



HAL
open science

Genotypic Characteristics and Antimicrobial Resistance of Escherichia coli ST141 Clonal Group

Audrey Emery, Didier Hocquet, Richard Bonnet, Xavier Bertrand

► **To cite this version:**

Audrey Emery, Didier Hocquet, Richard Bonnet, Xavier Bertrand. Genotypic Characteristics and Antimicrobial Resistance of Escherichia coli ST141 Clonal Group. *Antibiotics*, 2023, 12 (2), pp.382. 10.3390/antibiotics12020382 . hal-04157615

HAL Id: hal-04157615

<https://hal.science/hal-04157615>

Submitted on 6 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Genotypic Characteristics and Antimicrobial Resistance of *Escherichia coli* ST141 Clonal Group

Audrey Emery^{1,2,*} , Didier Hocquet^{1,3}, Richard Bonnet^{4,5} and Xavier Bertrand^{1,3} ¹ Hygiène Hospitalière, CHU, 25030 Besançon, France² Bactériologie, CHU Nice, 06202 Nice, France³ Chrono-Environnement, UMR 6249, CNRS Université de Franche-Comté, 25000 Besançon, France⁴ UMR 1071, Université d'Auvergne, 63000 Clermont-Ferrand, France⁵ Bactériologie, CHU Clermont-Ferrand, 63000 Clermont-Ferrand, France

* Correspondence: emery.a@chu-nice.fr; Tel.: +33-370-632-136

Abstract: *Escherichia coli* ST141 is one of the ExPEC lineages whose incidence is rising in France, even if no epidemic situation involving multidrug resistant isolates has been reported so far. Nonetheless, in a 2015–2017 monocentric study conducted in our French University hospital, ST141 was the most frequent lineage after ST131 in our collection of phylogroup B2 ESBL-producing *E. coli*. The genomes of 187 isolates representing ST141 group, including 170 genomes from public databases and 17 from our local collection, of which 13 produced ESBL, were analyzed to infer the maximum likelihood phylogeny SNP-based (Single Nucleotide Polymorphism) free-recombinant tree defining the ST141 population structure. Genomes were screened for genes encoding virulence factors (VFs) and antimicrobial resistance (AMR). We also evaluated the distribution of isolates according to their origin (host, disease, country) and the distribution of VFs or AMR genes. Finally, the phylogenetic tree revealed that ST141 isolates clustered into two main sublineages, with low genetic diversity. Contrasting with a highly virulent profile, as many isolates accumulated VFs, the prevalence of AMR was limited, with no evidence of multidrug resistant emerging lineage. However, our results suggest that surveillance of this clonal group, which has the potential to spread widely in the community, would be essential.

Keywords: *Escherichia coli*; ESBL; epidemic lineage; WGS; O2:H6; ST141



Citation: Emery, A.; Hocquet, D.; Bonnet, R.; Bertrand, X. Genotypic Characteristics and Antimicrobial Resistance of *Escherichia coli* ST141 Clonal Group. *Antibiotics* **2023**, *12*, 382. <https://doi.org/10.3390/antibiotics12020382>

Academic Editor: William R. Schwan

Received: 15 January 2023

Revised: 6 February 2023

Accepted: 7 February 2023

Published: 13 February 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

Escherichia coli is the predominant aerobic bacterium in the normal gut microbiota of humans and vertebrates as well as a major human pathogen [1]. Indeed, *E. coli* strains can cause both extra-intestinal pathologies (urinary tract infection, intra-abdominal or pulmonary infection as well as newborn meningitis or bacteraemia) and intestinal infections [2]. Currently, unlike these latter one (Intestinal Pathogenic *E. coli*, InPEC), for which virulence factors have been clearly identified and for which pathovar classification seems easy, the classification of ExPEC (Extraintestinal Pathogenic *E. coli*) strains is still subject to discussion, as no disease-specific virulence genes have been identified. It is the combination of different virulence factors, which may be numerous in these strains, that may explain their pathogenicity. Indeed, most, if not all, strains that may cause extra-intestinal infections have different genes encoding for virulence factors, such as adhesins, toxins, protectins and iron capture systems. However, the conditions under which *E. coli* strains emerge from their intestinal reservoir to cause extraintestinal infections, remain largely unknown as well as the reasons for the success of some pandemic lineages. Indeed, phylogenomic approaches have demonstrated that only four sequence types or sequencing type complexes (STCs), responsible of extraintestinal infections (STc131, STc73 and STc95 belonging to phylogroup B2, and STc69 belonging to phylogroup D) were always observed in epidemiological studies and were thus named “the big four ExPEC clones”. Although

not part of the big four ExPEC clones, *E. coli* ST141 is regularly reported as one of the most represented ExPEC [3]. Furthermore, among ExPEC, *E. coli* ST141 is singular: it has recently been hypothesized that the ST141 *E. coli* lineage genome would be very susceptible to recombinations, making it capable of acquiring and expressing specific genes of InPEC thus conferring a heteropathogen status [4]. The appearance of such hybrid clones could reshuffle the deck within the pathotype classification widely used to describe *E. coli* populations. Another matter of concern is the ability of *E. coli* to gain antibiotic resistance determinants given that *E. coli* may easily spread, as occurred with the pandemic of highly drug-resistant *E. coli* sequence type 131 (ST131) H30 sublineage [5]. To date, no epidemic situation involving multidrug resistant ST141 isolates has been reported [6]. Nonetheless, in a 2015–2017 monocentric study conducted in our French University hospital, consisting of collecting all non-duplicate extended-spectrum beta-lactamase producing *E. coli* (ES-BLEc), we identified that ST141 was the second most frequent lineage in our collection of phylogroup B2 ESBLec, accounting for 8.7% of ESBL-producing isolates, while ST131 lineage, the most frequent lineage, accounted for 55.3%, and isolates belonging to other STs each accounted for less than 3% of isolates [7].

