USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ESCOLA DE DOUTORAMENTO INTERNACIONAL (EDIUS)

Saskia Camille
Flament Simon

Tese de doutoramento

Molecular Characterization of Extended-Spectrum β-Lactamase-Producing Escherichia coli and Clones Causing Extraintestinal Infections in Humans. Animals as Reservoir of High-Risk E. coli Clones Pathogenic for Humans

Lugo, 2020

PD en Avances en Bioloxía Microbiana e Parasitaria

CENTRO INTERNACIONAL DE ESTUDOS DE DOUTORAMENTO E AVANZADOS DA USC (CIEDUS)

DOCTORAL THESIS

MOLECULAR CHARACTERIZATION OF EXTENDED-SPECTRUM B-LACTAMASE-PRODUCING *ESCHERICHIA COLI* AND CLONES CAUSING EXTRAINTESTINAL INFECTIONS IN HUMANS. ANIMALS AS RESERVOIR OF HIGH-RISK *E. COLI* CLONES PATHOGENIC FOR HUMANS.

SASKIA CAMILLE FLAMENT SIMON

ESCOLA DE DOUTORAMENTO INTERNACIONAL

PROGRAMA DE DOUTORAMENTO EN AVANCES EN BIOLOGÍA MICROBIANA Y PARASITARIA

LUGO

2020

RESPONSIBLE DECLARATION OF THE AUTHOR OF THE THESIS BY COMPENDIUM OF PUBLICATIONS

Ms. **Saskia Camille Flament Simon**

Thesis title: Molecular Characterization of Extended-Spectrum β-Lactamase-Producing *Escherichia coli* and Clones Causing Extraintestinal Infections in Humans. Animals as Reservoir of High-Risk *E. coli* Clones Pathogenic for Humans.

I declare that:

I present my thesis as a compendium of publications of which I am a co-author.

Flament-Simon SC *et al.* 2020 Antibiotics 9 (4) 161 doi: 10.3390/antibiotics9040161.
Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125.
Flament-Simon SC *et al.* 2019 Front Microbiol 10:1183 doi: 10.3389/fmicb.2019.01183.
Flament-Simon SC *et al.* 2020 Microorganims 8 (11):E1712 doi: 10.3390/microorganisms8111712.
Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387.

I DECLARE RESPONSIBLY:

The use of these publications / articles following the requirements established in article 37 c) of the USC Doctoral Studies Regulations.

In Lugo, 15th December 2020

Signed. Saskia Camille Flament Simon

DECLARATION OF THE AUTHOR

Ms. **Saskia Camille Flament Simon**

Thesis title: Molecular Characterization of Extended-Spectrum β-Lactamase-Producing *Escherichia coli* and Clones Causing Extraintestinal Infections in Humans. Animals as Reservoir of High-Risk *E. coli* Clones Pathogenic for Humans.

I present my thesis, following the appropriate Regulation procedure and I declare that:

1) The thesis covers the results of the elaboration of my work.
2) This thesis refers to the collaborations realised in this piece of work.
3) I confirm that the thesis does not incur any type of plagiarism from other authors or work presented by myself in order to obtain other titles.

In Lugo, 15th December 2020

Signed. Saskia Camille Flament Simon

AUTHORIZATION OF THE THESIS DIRECTOR / TUTOR

Molecular Characterization of Extended-Spectrum β-Lactamase-Producing *Escherichia coli* and Clones Causing Extraintestinal Infections in Humans. Animals as Reservoir of High-Risk *E. coli* Clones Pathogenic for Humans

Dr. **Jorge Blanco Álvarez**

REPORT:

That the present thesis corresponds to the work done by Ms. **Saskia Camille Flament Simon**, under my direction, and I authorize its presentation, considering that it meets the requirements required in the Regulations of Doctoral Studies of the USC, and that as its director it does not incur the causes of abstention established in Law 40/2015.

In accordance with article 37 of the Regulations for Doctoral Studies, it also declares that this doctoral thesis is suitable to be defended on the basis of the COMPENDIUM OF PUBLICATIONS, in which the participation of the doctoral student was decisive for its elaboration and the publications conform to the Research Program.

In Lugo, 15th December 2020

Signed. Jorge Blanco Álvarez

FUNDING

This study was supported by grants received by the Spanish team from projects: PI16/01477 from Plan Estatal de I+D+I 2013–2016, Instituto de Salud Carlos III, Subdirección General de Evaluación y Fomento de la Investigación, and AGL2016-79343-R from the Spanish Agencia Estatal de Investigación (AEI) from the Spanish Ministerio de Economía y Competitividad (MINECO) and Fondo Europeo de Desarrollo Regional (FEDER); and ED431C 2017/57 (Grupos de referencia competitiva) from the Consellería de Cultura, Educación e Ordenación Universitaria and Consellería de Economía, Emprego e Industria (Xunta de Galicia) and FEDER, and by the Strategic Researcher Cluster BioReDeS funded by the Regional Government Xunta de Galicia under the project no. ED431E 2018/09. It was also supported by a grant received by the French team from Project JPI-EC-AMR 2016, with l'Agence Nationale de la Recherche (France) as sponsor (No. ANR-16-JPEC-0002-04).

Saskia-Camille Flament-Simon thanks the Secretaría General de Universidades, Spanish Ministerio de Educación, Cultura y Deporte, Gobierno de España for providing her with a grant from the FPU program (FPU15/02644).

ACKNOWLEDGMENTS

This thesis could not have been achieved without the invaluable contribution and help of many people that I would like to thank.

Dr. Jorge Blanco, the director of my thesis, for all the hours he has dedicated to my attention and for the opportunity, enlightenment and achievements that I have reached.

Dr. Marie-Hélène Nicolas-Chanoine from the Université de Paris and Hőpital Beaujon, for being my international tutor, her dedication and instruction, her enthusiasm and energy and for her very important contribution to this thesis.

Dr. Maria de Toro from the Genomics and Bioinformatics Platform of the Centro de Investigación Biomédica de La Rioja (CIBIR**),** for receiving me, her instruction, introducing me to bioinformatics and for her very important contribution to this thesis.

All the members of the LREC-USC laboratory where I have been working all these years: instructors, guides, colleagues and friends.

Dr. Jesús Blanco, Dr. Miguel Blanco and **Dr. Azucena Mora**, for their instruction, not only with my laboratory skills but also related to my teacher training and for their important contribution to this thesis.

Dr. Vanesa García, for her instruction, help and important contribution to this thesis and for all the hours we have worked together.

Dr. Isidro García-Meniño and **Ms. Dafne Díaz-Jiménez**, my daily colleagues, which has made my work a pleasure and for their help and contribution to this thesis.

Dr. Liliana Costoya, Dr. Susana Viso, Dr. Alexandra Herrera, Dr. Juan Marzoa, Dr. Rosalía Mamani, Dr. Cecilia López y **Dr. Ghizlane Dahibi**, for their help and instruction at the very beginning of my phD.

I also have to thank **Dr. María Pilar Alonso** from the Hospital Universitario Lucus Augusti and **Dr. Ana Goica** from the Hostital Veterinario Rof Codina for their help and important contribution to this thesis.

Finally, **Mr. Juan Díaz-González**, for his help and contribution to this thesis.

Lastly, I would like to thank my family, my friends and Alfonso, for their love and support all these years.

INDEX

LIST OF PUBLICATIONS AND QUALITY INDEXES

The present doctoral thesis includes six studies (Flament-Simon SC *et al.* 2020 Antibiotics 9 (4) 161 doi: 10.3390/antibiotics9040161; Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125; Flament-Simon SC *et al.* 2019 Front Microbiol 10:1183 doi: 10.3389/fmicb.2019.01183; Flament-Simon SC *et al.* 2020 Microorganims 8 (11):E1712 doi: 10.3390/microorganisms8111712; Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387 and unpublished data).

