren, C. Subsequent infection with extended-spectrum beta-lactamase-producing Enterobacteriaceae in patients with prior infection or fecal colonization. *Eur. J. Clin. Microbiol. Infect. Dis.* **2018**, *37*, 1491–1497. [[CrossRef](#)] [[PubMed](#)]
5. Peirano, G.; Pitout, J.D. Molecular epidemiology of *Escherichia coli* producing CTX-M beta-lactamases: The worldwide emergence of clone ST131 O25, H4. *Int. J. Antimicrob. Agents* **2010**, *35*, 316–321. [[CrossRef](#)] [[PubMed](#)]
6. Peirano, G.; Lynch, T.; Matsumara, Y.; Nobrega, D.; Finn, T.J.; Devinney, R.; Pitout, J.D. Trends in Population Dynamics of *Escherichia coli* Sequence Type 131, Calgary, Alberta, Canada, 2006–2016(1). *Emerg. Infect. Dis.* **2020**, *26*, 2907–2915. [[CrossRef](#)]
7. Pitout, J.D.D.; Finn, T.J. The evolutionary puzzle of *Escherichia coli* ST131. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **2020**, *81*, 104265. [[CrossRef](#)]

8. Petty, N.K.; Ben Zakour, N.L.; Stanton-Cook, M.; Skippington, E.; Totsika, M.; Forde, B.M.; Phan, M.-D.; Moriel, D.G.; Peters, K.M.; Davies, M.; et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5694–5699. [[CrossRef](#)]
9. Biggel, M.; Moons, P.; Nguyen, M.N.; Goossens, H.; Van Puyvelde, S. Convergence of virulence and antimicrobial resistance in increasingly prevalent *Escherichia coli* ST131 papGII+ sublineages. *Commun. Biol.* **2022**, *5*, 752. [[CrossRef](#)]
10. Pajand, O.; Rahimi, H.; Darabi, N.; Roudi, S.; Ghassemi, K.; Aarestrup, F.M.; Leekitcharoenphon, P. Arrangements of Mobile Genetic Elements among Virotype E Subpopulation of *Escherichia coli* Sequence Type 131 Strains with High Antimicrobial Resistance and Virulence Gene Content. *mSphere* **2021**, *6*, e0055021. [[CrossRef](#)]
11. Hojabri, Z.; Darabi, N.; Mirmohammadkhani, M.; Rahimi, H.; Hemmati, R.; Saeedi, Z.; Roustaeae, K.; Leekitcharoenphon, P.; Pajand, O.; Aarestrup, F.M. Expansion of a Subset Within the C2 Subclade of *Escherichia coli* Sequence Type 131 (ST131) Is Driving the Increasing Rates of Aminoglycoside Resistance. *Open Forum Infect. Dis.* **2020**, *7*, ofaa410. [[CrossRef](#)]
12. Karami, N.; Lindblom, A.; Yazdanshenas, S.; Linden, V.; Ahren, C. Recurrence of urinary tract infections with extended-spectrum beta-lactamase-producing *Escherichia coli* caused by homologous strains among which clone ST131-O25b is dominant. *J. Glob. Antimicrob. Resist.* **2020**, *22*, 126–132. [[CrossRef](#)]
13. Lindblom, A.; Kiszakiewicz, C.; Kristiansson, E.; Yazdanshenas, S.; Kamenska, N.; Karami, N.; Åhrén, C. The impact of the ST131 clone on recurrent ESBL-producing *E. coli* urinary tract infection: A prospective comparative study. *Sci. Rep.* **2022**, *12*, 10048. [[CrossRef](#)]
14. Salvà-Serra, F.; Svensson-Stadler, L.; Busquets, A.; Jaén-Luchoro, D.; Karlsson, R.; Moore, E.R.; Gomila, M. A protocol for extraction and purification of high-quality and quantity bacterial DNA applicable for genome sequencing: A modified version of the Marmur procedure. *Protoc. Exch.* **2018**. [[CrossRef](#)]
15. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **2017**, *13*, e1005595. [[CrossRef](#)]
16. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [[CrossRef](#)]
17. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)]
18. Tatusova, T.; Dicuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **2016**, *44*, 6614–6624. [[CrossRef](#)]
19. Bessonov, K.; Laing, C.; Robertson, J.; Yong, I.; Ziebell, K.; Gannon, V.P.J.; Nichani, A.; Arya, G.; Nash, J.H.E.; Christianson, S. ECTyper: In silico *Escherichia coli* serotype and species prediction from raw and assembled whole-genome sequence data. *Microb. Genom.* **2021**, *7*, 000728. [[CrossRef](#)]