Here, we used a comprehensive data set of ST141 genomes available in databases to characterize population structure of this clonal group with the aim of identifying the emergence of a multidrug resistant sublineage.

2. Results

2.1. ST141 Collection (Table S1)

- Among the 187 isolates analyzed, the geographic origin was known for 154 (82%): 108 isolates originated from European continent (mainly from France, United Kingdom, and Germany), 40 from American continent (North and South) and 6 from Asian continent. Most of the isolates (72%, 135/187) have been collected from humans, while non-human strains (i.e., environment, animal, and food) accounted for 15% (28/187) of the collection.
- Available genomes came from strains isolated between 1988 and 2019 and the sample date was known for 139 of which. A majority of strains were collected after 2010's (71.2%, 99/139).
- Among the 135 human-isolated strains, information about source sample was known for 100 isolates: 36% (36/100) came from urine, 33% (33/100) from blood culture, 23% (23/100) from feces (of which 13/23 were responsible of diarrhea) and 8% (8/100) from various clinical samples (of which three came from neonatal infections).

2.2. Whole Genome-Based Typing and SNP-Based Recombination-Free Phylogenetic Tree of ST141 *E. coli* Strains

- In silico MLST (Multi-Locus Sequence Typing) and phylogrouping confirmed that all selected isolates belonged to ST141 and B2 phylogroup. Within this ST141 clone, we wanted to assess the diversity of serogroup O and the *fimH* allele, which are two hotspots of recombinations, to be more discriminating [2]. Surprisingly, the ST141 clone has a low O-serogroup diversity, as all strains tested belonged to O2:H6 serotype. The *fimH* allele typing, which is based on minor sequence variations, allowed to identify *fimH5* (115/187, 61.5%), *fimH14* (48/187, 25.7%) and *fimH350* (10/187, 5.3%) as major *fimH* subtypes. Then, 16 minor *fimH* subtypes (i.e., shared by two or less isolates) were identified (Table S1, Figure 1) [8].
- Sublineages were indicated by edge coloring. Color strip corresponded to source origin. For easy readability, branches with fewer leaf nodes have been displayed on the top. Only bootstraps with value > 80% of confidence were displayed (blue burred nodes). SNP-based analysis displayed two main sublineages (S1 and S2). Whereas S1 (green edges) appeared relatively homogeneous, with highly conserved *fimH5* subtype (115/118), S2 (blue edges) *fimH* subtyping appeared less homogeneous, with predominant *fimH14* (48/69) and *fimH350* (10/69) subtypes. Isolates from local

collection (colored in orange) are scattered throughout the phylogenetic tree. Most of them (15/17) belonged to S1 while only two strains (S89 and S96) belonged to S2.