1. 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. Flament-Simon SC, Nicolas-Chanoine MH, García V, Duprilot M, Mayer N, Alonso MP, García-Meniño I, Blanco JE, Blanco M, Blanco J.
   Antibiotics (Basel). 2020 Apr 4;9(4):161. doi: 10.3390/antibiotics9040161.
   2019 Journal Impact Factor: 3.893
   2019 Percentile Rank in Category PHARMACOLOGY & PHARMACY: 76.57%
   2019 Percentile Rank in Category INFECTIOUS DISEASES: 75.54%
2. High Prevalence of ST131 Subclades C2-H30Rx and C1-M27 Among Extended-Spectrum beta-Lactamase-Producing *Escherichia coli* Causing Human Extraintestinal Infections in Patients From Two Hospitals of Spain and France During 2015. Flament-Simon SC, García V, Duprilot M, Mayer N, Alonso MP, García-Meniño I, Blanco JE, Blanco M, Nicolas-Chanoine MH, Blanco J.
   Front Cell Infect Microbiol. 2020 Mar 24;10:125. doi: 10.3389/fcimb.2020.00125.
   2019 Journal Impact Factor: 4.123
   2019 Percentile Rank in Category MICROBIOLOGY: 72.43%
   2019 Percentile Rank in Category IMMUNOLOGY: 64.47%
3. Association Between Kinetics of Early Biofilm Formation and Clonal Lineage in *Escherichia coli*. Flament-Simon SC, Duprilot M, Mayer N, García V, Alonso MP, Blanco J, Nicolas-Chanoine MH.
   Front Microbiol. 2019 May 31;10:1183. doi: 10.3389/fmicb.2019.01183.
   2019 Journal Impact Factor: 4.236
   2019 Percentile Rank in Category MICROBIOLOGY: 75.37%
4. Molecular Characteristics of Extraintestinal Pathogenic *E. coli* (ExPEC), Uropathogenic *E. coli* (UPEC), and Multidrug Resistant *E.coli* Isolated from Healthy Dogs in Spain. Whole Genome Sequencing of Canine ST372 Isolates and Comparison with Human Isolates Causing Extraintestinal Infections. Flament-Simon SC, Toro M, García V, Blanco JE, Blanco M, Alonso MP, Goicoa A, Díaz-González J, Nicolas-Chanoine MH, Blanco J.
   Microorganisms. 2020 Oct 31;8(11):E1712. doi: 10.3390/microorganisms8111712.
   2019 Journal Impact Factor: 4.152
   2019 Percentile Rank in Category MICROBIOLOGY: 73.16%
5. Whole Genome Sequencing and Characteristics of mcr-1-Harboring Plasmids of Porcine *Escherichia coli* Isolates Belonging to the High-Risk Clone O25b:H4-ST131 Clade B. Flament-Simon SC, de Toro M, Mora A, García V, García-Meniño I, Díaz-Jiménez D, Herrera A, Blanco J.
   Front Microbiol. 2020 Mar 24;11:387. doi: 10.3389/fmicb.2020.00387.
   2019 Journal Impact Factor: 4.236
   2019 Percentile Rank in Category MICROBIOLOGY: 75.37

SUMMARY [RESUMEN/RESUMO/RÉSUMÉ]

SUMMARY

Escherichia coli is the leading cause of urinary tract (UTI) and bloodstream infections in humans and animals. Most infections like this are due to strains of pathotypes known as extraintestinal pathogenic *E. coli* (ExPEC) or uropathogenic *E. coli* (UPEC). Numerous virulence genes have been associated with strains causing extraintestinal infections, such as adhesins, toxins, siderophores and capsular antigens, that enable them to colonize host surfaces, capture available iron, injure host tissues and avoid host defence systems. The treatment of these infections has been seriously complicated by the appearance of multidrug-resistant (MDR) strains and especially by the rapid dissemination of extended-spectrum β-lactamase-producing *E. coli* (ESBLEC).

The present doctoral thesis includes five publications (Flament-Simon SC et al. 2020 Antibiotics 9 (4) 161 doi.org/10.3390/antibiotics9040161; Flament-Simon SC et al. 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125; Flament-Simon SC et al. 2019 Front Microbiol 10:1183 doi: 10.3389/fmicb.2019.01183; Flament-Simon SC et al. 2020 Microorganims 8 (11):E1712 doi: 10.3390/microorganisms8111712; Flament-Simon SC et al. 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387) and unpublished data with the following aims:

I. Evaluate the prevalence of sequence type 131 (ST131) and its subclones among *E. coli* strains consecutively isolated from patients with extraintestinal infections in 2016 and among ESBLEC isolated in 2015, in two hospitals (Lucus Augusti University hospital and AP-HP Beaujon hospital), paying special attention to the emergence of new highly virulent and/or MDR clones.
II. Gain knowledge of phenotypic and genotypic characteristics of the ST131 strains and other high-risk clones isolated from humans and animals, determining phylogroups, clonotypes, sequence types (STs), O:H serotypes, Virulence Factor (VF)-encoding genes, antibiotic resistance including the whole genome sequencing (WGS) of a subset of strains to perform core genome (CG) and plasmidome analysis.
III. Gain knowledge about early biofilm formation (EBF) in *E. coli* strains implicated in extraintestinal infections and contrast the biofilm-forming ability with genotypic characteristics of the strains.
IV. Study production animals (pigs) and companion animals (dogs) as reservoirs for ST131 strains and other high-risk clones for humans and investigate potential clone transfer.

Overall, we undertook a total of six studies of molecular epidemiology as well as a genomic analysis of *E. coli* strains. Within the four first studies, we analysed a total of 394 strains which were producers of extraintestinal infections retrieved from two different hospitals (Lucus Augusti hospital in Lugo and the Beaujon hospital in Paris), whereas in the remaining two studies we analysed strains from canine and porcine origin. The phenotypical characterization includes: (I) serotyping, (II) antimicrobial susceptibility testing and (III) biofilm determination. The molecular characterisation includes: (I) the description of virulence profiles through the identification of VF-encoding genes by Polymerase Chain Reaction (PCR) followed by the definition of the virulence profile status liked to pathotypes, (II) the description of the antimicrobial resistances profiles

through PCR followed by amplicon sequencing, (III) the determination of the phylogenetic groups of *E. coli* through PCR, (IV) the determination of clonotypes (*fum*C-*fim*H) and sequence types (ST) through PCR followed by amplicon sequencing, (VI) the identification of the clades and subclades of ST131 *E. coli* by PCR and (VII) the identification of the *fim*B gene by PCR followed by amplicon sequencing in a subset of strains from the human origin collection of bacteria and finally, (VIII) we performed the whole genome sequencing (WGS) of a total of 109 *E. coli* strains belonging to three different studies, allowing us to perform the core genome (CG) and plasmidome analysis. Statistical analysis include Shapiro-Wilk normality test, non-parametric Mann-Whitney-Wilcoxon test, two-tailed Fisher's exact test and analysis of variance (ANOVA) test. P-values < 0.05 were considered statistically significant. Correlation between quantitative variables was assessed by Pearson correlation coefficient, and the assessment of repeatability was via the Bland and Altman, (1986) statistical method.

In the first study, we undertake a molecular analysis of a total of 196 *E. coli* strains derived from consecutive cases of extraintestinal infections throughout 2016 in Spain (100 from Lucus Augusti hospital in Lugo) and France (96 from Beaujon hospital in Clichy). We determined phylogroups, clonotypes, sequence types (STs), O:H serotypes, virulence factor (VF)-encoding genes and antibiotic resistance. We observed that approximately 10% of the infections were caused by ST131 strains in both hospitals and approximately 60% of these infections were caused by strains belonging to only 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 and ST141). ST88 strains were frequent, especially in Spain, while ST141 strains significantly predominated in France. The 23 ST131 strains displayed four clonotypes: CH40-30, CH40-41, CH40-22 and CH40-298. Only 13 (6.6%) strains were carriers of extended-spectrum beta-lactamase (ESBL) enzymes. However, 37.2% of the strains were multidrug-resistant (MDR). Approximately 40% of the MDR strains belonged to only four of the dominant clones (B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 and D-CH35-27-ST69). Among the remaining MDR strains, two strains belonged to B2-CH14-64-ST1193, i.e., the new global emergent MDR clone. Moreover, a hybrid ExPEC-enteroaggregative strain belonging to the A-CH11-54-ST10 clone was identified.

In the second study, we completed a molecular analysis of a total of 188 extended-spectrum β-lactamase-producing *Escherichia coli* (ESBLEC) collected throughout 2015 in the Lucus Augusti University hospital and the AP-HP Beaujon hospital, with regard to other STs and to characterize the types of ESBL produced, serotypes, virulence factor (VF)-encoding genes and the ST131 clades and subclades**.** ST131 was detected in 33 (39.1%) and 46 (47.9%) of the strains in Lucus Augusti and Beaujon, respectively. The 109 remaining strains displayed 57 other STs, the following STs being displayed by at least three strains: ST10 (8 strains), ST23 (3), ST38 (4), ST58 (3), ST88 (5), ST95 (4), ST167 (3), ST354 (5), ST361 (3), ST410 (6), ST648 (4), ST744 (3), and ST1615 (6). ST354, ST410 and ST1615 were significantly (P< 0.05) more frequent in Lucus Augusti (5.4%, 6.5% and 6.5%) than in Beaujon (0% for the three STs). Furthermore, we described two strains belonging to the emergent clone B2-CH14-64-ST1193. CTX-M-15 was the commonest ESBL detected in the two hospitals (44.6% in Lucus Augusti and 50.0% in Beaujon). CTX-M-14 was significantly (P=0.0003) more frequent in Lucus Augusti (31.5%) than in Beaujon (10.4%), whereas CTX-M-1 (20.8% *versus* 7.6%; P = 0.008) and CTX-M-27 (15.6% *versus* 6.5%; P=0.0389) were more frequent in Beaujon than in Lucus Augusti. The ST131 strains showed a higher virulence score (mean 13.367) compared with the non-ST131 strains (mean 7.661) (P<0.001). Among the 79 ST131

strains, most of them (52; 65.8%) belonged to subclade C2 (also known as subclone *H*30Rx) followed by those belonging to subclade C1 (cluster C1-M27: 16 strains, 20.3%; cluster non-C1-M27: 6 strains, 7.6%) and clade A (4 strains; 5.1%). Besides, the C2 subclade strains showed a higher VF-encoding gene score (mean 14.250) compared with the C1-M27 cluster strains (mean 10.875) (P<0.001).