20. Waters, N.R.; Abram, F.; Brennan, F.; Holmes, A.; Pritchard, L. Easy phylotyping of *Escherichia coli* via the EzClermont web app and command-line tool. *Access Microbiol.* **2020**, *2*, acmi000143. [[CrossRef](#)]
21. Roer, L.; Tchesnokova, V.; Allesoe, R.; Muradova, M.; Chattopadhyay, S.; Ahrenfeldt, J.; Thomsen, M.C.; Lund, O.; Hansen, F.; Hammerum, A.M.; et al. Development of a Web Tool for *Escherichia coli* Subtyping Based on fimH Alleles. *J. Clin. Microbiol.* **2017**, *55*, 2538–2543. [[CrossRef](#)] [[PubMed](#)]
22. McArthur, A.G.; Waglechner, N.; Nizam, F.; Yan, A.; Azad, M.A.; Baylay, A.J.; Bhullar, K.; Canova, M.J.; De Pascale, G.; Ejim, L.; et al. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **2013**, *57*, 3348–3357. [[CrossRef](#)] [[PubMed](#)]
23. Chen, L.; Zheng, D.; Liu, B.; Yang, J.; Jin, Q. VFDB 2016: Hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* **2016**, *44*, D694–D697. [[CrossRef](#)] [[PubMed](#)]
24. Gonzales-Siles, L.; Karlsson, R.; Schmidt, P.; Salva-Serra, F.; Jaen-Luchoro, D.; Skovbjerg, S.; Moore, E.R.; Gomila, M. A Pangenome Approach for Discerning Species-Unique Gene Markers for Identifications of *Streptococcus pneumoniae* and *Streptococcus pseudopneumoniae*. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 222. [[CrossRef](#)] [[PubMed](#)]
25. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849. [[CrossRef](#)]
26. Clausen, P.; Aarestrup, F.M.; Lund, O. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinform.* **2018**, *19*, 307. [[CrossRef](#)]
27. Zhou, Z.; Alikhan, N.F.; Mohamed, K.; Fan, Y.; Agama Study, G.; Achtman, M. The Enterobase user’s guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia coli* core genomic diversity. *Genome Res.* **2020**, *30*, 138–152. [[CrossRef](#)]
28. Brynildsrud, O.; Bohlin, J.; Scheffer, L.; Eldholm, V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol.* **2016**, *17*, 238. [[CrossRef](#)]
29. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [[CrossRef](#)]
30. Blanco, J.; Mora, A.; Mamani, R.; Lopez, C.; Blanco, M.; Dahbi, G.; Herrera, A.; Marzoa, J.; Fernández, V.; De La Cruz, F.; et al. Four main virotypes among extended-spectrum-beta-lactamase-producing isolates of *Escherichia coli* O25b:H4-B2-ST131: Bacterial, epidemiological, and clinical characteristics. *J. Clin. Microbiol.* **2013**, *51*, 3358–3367. [[CrossRef](#)]

31. Forde, B.M.; Roberts, L.W.; Phan, M.-D.; Peters, K.M.; Fleming, B.A.; Russell, C.W.; Lenherr, S.M.; Myers, J.B.; Barker, A.P.; Fisher, M.A.; et al. Population dynamics of an *Escherichia coli* ST131 lineage during recurrent urinary tract infection. *Nat. Commun.* **2019**, *10*, 3643. [[CrossRef](#)]
32. Soto, S.M.; Smithson, A.; Martinez, J.A.; Horcajada, J.P.; Mensa, J.; Vila, J. Biofilm formation in uropathogenic *Escherichia coli* strains: Relationship with prostatitis, urovirulence factors and antimicrobial resistance. *J. Urol.* **2007**, *177*, 365–368. [[CrossRef](#)]
33. Thanert, R.; Reske, K.A.; Hink, T.; Wallace, M.A.; Wang, B.; Schwartz, D.J.; Seiler, S.; Cass, C.; Burnham, C.-A.D.; Dubberke, E.R.; et al. Comparative Genomics of Antibiotic-Resistant Uropathogens Implicates Three Routes for Recurrence of Urinary Tract Infections. *mBio* **2019**, *10*, e01977-19. [[CrossRef](#)]
34. Moreno, E.; Andreu, A.; Pigrau, C.; Kuskowski, M.A.; Johnson, J.R.; Prats, G. Relationship between *Escherichia coli* strains causing acute cystitis in women and the fecal *E. coli* population of the host. *J. Clin. Microbiol.* **2008**, *46*, 2529–2534. [[CrossRef](#)]
