Virulence potential of a multidrug-resistant Escherichia coli belonging to the emerging clonal group ST101-B1 isolated from bloodstream infection

Virulence of a multidrug-resistant *Escherichia coli* ST101 causing bloodstream infection

Ana Carolina de Mello Santos[1*], Rosa Maria Silva[1], Tiago Barcelos Valiatti[1], Fernanda Fernandes dos Santos[1], Rodrigo Cayô[1,2], Ana Paula Streling[3], Carolina Silva Nodari[3], Ana Cristina Gales[3], Milton Yutaka Nishiyama-Junior[4], Eneas Carvalho[5], Tânia Aparecida Tardelli Gomes[1].

[1] Universidade Federal de São Paulo - Escola Paulista de Medicina (UNIFESP-EPM), Disciplina de Microbiologia, Departamento de Microbiologia, Imunologia e Parasitologia - DMIP, São Paulo - SP, Brazil.

[2] Universidade Federal de São Paulo - UNIFESP, Laboratório de Imunologia e Bacteriologia - LIB, Setor de Biologia Molecular, Microbiologia e Imunologia, Departamento de Ciências Biológicas - CDB, Instituto de Ciências Ambientais, Químicas e Farmacêuticas - ICAQF, Diadema - SP, Brazil.

[3] Universidade Federal de São Paulo - UNIFESP, Laboratório Alerta, Disciplina de Infectologia, Departamento de Medicina, Escola Paulista de Medicina - EPM, São Paulo - SP, Brazil.

[4] Laboratório Especial de Toxinologia Aplicada – Instituti Butantan, São Paulo – SP, Brazil.

[5] Laboratório de Bacteriologia - Instituto Butantan, São Paulo - SP, Brazil.

*corresponding author: carolina.mello@unifesp.br(ACMS)

Abstract

Escherichia coli EC121 is a multidrug-resistant (MDR) strain isolated from bloodstream infection of an inpatient with persistent gastroenteritis and Zone T lymphoma that died due to septic shock. Despite causing an extraintestinal infection, strain EC121 harbors very few known virulence factors associated with extraintestinal pathogenic *E. coli* (ExPEC). Furthermore, this strain was assigned to phylogenetic group B1, which is usually related to commensals or diarrheagenic *E. coli* (DEC) strains, being rarely linked to extraintestinal infections. The question rises if *E. coli* EC121 is opportunistic or does have a true virulence potential. To address this question the genome of strain EC121 was sequenced, and an *in vitro* characterization of some pathogenic-associated properties was performed. The data retrieved from genome analyses showed that *E. coli* strain EC121 belongs to the O154:H25 serotype, and to the sequence type 101-B1 (ST101-B1), which has been epidemiologically linked to extraintestinal infections and antimicrobial resistance spread as well. Moreover, it is shown to be closely related to Shiga-toxin producing *E. coli* (STEC). Besides, strain EC121 is an MDR strain harboring 14 antimicrobial resistance genes, including $bla_{\text{CTX-M-2}}$, and more than 50 complete virulence genetic clusters, which are reported to be associated either with DEC or ExPEC, or both. Strain EC121 also displays the capacity to adhere to and invade HeLa cells, intestinal Caco-2 cells and bladder T24 cells, as well as the ability to form biofilms on abiotic surfaces, and survive the bactericidal serum complement activity. Altogether, the genetic and phenotypic traits presented by *E. coli* EC121 may be unveiling a pathogen powered by its multi-drug resistance characteristic. The recognition of such MDR pathogens makes it essential to carry out studies providing accurate information about their virulence potential. Such studies will help in the development of alternative therapies of infection management and spread control of MDR strains.

Keywords: Extraintestinal, intestinal, pathogen, *Escherichia coli*, Multidrug resistance, hybrid pathogen, bloodstream infection, ExPEC, DEC, MDR, genome, virulence, ST101

Author summary

Extraintestinal pathogenic *Escherichia coli* are mostly associated with phylogenetic group B2 and the majority of the studies regarding extraintestinal infection focus on the most virulent strains, which might also possess multidrug-resistant phenotype, like the ST131 strains. On the other hand, strains that belong to phylogroup B1 and are isolated from extraintestinal infections are almost neglected, being considered as merely opportunist pathogens, and the majority of studies reporting strains from this phylogroup focus on antimicrobial resistance. We concentrated our research on one multidrug-resistant strain isolated from bloodstream infection that belongs to phylogenetic group B1 to enlarge the knowledge about the virulence of this kind of strain. We demonstrated that strain EC121 has full potential to be considered a pathogenic strain, because it adheres to and invades intestinal and bladder human cells and survives in human serum; moreover, our data highlight some important EC121 features, which are typical of ST101 strains, like its involvement in the spread of antimicrobial resistance genes, its relationship with extraintestinal infections from diverse sources, and its close relatedness to Shiga toxin-producing *E. coli*. All these data are important for the development of global actions concerning the spread of antimicrobial resistance, as well as to elucidate the virulence potential of strains that at first glance may be considered as a mere opportunist.

Introduction

Escherichia coli is one of the most frequent pathogens isolated from bloodstream infections (BSI) around the world [1–5]. Despite all knowledge of extraintestinal infections due to pathogenic *E. coli,* the number of severe infections and outbreaks caused by these pathogens is rising [2,6]. Moreover, many of these infections are caused by multidrug-resistant (MDR) strains, leading to higher burden of disease [7–10].

The term extraintestinal pathogenic *Escherichia coli* (ExPEC) is used to define strains recovered from any extraintestinal infection in humans or animals. Although many virulence factors are associated with the pathogenicity of this group, it is difficult to identify or classify ExPEC strains based on a specific group of virulence genes [11]. Some studies developed molecular virulence patterns that define strains that harbor intrinsic extraintestinal virulence potential [12] or are potentially capable of causing urinary tract infection [13]. These molecular patterns are useful tools to track ExPEC both in the gastrointestinal tract or environment (soil, water, food), enabling the search for ExPEC reservoirs.

Even though such methods may recognize the most virulent strains, they fail in identifying a considerable part of isolates recovered from clinical samples [13,14]. The reason is that infections take place as a result of an imbalance between the virulence potential of the pathogen and the immune defenses of the host what makes it sometimes unclear whether the infection is being caused by a true pathogen or by an opportunistic strain. Considering this, the use of epidemiologic data and multi-locus sequence typing (MLST) for the identification of strains belonging to major pathogenic clonal groups could help in the determination of the potential pathogenic role played by an *E. coli* strain [15–17].

The emergence of MDR *E. coli* strains calls attention to the spread of clones carrying virulence along with resistance-encoding genes, making the control of these pathogens potentially difficult [2,18–22]. In this context, ST131 is the MDR high-risk clonal group widely disseminated and studied worldwide nowadays. Other clonal groups, presenting MDR phenotype, like ST405, ST38, and ST648, have also emerged and are already considered as of global risk [23]. On the other hand, some STs presenting MDR phenotype, although being isolated around the world, have not their pathogenic potential determined yet.

The spread of MDR pathogens is a major Public Health concern that needs to be adequately addressed towards efficient control. Based on that, the World Health Organization (WHO) called attention to this problem and the need for alternative therapeutic options for treatment of MDR infections. The development of vaccines and anti-virulence compounds could be alternative approaches to combat MDR strains, especially those showing pan drug resistance (PDR) phenotype [24]. However, for these alternatives to be effective, advanced knowledge is necessary, since not all pathogenic strains share the same virulence factors, and the use of the most prevalent virulence factors as targets can be problematic as they can adversely affect the gut microbiota.

It is well accepted that the phylogenetic grouping of *E. coli* keeps a very good correlation with the virulence potential of bacterial isolates. More recent epidemiological data have shown that non-virulent strains are mostly classified in the phylogenetic groups A, while diarrheagenic *E. coli* (DEC) are B1, and ExPEC are mainly B2. However, the fact that most *E. coli* virulence factors are carried on mobile genetic elements (e.g. plasmids and pathogenicity islands) may eventually cross these phylogenetic boundaries and promote the appearance of potential pathogens in atypical phylogroups.

Thus, a global analysis of virulence and resistance characteristics of ExPEC isolates, especially those escaping typical classification, is essential for the understanding of the infections they cause as well as for devising therapeutic alternatives to cope with MDR *E. coli* strains.

Aiming to provide information about the virulence determinants in non-typical ExPEC strains, we performed an extensive genotypic and phenotypic characterization of the MDR *E. coli* EC121 strain, which was characterized as belonging to phylogenetic group B1 and carrying only a few known virulence factors.

Results

Genetic characterization and *in silico* analysis.

The EC121 genome sequence generated 143 contigs. PlasFlow algorithm identified that 66 of those were contained in plasmids while 43 were chromosomal. The predicted genome size was 5,119,556 bp, with 50.34 % of GC content. It was composed by one chromosome and four plasmids that, together, would comprise at least 426,419 bp

EC121 belongs to serotype O154:H25, ST101-B1, and is related to diarrheagenic *E. coli*. The MLST analyses showed that the EC121 strain belonged to the ST101/ST88, according to the Warwick and Pasteur MLST schemes, respectively. The determination of O antigen type by sequencing analysis was inconclusive, because two possible O types, O154 and O100, could be assigned. Using specific serum agglutination assays it was determined that strain EC121 expressed the O154 antigen. Together with the *in silico* analysis of the *fliC* gene, which identified the H25 type flagella, EC121 was characterized as belonging to serotype O154:H25. Moreover, group IV capsule encoding genes were identified on its genome.

A phylogenetic tree built using reference *E. coli* strains from all pathotypes showed that EC121 was related to diarrheagenic *E. coli* (DEC) strains, since it was positioned in a clade closely associated with Shiga toxin-producing *E. coli* (STEC) (Fig 1).

A second phylogenetic tree was built with 93 strains available at the NCBI, which belonged to the ST101 complex and were recovered from distinct sources (Fig 2, Table S1). Analysis of the tree showed that the majority of the strains of this ST are MDR, many of them carrying *mcr*-1 (mobile colistin resistance gene), a variety of β-lactamases (*bla*$_{\text{CTX-M}}$-like, *bla*$_{\text{OXA}}$-like, *bla*$_{\text{NDM}}$-like) and genes related to fosfomycin resistance

(*fosA3*). Interestingly, these strains were isolated from food, environment, animals, and humans, as part of the microbiota or involved in both intestinal and extraintestinal infections. Among the strains isolated from human infections (Fig S1 and Table S1), most were diagnosed as extraintestinal pathogens (28 strains), while three were intestinal pathogens; among the latter, one was Shiga toxin-producing *E. coli* (STEC), and one was enterotoxigenic *E. coli* (ETEC). Regarding the isolates from food and animals, seven strains were identified as STEC (Fig 2, Fig S1 and Table S1). Showing that ST101 is associated with intestinal and extraintestinal infection and it is clonal group associated with MDR phenotype

EC121 harbors genes involved in virulence and stress response. The EC121 genome annotation showed that strain EC121 contained 5,175 coding sequences (CDS), 82 tRNA, and 13 rRNA. One CRISPR *locus* was identified as type 1-IE and presented two arrays and 30 CRISPR-repeat regions (Fig 2A). Among the CDSs annotated, 702 corresponded to putative proteins designated as hypothetical proteins, and 4,473 CDS to putative proteins with functional assignments. Of interest, 221 genes were reported as belonging to systems involved in response to stress, virulence, and defense (Table 1 and Figs 2B-C).

173 **Table 1 -** EC121 annotation overview.[a]

Protein features	Occurrence
Hypothetical proteins	702
Proteins with functional assignments	4,473
Proteins with E.C.[b] number assignments	1,300
Proteins with G.O.[b] assignments	1,076
Proteins with Pathway assignments	911
Proteins with PLfam[b] assignments	5,063
Proteins with PGfam[b] assignments	5,064

174 *a*. Results obtained using PATRIC annotation service;

175 *b*. Abbreviations: E.C. number - Enzyme commission universal G.O. - Gene Ontology
176 Consortium; PLfam- PATRIC genus-specific family; PGfam- PATRIC cross-genus
177 family.

178

179 Analysis using the MacSyFinder tool identified 10 type V secretion system
180 proteins, nine of which were from type T5aSS and one from T5cSS (EhaG) (Table S2).
181 One type III secretion system similar to *Salmonella* T3SS, and an incomplete type VI
182 secretion system that carried only *tssB*, *tssD*, *tssE*, *tssH*, and *tssI* genes were also present
183 (Table S2).