In the third study, we analyzed the biofilm formation of *E. coli* strains that produced extraintestinal infections in humans. *E. coli* biofilm formation has mostly been assessed in specific pathogenic *E coli* groups. Here, we assessed the early biofilm formation (EBF), i.e. adhesion stage, using the BioFilm Ring Test® on 394 *E. coli* clinical isolates (EC) [196 consecutively isolated (CEC) in 2016 and 198 ESBL-producing *E. coli* (ESBLEC) isolated in 2015]. Then, biofilm-forming ability was contrasted with phylogroups, clonotypes (CH) and sequence types (ST), all being used to define clones, virulence factors (VF) and FimB. According to both biofilm production levels at 2h, 3h and 5h, and EBF kinetics over 5h, CEC and ESBLEC strains segregated into three EBF groups: strong (G1), moderate (G2) and weak (G3) producers. At 2h, strong producers were more frequent among CEC (n=28; 14.3%) than among ESBLEC (n=8; 4%) (P=0.0004). As CEC and ESBLEC strains showed similar individual EBF kinetics in each group, comparison of strain features between each group was applied to gathered CEC and ESBLEC strains and after 2h incubation, 2h being the most representative point time of the CEC and ESBLEC strain segregation into the three groups. Phylogroup B2 displayed by 51.3% of the 394 strains was more frequent in G1 (77.8%) than G3 (47.6%) (P=0.0006). The 394 strains displayed 153 clones of which 31 included at least three strains. B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5/14-ST141 and E-CH100-96-ST362 clones were associated with G1 (P<0.03) and accounted for 41.7% of G1 strains. B2-CH40-30-ST131 clone was associated with G3 (P<0.0001) and accounted for 25.5% of G3 strains. VF mean was higher among G1 than among G3 strains (P<0.001). FimB-P2 variant was associated with G1 (P=0.0011) and FimB-P1 variant with G3 (P=0.0023). Clone, some VF, and FimB were associated with EBF, clonal lineage being able to explain 72 % of the variability of EBF. Among our 394 strains, less than a 10% of the strains were able to quickly and persistently produce high levels over 5h. These strains belong to a few clones previously described in various studies as dominant gut colonisers in mammalians and birds and comprised B2-CH40-22-ST131clone i.e. the ancestor of the globally disseminated B2-CH40-30-ST131 clone that is the dominant clone among the weak biofilm producers (G3).

In the fourth study, we investigate the insights of ST131 clonal complex ESBLEC (STC131 group constituted by 30 genomes) and other emerging ESBL lineages (non-STC131 group constituted by 45 genomes) through the analysis of a total of 75 ESBLEC whole sequenced genomes (51 genomes from Lucus Augusti hospital and 24 genomes from Beaujon hospital). These genomes come from different clinical samples: 57 isolates from urine, 11 from blood and 7 from various other sources including one from bile, one from ascitic fluid, one from a bone, two from wounds and two from digestive infections. We carried out a phylogenetic research that includes the core genome (CG) and plasmidome analysis. Through the CG analysis, we observed correlation with the phylogenetic groups of the genomes and described two new clades among the ST131 clonal complex represented by strains with the *fim*H5 and *fim*H35 alleles. We *in silico* investigated the presence of 164 VF-encoding genes and 13 point mutations (PMs) related to the expression of resistance to antibiotics in the chromosome of 75 ESBLEC genomes. Comparatively, STC131 group of genomes showed a higher average of virulence genes

than the non-STC131 group of genomes (62 *versus* 47) and a higher mean of PMs than the non-STC131 group (5 *versus* 3). Besides, there was a clear association between ExPEC status, UPEC status and the STC131 group of genomes (P<0.0001). The 75 ESBLEC genomes harbours 21 presumptive integrative conjugative elements (ICEs) and a total of 228 plasmids, of which 86 (32%) were found within the STC131 group and 142 (62%) were found within the non-STC131 group. The 92% of the plasmid belonged to MOB_F, MOB_P and MOB_Q families of relaxases. Besides, we described a wide variety of incompatibility groups (Inc.) and this diversity was higher within the non-STC131 than the STC131 group (12 Inc. including Col-like, F, Q, I, X, Y, P, 13, HI2, N, R and B/O/K/Z *versus* 6 Inc. including Col-like, F, X, 13, N, and B/O/K/Z, respectively). We noticed that Col-like and IncF plasmids were equally dominant within both groups of genomes. The IncF plasmids vary greatly in their composition and carried almost all the virulence and resistance genes found in the plasmidome of the 75 ESBLEC analysed genomes. Moreover, we localized the genes encoding for ESBL enzymes in 67 out of the 75 genomes, of which 21 were integrated in the chromosome (31%) and 46 of them were carried by plasmids (69%). Furthermore, we observed a correlation between the presence of specific types of IncF and the different ST131 clades. We found two *bla*$_{\text{CTX-M-27}}$ codifying genes carried by [F1:A2:B20]-IncF plasmids and five *bla*$_{\text{CTX-M-15}}$ codifying genes carried by [F2:A1:B-]-IncF plasmids associated with subclades C1 and C2, respectively. Thus, that argues for the existence of a plasmid-clade association. Regarding clade A, we describe the presence of [F29:A-:B10]-IncF plasmids and ICEs as distinctive features from genomes from clade C.

In the fifth study, under a "one health" perspective and the worldwide antimicrobial resistance concern, we investigated ExPEC, UPEC, and MDR *E. coli* from 197 strains recovered from healthy dogs in Spain between 2013 and 2017. A total of 91 (46.2%) strains were molecularly classified as ExPEC and/or UPEC, including 50 clones, among which four clones were dominant (B2-CH14-180-ST127, B2-CH52-14-ST141, B2-CH103-9-ST372 and F-CH4-58-ST648) and 15 had been identified among strains causing extraintestinal infections in Spanish and French humans in 2015 and 2016. A total of 28 (14.2%) strains were classified as MDR, associated with B1, D, and E phylogroups, and included 24 clones, of which eight had been identified among the human clinical isolates. We selected 23 ST372 strains, 21 from healthy dogs, and two from human clinical isolates for WGS and built an SNP-tree with these 23 genomes and 174 genomes obtained from public databases (128 from canine origin strains and 46 from human origin strains). These 197 genomes were segregated into six clusters. Cluster 1 comprised 74.6% of the strain genomes, mostly composed of canine origin strain genomes (P< 0.00001). Clusters 4 and 6 also included canine origin strain genomes, while clusters 2, 3, and 5 were significantly associated with human origin strain genomes. Finding several common clones and clone-related serotypes in dogs and humans suggests a potentially bidirectional clone transfer that argues for the one health perspective.

In the sixth study, we analysed strains belonging to the clone ST131 from porcine origin, which are scarcely documented. We undertook the WGS, CG and plasmidome analysis of seven isolates collected from diarrhoeic piglets and four from pork meat. All of the 11 ST131 strains belonged to serotype O25b:H4 and clade B, and showed *fimH*22 allele or mutational derivatives. The 11 porcine strains possessed virulence traits that classified the strains as avian pathogenic, uropathogenic and extraintestinal pathogenic *E. coli* like status (APEC, UPEC and ExPEC like) and constituted virotype D. The CG was performed for all porcine strains in addition to 73 ST131 reference strains from different

origins. Within clade B the CG showed nine sub-clusters, allowing us to describe five new subclades (B6, B6-like, B7, B8, B9). There was an association between subclade B6, PST43, virotype D2 and food origin, while subclade B7 included PST9 strains with virotype D5 from diarrhoeic piglets (P=0.007). The distance between human and porcine origin strains from subclades B6 and B7 had an average of 20 SNP/Mb and 15 SNP/Mb, respectively. [F2:A-:B1]-IncF, ColE1-like and IncX plasmids were the most prevalent. Besides, IncF plasmids harboured a ColV region frequent among APEC strains. Antimicrobial resistance genes conferring resistance to penicillin, tetracycline, quinolones and colistin were the most common. The *mcr*-1.1 gene was detected in 5 of 11 porcine origin strains, integrated into the chromosome of one strain and into plasmids in the remainder strains (two MOB_{H11}/IncHI2-ST4, one MOB_{P3}/IncX4 and one MOB_{F12}/IncF [F2:A-:B1] supposedly cointegrated with an IncHI2). The surrounding environments of the *mcr*-1 cassette showed variability. However, there were conserved structures within the same plasmid family. In conclusion, CG analysis defined five new subclades. The ST131 porcine strains belonged to new subclades B6 and B7. Moreover, porcine and clinical human strains were strongly related. The 11 porcine ST131 strains harboured a wide variety of plasmids, virulence and resistance genes. Furthermore, epidemic plasmids IncX4 and IncHI2 are responsible for the acquisition of *mcr*-1.1 gene. We hypothesize that the APEC-IncF plasmid acquired the *mcr*-1.1 gene via cointegrating an IncHI2 plasmid, which is worrying due to combination of virulence and resistance attributes in a single mobile genetic element.