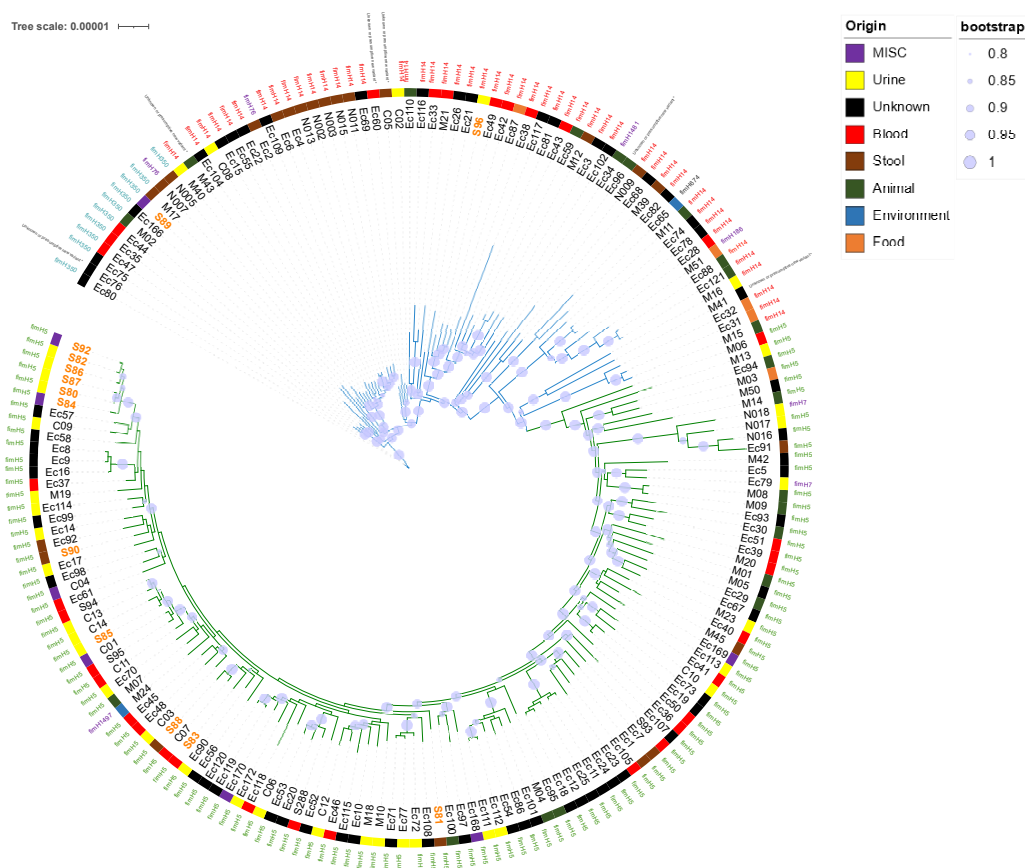


Figure 1. Representation of the two main *E. coli* ST141 sublineages obtained from maximum likelihood phylogeny SNP-based free-recombinant tree inferred from alignment to a fully sequenced reference strain of *E. coli* ST141.

- The SNP-based recombination-free phylogenetic tree depicted two main sublineages S1 and S2, clustering 118 and 69 isolates, respectively (Figure 1). S1 potentially appeared recently as a result of a *fimH* subtype conversion (Figure 1). Indeed, all isolates belonging to *fimH*5 clustered into S1 (115/118, 97.5%). The three remaining S1 isolates belonged to *fimH*7 (2/118, 1.7%) and *fimH*1497 (1/118, 0.8%) subtypes.
- Isolates from our local collection mostly belonged to S1 (15/17) and six of them exclusively grouped together. Isolates S86, S87, S80, S82, and S92 were indeed separated to each other by a maximum of 14 SNPs while S84 was distant with 26 SNPs. Finally, only two isolates, S89 and S96, did not belong to S1 (Table S1, Figure 2).
- Isolates from urines were more represented in S1 than in S2 (S1: 38%—31/81 vs. S2: 11%—5/47, $p < 0.01$). Conversely, isolates responsible for diarrhea (S1: 4%, 3/81 vs. S2: 22%—10/47) preferentially belonged to S2 ($p < 0.01$) (Table 1).
- Interestingly, 33 isolates of collection and previously studied by Gati et al., we noted a good correlation between the L1/L2 Gati's lineages and our S1/S2 distribution, as all isolates belonging to L2 clustered in the S1 sublineage and possessed the *fimH*5 subtype. Conversely, all isolates belonging to L1 clustered in the S2 sublineage, with different *fimH* subtypes (7 of subtype *fimH*14; three of subtype *fimH*350, two of subtype *fimH*76 and one of subtype *fimH* 674) [4] (Table S1, Figure 1).

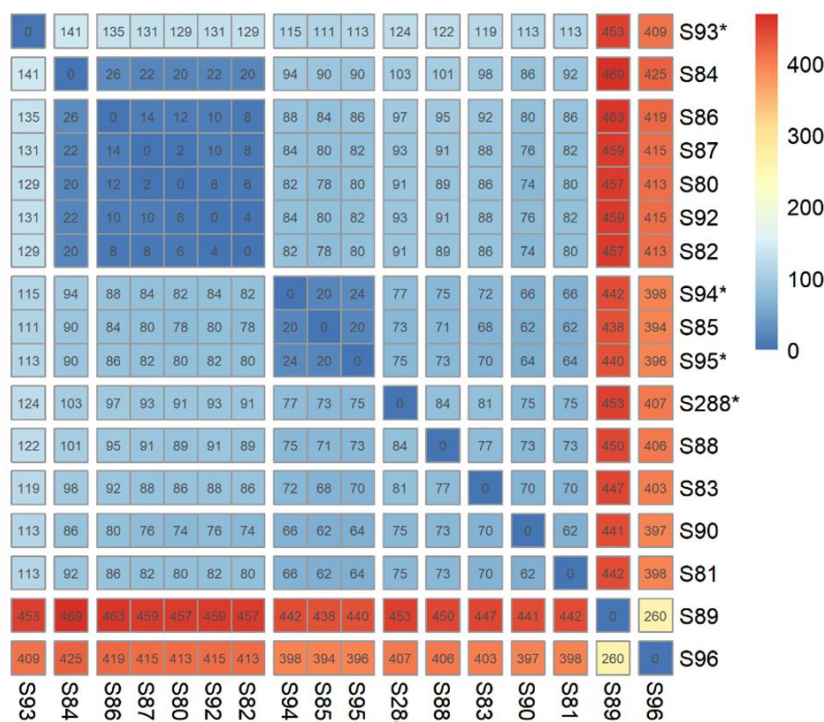


Figure 2. Distance matrix (SNP) of strains from our local collection ($n = 17$). Isolates identified with «*» symbol had no bla_{ESBL} gene. Five of our isolates were less than 15 SNP (S86, S87, S80, S92 and S82) or 26 (S84) from the nearest isolate.

Table 1. Origin and sublineage distribution of *E. coli* ST141 isolates ($n = 128$).