35. Kerrn, M.B.; Struve, C.; Blom, J.; Frimodt-Møller, N.; Krogfelt, K.A. Intracellular persistence of *Escherichia coli* in urinary bladders from mecillinam-treated mice. *J. Antimicrob. Chemother.* **2005**, *55*, 383–386. [[CrossRef](#)]
36. Johnson, J.R.; O'bryan, T.T.; Delavari, P.; Kuskowski, M.; Stapleton, A.; Carlino, U.; Russo, T.A. Clonal relationships and extended virulence genotypes among *Escherichia coli* isolates from women with a first or recurrent episode of cystitis. *J. Infect. Dis.* **2001**, *183*, 1508–1517. [[CrossRef](#)]
37. Nielsen, K.L.; Stegger, M.; Kiil, K.; Lilje, B.; Ejrnaes, K.; Leihof, R.F.; Skjøt-Rasmussen, L.; Godfrey, P.; Monsen, T.; Ferry, S.; et al. *Escherichia coli* Causing Recurrent Urinary Tract Infections: Comparison to Non-Recurrent Isolates and Genomic Adaptation in Recurrent Infections. *Microorganisms* **2021**, *9*, 1416. [[CrossRef](#)]
38. Ejrnaes, K.; Stegger, M.; Reinsner, A.; Ferry, S.; Monsen, T.; Holm, S.E.; Lundgren, B.; Frimodt-Møller, N. Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections: Phylogenetic groups, virulence factors and biofilm formation. *Virulence* **2011**, *2*, 528–537. [[CrossRef](#)]
39. Johnson, J.R.; Porter, S.; Johnston, B.; Kuskowski, M.A.; Spurbeck, R.R.; Mobley, H.L.; Williamson, D.A. Host Characteristics and Bacterial Traits Predict Experimental Virulence for *Escherichia coli* Bloodstream Isolates from Patients with Urosepsis. *Open Forum Infect. Dis.* **2015**, *2*, ofv083. [[CrossRef](#)]
40. Duprilot, M.; Baron, A.; Blanquart, F.; Dion, S.; Pouget, C.; Letteron, P.; Flament-Simon, S.-C.; Clermont, O.; Denamur, E.; Nicolas-Chanoine, M.-H. Success of *Escherichia coli* O25b:H4 Sequence Type 131 Clade C Associated with a Decrease in Virulence. *Infect. Immun.* **2020**, *88*, e00576-20. [[CrossRef](#)]
41. Hertz, F.B.; Marvig, R.L.; Frimodt-Møller, N.; Nielsen, K.L. In vitro Relative Fitness, in vivo Intestinal Colonization and Genomic Differences of *Escherichia coli* of ST131 Carrying bla (CTX-M-15). *Front. Microbiol.* **2021**, *12*, 798473. [[CrossRef](#)] [[PubMed](#)]
42. McNally, A.; Kallonen, T.; Connor, C.; Abudahab, K.; Aanensen, D.M.; Horner, C.; Peacock, S.J.; Parkhill, J.; Croucher, N.J.; Corander, J. Diversification of Colonization Factors in a Multidrug-Resistant *Escherichia coli* Lineage Evolving under Negative Frequency-Dependent Selection. *mBio* **2019**, *10*, e00644-19. [[CrossRef](#)] [[PubMed](#)]
43. Li, D.; Reid, C.J.; Kudinha, T.; Jarocki, V.M.; Djordjevic, S.P. Genomic analysis of trimethoprim-resistant extraintestinal pathogenic *Escherichia coli* and recurrent urinary tract infections. *Microb. Genom.* **2020**, *6*, e000475. [[CrossRef](#)] [[PubMed](#)]
44. Chen, S.L.; Ding, Y.; Apisarnthanarak, A.; Kalimuddin, S.; Archuleta, S.; Omar, S.F.S.; De, P.P.; Koh, T.H.; Chew, K.L.; Atiya, N.; et al. The higher prevalence of extended spectrum beta-lactamases among *Escherichia coli* ST131 in Southeast Asia is driven by expansion of a single, locally prevalent subclone. *Sci. Rep.* **2019**, *9*, 13245. [[CrossRef](#)]
45. Nobrega, D.; Peirano, G.; Lynch, T.; Finn, T.J.; Deviney, R.; Pitout, J.D.D. Spatial distribution of *Escherichia coli* ST131 C subclades in a centralized Canadian urban region. *J. Antimicrob. Chemother.* **2021**, *76*, 1135–1139. [[CrossRef](#)]
46. Jaen-Luchoro, D.; Karlsson, R.; Busquets, A.; Pineiro-Iglesias, B.; Karami, N.; Marathe, N.P.; Moore, E.R. Knockout of Targeted Plasmid-Borne beta-Lactamase Genes in an Extended-Spectrum-beta-Lactamase-Producing *Escherichia coli* Strain: Impact on Resistance and Proteomic Profile. *Microbiol. Spectr.* **2023**, *11*, e0386722. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.