Of the results obtained, we concluded:

1. Approximately 10% of the extraintestinal *E. coli* infections that had occurred in 2016 in the two studied hospitals were caused by ST131 strains, and approximately 60% of these infections were caused by strains belonging to only 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 and ST141).
2. The majority of ST12, ST73, ST95 and ST141 human extraintestinal strains were susceptible to most antibiotics, indicating that MDR was not the reason for their success. The results of the present study support the idea that their success is mainly due to the high number of VF-encoding genes that they possess.
3. Among the MDR human extraintestinal strains, four clones are predominant, especially: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 and D-CH35-27-ST69.
4. Despite the enormous genetic diversity observed in our ESBLEC collection (71 clones amongst 188 ESBLEC), it can be concluded that the majority of the strains belong to only three clonal complexes (CC10, CC23, and CC131) and that ST131 subclade C2 associated with the production of CTX-M-15 remains the most prevalent *E. coli lineage* among the ESBLEC strains identified in the studied Spanish and French hospitals.
5. Among our 394 extraintestinal *E. coli* strains, <10% are able to quickly and persistently produce high biofilm levels over 5 h. These strains belong to a few clones previously described in various studies as dominant gut colonizers in mammalians and birds, and comprised the B2-CH40-22-ST131 clone, i.e., the ancestor of the globally disseminated B2-CH40-30-ST131 clone that is the dominant clone among the weak biofilm producers.

6. The core genome analysis of ESBLEC highlights genomic diversity among ESBL *E. coli* population. That includes representatives from phylogroups A, B1, B2, C, E and F. We described two new clades in ST131 clonal complex represented by genomes harbouring the *fim*H5 and *fim*H35 alleles. Clade C is characterised by showing less SNPs distance in the core genome than its sister clades A and B.
7. There is a correlation between VF-encoding gene profiles and STs.
8. There is an impressive diversity of plasmids among ESBLEC. However, MOB_F, MOB_P and MOB_Q families of relaxases represent 92% of the plasmidome and Col-like and IncF are the most frequent incompatibility groups. Besides, most of the virulence and resistance associated encoding genes are carried by plasmids belonging to the IncF family.
9. In ST131 lineage, the plasmid incompatibility group diversity is higher among subclade C2 than among clade A and cluster C1-M27 subclade C1. Clade A is related with the presence of ICEs and [F29:A-:B10]-IncF plasmids.
10. The intestinal tract of healthy dogs appears as an important reservoir of ExPEC and/or UPEC, and, in a lesser extent, of MDR *E. coli* strains.
11. Among the canine strains displaying an ExPEC and/or UPEC status, clone B2-CH103-9-ST372 was dominant.
12. WGS analysis suggests that canine strains of clone B2-CH103-9-ST372, belonging to cluster 1 (serotypes O4:H31, O21:H31 and O83:H31) and cluster 6 (serotype O6:H31) might cause extraintestinal infections in both humans and dogs, whereas those strains of this clone belonging to cluster 2 and serotypes O18:H31 and O45:H31 might cause only human infections.
13. Clade B of ST131 showed a huge genetic diversity, and five new subclades were defined (B6, B6-like, B7, B8, and B9). The majority of ST131 porcine strains belong to new subclades B6 and B7.
14. Most porcine ST131 strains are MDR (91%) and carry many antibiotic resistance genes. Colistin resistance was introduced through MGEs and had been able to stabilize chromosomally. IncX4 and IncHI2 epidemic *mcr-1*–harboring plasmids are responsible for acquired colistin resistance encoded by *mcr-1.1* gene. The surrounding environment of the *mcr-1* cassette is variable but within the same family of plasmids, insertion, and stabilization had common structures.

Keywords: *Escherichia coli, E. coli*, clonal structure, biofilm formation, virulence genes, antibiotic resistance, β-lactamases, CTX-M, ESBL, ExPEC, UPEC, high-risk clones, ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131, ST141, ST372, ST1193, dogs, swine, colistin, *mcr-1.1*, WGS, core genome, plasmidome.

RESUMEN

Escherichia coli es la principal causa de infecciones del tracto urinario (ITU) y del torrente sanguíneo en seres humanos y animales. La mayoría de estas infecciones se deben a cepas de patotipos conocidos como *E. coli* patógenos extraintestinales (ExPEC) o *E. coli* uropatógenos (UPEC). Se han asociado numerosos genes de virulencia con cepas que causan infecciones extraintestinales, como adhesinas, toxinas, sideróforos y antígenos capsulares, que les permiten colonizar las superficies del huésped, capturar el hierro disponible, dañar los tejidos del hospedador y evitar los sistemas de defensa del hospedador. El tratamiento de estas infecciones se ha complicado seriamente por la aparición de cepas multirresistentes (MDR) y especialmente por la rápida diseminación de cepas de *E. coli* productoras de β-lactamasas de espectro extendido (ECBLEE).

La presente tesis doctoral incluye cinco publicaciones (Flament-Simon SC *et al.*2020 Antibiotics 9 (4) 161 doi.org/10.3390/antibiotics9040161; Flament-Simon SC *et al.*2020 Front Cell Infect Microbiol 10: 125 doi: 10.3389/fcimb .2020.00125; Flament-Simon SC *et al.*2019 Front Microbiol 10: 1183 doi: 10.3389/fmicb.2019.01183; Flament-Simon SC *et al.*2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms8111712; Flament-Simon SC *et al*.2020 Front Microbiol 2020 11: 387 doi: 10.3389/fmicb.2020.00387) y datos sin publicar con los siguientes objetivos:

I. Evaluar la prevalencia de la secuencia tipo 131 (ST131) y sus subclones entre cepas de *E. coli* causantes de infecciones extraintestinales en seres humanos aisladas de casos consecutivos en 2016 y entre cepas ECBLEE aisladas en 2015, en dos hospitales (hospital universitario Lucus Augusti y hospital AP-HP Beaujon), prestando especial atención a la aparición de nuevos clones altamente virulentos y/o MDR.
II. Adquirir conocimiento de las características fenotípicas y genotípicas de las cepas ST131 y otros clones de alto riesgo aislados de seres humanos y animales, determinando los filogrupos, clonotipos, secuencias tipo (ST), serotipos O:H, genes que codifican Factores de Virulencia (FV), resistencia a los antibióticos, además de la secuenciación del genoma completo (*whole genome sequencing*) (WGS) de un subconjunto de cepas para realizar el análisis de su genoma esencial (*core genome*) (CG) y de su plasmidoma.
III. Adquirir conocimientos sobre la formación temprana de biopelículas (*early biofilm formation*) (EBF) en cepas de *E. coli* implicadas en infecciones extraintestinales y comparar la capacidad de formación de biopelículas con las características genotípicas de las cepas.
IV. Estudiar animales de producción (cerdos) y animales de compañía (perros) como reservorios de cepas ST131 y otros clones de alto riesgo para seres humanos e investigar la posible transferencia de clones.

En total realizamos seis estudios de epidemiología molecular, así como un análisis genómico de cepas de *E. coli*. En los cuatro primeros estudios analizamos un total de 394 cepas productoras de infecciones extraintestinales extraídas de dos hospitales (hospital Lucus Augusti de Lugo y hospital Beaujon de París), mientras que en los dos estudios restantes analizamos cepas de origen canino y porcino. La caracterización fenotípica incluye: (I) la serotipificación, (II) pruebas de susceptibilidad antimicrobiana y (III) la

determinación de biopelículas. La caracterización molecular incluye: (I) la descripción de los perfiles de virulencia a través de la identificación de genes que codifican FV mediante la reacción en cadena de la polimerasa (PCR) seguida de la definición de los estatus de virulencia vinculados a patotipos, (II) la descripción de perfiles de resistencias a los antibióticos mediante PCR seguida de la secuenciación de amplicones, (III) la determinación de los grupos filogenéticos de *E. coli* mediante PCR, (IV) la determinación de clonotipos (*fum*C-*fim*H) y secuencias tipo (ST) mediante PCR seguida de la secuenciación de amplicones, (VI) la identificación de los clados y subclados de *E. coli* ST131 por PCR, (VII) la identificación del gen *fim*B por PCR seguida de la secuenciación de amplicones en un subconjunto de cepas de la colección de bacterias de origen humano y finalmente, (VIII) realizamos la secuenciación del genoma completo (WGS) de un total de 109 cepas de *E. coli* pertenecientes a tres estudios diferentes, permitiéndonos realizar el análisis del CG y del plasmidoma. El análisis estadístico incluye la prueba de normalidad de Shapiro-Wilk, la prueba no paramétrica de Mann-Whitney-Wilcoxon, la prueba exacta de Fisher y la prueba de análisis de varianza (ANOVA). Los valores $P<0,05$ se consideraron estadísticamente significativos. La correlación entre variables cuantitativas se evaluó mediante el coeficiente de correlación de Pearson, y la evaluación de la reproductibilidad se realizó mediante el método estadístico de Bland y Altman, (1986).

En el primer estudio realizamos el análisis molecular de un total de 196 cepas de *E. coli* derivadas de casos consecutivos de infecciones extraintestinales aisladas a lo largo del año 2016 en España (100 cepas procedentes del hospital Lucus Augusti de Lugo) y Francia (96 cepas procedentes del hospital Beaujon de Clichy). Determinamos los filogrupos, clonotipos, secuencias tipo (ST), serotipos O:H, genes que codifican FV y genes que codifican resistencia a los antibióticos. Observamos que aproximadamente el 10% de las infecciones fueron causadas por cepas del clon ST131 en ambos hospitales y que aproximadamente el 60% de estas infecciones fueron causadas por cepas pertenecientes a solo 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 y ST141). Las cepas ST88 fueron frecuentes, especialmente en España, mientras que las cepas ST141 fueron significativamente predominantes en Francia. Identificamos 23 cepas ST131 que representaron cuatro clonotipos: CH40-30, CH40-41, CH40-22 y CH40-298. Sólo 13 (6,6%) cepas fueron portadoras de β-lactamasas de espectro extendido (BLEE). Sin embargo, el 37,2% de las cepas fueron MDR. Aproximadamente el 40% de las cepas MDR pertenecieron a tan sólo cuatro de los clones dominantes (B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 y D-CH35-27-ST69). Entre las cepas MDR restantes, dos pertenecieron al clon B2-CH14-64-ST1193, es decir, el nuevo clon MDR emergente a nivel mundial. Además, identificamos una cepa híbrida ExPEC-enteroagregativa perteneciente al clon A-CH11-54-ST10.