184 All genes that were reported by the PATRIC virulence factor database were
185 manually curated to provide information about their full sequence. As shown in Table 2
186 and Table S3 and S4, the genome of the EC121 strain encodes multiple adhesins, invasins,
187 iron uptake systems, and genes involved with evasion of the host immune system. By the
188 position of strain EC121 in the pathotype phylogenetic tree, some of these virulence
189 factors are related to the pathogenesis of DEC, specifically of STEC and ETEC (Tables

2 and S3). Other virulence factors found were related to *Salmonella* spp. (PagN adhesin and systems associated with immune evasion and macrophage survival) and *Shigella* spp. (genes associated with intracellular survival and spread). Moreover, many accessory genetic clusters associated with the bacterial ability to cause extraintestinal infections, i.e., genes involved with biofilm formation, adherence to extraintestinal cells, iron acquisition and immune evasion were also detected in the EC121 genome. Furthermore, other clusters associated with Urinary Tract Infections (UTI) were detected (Tables 2 and S3).

200 Table 2 – Complete virulence factors identified

Virulence traits	Associated with virulence in			
	Other Genera	DEC	ExPEC	Various
Colonization and invasion	MlaA, MsbB, Sfm, MisL, IcsP	Lpf-1$_{O26}$, EhaG, Hcp, Elf, EhaA, EhaB, CFA-I	FdeC, Pix, Ygi, Yad, Yeh, Yra, Yfc, YchO, IbeB, IbeC, EptC, OmpA	Type 1 Fimbriae (H190), Ecp, Curli
Immune evasion	SodB, TrxA, SirA, FpkA		Iss, OmpTp[a], OmpTc[a], Mig-14, HlyF	RelA
Iron acquisition				Sit, Iro, Fhu, Enterobactin
Regulators	DsbA, DegP, SlyA, CpxAR	EvgAS	DsbAB, PhoPQ	QseBC, RcsAB
Toxins and Bacteriocin			Microcin V, Colicin B, Colicin M	ClyA, Hly III
Non-LEE effectors		EspL1, EspL4, EspX1, EspX4, EspX5, EspR1		

201 [a] OmpTc for chromosomal variant of OmpT protein, and OmpTp for plasmid variant of OmpT protein.

202

203

204

205

206

No complete phage sequences were detected in the EC121 strain. The search for phage sequences in strain EC121 identified 12 regions containing genes from a variety of different phages, ranging from 6 to 31 kb (Table 3, Table S4). Although the database considered a predicted phage sequence in region 6 as intact, based on their score criteria, it´s size (14,400 bp) was not compatible with the size of the predicted "*Salmonella* phage Fels-2", whose complete genome sequence deposited in NCBI database is 33,693 bp. Remarkably, parts of the cytolethal distending toxin (Cdt-I and Cdt-V), and Shiga-like toxin (Stx1a and Stx2c) converting phages, as well as of *Shigella* serotype-converting phages SfI, SfII, and SfV were detected among these regions (Table 3 and Table S4).

Table 3 - Predicted phages detected in the EC121 genome.

Region	Length	Completeness	Most Common Phage/accession number	GC %
1	6.3 kb	incomplete	*Bacillus* phage G/NC_023719	51.27
2	30.9 kb	incomplete	*Salmonella* phage 118970_sal3/NC_031940	49.82
3	14.2 kb	questionable	Enterobacteria phage P88/NC_026014	49.75
4	20 kb	incomplete	*Shigella* phage POCJ13/NC_025434	47.95
5	28.5 kb	incomplete	*Shigella* phage SfII/NC_021857	45.37
6	14.4 kb	intact	Enterobacteria phage Fels-2/NC_010463	49.56
7	9.4 kb	incomplete	*Bacillus* phage Shanette/NC_028983	49.50
8	7.5 kb	incomplete	Enterobacteria phage phi92/NC_023693	44.44
9	10.2 kb	questionable	Enterobacteria phage P2/NC_001895	55.39
10	3.8 kb	incomplete	Bacteriophage WPhi/NC_005056	50.16
11	8.4 kb	questionable	Phage cdtI/NC_009514	46.05
12	6.7 kb	incomplete	*Shewanella* sp. phage¼/NC_025436	47.92

a. Phage completeness was determined by scores obtained in the search algorithm (PHASTER) based on the number of the phage's specifics CDSs detected in the analyzed region.

EC121 has multiple resistance genes, and several efflux-pumps compatible with its antimicrobial susceptibility profile. The software ResFinder identified 15 resistance genes in strain EC121, which are involved in reduced susceptibility to aminoglycosides, β-lactams, macrolides, phenicols, sulphonamides, trimethoprim, and tetracyclines (Table 4). Mutations in *parE*, *parC*, and in *gyrA* genes, which confer resistance to fluoroquinolones, were also observed in the EC121 genome. Additionally, several efflux-pumps related to resistance to heavy metals (copper and mercury), arsenic, disinfectants

(QacE), and antimicrobials (AcrAB-TolC, AcrAD-TolC, AcrEF-TolC, AcrZ, EmrAB-TolC, EmrD, EmrKY-TolC, MacA, MacB, MdfA/Cmr, MdtABC-TolC, MdtEF-TolC, MdtL, MdtM, and SugE) were also detected. The MDR phenotype profile of strain EC121 assessed by the microdilution method was consistent with the genomic findings (Table 4).

Table 4 -Antimicrobial resistance genotype and phenotype observed in strain EC121.

Antimicrobial Class	**Genotype**	**Phenotype**	
		Antimicrobial agent	**MIC[b](µg /mL)**
Aminoglycosides	*aph(3'')-Ib*, *aph(6)-Id*, *aph(4)-Ia*, *aac(3)-Iva*	Amikacin	2
		Gentamicin	**32**
β-lactams	bla_{TEM-1B}, $bla_{CTX-M-2}$	Ampicillin	**≥256**
		Ceftazidime	**64**
		Ceftriaxone	**≥512**
		Cefepime	**16**
		Piperacillin/Tazobactam	**>256/4**
		Meropenem	≤0.5
		Ertapenem	≤0.5
		Imipenem	≤0.5
		Aztreonam	**>32**
Sulfonamides/ Trimethoprim	*sul2*, *sul1*, *sul1'*, *dfrA14*, *dfrA7*	Trimethoprim/ Sulfamethoxazole	**>128/ 2,432**
Phenicols	*catA1*	Chloramphenicol	**>64**
Tetracyclines	*tet(A)*	Minocycline	**16**
		Tigecycline	1
Fluoroquinolones	*parE*(S_{458}A); *parC* (S_{80}I); *gyrA* (S_{83}L and D_{87}Y)	Ciprofloxacin	**>64**
Polymyxins	-	Colistin	≤0.25
		Polymyxin B	≤0.25
Fosfomycin	-	Fosfomycin	NT
Macrolides	*mph(A)*, *mdf(A)*	Azithromycin	NT

a. chromosomal mutations and mobile genes related to antimicrobial resistance identified;

b. Resistance is highlighted in bold, following EUCAST (2019) breakpoints.

NT, not tested.

The EC121 strain harbors four circular plasmids. Four plasmid bands were detected by agarose gel electrophoresis, three of which presented high molecular weight masses (approximately 67 kb, 110 kb, and 180 kb), and one was smaller than 6.9 kb (Fig 4). Seven different replicons were found in EC121 using PlasmidFinder 2.0 (IncHI2A, IncHI2, IncQ, IncFII, IncFIB, IncN, and IncL/M). This information was associated with PlasFlow and Blast analyses to manually determine plasmid sequences.

The pEC121.A, which was the largest plasmid detected by gel electrophoresis consists of a 189,202 bp plasmid carrying three basic replicons belonging to incompatibility groups IncHI2A, IncHI2. This plasmid shows high identity with the pYps.F1 plasmid from *Yersinia pseudotuberculosis* strain Yps.F1 (cover: 99%, e-value: 0.0, identity: 99.77%) (Fig 4 and Fig S2). More important, plasmid pEC121.A carries the antimicrobial resistance genes *sul2, tet(A),* $bla_{\text{TEM-1B}}$, *aph*-3'-*Ib*, and *aph-6-Id* genes.

Plasmid pEC121.B was an IncFIB/IncFII plasmid with about 104,053 bp that is similar to pAPEC plasmids that carry virulence and resistance genes simultaneously (Fig S3). Such plasmid carries the beta-lactamase encoding gene $bla_{\text{CTX-M-2,}}$ mercury and copper resistance-encoding genes, and a class 1 integron, as well as the *iro, sit, tra,* and *hlyF mig-14* operons; and *ompT*. The resistance genes were inserted close to a Tn21 transposon.

Plasmid pEC121.C is an IncL/M plasmid of 65,565 bp with high homology (100% of identity and coverage) to the pASM2 plasmid (accession nº NZ_CP019841.1) of *Enterobacter roggenkampii* strain R11 (Fig S4). Finally, the smallest plasmid, pEC121.D, carries an IncN replicon and is 3,205 bp in size.

Virulence phenotype.

263 To evaluate the expression of the virulence-encoding genes detected, *in vitro*
264 assays were performed to analyze the ability of the strain to: (i) resist to the bactericidal
265 activity of the serum complement system; (ii) attach to abiotic surfaces and form biofilms;
266 (iii) adhere to and invade eukaryotic cells.
267
268 **The EC121 strain was resistant to the bactericidal activity of the complement system**
269 **and forms biofilm on abiotic surfaces.** To disseminate in the host, extraintestinal
270 pathogenic bacteria must be able to survive the serum bactericidal activity. To identify
271 such a feature in EC121, we determined the lowest serum-resistant bacterial inoculum
272 using a pool of normal human sera (NHS). The lowest inoculum of EC121 strain that
273 resisted serum activity after two hours was 10^2 CFU/mL, which was similar to the one
274 obtained for the resistant control strain J96. *E. coli* strain C600, used as the susceptible
275 control, barely resisted to 30 min-exposition period in the highest inoculum tested (10^8
276 CFU/mL). To validate if the bacterial survival was associated with the resistance to
277 complement activity, assays were repeated with heat-inactivated serum. In this condition,
278 all strains survived the challenge with similar inoculum (Table 5).
279

Table 5 -Estimated bacterial inoculum resistant to serum activity (CFU/mL).[a]

	Strains					
	J96		EC121		C600	
Challenge period	**NHS**	**inHS**	**NHS**	**inHS**	**NHS**	**inHS**
30 min	10^1	10^1	10^2	10^1	10^8	10^1
1 h	10^2	10^1	10^2	10^1	NG	10^1
2 h	10^2	10^1	10^2	10^1	NG	10^1

a. Values represent the approximate relative mean of the lowest bacterial inoculum that remained viable after the challenge. All assays were performed in triplicate using 50% serum diluted in PBS (v/v);

NHS - Normal human serum; inHS - inactivated human serum; NG - no growth after the challenge.

The ability to form biofilm can confer many advantages to any pathogen, including persistence in particular niches and tolerance against antimicrobials and the host immune system. The EC121 strain was able to adhere to and form biofilms on a glass surface when grown in Dulbecco's Modified Eagle Medium (DMEM), as shown in Fig 5A. Although its adherence to the abiotic surface was not massive as the adherence presented by the positive control (EAEC 042 strain), it was significantly more intense than that of the negative control (HB101 strain), as shown by results obtained using the crystal violet absorption index (Fig 5B).

The EC121 strain adheres to and invades eukaryotic cell lineages. At first, we accessed the ability of strain EC121 to adhere to HeLa cells using a classic 3-h adherence assay, on which it was able to adhere to, both in the presence or absence of D-mannose,

299 which abolishes the adherence mediated by type-1 fimbriae. The presence of D-mannose
300 reduced the adherence ability of strain EC121 in more than 90% (Fig 6), from $2.1x10^8$
301 CFU/mL to $1.4x10^7$ CFU/mL. However, in such condition, the mean adherence index in
302 the presence of D-mannose was similar to the diffusely adherent *E coli* C1845 strain
303 ($1.4x10^7$ CFU/mL), used as a positive control (Fig 6).