Source	Data Source	S1 ¹	S2 ¹	p-Value	Total
Human	Urine	31 (38%)	5 (11%)	0.0009 *	36 (28%)
Human	Blood	22 (27%)	11 (23%)	0.6806	33 (26%)
Human	Diarrhea	3 (4%)	10 (21%)	0.0041 *	13 (10%)
Human	Stool	5 (6%)	5 (11%)	0.4962	10 (8%)
Human	Misc. ¹	7 (9%)	1 (2%)	0.2564	8 (6%)
Non-human	Animal	11 (14%)	10 (21%)	0.3231	21 (16%)
Non-human	Food	1 (1%)	4 (9%)	0.0604	5 (4%)
Non-human	Environment	1 (1%)	1 (2%)	1.000	2 (2%)
Total		81 (100%)	47 (100%)	-	128 (100%)

¹ S1, Sublineage 1; S2, Sublineage 2; Misc., Miscellaneous. * Significantly different ($p < 0.01$). p-values were calculated with Fisher’s exact test. Strains with missing data (source and/or data source, $n = 59/187$) were omitted and human and non-human isolates differentiated.

2.3. Antimicrobial Resistance Genes in *E. coli* ST141 (Table S1, Figure 3)

Regarding resistance to β -lactams associated with extended-spectrum β -lactamase, a total of 35 strains (35/187, 18.7%) carried various bla_{ESBL} genes. Class A bla_{CTX_m} genes were the most prevalent ESBL genes (80.0%, 28/35) with bla_{CTX_M-14} being the most frequent (7.0%, 13/187) followed by bla_{CTX_M-15} (3.7%, 7/187), which was most often associated with the broad spectrum β -lactamase bla_{TEM-1} gene (3/7, 43%), and bla_{CTX_M-1} 2.1% (4/187). Of note, 13 out of these 28 ESBL-producing isolates came from our local collection. No statistical difference relating to the ESBL distribution was observed between S1 (22.9%, 27/118), and S2 (11.2% 8/69, $p = 0.08$). We did not retrieve bla_{ESBL} genes in ST141 isolates recovered from the environment. Additionally, two isolates harbored carbapenemase-encoding genes bla_{OXA-48} (Ec81) and bla_{VIM-1} (C04) and one strain (Ec110) isolated from poultry had colistin resistance gene mcr-1 (Table S1, Figure 2).

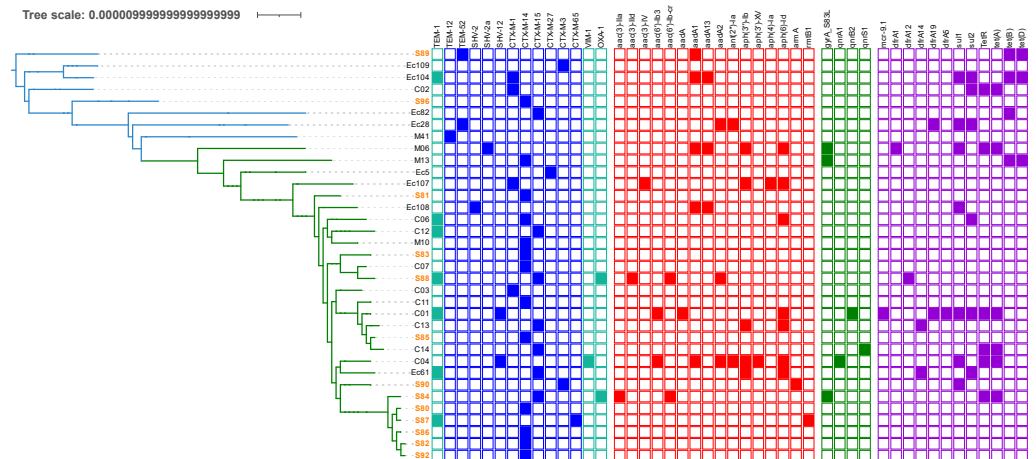


Figure 3. AMR genes distribution among *Escherichia coli* ST141 genomes with ESBL gene. Sublineages were indicated by edge coloring (S1, green edges, S2, blue edges). The presence/Absence of genes encoding AMR were indicated by colored squares (blue: β -lactam resistance—ESBL were shown in dark blue; red: aminoglycosides resistance; green: quinolone resistance; purple: miscellaneous).

Resistance genes to aminoglycosides were found in 21.9% (41/187) of the isolates and many of them (82.9%, 34/41) accumulated two or more resistance genes. The *aph(6)*-Id and *aph(3'')*-Ib, genes were the most common genes in our collection, as they were present in 15.0% (28/187) and 12.3% (23/187) of the isolates, respectively. We found no difference in the distribution of resistance genes to aminoglycosides between isolates of human and non-human origin ($p = 0.6179$) or between urine, blood or stool origin. Resistance genes to aminoglycosides were equally distributed between S1 (25.4%, 30/118) and S2 (15.9%, 11/69, $p = 0.1462$) (Table S1).