En el segundo estudio, completamos el análisis molecular de un total de 188 cepas de *E. coli* productoras BLEE recolectadas a lo largo del año 2015 en el hospital universitario Lucus Augusti y el hospital AP-HP Beaujon, con objeto de analizar los clados y subclados del clon ST131, otros STs y caracterizar los tipos de enzimas BLEE producidos, los serotipos y los genes que codifican FV. ST131 se detectó en 33 (39,1%) y 46 (47,9%) de los aislamientos procedentes del hospital Lucus Augusti y del hospital Beaujon, respectivamente. Las 109 cepas restantes constituyeron 57 STs diferentes, donde las siguientes STs incluyeron al menos tres representantes: ST10 (8 cepas), ST23 (3), ST38 (4), ST58 (3), ST88 (5), ST95 (4), ST167 (3), ST354 (5), ST361 (3), ST410 (6), ST648 (4), ST744 (3) y ST1615 (6). ST354, ST410 y ST1615 fueron significativamente ($P<0,05$)

más frecuentes en Lucus Augusti (5,4%, 6,5% y 6,5%) que en Beaujon (0% para las tres STs). Además, describimos dos cepas pertenecientes al clon emergente B2-CH14-64-ST1193. CTX-M-15 fue la enzima BLEE detectada con mayor frecuencia en los dos hospitales (44,6% en Lucus Augusti y 50,0% en Beaujon). CTX-M-14 fue significativamente (P=0,0003) más frecuente en Lucus Augusti (31,5%) que en Beaujon (10,4%), mientras que CTX-M-1 (20,8% *versus* 7,6%; P=0,008) y CTX-M-27 (15,6% *versus* 6,5%; P=0,0389) fueron más frecuentes en Beaujon que en Lucus Augusti. Las cepas ST131 presentaron una media de genes que codifican FV superior (media 13,367) en comparación con las cepas no-ST131 (media 7,661) (P <0,001). De entre las 79 cepas ST131, la mayoría (52; 65,8%) pertenecieron al subclado C2 (también conocido como subclón H30Rx), seguidas por cepas pertenecientes al subclado C1 (grupo C1-M27: 16 cepas, 20,3%; grupo no C1-M27: 6 cepas, 7,6%) y al clado A (4 cepas; 5,1%). Además, las cepas del subclado C2 mostraron una media de genes que codifican FV superior (media 14,250) que las cepas pertenecientes al grupo C1-M27 (media 10,875) (P <0,001).

En el tercer estudio, analizamos la formación biopelículas en cepas de *E. coli* productoras de infecciones extraintestinales en seres humanos. La formación de biopelículas en *E. coli* ha sido evaluada principalmente en grupos específicos de *E. coli* patógenos. Aquí evaluamos la formación temprana de biopelículas (EBF), es decir, la etapa de adhesión, utilizando el BioFilm Ring Test® en 394 aislados de origen clínico de *E. coli* (EC) [196 casos consecutivos (ECC) en 2016 y 198 *E. coli* productores de BLEE (ECBLEE) aislados en 2015]. Tras lo cual, la capacidad de formación de biopelículas se contrastó con los filogrupos, clonotipos (CH) y secuencias tipo (STs) utilizados para definir los clones, los factores de virulencia (FV) y FimB. De acuerdo con los niveles de producción de biopelículas a las 2h, 3h y 5h, y la cinética de EBF durante 5h, las cepas pertenecientes a ECC y ECBLEE se segregaron en tres grupos de EBF: fuertes productores (G1), moderados productores (G2) y débiles productores (G3). A las 2h, los fuertes productores fueron más frecuentes entre ECC (n=28; 14,3%) que entre ECBLEE (n=8; 4%) (P=0,0004). Como las cepas de ECC y ECBLEE mostraron cinéticas individuales de EBF similares por grupo, la comparación de las características de las cepas pertenecientes a cada grupo en las colecciones ECC y ECBLEE se aplicaron de forma conjunta tras 2h de incubación, siendo las 2h el tiempo de lectura más representativo de la segregación de cepas de ECC y ECBLEE en tres grupos. El filogrupo B2 representado por el 51,3% de las 394 cepas fue más frecuente en G1 (77,8%) que en G3 (47,6%) (P=0,0006). Las 394 cepas constituyeron 153 clones de los cuales 31 incluyeron al menos tres representantes. Los clones B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5/14-ST141 y E-CH100-96-ST362 se asociaron con G1 (P<0,03) representando el 41,7% de las cepas de G1. El clon B2-CH40-30-ST131 se asoció con G3 (P <0,0001) representando el 25,5% de las cepas de G3. La media de FV fue mayor entre G1 que entre G3 (P<0,001). La variante proteica FimB-P2 se asoció con G1 (P=0,0011) y la variante proteica FimB-P1 con G3 (P=0,0023). Los clones, algunos FV y FimB se asociaron con EBF, siendo el grupo clonal explicativo del 72% de la variabilidad de EBF. De las 394 cepas analizadas, menos del 10% fueron capaces de producir niveles altos de biopelícula de forma rápida y persistente durante 5 h. Estas cepas pertenecieron a un número limitado de clones descritos en varios estudios anteriores como colonizadores intestinales dominantes en mamíferos y aves, incluyendo el clon B2-CH40-22-ST131, es decir, el antepasado del globalmente extendido clon B2-CH40-30-ST131 dominante en el grupo de cepas débiles productoras de biopelículas (G3).

En el cuarto estudio, buscamos profundizar en los conocimientos sobre el complejo clonal ST131 productor de BLEE de *E. coli* (grupo STC131 constituido por 30 genomas) y otros linajes de *E. coli* productores de BLEE emergentes (grupo no-STC131 constituido por 45 genomas), a través del análisis de la secuenciación del genoma completo de un total de 75 cepas BLEE de *E. coli* (51 genomas del hospital Lucus Augusti y 24 genomas del hospital de Beaujon). Estos genomas provienen de diferentes muestras clínicas: 57 aislados de orina, 11 de sangre y 7 de diversas fuentes distintas, incluyendo una de bilis, una de líquido ascítico, una de hueso, dos de heridas y dos de infecciones digestivas. Llevamos a cabo una investigación filogenética que incluyó el análisis del CG y del plasmidoma. A través del análisis del CG, observamos correlación con los grupos filogenéticos de los genomas y describimos dos nuevos clados pertenecientes al complejo clonal ST131 representados por cepas con los alelos *fim*H5 y *fim*H35. Investigamos *in silico* la presencia de 164 genes que codifican FV y 13 mutaciones puntuales (MP) relacionadas con la expresión de resistencias a los antibióticos en el cromosoma de los 75 genomas *E. coli* BLEE. El grupo de genomas STC131 presentó un promedio mayor de genes que codifican FV que el grupo de genomas no-STC131 (62 *versus* 47) y un promedio mayor de MP que el grupo no-STC131 (5 *versus* 3). Además, observamos una clara asociación entre el estatus ExPEC, el estatus UPEC y el grupo de genomas STC131 ($P<0,0001$). Los 75 genomas BLEE albergaron 21 presuntos elementos conjugativos integradores (*integrative conjugative elements*) (ICEs) y un total de 228 plásmidos, de los cuales 86 (32%) hallados en el grupo STC131 y 142 (62%) hallados en el grupo no-STC131. El 92% de los plásmidos pertenecieron a las familias de relaxasas MOB_F, MOB_P y MOB_Q. Además, describimos una amplia diversidad de grupos de incompatibilidad (Inc.) siendo esta diversidad mayor en el grupo no-STC131 frente al grupo STC131 (12 Inc. incluyendo Col-like, F, Q, I, X, Y, P, 13, HI2, N, R y B/O/K/Z *versus* 6 Inc. incluyendo Col-like, F, X, 13, N y B/O/K/Z, respectivamente). Observamos que los plásmidos de tipo Col e IncF fueron dominantes en ambos grupos de genomas. Los plásmidos IncF presentaron una composición variable siendo portadores de casi todos los genes de virulencia y resistencia hallados en el plasmidoma de los 75 genomas BLEE analizados. Además, localizamos los genes que codifican las enzimas BLEE en 67 de los 75 genomas, de los cuales 21 estaban integrados en el cromosoma (31%) y 46 estaban vehiculados por plásmidos (69%). También observamos una correlación entre la presencia de tipos específicos de IncF y los diferentes clados de ST131. Encontramos dos genes codificadores de $bla_{\text{CTX-M-27}}$ vehiculados por plásmidos [F1:A2:B20]-IncF y cinco genes codificadores de $bla_{\text{CTX-M-15}}$ vehiculados por plásmidos [F2:A1:B-]-IncF asociados con los subclados C1 y C2, respectivamente. Todo ello aboga la existencia de una asociación plásmido-clado. Con respecto al clado A, describimos la presencia de plásmidos [F29:A-:B10]-IncF y la presencia de ICEs como características distintivas de los genomas del clado C.