304 We also investigated whether the EC121 strain was able to invade HeLa cells. In
305 a short-period invasion assay, the EC121 strain interacted with HeLa cells and invaded
306 with higher efficiency than reference invasive control *Shigella flexneri* strain M90T.
307 Indeed, the total number of internalized EC121 bacteria (mean of $7.11x10^4$ CFU/mL) was
308 22-fold higher than the number of internalized M90T cells (Figs 7 and S5). On the other
309 hand, the proportion of invasion/interaction of strain EC121, determined as the invasion
310 index, was 2-fold lower than strain M90T (0.17% and 0.40%, respectively). The
311 difference between the invasion indexes of M90T and EC121 might be due to the large
312 difference in the adherence efficiency between them; the interaction of EC121 with HeLa
313 cells was 50-fold higher than that of the M90T strain (Fig 7). Invasiveness in the presence
314 of D-mannose also was assessed to evaluate whether the invasion ability of EC121 was
315 dependent on type-1 fimbriae or another mannose-dependent adhesin. EC121 remained
316 invasive in the presence of D-mannose but with reduced bacterial counts (Fig 7 and S5),
317 thus showing that type-1 fimbriae contributed to both adherence and invasion of EC121,
318 although it was not the only factor associated to these traits.

319 The capacity to interact (adhere to and invade) with differentiated intestinal Caco-
320 2 cells and bladder T24 cells was also assessed. The EC121 strain also adhered to and
321 invaded both cell lineages. Even though the adherence efficiency was similar, the
322 invasion index was lower in Caco-2 cells than in HeLa cells (Fig 7 and S5). On the other
323 hand, in T24 cells, these parameters were similar to those found in HeLa cells. A classic

qualitative adherence assay revealed that the EC121 strain interacted abundantly with all tested cell lines, avoiding the coverslip surface even in the absence of D-mannose (Fig S6).

Discussion

The *E. coli* EC121 strain was isolated in 2007 from a bloodstream infection of an inpatient that presented persistent gastroenteritis and T-zone lymphoma. Since initial analyses showed that it belonged to phylogenetic group B1 and carried few virulence markers commonly related to extraintestinal pathogenic *E. coli,* it was classified as ExPEC negative (ExPEC-) [25] and therefore considered as an opportunistic pathogen. However, considering that about 40% of extraintestinal infections are caused by strains devoid of virulence factors [13,14] and that EC121 was an MDR strain, its entire genome was sequenced to further understand its virulence potential.

Interestingly, the EC121 strain belongs to ST101, which has been previously reported to be involved in nosocomial outbreaks caused by Metallo-β-lactamases-producing strains in many countries from Europe, Asia, and Oceania [26–30]. Furthermore, ST101 has also been detected among strains of non-outbreak related extraintestinal infections [31–38], water [39], poultry infection [40], retail food [31,41–43], and healthy human and animal intestinal microbiota [38,44–47], mostly presenting an MDR phenotype. Shrestha et al. [48] drew attention to ST101 due to the PDR phenotype presented by some strains of this ST, and mainly because it is not considered a pandemic clone, although it has been isolated worldwide. We, therefore, analyzed data about the infection type, isolation source, and resistance genetic markers presented by the strains of the ST101 complex that were previously deposited in the NCBI (Table S1). Such analysis evidenced that MDR strains of this complex were spread worldwide. In

addition, such ST complex is related to Shiga toxin-producing strains as well as strains isolated from extraintestinal infections, human and animal microbiota, retail food, and environment. Moreover, many strains simultaneously carry the $bla_{\text{CTX-M-55}}$, *mcr-1*, *fosA3*, and *qnrS1* genes. Interestingly, one strain present simultaneously *aac(6')-Ib-cr*, $bla_{\text{CTX-M-55}}$, $bla_{\text{NDM-5}}$, $bla_{\text{OXA-1}}$, *mcr-1*, *fosA3*, and other 19 resistance genes.

Likewise, the EC121 strain showed multiple antimicrobial resistance genes, including genes that confer resistance to third-generation cephalosporins ($bla_{\text{CTX-M-2}}$). It is worth to mention that the EC121 strain was isolated in 2007, and its MDR phenotype was relevant since at that period it was susceptible only to carbapenems, polymyxins, and amikacin. Recently, the *E. coli* strain ICBEC72H, which belongs to ST101 and carried only $bla_{\text{CTX-M-8}}$ and *mrc-1*[32] was isolated from a human extraintestinal infection in Brazil. Similarly, the ST101 *E. coli* strain 200H (Table S1) was isolated from a human urinary tract infection and carried $bla_{\text{OXA-9}}$, *mcr-1*, and *aac(6')-Ib-cr*. These reports show that MDR strains belonging to the ST101 complex have been circulating in Brazil for a long time.

Some authors showed that the *E. coli* strain 912 (ST101) was selected by the usage of antimicrobial agents in animals and that it was able to colonize human and pig gut and spread through the environment, reaching and colonizing animals that were not under antimicrobial treatment [49,50]. These same authors have also shown that ST101 strains can naturally acquire and transfer plasmid-borne antimicrobial resistance genes in the gut [49,50]. Such a feature is important for various reasons. First, strain EC121 carried three large plasmids, two of which (pEC121.A and pEC121.B) harbored different antimicrobial resistance genes. Besides, *E. coli* strains belonging to ST101 were recovered from retail meat in Europe and Asia [42,43], and from extraintestinal infections in Brazil and USA in the same regions in which they were detected from retail meat [31,40].

Additionally, strains belonging to the ST101 complex carrying multiple resistance genes were recovered from the intestine of healthy humans and animals in many countries. Therefore, even if these strains do not cause infection directly, they could potentially transfer plasmids to other bacteria, even from distinct genera. Such cross genera plasmid transfer could be easily identified in the plasmids reported in the present study; pEC121.C, for example, is closely related to plasmids found in *Klebsiella* spp. and *Enterobacter* spp., while pEC121.A is related to *Salmonella* spp. and *Yersinia* spp. plasmids. Together, these findings reinforce the high risks associated with strains belonging to the ST101complex due to their ability to colonize humans and animals' gut, to easily disseminate via retail food and water, being able to acquire and spread antimicrobial resistance-encoding genes.

Strains from the ST101complex are included in the phylogenetic group B1, which implies that they do not have all the classical virulence factors that are usually associated with the most virulent ExPEC strains [51,52]. Many studies reported phylogroup B1 *E. coli* strains as commensals or as intestinal pathogens, but not as extraintestinal pathogens [51–53].

The genomic analysis of the EC121 strain showed a high number of virulence genes, demonstrating that it presents all the traits necessary to be considered as an extraintestinal pathogenic agent. Moreover, like other ExPEC strains, EC121 displayed multiple virulence genes related to each feature, reflecting the redundant phenotype that ensures its pathogenicity. However, even considering the completeness of each sequence and each operon, which was manually checked, the presence of virulence genes "per se" does not guarantee that all of them are expressed. Therefore, to evaluate the expression of such genes, distinct phenotypic assays were performed and confirmed the virulence genetic background of EC121.

399

400 To test the bacterial ability to resist the serum complement activity, a two hours
401 challenge assay was performed, in which one particle can traverse all circulatory system
402 at least twice, so a pathogen that resists complement's activity during this period, even
403 with a small bacterial load, is, in theory, more capable of reaching different niches and
404 spread through the bloodstream or cause a bloodstream infection. The EC121 strain
405 resisted the NHS for two hours with an inoculum similar to the resistant *E. coli* J96 control
406 strain, thus confirming the EC121 serum resistant phenotype. Serum complement is the
407 first immunological barrier to control pathogens that reach the bloodstream. Complement
408 resistance confers the possibility to spread to different body sites through the
409 bloodstream. Hallström et al. [54] reported the relationship of bacterial resistance to NHS
410 with sepsis severity, and other authors have associated it to different kinds of
411 extraintestinal infections [55–57].

412 The ability to colonize and attach to surfaces is also an important trait for any
413 pathogenic bacteria; in this way, the assays carried out showed not only that EC121 strain
414 was able to adhere to and invade different cell lineages, including bladder T24 cells but
415 that it could also attach and produce biofilm on abiotic surfaces. Peirano et al. [58] showed
416 that ExPEC negative ST101 MDR strains isolated from extraintestinal infections could
417 interact with HEp-2 and Caco-2 cells more efficiently than strains belonging to the
418 epidemic clones ST131 and ST405, which are ExPEC positive [58]. Moreover, the
419 capacity to produce biofilm could confer to EC121 strain protection against the
420 immunological system and antibiotics, assisting its persistence and spread in the
421 environment.

422 Interestingly, many of the EC121 virulence factors detected in the draft genome
423 are related to diarrheagenic E. coli, even though none of them is implicated in DEC

pathotype definition. The presence of many genetic features related to Shiga-toxin producing *E. coli* (STEC) strains, *e.g.*, Hcp, EhaG, and Lpf-1$_{O26}$, as well as the proximity of EC121 to the clade that contains STEC strains and *E. coli* O104:H4 strain 2011C-3493, draws attention to its potential to cause diarrhea.

Many features identified in the EC121 genome reinforce its linkage with STEC strains. Phenotypically, EC121 expressed the O154:H25 serotype, but it possesses the group IV capsule-encoding genes. This kind of capsular group is known to be thermoresistant and expressed as K_{LPS} or O-antigen capsule. This could explain the expression of the O154 instead of O100 antigen, despite of the presence of all genes related to the expression of the latter. Interestingly O100 is a STEC related serogroup.

Moreover, many of the phage remains detected in the EC121 strain were related to Stx-converting phages; besides, ST101 strains carrying the *stx*$_{1a}$ gene have been reported in food sources [41,59]. In humans, ST101 strains were already reported in a patient with Hemolytic Uremic Syndrome (HUSEC) [59–61] and in non-bloody diarrhea related to a Stx1a-producing *E. coli* strain [62]. Although only one of these strains had its genome sequenced, some ST101 *E. coli* strains recovered from animals and food were found to carry *stx1*. Interestingly, most Stx-converting phages remains found in EC121 were similar to those commonly related to Stx1a production, corroborating with the results presented here. The genome of three non-STEC strains from diarrheic patients was found in GenBank, one of which was devoid of DEC virulence factors. Likewise, EC121 was isolated from bloodstream infection of one inpatient with persistent infectious gastroenteritis which was probably the source of EC121 infection. Unfortunately, the *E. coli* isolated from stool was not stored not allowing further comparison between bloodstream and stool *E. coli* isolates.

In summary, our extensive *in silico* and *in vitro* analyses of virulence and resistance properties of *E. coli* strain EC121, an O154:H25 B1-ST101 strain isolated from a human bloodstream infection, confirmed its virulence potential in both intestinal and extraintestinal infections and increased the knowledge on the complex scenario of virulence traits present in the MDR *E. coli* ExPEC negative group, contributing to the potential development of strategies to control the spread of such pathogens.

454 # Material and Methods

455 **Bacterial strain.**

456 The *E. coli* strain EC121 was isolated from the blood of a patient diagnosed with
457 T-zone lymphoma and persistent infectious gastroenteritis, who had been hospitalized in
458 a tertiary hospital located in the city of São Paulo, Brazil, in 2007. The patient died due
459 to septic shock two days after isolation of the EC121. The EC121 strain was kept frozen
460 in glycerol at -80ºC in ENTEROBACTERIALES-EXTRAINTESTINAL-EPM-DMIP
461 collection nº A27A7C3. The initial virulence and resistance characterization showed that
462 EC121 strain belonged to phylogroup B1, presented an MDR phenotype by routine
463 susceptibility testing, and harbored few known virulence genes (*fim, hra, cvaC, ompA,*
464 *ompT, sitA, iroN*). Furthermore, it was not considered a pathogenic strain because it
465 harbored none of the virulence factors commonly involved in the characterization of
466 ExPEC (presence of two of the following genes: *papA/C*, *sfaDE*, *afaBC*,*iuc/iut*,*kpsMT*II)
467 [25].

468

469 **Total DNA extraction, whole-genome sequencing (WGS), and genome assembly.**

470 The total bacterial DNA extraction was done using Wizard® Genomic DNA
471 Purification Kit (Promega - USA) following the manufacturer's protocol. The extracted
472 DNA was sequenced in an Illumina® Hiseq1500 (Illumina-USA), using the Rapid
473 protocol to obtain 2x250 paired-end reads, according to the manufacturer´s
474 recommendations. Raw data were processed with Trimmomatic, and then the paired-end
475 reads were assembled using SPAdes (version 3.12.0), with default parameters, and careful
476 mode on [63].