Resistance to quinolones was rare, as mutations in quinolone resistance-determining regions (QRDR) were found in nine isolates (4.8%, 9/187) and Plasmid-Mediated Quinolone-Resistance (PMQR) genes were present in only 2.1% (4/187) for *qnr* genes and 1.1% (2/187) for *aac(6')*-Ib-cr gene. One isolate (S84) harbored both *aac(6')*-Ib-cr and *gyrA* S83L mutation. Globally, the frequency of quinolone resistance (QRDR and PMQR) was not statistically different between S1 (10.2%, 12/118) and S2 (2.9%, 2/69), ($p = 0.09$) (Table S1).

To note, all 13 ESBL *E. coli* strains in our collection were tested in vitro (diffusion susceptibility testing according to EUCAST/CA-SFM recommendations in force at the time of the establishment of the strain collection) with a perfect correlation with results obtained in silico.

2.4. Virulence Factors in *E. coli* ST141 (Tables 2 and S1)

- Overall, we did not find a strong association between the distribution of VFs and the geographical origin, the origin (human or non-human), and the type of clinical sample. However, some differences in the distribution of VFs in S1 and S2 were statistically significant and displayed in Table 2. Thus, *stx2* gene was more frequent in the S2 sublineage (due to the predominance of strains causing diarrhea in this sublineage (S1 = 5/118; S2 = 12/118, $p = 0.004$)), while the Enteraggregative *E. coli* (EAEC) *astA* gene previously described to exclusively belong to strains of the L2 (i.e., S1 in this study) sublineage was significantly predominant in S1 sublineage [4]. However, the *astA* gene was also recovered in two isolates (Ec116 and C08) from our collection, classified as lineage S2, as they belonged to *fimH14* subtype. Of note, while this gene is supposed to be typical of EAEC (Enteroadhesive *E. coli*, responsible of diarrhea), we showed that *astA* gene was statistically more present in isolates from urine (23/36, 63.9%) than in isolates from stool (5/18, 21.7%, $p = 0.0028$). Finally, the invasion gene *ibeABC* was exclusively present in S2 sublineage.

- Assuming that the association of *hek*, *pap*, *cnf1* and *hlyA* genes could lead to the hypothesis of the presence of PAI II_{536-like}, we sought to highlight the strains in our collection that possessed these four genes. Finally, a total of 102 isolates (S1 = 69/118, 58.5% and S2 = 33/69, 47.8%, $p = 0.17$), from various origins (blood = 15/33, 45.5%, urine = 28/36, 77.8%, stool = 17/23, 73.9%, other or unknown = 42/95) could possess the PAI II_{536-like} UPEC virulence factor [4]. Statistically, PAI II_{536-like} was less frequent in strains isolated in blood than those isolated from urine ($p = 0.007$). No such difference could be demonstrated between stool and urinary strains (statistically insignificant, $p > 0.01$) (Table S1).
- Based on the evolutionary model of STEC/UPEC hybrid proposed by Gati et al., we then searched for an association of the gene *stx2* with other specific genes (PAI II_{536-like} or EHEC-*hly* gene) [4]. Among the four additional isolates owning *stx2* and not previously described by Gati et al. (Ec22, Ec23, Ec93 and Ec166), only one of them had both *stx2* and PAI II_{536-like} (i.e., Ec22), suggesting that the three others might have lost PAI II_{536-like} (Table S1). However, in our study collection, we have not identified any strain with EHEC-*hly* gene, except for the strain previously described in Gati's study (N011) [4].

Table 2. Virulence genes and lineage distribution of *E. coli* ST141 isolates.

Pathovar	Virulence Genes	S1 ¹ (%)	S2 ¹ (%)	p-Value
ExPEC ¹	<i>tia/hek</i>	80 (67.8%)	51 (73.9%)	0.41
ExPEC	alpha-hemolysin	73 (61.9%)	37 (53.6%)	0.28
ExPEC	<i>cnf1</i>	74 (62.7%)	36 (52.2%)	0.17
ExPEC	<i>vat</i>	99 (83.9%)	55 (79.7%)	0.55
ExPEC	Salmochelein system	114 (96.6%)	56 (81.2%)	<0.001 *
ExPEC	S pilus	1 (0.8%)	48 (69.6%)	<0.001 *
ExPEC	Type 1 pilus	118 (100.0%)	69 (100.0%)	1
ExPEC	P pilus	74 (62.7%)	36 (52.2%)	0.17
EXPEC	<i>ag43</i>	115 (97.5%)	62 (89.9%)	0.04
EXPEC	<i>ibeABC</i>	0 (0%)	69 (100%)	<0.001 *
EXPEC	PAI II _{536-like}	69 (58.5%)	33 (47.8%)	0.17
STEC ¹	Shiga like toxin 2	5 (4.2%)	12 (17.4%)	0.004 *
STEC	<i>iha</i>	7 (5.9%)	0 (0.0%)	0.09
EAEC ¹	<i>astA</i>	66 (55.9%)	2 (2.9%)	<0.001 *
EAEC	<i>aatA</i>	1 (0.8%)	1 (1.4%)	1
EAEC	<i>pic</i>	26 (30.5%)	46 (66.7%)	<0.001 *
Capsule	K1 capsule cluster	111 (94.1%)	49 (71.0%)	<0.001 *
	Total	118 (100%)	69 (100%)	-

¹ S1, Sublineage 1; S2, Sublineage 2; Misc., Miscellaneous; ExPEC, extra-intestinal pathogenic *E. coli*; STEC, shiga-toxin-producing *E. coli*; EAEC, Enteroaggregative *E. coli*. * Significantly different ($p < 0.01$). *p*-values were calculated with Fisher's exact test. S pilus, Shiga like toxin 2, *pic* and *ibeABC* virulence genes were more frequent in S2 (written in bold). Conversely, the frequency of Salmochelein system, *astA* and K1 capsule were more important in S1 (written in bold).