En el quinto estudio, bajo la perspectiva de "una sola salud" y la preocupación mundial por el incremento de las resistencias a los antimicrobianos, investigamos 197 cepas de *E. coli* aisladas de perros sanos en España entre 2013 y 2017 para establecer las que eran ExPEC, UPEC y MDR. Un total de 91 (46,2 %) cepas se clasificaron a nivel molecular como ExPEC y/o UPEC, incluyendo 50 clones, entre los cuales cuatro clones fueron dominantes (B2-CH14-180-ST127, B2-CH52-14-ST141, B2-CH103-9-ST372 y F-CH4-58-ST648) y 15 habían sido previamente identificados entre las cepas que causaron infecciones extraintestinales en seres humanos españoles y franceses en 2015 y 2016. Un total de 28 (14.2%) cepas se clasificaron como MDR, asociadas a los filogrupos

B1, D y E, y representadas por 24 clones, de los cuales ocho habían sido identificados previamente entre los aislados clínicos humanos. Seleccionamos 23 cepas ST372, 21 cepas procedentes de perros sanos y dos cepas procedentes de aislados clínicos humanos para WGS, tras lo cual construimos un árbol filogenético basado en (*single nucleotide polymorphisms*) SNPs con estos 23 genomas y otros 174 genomas obtenidos de bases de datos públicas (128 procedentes de cepas de origen canino y 46 procedentes de cepas de origen humano). Estos 197 genomas se segregaron en seis grupos. El grupo 1 comprendió el 74,6% de los genomas, constituido en su mayoría por genomas de cepas de origen canino ($P<0,00001$). Los grupos 4 y 6 también incluyeron genomas de cepas de origen canino, mientras que los grupos 2, 3 y 5 se asociaron significativamente con genomas de cepas de origen humano. Encontramos varios clones con serotipos comunes entre los presentes en perros y seres humanos, lo cual sugiere una potencial transferencia bidireccional de clones, que respalda la perspectiva de un mundo con "una sola salud".

En el sexto estudio analizamos cepas pertenecientes al clon ST131 de origen porcino, las cuales están escasamente documentadas. Realizamos WGS, analizamos el CG y el plasmidoma de siete cepas procedentes de lechones con diarrea y cuatro cepas procedentes de carne de cerdo. Las 11 cepas pertenecieron al serotipo O25b:H4 y al clado B de ST131 presentando el alelo *fim*H22 o mutaciones alélicas derivadas del mismo. Las 11 cepas codificaron FV que las clasificaron como equivalentes a patógenos aviares, uropatógenos y patógenos extraintestinales de *E. coli* (equivalentes a APEC, UPEC y ExPEC) y constituyeron el virotipo D. El CG se realizó incluyendo los genomas de todas las cepas de origen porcino además de 73 cepas ST131 de referencia con distintos orígenes. Dentro del clado B, el CG se segregó dando lugar a nueve subgrupos, lo que nos permitió describir cinco nuevos subclados (B6, B6-like, B7, B8, B9). Observamos una asociación entre las cepas pertenecientes al subclado B6, PST43, virotipo D2 y con origen alimentario, y una asociación entre cepas pertenecientes al subclado B7, PST9, virotipo D5 y procedentes de lechones con diarrea ($P=0,007$). La distancia entre los genomas de origen humano y porcino de los subclados B6 y B7 tuvo un promedio de 20 SNP/Mb y 15 SNP/Mb, respectivamente. Los plásmidos [F2:A-:B1]-IncF, tipo ColE1 e IncX fueron los más prevalentes. Además, los plásmidos IncF codificaron una región ColV frecuente entre cepas APEC. Los genes de resistencia antimicrobiana hallados con mayor frecuencia fueron aquellos que confieren resistencia a la penicilina, tetraciclina, quinolonas y colistina. El gen *mcr*-1.1 se detectó en 5 de las 11 cepas de origen porcino, integrado en el cromosoma de una cepa y vehiculado por plásmidos en el resto (dos MOB_{H11}/IncHI2-ST4, un MOB_{P3}/IncX4 y un MOB_{F12}/IncF [F2:A-:B1] supuestamente cointegrado con un IncHI2). Los entornos del casete *mcr*-1 mostraron variabilidad. Sin embargo, observamos estructuras conservadas dentro de una misma familia de plásmidos. En conclusión, el análisis del CG definió cinco nuevos subclados. Las cepas ST131 de origen porcino pertenecen a los nuevos subclados B6 y B7. Además, las cepas de origen clínico humanas y las cepas de origen porcino están estrechamente relacionadas. Las 11 cepas ST131 porcinas albergaron una amplia variedad de plásmidos, genes de virulencia y resistencia. Además, los plásmidos epidémicos IncX4 e IncHI2 son responsables de la adquisición del gen *mcr*-1.1. Establecemos la hipótesis de que el plásmido APEC-IncF adquirió el gen *mcr*-1.1 mediante la cointegración de un plásmido IncHI2, lo cual es preocupante debido a la combinación de atributos de virulencia y resistencia en un solo elemento genético móvil.

De los resultados obtenidos, concluimos:

1. Aproximadamente el 10% de las infecciones extraintestinales causadas por *E. coli* acontecidas en 2016 en los dos hospitales estudiados fueron causadas por cepas del clon ST131, y aproximadamente el 60% de estas infecciones fueron causadas por cepas pertenecientes a solo 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 y ST141).
2. La mayoría de las cepas ST12, ST73, ST95 y ST141 causantes de infecciones extraintestinales en seres humanos fueron susceptibles a la mayoría de los antibióticos, lo que indica que la MDR no fue la razón de su éxito. Los resultados del presente estudio apoyan la idea de que su éxito se debe principalmente a que albergan una gran cantidad de genes que codifican FV.
3. Entre las cepas MDR causantes de infecciones extraintestinales en seres humanos, predominan cuatro clones: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 y D-CH35-27-ST69.
4. A pesar de la enorme diversidad genética observada en nuestra colección de cepas de *E. coli* productoras de BLEE (71 clones presentes en 188 cepas BLEE), se puede concluir que la mayoría de las cepas pertenecen a solo tres complejos clonales (CC10, CC23 y CC131) y que el subclado C2 de ST131 asociado a la producción de CTX-M-15 sigue siendo el linaje de *E. coli* más prevalente entre las cepas de *E. coli* productoras de BLEE identificadas en los hospitales español y francés estudiados.
5. Entre nuestras 394 cepas de *E. coli* procedentes de infecciones extraintestinales, <10% son capaces de producir niveles altos de biopelícula de forma rápida y persistente durante 5 h. Estas cepas pertenecen a unos pocos clones descritos previamente en varios estudios como colonizadores intestinales dominantes en mamíferos y aves, e incluyen al clon B2-CH40-22-ST131, es decir, el antepasado del globalmente extendido clon B2-CH40-30-ST131, que es el clon dominante entre las cepas débiles productores de biopelículas.
6. En el análisis del CG de cepas de *E. coli* productoras de BLEE destaca la diversidad genómica, incluyendo representantes de los filogrupos A, B1, B2, C, E y F. Describimos dos nuevos clados en el complejo clonal ST131 representados por genomas que albergan los alelos *fim*H5 y *fim*H35. El clado C se caracteriza por mostrar una menor distancia de SNP en su genoma esencial en comparación a sus clados hermanos A y B.
7. Existe una correlación entre los perfiles de genes que codifican FV y las STs.
8. Existe una impresionante diversidad de plásmidos entre los genomas de *E. coli* productores de BLEE. Sin embargo, las familias de relaxasas MOB_F, MOB_P y MOB_Q representan el 92% del plasmidoma y Col-like e IncF son los grupos de incompatibilidad más frecuentes. Además, la mayoría de los genes que codifican factores asociados a la virulencia y a la resistencia son vehiculados por plásmidos pertenecientes a la familia IncF.
9. En el linaje ST131 la diversidad de grupos de incompatibilidad de plásmidos es mayor en el subclado C2 que en el clado A y el subclado C1-M27. El clado A está relacionado con la presencia de ICEs y plásmidos [F29:A-:B10]-IncF.
10. El tracto intestinal de los perros sanos constituye un importante reservorio de cepas ExPEC y/o UPEC y, en menor medida, de cepas de *E. coli* MDR.