477

478 **Genomic analyses and annotation.**

The obtained draft genome was submitted to various online bioinformatics platforms of the Center of Genomic Epidemiology (CGE) pipeline to determine (i) sequence types [64] for both *E. coli* MLST schemes;(ii) serotype (SerotypeFinder 2.0)[65];(iii) presence and types of plasmid replicons (PlasmidFinder 2.0) [66];(iv) presence of resistance genes (ResFinder 3.1)[67];and (v) STEC virulence factors (VirulenceFinder 2.0) [68]. PHASTER [69] and PHAST [70] were used to detect bacteriophage sequences in each contig of the draft genome.

The genome was annotated using Pathosystems Resource Integration Center (PATRIC) Comprehensive Genome Analysis service that uses RASTtk[71]. Each sequence that was assigned as a virulence factor in PATRIC's database was manually submitted to BLAST/NCBI [72] and UniProt [73] to validate the virulence factors, to get all information about the virulence genes detected, to evaluate the completeness of the sequence and to determine its homology in relation to the RefSeq protein in Swiss-Prot. PATRIC [71] service was also used to build a phylogenetic tree using RAxML-VI-HPC or Fast tree 2, where all representative *E. coli* genomes from different pathotypes were used to construct the tree, as well as the deposited genomes of *E. coli* strains belonging to ST101 complex from diverse sources. The tree was built based on the concatenated sequence of all shared proteins among all genomes using RAxML or FastTree2. To construct the phylogenetic tree of EC121 and representative *E. coli* pathotypes, two *Escherichia fergusonni* strains ATCC35469 and NCTC12128 were used as outgroups. To build ST101 complex phylogenetic trees*, E. fergusonni* strains ATCC35469, *E. coli* str IAI1, *E. coli* O157:H7 str Sakai, *E. coli* O104:H4 str 2011c-3493 were used as outgroup. All phylogenetic tree final layout and annotation were done using iTOL v.4 [74]. The annotated genome was submitted to MacSyFinder from Galaxy@Pasteur [75] to detect CAS-CRISPR sequence type and the presence of secretion systems [76,77].

504

505 **Serum agglutination assay for typing the O and H antigens.**

506 Serum agglutination assay was carried out following the standard methodology as
507 described by Orskov and Orskov [78] for serotyping, using O serum against O100 and
508 O154, and H serum against H25 provided by the Centers for Disease Control and
509 Prevention (CDC, USA).

510

511 **Plasmid DNA extraction and analysis.**

512 Bacteria were cultivated in Tryptic soy broth (TSB - Difco, USA) at 37ºC, in a
513 static stove for approximately 18 h, and 1 ml of the culture was submitted to plasmid
514 alkaline extraction protocol [79]. The *E. coli* strain 39R861 was used as a plasmid mass
515 reference ladder and as control of extraction [80]. The plasmid extract was submitted to
516 electrophoresis in an agarose gel (0.8%) in Tris-Borate-EDTA (TBE) buffer, stained with
517 ethidium bromide solution (5 µg/mL), analyzed using Molecular Imager®Gel Doc™ XR+
518 with Image Lab™ Software System from Bio-Rad (USA).

519

520 ***In silico* plasmid analysis.**

521 To accomplish the plasmid analysis, the following strategies were used. First, the
522 draft genome was submitted to CGE to identify the contigs that contained replicons;
523 subsequently, the assembled genome was analyzed by PlasFlow [81] to classify the
524 possible source of each contig (as chromosomal or plasmid). The contigs containing
525 replicons were analyzed using the Standard Nucleotide BLAST in NCBI, all the retrieves
526 that had returned an e-value of 0.0 and with more than 90% of identity and coverage were
527 BLASTed against EC121 draft genome to find the best complete sequenced plasmid
528 related to those found in EC121. When the sum of the contigs identified was consistent

with plasmid predicted size determined in the agarose gels electrophoresis, the contigs were marked and signed as belonging to the same plasmid. In parallel, using Bowtie2 version2.2.6, raw reads were mapped to the best complete plasmid that better matched with the EC121 draft to validate the analyses.

Data availability.

The EC121 Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession VYQD00000000. The version described in this paper is version VYQD01000000.

Determination of the lowest bacterial inoculum which was resistant to human serum complement.

To access the bacterial serum-resistance, the lowest bacterial inoculum resistant to serum was assessed. Lyophilized human complement serum (Sigma, USA) was reconstituted in sterile phosphate-buffered saline (PBS). The assay was performed in 96-wells plates, where complement serum was distributed in each well (90 uL per well). Bacteria were grown overnight at 37°C, serially diluted (1:10) in complement serum until 10^{-10} and incubated at 37°C. Aliquots of 10 µL of each well were seeded onto MacConkey agar plates after 30 min, 1 h and 2 h of incubation. Simultaneously, another assay was performed with previously heat-inactivated serum as control. The *E. coli* strains J96 and C600 were used as resistant and susceptible controls, respectively [82]. The lowest bacterial inoculum resistant to human complement was determined by the last bacterial dilution which had bacterial growth onto MacConkey after the challenge. For each assay, the initial bacterial inoculum was determined by diluting bacteria in PBS, plating in

553 MacConkey agar and CFU counting. The data was reported in CFU/mL. Biological
554 assays were performed in triplicates.
555
556 **Antimicrobial susceptibility testing.**
557 The minimum inhibitory concentration (MIC) was determined using the broth
558 microdilution method, following the European Committee on Antimicrobial
559 Susceptibility Testing (EUCAST) recommendations and breakpoints [83]. The following
560 antimicrobials (Sigma - USA) were tested: ampicillin, piperacillin/tazobactam,
561 ceftriaxone, ceftazidime, cefepime, aztreonam, ertapenem, imipenem, meropenem,
562 ciprofloxacin, amikacin, gentamicin, tigecycline, colistin, polymyxin B,
563 trimethoprim/sulfamethoxazole, and chloramphenicol. *E. coli* ATCC 25922 and
564 *Pseudomonas aeruginosa* ATCC 27853 were used as quality control strains.
565
566 **Biofilm formation on abiotic surfaces.**
567 Biofilm formation was evaluated on polystyrene and glass surfaces as described
568 by Lima et al [84] in an 24 hours assay. Each assay was performed in biological and
569 experimental triplicates. The *E. coli* strain EAEC 042 and laboratory *E. coli* strain HB101
570 were used as positive and negative controls, respectively; in all assays, a non-inoculated
571 well was used as control of dye retention.
572
573 **Cell culture and maintenance.**
574 HeLa (ATCC® CCL-2™), intestinal Caco-2 (ATCC® HTB-37™) and bladder T24
575 (ATCC® HTB-4™) cell lineages were used to evaluate the ability of strain EC121 to
576 interact with eukaryotic cells. HeLa and Caco-2 cells were cultured in DMEM, High
577 Glucose, GlutaMax™ (Gibco- ThermoFisher Scientific, USA), supplemented with 10%

bovine fetal serum (BFS) (Gibco, USA), 1% non-essential amino acids (Gibco, USA), and 1x Penicillin-Streptomycin-Neomycin (PSN) antibiotic mixture (Gibco, USA), while T24 cells (ATCC HTB-4) were cultured in McCoy 5A (modified) media (Gibco, USA), supplemented with 10% of BFS and 1x PSN antibiotic mixture. All lineages were kept at 37°C in an atmosphere of 5% CO_2. For all assays, cell suspensions containing 10^5 cells/mL were seeded in 24-well plates, with or without glass coverslips for qualitative or quantitative assays, and cultured for 48 h (HeLa cells) or ten days (Caco-2 cells).

Adherence assay in HeLa, Caco-2, and T24 cells.

Adherence, fixing, and staining procedures were performed as described by Garcia et al. [85]. All cell lineages were washed three times with PBS and 1 mL of proper media, supplemented with 2% BFS and 2% D-mannose, was added. The assay was also performed without D-mannose to assess the impact of abolishment of the mannose sensitive-adherence via Type 1-fimbria. To evaluate the efficiency of adherence to the cells, the assay was performed as described previously [85], except that after 3 h of incubation the epithelial cells were washed thrice with sterile PBS, lysed with 1 mL of sterile bi-distilled water for 1 h, collected, diluted and plated onto MacConkey agar for quantification. The assays were performed in biological and experimental triplicates and the data were expressed as SEM. The *E. coli* C1845 and laboratory *E. coli* MA3456 strains were used as adherent and non-adherent control strains, respectively.

Short period invasion assay in HeLa, Caco-2, and T24 cells.

The invasion assays were carried out as described by Martinez et al. [86], with modifications, in two sets of plates simultaneously. Aliquots of 20 µL or 40 µL of overnight bacterial cultures were inoculated in each well-containing epithelial cells, as

described above, and were incubated for 2 h, at 37°C, in a normal atmosphere. After this period, one plate set was washed three times with PBS and incubated again with PBS containing 100 µg/mL of amikacin for 1 hour at 37°C, to kill all extracellular bacteria. After the incubation period, the assay was washed three times to remove all antibiotics, cells were lysed with water, and the well contents were collected, diluted and plated onto MacConkey agar to obtain the number of internalized bacteria. The other set was washed with PBS three times, the cells were lysed, and contents of each well were collected, diluted and plated to obtain the total number of bacteria interacting with the cells in the period. An aliquot of the PBS recovered from the last wash after incubation with amikacin was collected and plated without dilution, to ensure that the treatment had killed all extracellular bacteria. The initial inoculum was determined for each strain by serial dilutions, plating, and counting the colony-forming units per milliliter (CFU/mL) before starting the assay. The invasion index was determined by the ratio between the number of internalized bacteria and the total number of interacting bacteria multiplied per 100. The *Shigella flexneri* strain M90T was used as an invasive control [87] and the *Escherichia albertii* strain 1551::*eae* was used as adherent and non-invasive control [88]. The assays were performed in biological and experimental triplicates, and data were reported as SEM.

Acknowledgments

We are grateful to the Laboratório Especial de Microbiologia Clínica (LEMC) of the Federal University of São Paulo (UNIFESP), São Paulo - SP, Brazil, for providing the *E. coli* strain EC121.

Financial support

This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) that provided research grants to E.C. (Process Number: 16/01656-5), R.M.S. (Process Number: 2009/00402-6), T.A.T.G. (Process Number: 2018/17353-7), and master scholarship to T.B.V. (Process Number: 2017/21947-7). We are also grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for providing grants under finance code 001 to A.C.M.S. (PNPD), A.P.S. (PhD), C.S.N. (PhD), and to the National Council for Science and Technological Development (CNPq) for providing grant to A.C.G. (Process number: 305535/2014-5).

Ethics

The strain EC121 used in this research was obtained from clinical routine after laboratory procedures. No additional procedure was performed to acquire any bacterial strain, so the consent form was not required as determinate by the Brazilian National Health Council nº 466/12 and 510/16. All patient information was obtained from medical records, and the research was done with the approval of the local Research Ethics Committee of the Federal University of São Paulo - UNIFESP/São Paulo Hospital (CEP 2031/08 and CEP N 7140160317).

Competing interests

A.C.G. has recently received research funding and/or consultation fees from Eurofarma, MSD, Pfizer, and Zambon. Other authors have nothing to declare. This study was not financially supported by any Diagnostic/Pharmaceutical company.

Supporting information captions

Table S1 – ST101 complex strains, NCBI accession, multidrug resistance genotype, source, and country of isolation.

Table S2 – Type V Secretion system protein identified by MaSyFinder.

Table S3 – Complete Genetic cluster identified in EC121 strain and their predicted association with virulence.

Table S4 -Virulence genes identification and validation in NCBI and Swiss-Prot.

Table S5 - Detailed information of all predicted phages detected in strain EC121.

Fig S1 – Phylogenetic relationship among 62 human isolates from the ST101 complex

Fig S2 – Phylogenetic relationship obtained from sequence alignment to plasmid IncHI2 pEC121.A, using BLASTn

Fig S3 – Phylogenetic relationship obtained from sequence alignment to plasmid IncFIB/IncFII pEC121.B, using BLASTn

Fig S4 – Phylogenetic relationship obtained from sequence alignment to plasmid IncL/M pEC121.C, using BLASTn

Fig S5 – EC121 interaction and invasion assay in Eukaryotic cell lineages

Fig S6 – EC121 adherence in eukaryotic cell linages HeLa, Caco-2, and T24.