3. Discussion

We have demonstrated here that the population of *E. coli* ST141 is organized in two main sublineages S1 and S2, one predominantly associated with urinary tract infection (S1) and the other, more frequently associated with intestinal infections (S2). This confirms results generated from a smaller collection [4]. Indeed, for 41 isolates of our collection and previously studied by Gati et al., we noted a good correlation between the L1/L2 Gati's lineages and our S1/S2 distribution, as all isolates belonging to L2 clustered in the S1 sublineage and possessed the *fimH5* subtype (Figure 1). Conversely, all isolates belonging to L1 clustered in the S2 sublineage, with different *fimH* subtypes (seven of subtype *fimH14*; three of subtype *fimH350*, two of subtype *fimH76* and one of subtype *fimH674*) [4] (Table S1, Figure 1).

The ST141 lineage was highly conserved as all the isolates belonged to B2-phylogroup and had O2:H6 serotype with a restricted number of *fimH* subtypes. This contrasts with other B2 ExPEC successful lineages such as ST95, ST117 and ST131, which display a diversity of O-serogroups [2]. Although, *E. coli* ST141 isolates show various combinations of numerous VFs, as previously observed by Flament et al., we found no clear genomic signature that would indicate an ecological adaptation to a host species (i.e., humans and animals), infections sites, and countries of origin (Table S1, Figure 1) [9]. However, such diversity of VFs, consistent with the hypothesis that ST141 acts as one of the melting pots within the *E. coli* population, could be evidence of numerous recombinations and the presence of different PAI (pathogenicity islands), which could be interesting to explore in more details [4].

Previously, ST141 had been described as a STEC/UPEC hybrid or an EAEC/UPEC hybrid as some isolates carried *stx2* or *pic* and *astA* genes, respectively [4]. However, the majority (13/17) of strains in our ST141 collection ($n = 187$) which genomes contained *stx2* were those described by Gati et al., and only one of the four additional *stx2* gene positive strains also possibly had PAI II₅₃₆-like UPEC virulence factor, questioning the systematic and specifically heteropathogenicity of this clonal group. Moreover, Lindstedt et al. found in a Norwegian collection, a high frequency (64.3%) of *E. coli* strains combining IPEC and ExPEC virulence-associated genes [10]. Another German study revealed that 10.6% of strains isolated from UTIs harbored at least one IPEC virulence factor [11]. Given these conflicting data, the frequency of heteropathogenicity in *E. coli* remains to be clarified, as well as the involvement of different lineages (notably ST141) in this trait.

The fact remains that the ST141 clone has many virulence factors identified in silico. Taking into account the concept of antagonistic pleiotropy and epistatic interactions, a study aiming to characterize the expression of these, in particular in vivo, in an animal model or in vitro, by the capacity of the strains of our collection to produce biofilm could be envisaged in a future work [2].

Nevertheless, *E. coli* ST141 is commonly involved in human diseases and its incidence may reach a significant level. Indeed, ST141 lineage belonged to the 12 most frequent STs involved in bacteremia in France in 2014 [3]. In 2020, ST141 was quoted as one of the most common B2-ExPEC in France [9]. However, incidence of *E. coli* ST141 may vary between countries since a higher incidence of ST141-associated infections was found in France, compared to Spain [9]. In addition, *E. coli* ST141 has been found as the most frequent *E. coli* lineage responsible for ventilator-associated pneumonia over the 2012–2014 study period [12]. Interestingly, in a previous study, Philipps-Houlbracq et al., identified the antigen-43 (Ag43) as significantly associated in pneumonia pathogenesis [13]. It is important to note that Ag43 is widely represented among our collection (like most of B2 *E. coli*) as 97.5% (115/118) and 89.9% (62/69) of strains from sublineage 1 and 2, respectively, harbored this gene (Table 2).

Similarly, the panel of antibiotic resistance genes in the lineage ST141 was large with many combinations (Table S1, Figure 3). However, we could not identify a multi-resistant epidemic sub-lineage, such as that was observed with ST131 (i.e., C2-H30 producing CTX-M-15) [5]. Indeed, our study depicted a situation where *E. coli* ST141 might only contribute to limited locally spread of ESBL-encoding genes in the community, since the majority of ESBL strains isolated in our hospital form a cluster of strain separated by less than 30 SNPs, suggesting a local limited dissemination of an ESBL clone, although these strains have been isolated from patients with non-obvious epidemiological links (Figure 2) [7].

Likewise, although *E. coli* ST141 was capable of acquiring resistance genes to last resort antibiotics such as carbapenems (*bla*_{OXA-48}, *bla*_{VIM-1}) and colistin (*mcr-1*), we did not find evidence of the spread of such MDR clones.

4. Materials and Methods

4.1. ST141 Genome Collection

The genomes of 187 isolates representing ST141 *E. coli* group were obtained from various sources collected over a 30-year period (Table S1).