11. Entre las cepas de origen canino que presentaron el estatus ExPEC y/o UPEC, el clon B2-CH103-9-ST372 fue dominante.
12. El análisis de WGS sugiere que las cepas de origen canino del clon B2-CH103-9-ST372, pertenecientes al grupo 1 (serotipos O4:H31, O21:H31 y O83:H31) y al grupo 6 (serotipo O6:H31) podrían causar infecciones extraintestinales tanto en seres humanos como en perros, mientras que las cepas de este clon pertenecientes al grupo 2 y los serotipos O18:H31 y O45:H31 podrían causar solo infecciones en seres humanos.
13. El clado B de ST131 mostró una enorme diversidad genética y definimos cinco nuevos subclados (B6, B6-like, B7, B8 y B9). La mayoría de las cepas ST131 de origen porcino pertenecen a los nuevos subclados B6 y B7.
14. La mayoría de las cepas de ST131 de origen porcino son MDR (91%) y albergan muchos genes de resistencia a los antibióticos. La resistencia a la colistina se introdujo a través de elementos genéticos móviles y logra estabilizarse en el cromosoma. Los plásmidos epidémicos IncX4 e IncHI2 que albergan el gen *mcr*-1 son responsables de la resistencia adquirida a la colistina codificada por el gen *mcr*-1.1. El entorno circundante del casete *mcr*-1 es variable, pero dentro de la misma familia de plásmidos, la inserción y la estabilización conserva estructuras comunes.

Palabras clave: *Escherichia coli, E. coli,* estructura clonal, formación de biopelículas, genes de virulencia, resistencia a antibióticos; β-lactamasas, CTX-M, BLEE, ExPEC, UPEC, clones de alto riesgo, ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131, ST141, ST372, ST1193, perros, porcino, colistina, *mcr*-1.1, WGS, genoma esencial, plasmidoma.

RESUMO

Escherichia coli é a principal causa de infeccións do tracto urinario (ITU) e do torrente sanguíneo en seres humanos e animais. A maioría destas infeccións débense a cepas de patotipos coñecidos como *E. coli* patóxeno extraintestinal (ExPEC) ou *E. coli* uropatóxeno (UPEC). Asociáronse numerosos xenes de virulencia con cepas que causan infeccións extraintestinales, como adhesinas, toxinas, sideróforos e antíxenos capsulares, que lles permiten colonizar as superficies do hóspede, capturar o ferro dispoñible, danar os tecidos do hospedador e evitar os sistemas de defensa do hospedador. O tratamento destas infeccións complicouse seriamente pola aparición de cepas multirresistentes (MDR) e especialmente pola rápida diseminación de cepas de *E. coli* produtoras de β-lactamasas de espectro estendido (ECBLEE).

A presente tese doutoral inclúe cinco publicacións (Flament-Simon SC *et al*..2020 Antibiotics 9 (4) 161 doi. org/10.3390/ antibiotics9040161; Flament- Simon SC *et al*.2020 Front Cell Infect Microbiol 10: 125 doi: 10.3389/ fcimb .2020.00125; Flament-Simon SC *et al*.2019 Front Microbiol 10: 1183 doi: 10.3389/ fmicb.2019.01183; Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/ microorganisms 8111712; Flament- Simon SC *et al.* 2020 Front Microbiol 2020 11: 387 doi: 10.3389/ fmicb.2020.00387) e datos sen publicar cos seguintes obxectivos:

I. Avaliar a prevalencia da secuencia tipo 131 (ST131) e os seus subclones entre cepas de *E. coli* causantes de infeccións extraintestinales en seres humanos illadas de casos consecutivos en 2016 e entre cepas ECBLEE illadas en 2015, en dous hospitais (hospital universitario Lucus Augusti e hospital AP-HP Beaujon), prestando especial atención á aparición de novos clons altamente virulentos e/ou MDR.
II. Adquirir coñecemento das características fenotípicas e xenotípicas das cepas ST131 e outros clons de alto risco illados de seres humanos e animais, determinando os filogrupos, clonotipos, secuencias tipo (STs), serotipos O:H, xenes que codifican Factores de Virulencia (FV), resistencia aos antibióticos, ademais da secuenciación do xenoma completo (*whole genome sequencing*) (WGS) dun subconxunto de cepas para realizar a análise do seu xenoma esencial (*core genome*) (CG) e do seu plasmidoma.
III. Adquirir coñecementos sobre a formación temperá de biopelículas (*early biofilm formation*) (EBF) en cepas de *E. coli* implicadas en infeccións extraintestinales e comparar a capacidade de formación de biopelículas coas características xenotípicas das cepas.
IV. Estudar animais de produción (porcos) e animais de compañía (cans) como reservorios de cepas ST131 e outros clons de alto risco para seres humanos e investigar a posible transferencia de clons.

En total realizamos seis estudos de epidemioloxía molecular, así como unha análise xenómica de cepas de *E. coli*. Nos catro primeiros estudos analizamos un total de 394 cepas produtoras de infeccións extraintestinales extraídas de dous hospitais (hospital Lucus Augusti de Lugo e hospital Beaujon de París), mentres que nos dous estudos restantes analizamos cepas de orixe canina e porcina. A caracterización fenotípica inclúe: (I) a serotipificación, (II) probas de susceptibilidade antimicrobiana e (III) a

determinación de biopelículas. A caracterización molecular inclúe: (I) a descrición dos perfís de virulencia a través da identificación de xenes que codifican FV mediante a reacción en cadea da polimerasa (PCR) seguida da definición dos status de virulencia vinculados a patotipos, (II) a descrición de perfís de resistencias aos antibióticos mediante PCR seguida da secuenciación de amplicones, (III) a determinación dos grupos filoxenéticos de *E. coli* mediante PCR, (IV) a determinación de clonotipos (*fum*C-*fim*H) e secuencias tipo (STs) mediante PCR seguida da secuenciación de amplicones, (VI) a identificación dos clados e subclados de *E. coli* ST131 por PCR, (VII) a identificación do xene *fim*B por PCR seguida da secuenciación de amplicones nun subconxunto de cepas da colección de bacterias de orixe humana e finalmente, (VIII) realizamos a secuenciación do xenoma completo (WGS) dun total de 109 cepas de *E. coli* pertencentes a tres estudos diferentes, permitíndonos realizar a análise do CG e do plasmidoma. A análise estatística inclúe a proba de normalidade de Shapiro- Wilk, a proba non paramétrica de Mann-Whitney- Wilcoxon, a proba exacta de Fisher e a proba de análise de varianza (ANOVA). Os valores $P<0,05$ consideráronse estatisticamente significativos. A correlación entre variables cuantitativas avaliouse mediante o coeficiente de correlación de Pearson, e a avaliación da reproductibilidad realizouse mediante o método estatístico de Bland e Altman, (1986).

No primeiro estudo realizamos a análise molecular dun total de 196 cepas de *E. coli* derivadas de casos consecutivos de infeccións extraintestinales illadas ao longo do ano 2016 en España (100 cepas procedentes do hospital Lucus Augusti de Lugo) e Francia (96 cepas procedentes do hospital Beaujon de Clichy). Determinamos os filogrupos, clonotipos, secuencias tipo (STs), serotipos O:H, xenes que codifican FV e xenes que codifican resistencia aos antibióticos. Observamos que aproximadamente o 10% das infeccións foron causadas por cepas do clon ST131 en ambos os hospitais e que aproximadamente o 60% destas infeccións foron causadas por cepas pertencentes a só 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 e ST141). As cepas ST88 foron frecuentes, especialmente en España, mentres que as cepas ST141 foron significativamente predominantes en Francia. Identificamos 23 cepas ST131 que representaron catro clonotipos: CH40-30, CH40-41, CH40-22 e CH40-298. Só 13 (6,6%) cepas foron portadoras de β-lactamasas de espectro estendido (BLEE). Con todo, o 37,2% das cepas foron MDR. Aproximadamente o 40% das cepas MDR pertenceron a tan só catro dos clons dominantes (B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39- ST88 e D- CH35-27- ST69). Entre as cepas MDR restantes, dúas pertenceron ao clon B2-CH14-64-ST1193, é dicir, o novo clon MDR emerxente a nivel mundial. Ademais, identificamos unha cepa híbrida ExPEC-enteroagregativa pertencente ao clon A-CH11-54-ST10.

No segundo estudo, completamos a análise molecular dun total de 188 cepas de *E. coli* produtoras BLEE colleitadas ao longo do ano 2015 no hospital universitario Lucus Augusti e o hospital AP- HP Beaujon, con obxecto de analizar os clados e subclados do clon ST131, outros STs e caracterizar os tipos de encimas BLEE producidos, os serotipos e os xenes que codifican FV. ST131 detectouse en 33 (39,1%) e 46 (47,9%) dos illamentos procedentes do hospital Lucus Augusti e do hospital Beaujon, respectivamente. As 109 cepas restantes constituíron 57 STs diferentes, onde as seguintes STs incluíron polo menos tres representantes: ST10 (8 cepas), ST23 (3), ST38 (4), ST58 (3), ST88 (5), ST95 (4), ST167 (3), ST354 (5), ST361 (3), ST410 (6), ST648 (4), ST744 (3) e ST1615 (6). ST354, ST410 e ST1615 foron significativamente ($P <0,05$) máis frecuentes en Lucus Augusti (5,4%, 6,5% e 6,5%) que en Beaujon (0% para as tres ST). Ademais, describimos dúas cepas pertencentes ao clon emerxente B2-CH14-64-ST1193.