667 # References

668 1. Biedenbach DJ, Moet GJ, Jones RN. Occurrence and antimicrobial resistance
669 pattern comparisons among bloodstream infection isolates from the SENTRY
670 Antimicrobial Surveillance Program (1997–2002). *Diagn Microbiol Infect Dis*.
671 2004;50(1):59-69. doi:10.1016/J.DIAGMICROBIO.2004.05.003
672 2. Temkin E, Fallach N, Almagor J, et al. Estimating the number of infections
673 caused by antibiotic-resistant Escherichia coli and Klebsiella pneumoniae in
674 2014: a modelling study. *Lancet Glob Heal*. 2018;6(9):e969-e979.
675 doi:10.1016/S2214-109X(18)30278-X
676 3. Tu B, Liu SX, Wu D, et al. [Clinical features of community-acquired
677 bloodstream infection due to Gram-negative bacilli in patients with liver
678 cirrhosis]. *Zhonghua Gan Zang Bing Za Zhi*. 2018;26(1):23-27.
679 doi:10.3760/cma.j.issn.1007-3418.2018.01.007
680 4. Satwani P, Freedman JL, Chaudhury S, et al. A Multicenter Study of Bacterial
681 Blood Stream Infections in Pediatric Allogeneic Hematopoietic Cell
682 Transplantation Recipients: The Role of Acute Gastrointestinal Graft-versus-Host
683 Disease. *Biol Blood Marrow Transplant*. 2017;23(4):642-647.
684 doi:10.1016/j.bbmt.2017.01.073
685 5. Nivesvivat T, Piyaraj P, Thunyaharn S, Watanaveeradej V, Suwanpakdee D.
686 Clinical epidemiology, risk factors and treatment outcomes of extended-spectrum
687 beta-lactamase producing Enterobacteriaceae bacteremia among children in a
688 Tertiary Care Hospital, Bangkok, Thailand. *BMC Res Notes*. 2018;11(1):624.
689 doi:10.1186/s13104-018-3729-3
690 6. GLASS | Global antimicrobial resistance surveillance system (GLASS) report.
691 *WHO*. 2018. https://www.who.int/glass/resources/publications/early-

692 implementation-report/en/. Accessed June 17, 2019.

693 7. Landman WJM, van Eck JHH. The incidence and economic impact of the
694 *Escherichia coli* peritonitis syndrome in Dutch poultry farming. *Avian Pathol*.
695 2015;44(5):370-378. doi:10.1080/03079457.2015.1060584

696 8. Kaye KS, Marchaim D, Chen T-Y, et al. Effect of Nosocomial Bloodstream
697 Infections on Mortality, Length of Stay, and Hospital Costs in Older Adults. *J*
698 *Am Geriatr Soc*. 2014;62(2):306-311. doi:10.1111/jgs.12634

699 9. Barriere SL. Clinical, economic and societal impact of antibiotic resistance.
700 *Expert Opin Pharmacother*. 2015;16(2):151-153.
701 doi:10.1517/14656566.2015.983077

702 10. Ciorba V, Odone A, Veronesi L, Pasquarella C, Signorelli C. Antibiotic
703 resistance as a major public health concern: epidemiology and economic impact.
704 *Ann Ig*. 27(3):562-579. doi:10.7416/ai.2015.2048

705 11. Johnson JR, Russo TA. Extraintestinal pathogenic Escherichia coli : "The other
706 bad E coli ." *J Lab Clin Med*. 2002;139(3):155-162.
707 doi:10.1067/mlc.2002.121550

708 12. Johnson JR, Murray AC, Gajewski A, et al. Isolation and molecular
709 characterization of nalidixic acid-resistant extraintestinal pathogenic Escherichia
710 coli from retail chicken products. *Antimicrob Agents Chemother*.
711 2003;47(7):2161-2168. doi:10.1128/aac.47.7.2161-2168.2003

712 13. Spurbeck RR, Dinh PC, Walk ST, et al. Escherichia coli isolates that carry vat,
713 fyuA, chuA, and yfcV efficiently colonize the urinary tract. *Infect Immun*.
714 2012;80(12):4115-4122. doi:10.1128/IAI.00752-12

715 14. Santos ACM, Zidko ACM, Pignatari AC, Silva RM. Assessing the diversity of
716 the virulence potential of Escherichia coli isolated from bacteremia in São Paulo,

717 Brazil. *Brazilian J Med Biol Res = Rev Bras Pesqui medicas e Biol*.
718 2013;46(11):968-973. doi:10.1590/1414-431X20133184
719 15. Zhao L, Zhang J, Zheng B, et al. Molecular Epidemiology and Genetic Diversity
720 of Fluoroquinolone-Resistant Escherichia coli Isolates from Patients with
721 Community-Onset Infections in 30 Chinese County Hospitals. *J Clin Microbiol*.
722 2015;53(3):766-770. doi:10.1128/JCM.02594-14
723 16. Izdebski R, Baraniak A, Herda M, et al. MLST reveals potentially high-risk
724 international clones of Enterobacter cloacae*. *J Antimicrob Chemother*.
725 2015;70(1):48-56. doi:10.1093/jac/dku359
726 17. Brauner J, Hallin M, Deplano A, et al. Community-acquired methicillin-resistant
727 Staphylococcus aureus clones circulating in Belgium from 2005 to 2009:
728 changing epidemiology. *Eur J Clin Microbiol Infect Dis*. 2013;32(5):613-620.
729 doi:10.1007/s10096-012-1784-6
730 18. Riley LW. Pandemic lineages of extraintestinal pathogenic Escherichia coli. *Clin*
731 *Microbiol Infect*. 2014;20(5):380-390. doi:10.1111/1469-0691.12646
732 19. Roer L, Overballe-Petersen S, Hansen F, et al. Escherichia coli Sequence Type
733 410 Is Causing New International High-Risk Clones. *mSphere*.
734 2018;3(4):e00337-18. doi:10.1128/mSphere.00337-18
735 20. Reineke K, Sevenich R, Hertwig C, et al. Comparative study on the high pressure
736 inactivation behavior of the Shiga toxin-producing Escherichia coli O104:H4 and
737 O157:H7 outbreak strains and a non-pathogenic surrogate. *Food Microbiol*.
738 2015;46:184-194. doi:10.1016/J.FM.2014.07.017
739 21. Kennedy CA, Walsh C, Karczmarczyk M, et al. Multi-drug resistant Escherichia
740 coli in diarrhoeagenic foals: Pulsotyping, phylotyping, serotyping, antibiotic
741 resistance and virulence profiling. *Vet Microbiol*. 2018;223:144-152.

742 doi:10.1016/J.VETMIC.2018.08.009

743 22. WHO | Global action plan on antimicrobial resistance. *WHO*. 2017.
744 https://www.who.int/antimicrobial-resistance/publications/global-action-plan/en/.
745 Accessed June 17, 2019.

746 23. Shaik S, Ranjan A, Tiwari SK, et al. Comparative Genomic Analysis of Globally
747 Dominant ST131 Clone with Other Epidemiologically Successful Extraintestinal
748 Pathogenic Escherichia coli (ExPEC) Lineages. *MBio*. 2017;8(5):e01596-17.
749 doi:10.1128/mBio.01596-17

750 24. Magiorakos A-P, Srinivasan A, Carey RB, et al. Multidrug-resistant, extensively
751 drug-resistant and pandrug-resistant bacteria: an international expert proposal for
752 interim standard definitions for acquired resistance. *Clin Microbiol Infect*.
753 2012;18(3):268-281. doi:10.1111/J.1469-0691.2011.03570.X

754 25. Santos AC de M. Potencial de virulência e suscetibilidade a antimicrobianos de
755 amostras de Escherichia coli isoladas de bacteremia: sua relação com o estado de
756 imunocompetência do paciente e a origem da infecção. 2013.
757 http://repositorio.unifesp.br/handle/11600/22975. Accessed June 20, 2019.

758 26. Ranjan A, Shaik S, Mondal A, et al. Molecular Epidemiology and Genome
759 Dynamics of New Delhi Metallo-β-Lactamase-Producing Extraintestinal
760 Pathogenic Escherichia coli Strains from India. *Antimicrob Agents Chemother*.
761 2016;60(11):6795-6805. doi:10.1128/AAC.01345-16

762 27. Pfeifer Y, Trifonova A, Pietsch M, et al. Clonal Transmission of Gram-Negative
763 Bacteria with Carbapenemases NDM-1, VIM-1, and OXA-23/72 in a Bulgarian
764 Hospital. *Microb Drug Resist*. 2017;23(3):301-307. doi:10.1089/mdr.2016.0059

765 28. Poirel L, Lagrutta E, Taylor P, Pham J, Nordmann P. Emergence of metallo-β-
766 lactamase NDM-1-producing multidrug-resistant Escherichia coli in Australia.

767 *Antimicrob Agents Chemother*. 2010;54(11):4914-4916.
768 doi:10.1128/AAC.00878-10
769 29. Poirel L, Savov E, Nazli A, et al. Outbreak caused by NDM-1- and RmtB-
770 producing Escherichia coli in Bulgaria. *Antimicrob Agents Chemother*.
771 2014;58(4):2472-2474. doi:10.1128/AAC.02571-13
772 30. Yoo JS, Kim HM, Koo HS, et al. Nosocomial transmission of NDM-1-producing
773 Escherichia coli ST101 in a Korean hospital. *J Antimicrob Chemother*.
774 2013;68(9):2170-2172. doi:10.1093/jac/dkt126
775 31. Yamaji R, Friedman CR, Rubin J, et al. A Population-Based Surveillance Study
776 of Shared Genotypes of Escherichia coli Isolates from Retail Meat and Suspected
777 Cases of Urinary Tract Infections. *mSphere*. 2018;3(4):e00179-18.
778 doi:10.1128/mSphere.00179-18
779 32. Fernandes MR, McCulloch JA, Vianello MA, et al. First Report of the Globally
780 Disseminated IncX4 Plasmid Carrying the mcr-1 Gene in a Colistin-Resistant
781 Escherichia coli Sequence Type 101 Isolate from a Human Infection in Brazil.
782 *Antimicrob Agents Chemother*. 2016;60(10):6415-6417.
783 doi:10.1128/AAC.01325-16
784 33. GERHOLD G, SCHULZE MH, GROSS U, BOHNE W. Multilocus sequence
785 typing and CTX-M characterization of ESBL-producing *E. coli* : a prospective
786 single-centre study in Lower Saxony, Germany. *Epidemiol Infect*.
787 2016;144(15):3300-3304. doi:10.1017/S0950268816001412
788 34. Khan ER, Aung MS, Paul SK, et al. Prevalence and Molecular Epidemiology of
789 Clinical Isolates of *Escherichia coli* and *Klebsiella pneumoniae* Harboring
790 Extended-Spectrum Beta-Lactamase and Carbapenemase Genes in Bangladesh.
791 *Microb Drug Resist*. 2018;24(10):1568-1579. doi:10.1089/mdr.2018.0063

792 35. Pfeifer Y, Witte W, Holfelder M, Busch J, Nordmann P, Poirel L. NDM-1-
793 producing Escherichia coli in Germany. *Antimicrob Agents Chemother*.
794 2011;55(3):1318-1319. doi:10.1128/AAC.01585-10
795 36. Williamson DA, Freeman JT, Roberts SA, et al. Rectal colonization with New
796 Delhi metallo- -lactamase-1-producing Escherichia coli prior to transrectal
797 ultrasound (TRUS)-guided prostate biopsy. *J Antimicrob Chemother*.
798 2013;68(12):2957-2959. doi:10.1093/jac/dkt266
799 37. Potron A, Poirel L, Rondinaud E, Nordmann P. Intercontinental spread of OXA-
800 48 beta-lactamase-producing Enterobacteriaceae over a 11-year period, 2001 to
801 2011. *Eurosurveillance*. 2013;18(31):20549. doi:10.2807/1560-
802 7917.ES2013.18.31.20549
803 38. Samuelsen Ø, Overballe-Petersen S, Bjørnholt JV, et al. Molecular and
804 epidemiological characterization of carbapenemase-producing Enterobacteriaceae
805 in Norway, 2007 to 2014. Butaye P, ed. *PLoS One*. 2017;12(11):e0187832.
806 doi:10.1371/journal.pone.0187832
807 39. Mantilla-Calderon D, Jumat MR, Wang T, Ganesan P, Al-Jassim N, Hong P-Y.
808 Isolation and Characterization of NDM-Positive Escherichia coli from Municipal
809 Wastewater in Jeddah, Saudi Arabia. *Antimicrob Agents Chemother*.
810 2016;60(9):5223-5231. doi:10.1128/AAC.00236-16
811 40. Braga JFV, Chanteloup NK, Trotereau A, et al. Diversity of Escherichia coli
812 strains involved in vertebral osteomyelitis and arthritis in broilers in Brazil. *BMC*
813 *Vet Res*. 2016;12(1):140. doi:10.1186/s12917-016-0762-0
814 41. Koo H-J, Kwak H-S, Yoon S-H, Woo G-J. Phylogenetic group distribution and
815 prevalence of virulence genes in Escherichia coli isolates from food samples in
816 South Korea. *World J Microbiol Biotechnol*. 2012;28(4):1813-1816.