Firstly, 13 non-duplicate ESBL-producing ST141 *E. coli* isolates, collected in our University hospital, in inpatients, during a previous prospective observational cohort study, between 02/2015 and 01/2017 have been paired-end sequenced with Illumina NextSeq at 2×150 bp and included [7]. Then, four supplemental non-ESBL-producing ST141 *E. coli* responsible for bloodstream infection, isolated in 2014 in our hospital, have also been fully sequenced and added [3]. Finally, we collected all the ST 141 *E. coli* genomes available in public databases in March 2020 (NCBI and ENA).

4.2. Genome Analysis

Raw reads were first trimmed using Sickle and the 187 genomes were de novo assembled from reads using SPADes, as previously described [14,15]. From these assemblies, we determined in silico: MultiLocus Sequence Typing (MLST) according to the Achtman scheme, O:H serotype, *fimH* type, and phylogroup [8,16]. In the same way, we searched and identified antibiotic resistance genes with ResFinder database, and putative virulence factors (VFs) genes using the VFDB database, which compiles most *E. coli* VFs related to adhesion/invasion, autotransporter system, *fimbria* or flagella expression, iron uptake, serum resistance and toxicity [17,18]. A SNP (Single Nucleotide Polymorphism) call variant was performed against a fully sequenced ST141 *E. coli* genome (NCBI biosample accession number SAMN10740161) used as reference genome using BACTSNP [19]. After recombination curation with Gubbins, a maximum likelihood phylogenetic tree was then inferred from the resulting SNP-based pseudogenomes using RaxML [20,21]. A Cluster Picker analysis was processed to identify phylogenetic clusters (Table S1) [22]. The tree and corresponding metadata information were visualized with iTOL [23].

4.3. Statistical Analysis

To evaluate the distribution of isolates according to their origin and the distribution of VF genes, variables were examined by univariate analysis using the Fisher's exact test. All statistical tests were two tailed, and p -value < 0.01 was considered statistically significant.

5. Conclusions

Using a genome-based methodology, we depicted in our study the population structure of *E. coli* ST141 group using a large collection of 187 genomes. Our findings did not confirm the initial hypothesis of an emerging ESBL-producing ST141 sublineage, but rather demonstrated the spread of a subgroup of isolates, showing a closer relatedness, on a regional scale. Nonetheless, we found that *E. coli* ST141 readily acquires and cumulates VF encoding genes and that its prevalence had recently increased either in both extra-intestinal and intestinal diseases. Considering this, there is a need to monitor the spread of this clonal group that has the potential to largely spread in the community and whose involvement in human disease seemed to increase.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/antibiotics12020382/s1>, Table S1: Genomes features (origin, VFs, AMR).

Author Contributions: Conceptualization, A.E. and X.B.; Data curation, A.E. and R.B.; Formal analysis, A.E. and R.B.; Methodology, A.E., D.H., R.B. and X.B.; Project administration, X.B.; Supervision, X.B.; Validation, D.H., R.B. and X.B.; Writing—original draft, A.E. and X.B.; Writing—review and editing, D.H., R.B., A.E. and X.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: All information associated with the genome data were anonymized, thus identification of individuals is not possible. Therefore, ethical approval was not required.

Informed Consent Statement: Not applicable.

Data Availability Statement: All ST141 genomes from strains recovered from Besançon (local collection, $n = 17$) are publicly available through the NCBI BioProject PRJNA667655.