817 doi:10.1007/s11274-011-0954-5

818 42. Müller A, Jansen W, Grabowski NT, Monecke S, Ehricht R, Kehrenberg C.
819 ESBL- and AmpC-producing Escherichia coli from legally and illegally imported
820 meat: Characterization of isolates brought into the EU from third countries. *Int J*
821 *Food Microbiol*. 2018;283:52-58. doi:10.1016/J.IJFOODMICRO.2018.06.009

822 43. Xie M, Lin D, Chen K, Chan EWC, Yao W, Chen S. Molecular Characterization
823 of Escherichia coli Strains Isolated from Retail Meat That Harbor blaCTX-M and
824 fosA3 Genes. *Antimicrob Agents Chemother*. 2016;60(4):2450-2455.
825 doi:10.1128/AAC.03101-15

826 44. Wu C, Wang Y, Shi X, et al. Rapid rise of the ESBL and *mcr-1* genes in
827 *Escherichia coli* of chicken origin in China, 2008–2014. *Emerg Microbes Infect*.
828 2018;7(1):1-10. doi:10.1038/s41426-018-0033-1

829 45. Boehmer T, Vogler AJ, Thomas A, et al. Phenotypic characterization and whole
830 genome analysis of extended-spectrum beta-lactamase-producing bacteria
831 isolated from dogs in Germany. Mokrousov I, ed. *PLoS One*.
832 2018;13(10):e0206252. doi:10.1371/journal.pone.0206252

833 46. Zhou X, García-Cobos S, Ruijs GJHM, et al. Epidemiology of Extended-
834 Spectrum β-Lactamase-Producing E. coli and Vancomycin-Resistant Enterococci
835 in the Northern Dutch–German Cross-Border Region. *Front Microbiol*.
836 2017;8:1914. doi:10.3389/fmicb.2017.01914

837 47. Wang Y, Tian G-B, Zhang R, et al. Prevalence, risk factors, outcomes, and
838 molecular epidemiology of mcr-1-positive Enterobacteriaceae in patients and
839 healthy adults from China: an epidemiological and clinical study. *Lancet Infect*
840 *Dis*. 2017;17(4):390-399. doi:10.1016/S1473-3099(16)30527-8

841 48. Shrestha B, Tada T, Miyoshi-Akiyama T, et al. Identification of a novel NDM

variant, NDM-13, from a multidrug-resistant Escherichia coli clinical isolate in Nepal. *Antimicrob Agents Chemother*. 2015;59(9):5847-5850. doi:10.1128/AAC.00332-15

49. Herrero-Fresno A, Zachariasen C, Hansen MH, et al. Apramycin treatment affects selection and spread of a multidrug-resistant Escherichia coli strain able to colonize the human gut in the intestinal microbiota of pigs. *Vet Res*. 2016;47(1):12. doi:10.1186/s13567-015-0291-z

50. Trobos M, Lester CH, Olsen JE, Frimodt-Moller N, Hammerum AM. Natural transfer of sulphonamide and ampicillin resistance between Escherichia coli residing in the human intestine. *J Antimicrob Chemother*. 2009;63(1):80-86. doi:10.1093/jac/dkn437

51. Picard B, Garcia JS, Gouriou S, et al. The link between phylogeny and virulence in Escherichia coli extraintestinal infection. *Infect Immun*. 1999;67(2):546-553. http://www.ncbi.nlm.nih.gov/pubmed/9916057. Accessed June 21, 2019.

52. Johnson JR, Clermont O, Menard M, Kuskowski MA, Picard B, Denamur E. Experimental Mouse Lethality of *Escherichia coli* Isolates, in Relation to Accessory Traits, Phylogenetic Group, and Ecological Source. *J Infect Dis*. 2006;194(8):1141-1150. doi:10.1086/507305

53. Duriez P, Clermont O, Picard B, et al. Commensal Escherichia coli isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology*. 2001;147(6):1671-1676. doi:10.1099/00221287-147-6-1671

54. Hallström T, Resman F, Ristovski M, Riesbeck K. Binding of complement regulators to invasive nontypeable Haemophilus influenzae isolates is not increased compared to nasopharyngeal isolates, but serum resistance is linked to disease severity. *J Clin Microbiol*. 2010;48(3):921-927. doi:10.1128/JCM.01654-

09

55. Hughes C, Phillips R, Roberts AP. Serum resistance among Escherichia coli strains causing urinary tract infection in relation to O type and the carriage of hemolysin, colicin, and antibiotic resistance determinants. *Infect Immun*. 1982;35(1):270-275. http://www.ncbi.nlm.nih.gov/pubmed/7033137. Accessed June 21, 2019.

56. Adler NRL, Stevens MP, Dean RE, et al. Systematic Mutagenesis of Genes Encoding Predicted Autotransported Proteins of Burkholderia pseudomallei Identifies Factors Mediating Virulence in Mice, Net Intracellular Replication and a Novel Protein Conferring Serum Resistance. Wooten RM, ed. *PLoS One*. 2015;10(4):e0121271. doi:10.1371/journal.pone.0121271

57. Desroches M, Clermont O, Lafeuillade B, et al. Genotypic and phenotypic characteristics of *Escherichia coli* involved in transfusion-transmitted bacterial infections: implications for preventive strategies. *Transfusion*. 2018;58(8):1940-1950. doi:10.1111/trf.14812

58. Peirano G, Mulvey GL, Armstrong GD, Pitout JDD. Virulence potential and adherence properties of Escherichia coli that produce CTX-M and NDM β-lactamases. *J Med Microbiol*. 2013;62(4):525-530. doi:10.1099/jmm.0.048983-0

59. Hauser E, Mellmann A, Semmler T, et al. Phylogenetic and molecular analysis of food-borne shiga toxin-producing Escherichia coli. *Appl Environ Microbiol*. 2013;79(8):2731-2740. doi:10.1128/AEM.03552-12

60. Mellmann A, Bielaszewska M, Köck R, et al. Analysis of Collection of Hemolytic Uremic Syndrome–associated Enterohemorrhagic *Escherichia coli*. *Emerg Infect Dis*. 2008;14(8):1287-1290. doi:10.3201/eid1408.071082

61. Bai X, Mernelius S, Jernberg C, et al. Shiga Toxin-Producing Escherichia coli

892 Infection in Jönköping County, Sweden: Occurrence and Molecular
893 Characteristics in Correlation With Clinical Symptoms and Duration of stx
894 Shedding. *Front Cell Infect Microbiol*. 2018;8:125.
895 doi:10.3389/fcimb.2018.00125
896 62. Ferdous M, Friedrich AW, Grundmann H, et al. Molecular characterization and
897 phylogeny of Shiga toxin–producing Escherichia coli isolates obtained from two
898 Dutch regions using whole genome sequencing. *Clin Microbiol Infect*.
899 2016;22(7):642.e1-642.e9. doi:10.1016/j.cmi.2016.03.028
900 63. Bankevich A, Nurk S, Antipov D, et al. SPAdes: A New Genome Assembly
901 Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol*.
902 2012;19(5):455-477. doi:10.1089/cmb.2012.0021
903 64. Larsen M V, Cosentino S, Rasmussen S, et al. Multilocus sequence typing of
904 total-genome-sequenced bacteria. *J Clin Microbiol*. 2012;50(4):1355-1361.
905 doi:10.1128/JCM.06094-11
906 65. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and
907 Easy In Silico Serotyping of Escherichia coli Isolates by Use of Whole-Genome
908 Sequencing Data. *J Clin Microbiol*. 2015;53(8):2410-2426.
909 doi:10.1128/JCM.00008-15
910 66. Carattoli A, Zankari E, García-Fernández A, et al. In silico detection and typing
911 of plasmids using PlasmidFinder and plasmid multilocus sequence typing.
912 *Antimicrob Agents Chemother*. 2014;58(7):3895-3903. doi:10.1128/AAC.02412-
913 14
914 67. Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial
915 resistance genes. *J Antimicrob Chemother*. 2012;67(11):2640-2644.
916 doi:10.1093/jac/dks261

68. Joensen KG, Scheutz F, Lund O, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. *J Clin Microbiol*. 2014;52(5):1501-1510. doi:10.1128/JCM.03617-13

69. Arndt D, Grant JR, Marcu A, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*. 2016;44(W1):W16-W21. doi:10.1093/nar/gkw387

70. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A Fast Phage Search Tool. *Nucleic Acids Res*. 2011;39(suppl):W347-W352. doi:10.1093/nar/gkr485

71. Wattam AR, Davis JJ, Assaf R, et al. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res*. 2017;45(D1):D535-D542. doi:10.1093/nar/gkw1017

72. Altschul S, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402. doi:10.1093/nar/25.17.3389

73. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47(D1):D506-D515. doi:10.1093/nar/gky1049

74. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019;47(W1):W256-W259. doi:10.1093/nar/gkz239

75. Mareuil F, Doppelt-Azeroual O, Ménager H, Mareuil F, Doppelt-Azeroual O, Ménager H. A public Galaxy platform at Pasteur used as an execution engine for web services. *F1000Research*. 2017;6. doi:10.7490/F1000RESEARCH.1114334.1

76. Abby SS, Néron B, Ménager H, Touchon M, Rocha EPC. MacSyFinder: A

942 Program to Mine Genomes for Molecular Systems with an Application to
943 CRISPR-Cas Systems. Torres N V., ed. *PLoS One*. 2014;9(10):e110726.
944 doi:10.1371/journal.pone.0110726
945 77. Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EPC.
946 Identification of protein secretion systems in bacterial genomes. *Sci Rep*.
947 2016;6(1):23080. doi:10.1038/srep23080
948 78. Orskov F, Orskov I. Escherichia coli O:H serotypes isolated from human blood.
949 Prevalence of the K1 antigen with technical details of O and H antigenic
950 determination. *Acta Pathol Microbiol Scand Suppl*. 1975;83(6):595-600.
951 http://www.ncbi.nlm.nih.gov/pubmed/1106116. Accessed June 23, 2019.
952 79. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening
953 recombinant plasmid DNA. *Nucleic Acids Res*. 1979;7(6):1513-1523.
954 doi:10.1093/nar/7.6.1513
955 80. Moran RA, Anantham S, Hall RM. An improved plasmid size standard,
956 39R861+. *Plasmid*. 2019;102:6-9. doi:10.1016/j.plasmid.2019.01.002
957 81. Krawczyk PS, Lipinski L, Dziembowski A. PlasFlow: predicting plasmid
958 sequences in metagenomic data using genome signatures. *Nucleic Acids Res*.
959 2018;46(6):e35-e35. doi:10.1093/nar/gkx1321
960 82. Houdouin V, Bonacorsi S, Brahimi N, Clermont O, Nassif X, Bingen E. A
961 uropathogenicity island contributes to the pathogenicity of Escherichia coli
962 strains that cause neonatal meningitis. *Infect Immun*. 2002;70(10):5865-5869.
963 doi:10.1128/iai.70.10.5865-5869.2002
964 83. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint*
965 *Tables for Interpretation of MICs and Zone Diameters. Version 9.0, 2019*.
966 http://www.eucast.org. Accessed June 20, 2019.

967 84. Lima MP, Yamamoto D, Santos AC de M, et al. Phenotypic characterization and
968 virulence-related properties of Escherichia albertii strains isolated from children
969 with diarrhea in Brazil. *Pathog Dis*. 2019;77(2). doi:10.1093/femspd/ftz014
970 85. Garcia BG, Ooka T, Gotoh Y, et al. Genetic relatedness and virulence properties
971 of enteropathogenic Escherichia coli strains of serotype O119:H6 expressing
972 localized adherence or localized and aggregative adherence-like patterns on HeLa
973 cells. *Int J Med Microbiol*. 2016;306(3):152-164.
974 doi:10.1016/j.ijmm.2016.02.008
975 86. Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ. Type 1
976 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J*.
977 2000;19(12):2803-2812. doi:10.1093/EMBOJ/19.12.2803
978 87. Sansonetti PJ, Kopecko DJ, Formal SB. Involvement of a plasmid in the invasive
979 ability of Shigella flexneri. *Infect Immun*. 1982;35(3):852-860.
980 http://www.ncbi.nlm.nih.gov/pubmed/6279518. Accessed June 23, 2019.
981 88. Hernandes RT, Silva RM, Carneiro SM, et al. The localized adherence pattern of
982 an atypical enteropathogenic Escherichia coli is mediated by intimin omicron and
983 unexpectedly promotes HeLa cell invasion. *Cell Microbiol*.
984 2008;0(0):071003010119002-??? doi:10.1111/j.1462-5822.2007.01054.x
985

FIGURE LEGENDS

Fig 1 - EC121 Phylogenetic tree.

A phylogenetic tree was built with genomes of reference *E. coli* strains, of relevant pathogenic strains from all *E. coli* pathotypes, and some strains from ST101, using Maximum Likelihood-based algorithm (RAxML) in PATRIC. When known, the strain's pathotype was in parenthesis, following the strain identification. Diarrheagenic *E. coli* strains were in orange; Extraintestinal pathogenic *E. coli* strains were in red, commensal or strains that the origin was not described were in black. In bold and with purple label background, the strain studied in present work. Bootstrap upper than 50 were informed in the tree.

Fig 2 - ST101 complex phylogenetic tree.

The phylogenetic tree was built using 95 *E. coli* strains from ST101 complex (ST101, ST359, ST2480, ST5957, and ST6388) from diverse source and countries using all shared protein among them on FastTree2 to build the tree. *E. coli* strain IAI1, *E. coli* O157:H7 str Sakai, and *E. coli* O104:H4 str 2011c-3493 are used as outgroups. Bootstrap upper than 50 were informed in the tree. Label colors are related with isolates origin or host diseases, being Shiga toxin-producing *E. coli* (STEC) in orange independently of the origin; Diarrheagenic *E. coli* (DEC) strains in yellow; Extraintestinal pathogenic *E. coli* (ExPEC) strains in red; isolates from microbiota in brown; isolates from retail food or environment in green; strains that the origin was not described in black. One strain isolated from Crohn's disease is in purple. In bold and with purple label background, the strain of present work. * All strains that harbor AMR for 3 or more antimicrobial classes were designed as MDR. Other FQ (Fluoroquinolone) resistance genes detected were the

mobile genes *qepA* and *aac(6')-Ib-cr*. Mutations that confer resistance to FQ were not considered to build the AMR information in this tree.

Fig 3 - CRISPR locus composition and biological systems assignment in EC121. **A –** Genomic architecture representation of CRISPR *locus* in the EC121 genome; the image was obtained in MacSyFinder from genome annotation; **B -** Representation of EC121 genome composition in subsystems based on protein biological data obtained *in silico*; **C -** EC121 genome schematic composition, based on annotation, ordered by contig size. In circle from outer to the inner, forward strand, reverse strand, RNA related genes, antimicrobial resistance, virulence factors, GC content, and GC skew. The colors in forward and reverse strands correspond to the subsystems presented in B. Figures 2B and 2C were obtained using the comprehensive genome analysis service at PATRIC.

Fig 4 - Plasmid content of the EC121strain. Negative image of the plasmid content of EC121 strain obtained by alkaline extraction, followed by electrophoresis on 0.8% agarose gel in TBE buffer. Approximate sizes were predicted based on the plasmid migration in agarose gel.

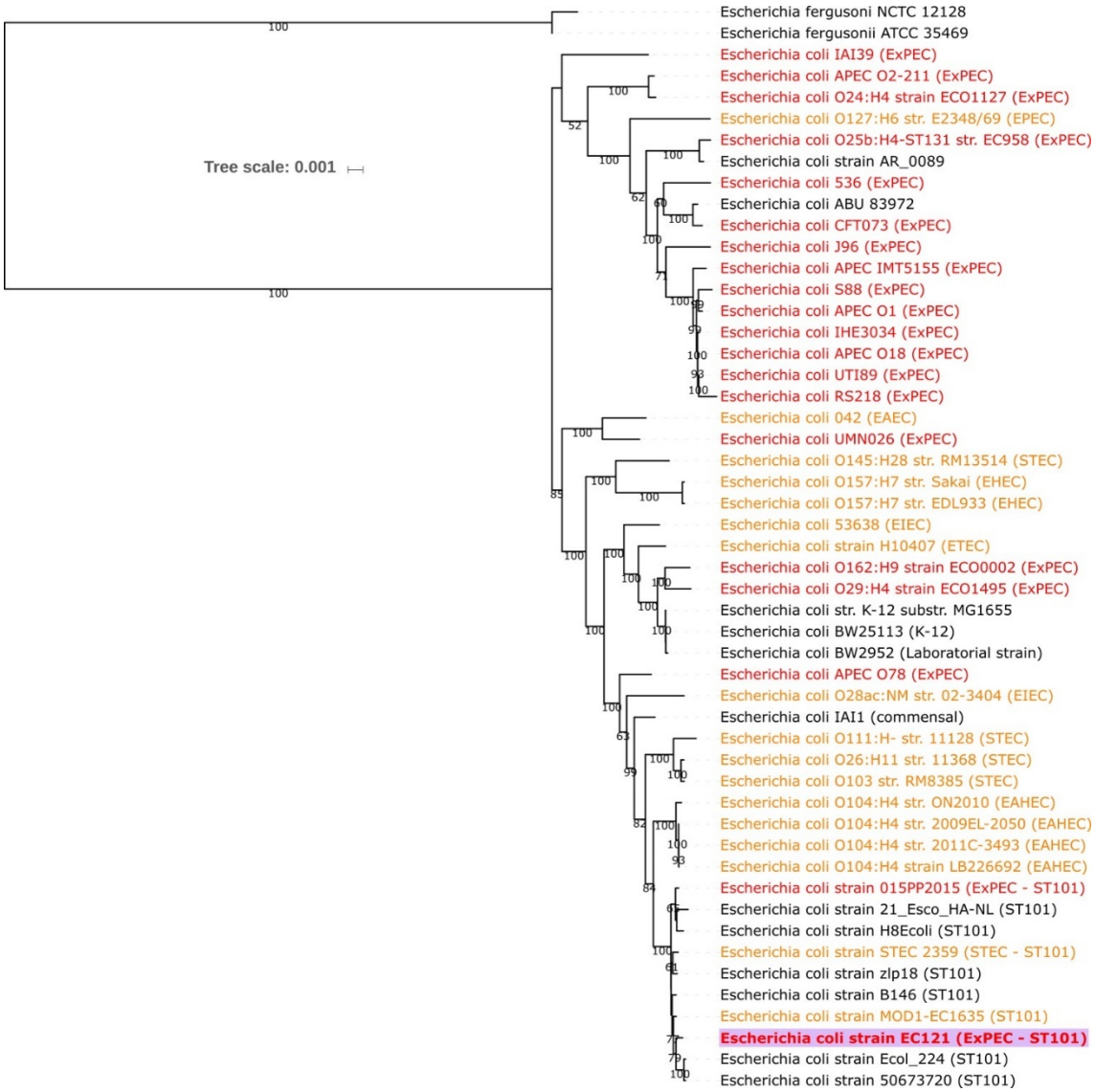
Fig 5 - Adherence and biofilm formation on abiotic surfaces. Biofilm assays were performed in DMEM, at 37°C, for 24 h. **A -** Qualitative adherence assay on a glass surface. Bacteria were stained with crystal violet, observed in optical microscopy (OM) 400 x; **B -** Quantitative adherence assay on polystyrene.

Figure 6 – Classical adherence assay on HeLa cell. The assays were performed in biological and experimental triplicates in the presence of 2% D-Mannose, except when

1034 indicated; M-, assay performed in the absence of 2% D-Mannose. C1845 strain was used
1035 as adherent control; *E. coli* strain C600 was used as weakly adherent control.
1036

1037 **Fig 7– Interaction and invasion in different eukaryotic cell lineages**. Short period
1038 invasion quantitative assay was performed in HeLa, Caco-2 and T24 cells. All assays
1039 were carried out in the absence of D-mannose, except for EC121 M+, where 2% D-
1040 mannose was added in the media; *S. flexneri* M90T was used as invasive control and *E.*
1041 *albertti*1551-2::eae was used as adherent non-invasive control in HeLa and Caco-2 cells.
1042 Data shown are the mean of the relative interaction and invasion comparing with strains
1043 against M90T in each cell lineage.