Acknowledgments: We thank B. Valot and A. Birer for technical assistance in WGS analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tenaillon, O.; Skurnik, D.; Picard, B.; Denamur, E. The population genetics of commensal *Escherichia coli*. *Nat. Rev. Genet.* **2010**, *8*, 207–217. [CrossRef]
2. Denamur, E.; Clermont, O.; Bonacorsi, S.; Gordon, D. The population genetics of pathogenic *Escherichia coli*. *Nat. Rev. Genet.* **2020**, *19*, 37–54. [CrossRef]
3. Van Der Mee-Marquet, N.L.; Blanc, D.S.; Gbaguidi-Haore, H.; Dos Santos Borges, S.; Viboud, Q.; Bertrand, X.; Quentin, R. Marked increase in incidence for bloodstream infections due to *Escherichia coli*, a side effect of previous antibiotic therapy in the elderly. *Front. Microbiol.* **2015**, *6*, 646. [CrossRef] [PubMed]
4. Gati, N.S.; Middendorf-Bauchart, B.; Bletz, S.; Dobrindt, U.; Mellmann, A. Origin and Evolution of Hybrid Shiga Toxin-Producing and Uropathogenic *Escherichia coli* Strains of Sequence Type 141. *J. Clin. Microbiol.* **2019**, *58*, e01309-19. [CrossRef] [PubMed]
5. Nicolas-Chanoine, M.-H.; Bertrand, X.; Madec, J.-Y. *Escherichia coli* ST131, an Intriguing Clonal Group. *Clin. Microbiol. Rev.* **2014**, *27*, 543–574. [CrossRef] [PubMed]
6. Brisse, S.; Diancourt, L.; Laouénan, C.; Vigan, M.; Caro, V.; Arlet, G.; Drieux, L.; Leflon-Guibout, V.; Mentré, F.; Jarlier, V.; et al. Phylogenetic Distribution of CTX-M- and Non-Extended-Spectrum- β -Lactamase-Producing *Escherichia coli* Isolates: Group B2 Isolates, Except Clone ST131, Rarely Produce CTX-M Enzymes. *J. Clin. Microbiol.* **2012**, *50*, 2974–2981. [CrossRef]
7. Muller, A.; Gbaguidi-Haore, H.; Cholley, P.; Hocquet, D.; Sauget, M.; Bertrand, X. Hospital-diagnosed infections with *Escherichia coli* clonal group ST131 are mostly acquired in the community. *Sci. Rep.* **2021**, *11*, 5702. [CrossRef]
8. Roer, L.; Tchesnokova, V.; Allesøe, R.; Muradova, M.; Chattopadhyay, S.; Ahrenfeldt, J.; Thomsen, M.C.F.; 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]
9. Flament-Simon, S.-C.; Nicolas-Chanoine, M.-H.; García, V.; Duprilot, M.; Mayer, N.; Alonso, M.P.; García-Meniño, I.; Blanco, J.E.; Blanco, M.; Blanco, J. Clonal Structure, Virulence Factor-encoding Genes and Antibiotic Resistance of *Escherichia coli*, Causing Urinary Tract Infections and Other Extraintestinal Infections in Humans in Spain and France during 2016. *Antibiotics* **2020**, *9*, 161. [CrossRef]
10. Lindstedt, B.-A.; Finton, M.D.; Porcellato, D.; Brandal, L.T. High frequency of hybrid *Escherichia coli* strains with combined Intestinal Pathogenic *Escherichia coli* (IPEC) and Extraintestinal Pathogenic *Escherichia coli* (ExPEC) virulence factors isolated from human faecal samples. *BMC Infect. Dis.* **2018**, *18*, 544. [CrossRef]
11. Toval, F.; Köhler, C.-D.; Vogel, U.; Wagenlehner, F.; Mellmann, A.; Fruth, A.; Schmidt, M.A.; Karch, H.; Bielaszewska, M.; Dobrindt, U. Characterization of *Escherichia coli* Isolates from Hospital Inpatients or Outpatients with Urinary Tract Infection. *J. Clin. Microbiol.* **2014**, *52*, 407–418. [CrossRef] [PubMed]
12. La Combe, B.; Clermont, O.; Messika, J.; Eveillard, M.; Kouatchet, A.; Lasocki, S.; Corvec, S.; Lakhali, K.; Billard-Pomares, T.; Fernandes, R.; et al. Pneumonia-Specific *Escherichia coli* with Distinct Phylogenetic and Virulence Profiles, France, 2012–2014. *Emerg. Infect. Dis.* **2019**, *25*, 710–718. [CrossRef] [PubMed]
13. Phillips-Houlbracq, M.; Ricard, J.-D.; Foucrier, A.; Yoder-Himes, D.; Gaudry, S.; Bex, J.; Messika, J.; Margetis, D.; Chatel, J.; Dobrindt, U.; et al. Pathophysiology of *Escherichia coli* pneumonia: Respective contribution of pathogenicity islands to virulence. *Int. J. Med. Microbiol.* **2018**, *308*, 290–296. [CrossRef] [PubMed]
14. Joshi, N.A.; Fass, J.N. Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files; Version 1.33; 2011. Available online: https://scholar.google.com/scholar?hl=zh-CN&as_sdt=0,5&cluster=15670407978335035877 (accessed on 14 January 2023).
15. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef]
16. Beghain, J.; Bridier-Nahmias, A.; Le Nagard, H.; Denamur, E.; Clermont, O. ClermonTyping: An easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb. Genom.* **2018**, *4*, e000192. [CrossRef]
17. Zankari, E.; Hasman, H.; Cosentino, S.; Vestergaard, M.; Rasmussen, S.; Lund, O.; Aarestrup, F.M.; Larsen, M.V. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **2012**, *67*, 2640–2644. [CrossRef]
18. Chen, L.; Yang, J.; Yu, J.; Yao, Z.; Sun, L.; Shen, Y.; Jin, Q. VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* **2005**, *33*, D325–D328. [CrossRef]

19. Yoshimura, D.; Kajitani, R.; Gotoh, Y.; Katahira, K.; Okuno, M.; Ogura, Y.; Hayashi, T.; Itoh, T. Evaluation of SNP calling methods for closely related bacterial isolates and a novel high-accuracy pipeline: BactSNP. *Microb. Genom.* **2019**, *5*, e000261. [[CrossRef](#)]
20. Croucher, N.J.; Page, A.J.; Connor, T.R.; Delaney, A.J.; Keane, J.A.; Bentley, S.D.; Parkhill, J.; Harris, S.R. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **2015**, *43*, e15. [[CrossRef](#)]
21. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
22. Ragonnet-Cronin, M.; Hodcroft, E.; Hué, S.; Fearnhill, E.; Delpech, V.; Brown, A.J.L.; Lycett, S. Automated analysis of phylogenetic clusters. *BMC Bioinform.* **2013**, *14*, 317. [[CrossRef](#)] [[PubMed](#)]
23. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, 256–259. [[CrossRef](#)] [[PubMed](#)]

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.