1044 **Fig 1.**



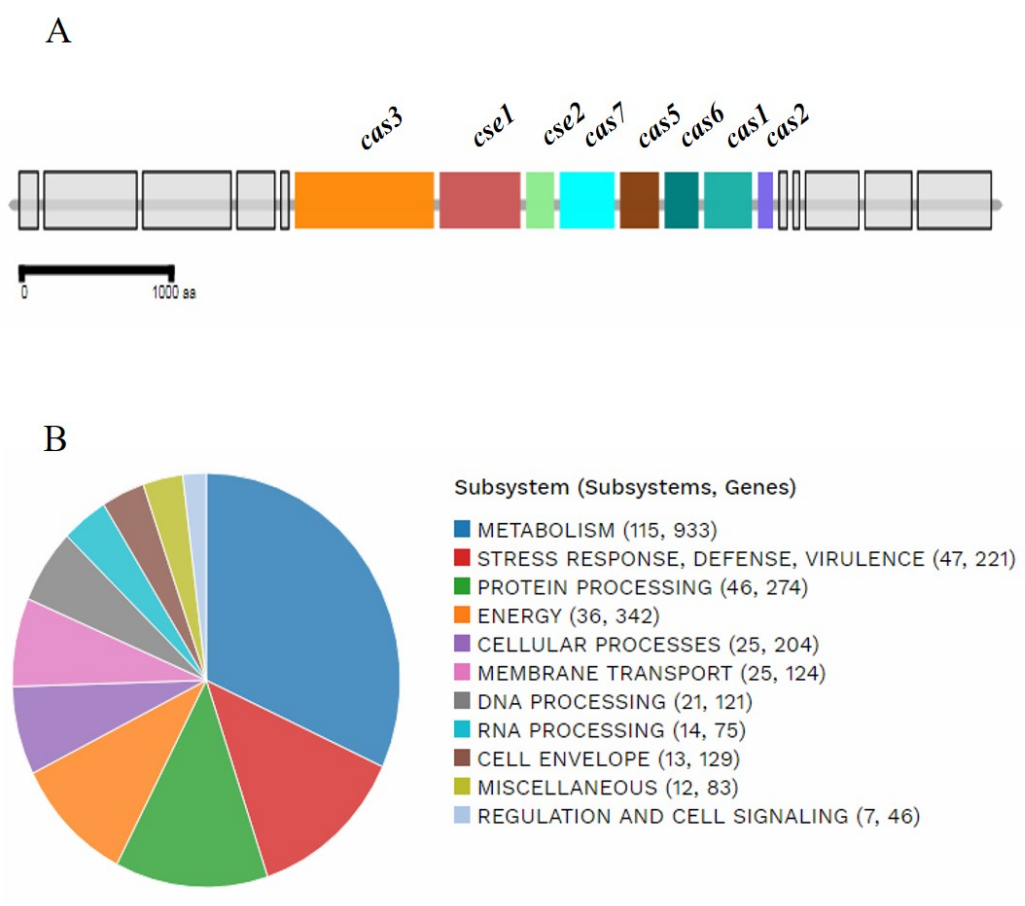
1045

1046

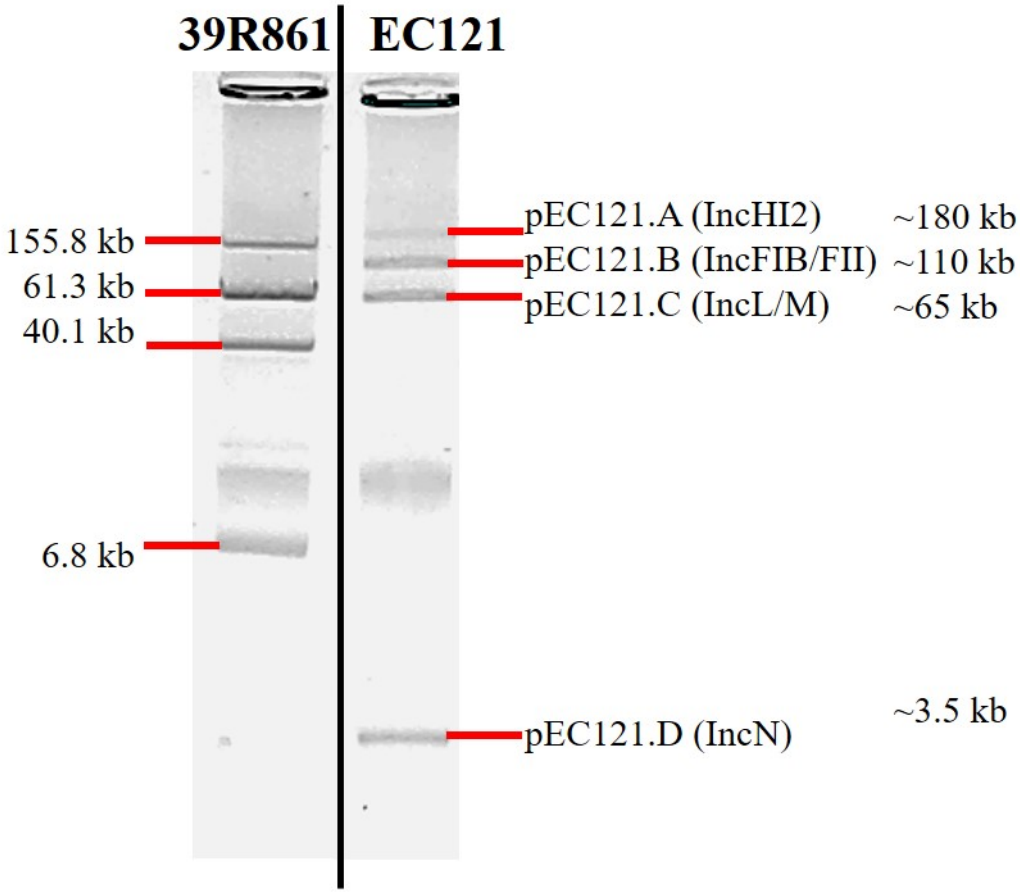
1047 **Fig 2.**

1048

Fig 3.

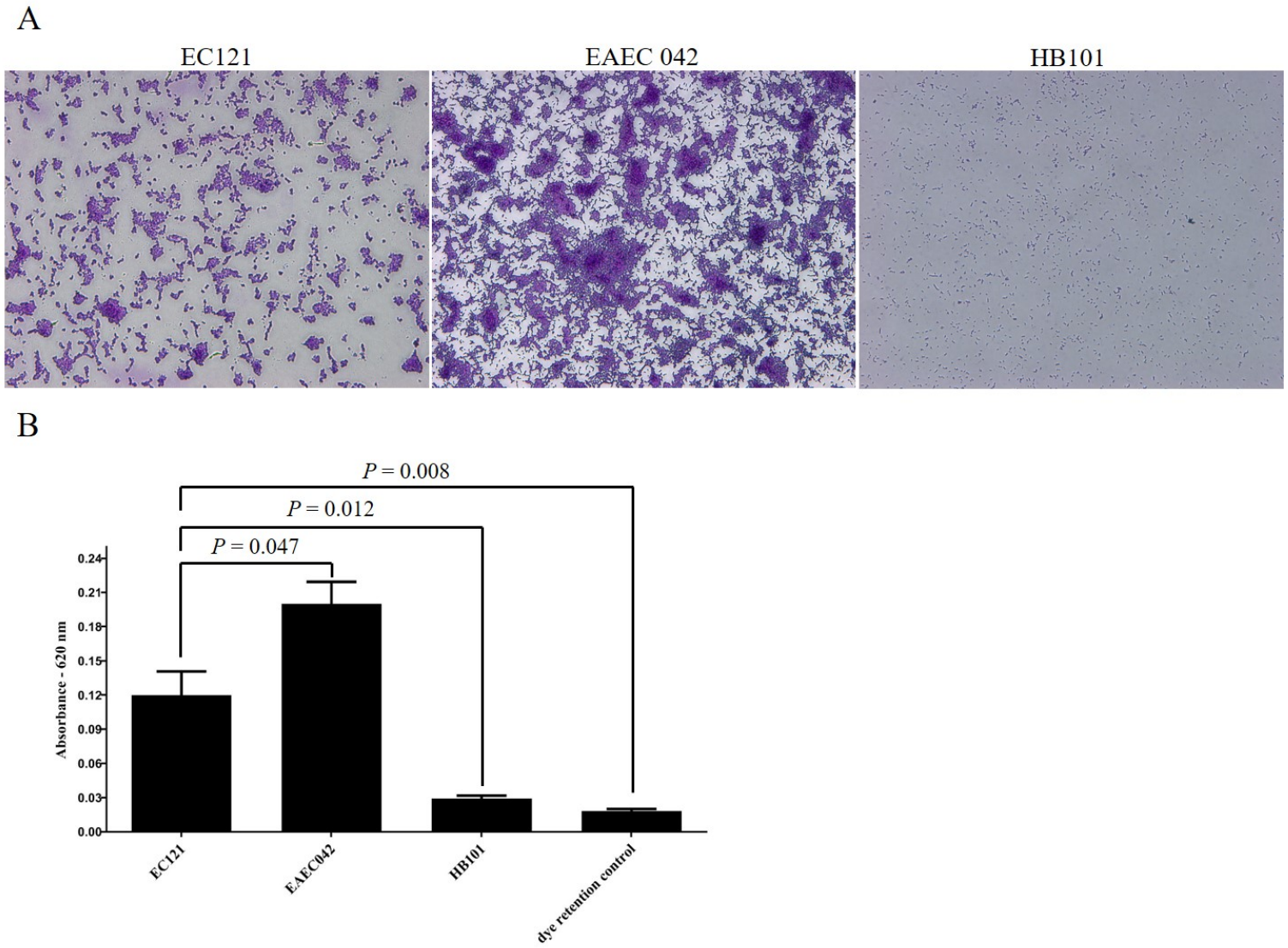

A
cas3
cse1
cse2
cas7
cas5
cas6
cas1
cas2
0
1000 aa
B
Subsystem (Subsystems, Genes)
METABOLISM (115, 933)
STRESS RESPONSE, DEFENSE, VIRULENCE (47, 221)
PROTEIN PROCESSING (46, 274)
ENERGY (36, 342)
CELLULAR PROCESSES (25, 204)
MEMBRANE TRANSPORT (25, 124)
DNA PROCESSING (21, 121)
RNA PROCESSING (14, 75)
CELL ENVELOPE (13, 129)
MISCELLANEOUS (12, 83)
REGULATION AND CELL SIGNALING (7, 46)
C

1052 **Fig 4.**



1053

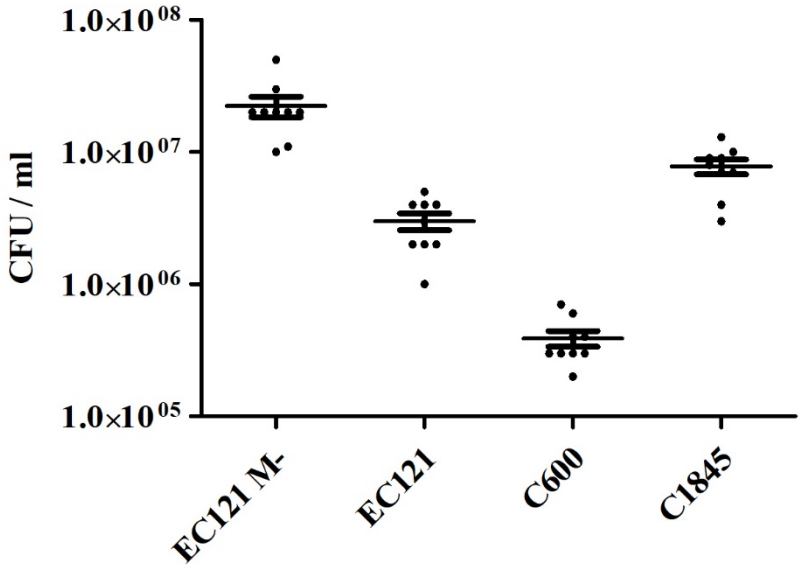
1054


A
EC121
EAEC 042
HB101
B
P = 0.008
P = 0.012
P = 0.047
Absorbance - 620 nm
0.24
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0.00
EC121
EAEC042
HB101
dye retention control

1055 **Fig 5.**

[Research article]

1056 **Fig 6.**



1057

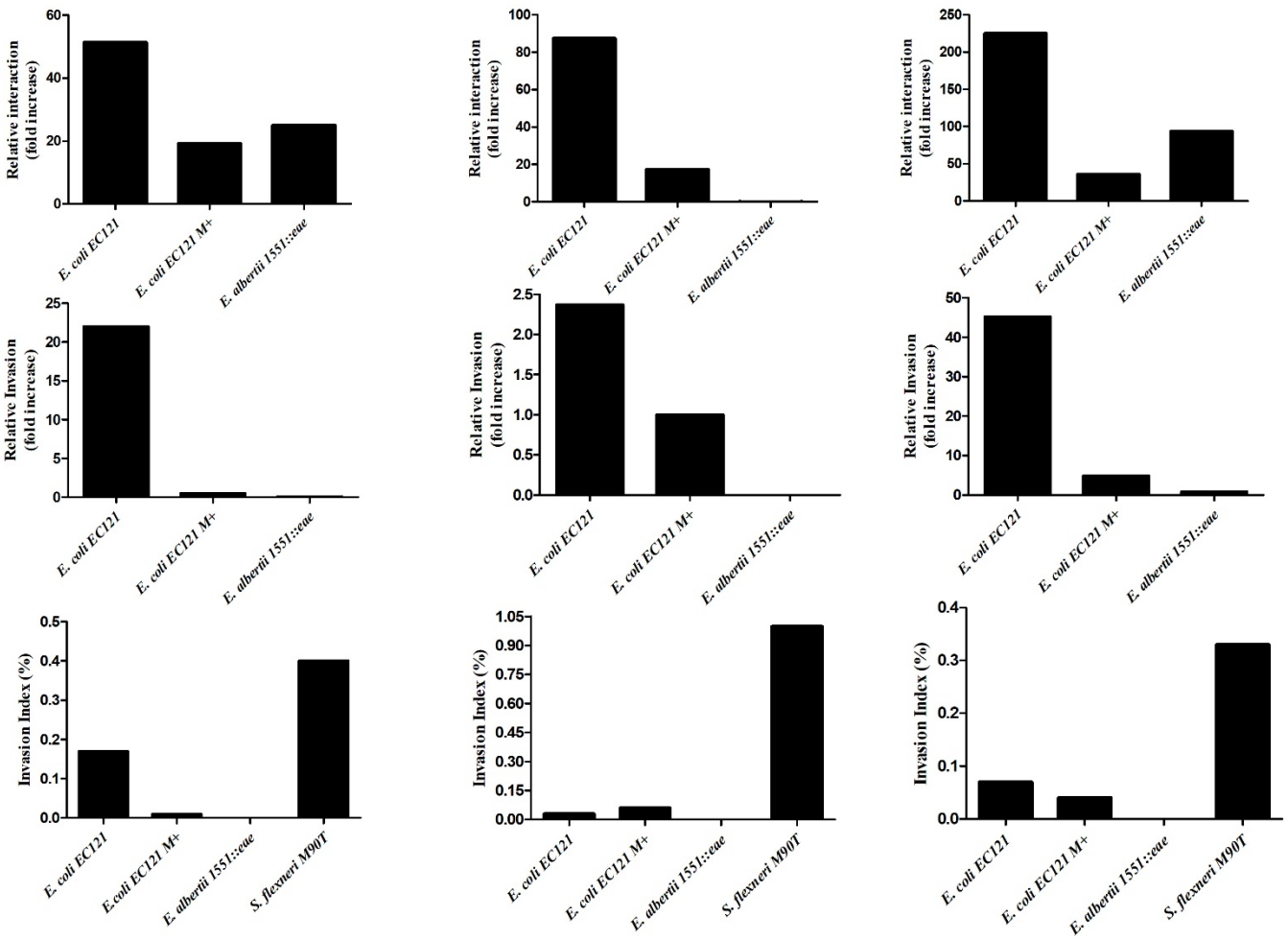
1058

1059

1060

1061

1062 **Fig7.**



1063