CTX- M-15 foi a encima BLEE detectada con maior frecuencia nos dous hospitais (44,6% en Lucus Augusti e 50,0% en Beaujon). CTX-M-14 foi significativamente (P=0,0003) máis frecuente en Lucus Augusti (31,5%) que en Beaujon (10,4%), mentres que CTX-M-1 (20,8% *versus* 7,6%; P=0,008) e CTX-M-27 (15,6% *versus* 6,5%; P=0,0389) foron máis frecuentes en Beaujon que en Lucus Augusti. As cepas ST131 presentaron unha media de xenes que codifican FV superior (media 13,367) en comparación coas cepas non- ST131 (media 7,661) (P <0,001). De entre as 79 cepas ST131, a maioría (52; 65,8%) pertenceron ao subclade C2 (tamén coñecido como subclón H30Rx), seguidas por cepas pertencentes ao subclade C1 (grupo C1-M27: 16 cepas, 20,3%; grupo non-C1-M27: 6 cepas, 7,6%) e ao clado A (4 cepas; 5,1%). Ademais, as cepas do subclade C2 mostraron unha media de xenes que codifican FV superior (media 14,250) que as cepas pertencentes ao grupo C1-M27 (media 10,875) (P <0,001).

No terceiro estudo, analizamos a formación de biopelículas en cepas de *E. coli* produtoras de infeccións extraintestinales en seres humanos. A formación de biopelículas en *E. coli* foi avaliada principalmente en grupos específicos de *E. coli* patóxenos. Aquí avaliamos a formación temperá de biopelículas (EBF), é dicir, a etapa de adhesión, utilizando o BioFilm Ring Test® en 394 illados de orixe clínica de *E. coli* (EC) [196 casos consecutivos (ECC) en 2016 e 198 *E. coli* produtores de BLEE (ECBLEE) illados en 2015]. Tras o cal, a capacidade de formación de biopelículas contrastouse cos filogrupos, clonotipos (CH) e secuencias tipo (ST) utilizados para definir os clons, os factores de virulencia (FV) e FimB. De acordo cos niveis de produción de biopelículas ás 2h, 3h e 5h, e a cinética de EBF durante 5h, as cepas pertencentes a ECC e ECBLEE se segregaron en tres grupos de EBF: fortes produtores (G1), moderados produtores (G2) e débiles produtores (G3). Ás 2h, os fortes produtores foron máis frecuentes entre ECC (n=28; 14,3%) que entre ECBLEE (n=8; 4%) (P=0,0004). Como as cepas de ECC e ECBLEE mostraron cinéticas individuais de EBF similares por grupo, a comparación das características das cepas pertencentes a cada grupo nas coleccións ECC e ECBLEE aplicáronse de forma conxunta tras 2h de incubación, sendo as 2h o tempo de lectura máis representativo da segregación de cepas de ECC e ECBLEE en tres grupos. O filogrupo B2 representado polo 51,3% das 394 cepas foi máis frecuente en G1 (77,8%) que en G3 (47,6%) (P=0,0006). As 394 cepas constituíron 153 clons dos cales 31 incluíron polo menos tres representantes. Os clons B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5/14-ST141 e E-CH100-96-ST362 asociáronse con G1 (P <0,03) representando o 41,7% das cepas de G1. O clon B2-CH40-30-ST131 asociouse con G3 (P<0,0001) representando o 25,5% das cepas de G3. A media de FV foi maior entre G1 que entre G3 (P <0,001). A variante proteica FimB-P2 asociouse con G1 (P = 0,0011) e a variante proteica FimB-P1 con G3 (P=0,0023). Os clons, algúns FV e FimB asociáronse con EBF, sendo o grupo clonal explicativo do 72% da variabilidade de EBF. Das 394 cepas analizadas, menos do 10% foron capaces de producir niveis altos de biopelícula de forma rápida e persistente durante 5h. Estas cepas pertenceron a un número limitado de clons descritos en varios estudos anteriores como colonizadores intestinais dominantes en mamíferos e aves, incluíndo o clon B2-CH40-22-ST131, é dicir, o antepasado do globalmente estendido clon B2-CH40-30-ST131 dominante no grupo de cepas débiles produtoras de biopelículas (G3).

No cuarto estudo, buscamos profundar nos coñecementos sobre o complexo clonal ST131 produtor de BLEE de *E. coli* (grupo STC131 constituído por 30 xenomas) e outras liñaxes de *E. coli* produtores de BLEE emerxentes (grupo non-STC131 constituído por 45 xenomas), a través da análise da secuenciación do xenoma completo dun total de 75

cepas BLEE de *E. coli* (51 xenomas do hospital Lucus Augusti e 24 xenomas do hospital de Beaujon). Estes xenomas proveñen de diferentes mostras clínicas: 57 illados de ouriños, 11 de sangue e 7 de diversas fontes, incluíndo unha de bilis, unha de líquido ascítico, unha de óso, dous de feridas e dúas de infeccións dixestivas. Levamos a cabo unha investigación filoxenética que incluíu a análise do CG e do plasmidoma. A través da análise do CG, observamos correlación cos grupos filoxenéticos dos xenomas e describimos dous novos clados pertencentes ao complexo clonal ST131 representados por cepas cos alelos *fim*H5 e *fim*H35. Investigamos *in silico* a presenza de 164 xenes que codifican FV e 13 mutacións puntuais (MP) relacionadas coa expresión de resistencias aos antibióticos no cromosoma dos 75 xenomas *E. coli* BLEE. O grupo de xenomas STC131 presentou unha media maior de xenes que codifican FV que o grupo de xenomas non-STC131 (62 *versus* 47) e unha media maior de MP que o grupo non-STC131 (5 *versus* 3). Ademais, observamos unha clara asociación entre o status ExPEC, o status UPEC e o grupo de xenomas STC131 ($P < 0,0001$). Os 75 xenomas BLEE albergaron 21 presuntos elementos conxugativos integradores (*integrative conjugative elements*) (ICEs) e un total de 228 plásmidos, dos cales 86 (32%) achados no grupo STC131 e 142 (62%) achados no grupo non-STC131. O 92% dos plásmidos pertenceron ás familias de relaxasas MOB_F, MOB_P e MOB_Q. Ademais, describimos unha ampla diversidade de grupos de incompatibilidade (Inc.) sendo esta diversidade maior no grupo non-STC131 fronte ao grupo STC131 (12 Inc. incluíndo Col- like, F, Q, I, X, E, P, 13, HI2, N, R e B/O/K/Z *versus* 6 Inc. incluíndo Col- like, F, X, 13, N e B/ O/K/Z, respectivamente). Observamos que os plásmidos de tipo Col e IncF foron dominantes en ambos os grupos de xenomas. Os plásmidos IncF presentaron unha composición variable sendo portadores de case todos os xenes de virulencia e resistencia achados no plasmidoma dos 75 xenomas BLEE analizados. Ademais, localizamos os xenes que codifican as encimas BLEE en 67 dos 75 xenomas, dos cales 21 estaban integrados no cromosoma (31%) e 46 estaban vehiculados por plásmidos (69%). Tamén observamos unha correlación entre a presenza de tipos específicos de IncF e os diferentes clados de ST131. Atopamos dous xenes codificadores de $bla_{CTX\text{-}M\text{-}27}$ vehiculados por plásmidos [F1:A2:B20]-IncF e cinco xenes codificadores de $bla_{CTX\text{-}M\text{-}15}$ vehiculados por plásmidos [F2:A1:B-]-IncF asociados cos subclados C1 e C2, respectivamente. Todo iso avoga a existencia dunha asociación plásmido- clado. Con respecto ao clado A, describimos a presenza de plásmidos [F29:A-:B10]-IncF e a presenza de ICEs como características distintivas dos xenomas do clado C.

No quinto estudo, baixo a perspectiva de "unha soa saúde" e a preocupación mundial polo incremento das resistencias aos antimicrobianos, investigamos 197 cepas de *E. coli* illadas de cans sans en España entre 2013 e 2017 para ExPEC, UPEC e MDR. Un total de 91 (46,2 %) cepas clasificáronse a nivel molecular como ExPEC e/ou UPEC, incluíndo 50 clons, entre os cales catro clons foron dominantes (B2-CH14-180-ST127, B2-CH52-14-ST141, B2-CH103-9-ST372 e F-CH4-58-ST648) e 15 foran previamente identificados entre as cepas que causaron infeccións extraintestinales en seres humanos españois e franceses en 2015 e 2016. Un total de 28 (14.2%) cepas clasificáronse como MDR, asociadas aos filogrupos B1, D e E, e representadas por 24 clons, dos cales oito foran identificados previamente entre os illados clínicos humanos. Seleccionamos 23 cepas ST372, 21 cepas procedentes de cans sans e dúas cepas procedentes de illados clínicos humanos para WGS, tras o cal construímos unha árbore filoxenética baseado en (*single nucleotide polymorphisms*) SNPs con estes 23 xenomas e outros 174 xenomas obtidos de bases de datos públicas (128 procedentes de cepas de orixe canina e 46

procedentes de cepas de orixe humana). Estes 197 xenomas se segregaron en seis grupos. O grupo 1 comprendeu o 74,6% dos xenomas, constituído na súa maioría por xenomas de cepas de orixe canina (P<0,00001). Os grupos 4 e 6 tamén incluíron xenomas de cepas de orixe canina, mentres que os grupos 2, 3 e 5 asociáronse significativamente con xenomas de cepas de orixe humana. Atopamos varios clons con serotipos comúns entre os presentes en cans e seres humanos, o cal suxire unha potencial transferencia bidireccional de clons, que apoia a perspectiva dun mundo con "unha soa saúde".

No sexto estudo analizamos cepas pertencentes ao clon ST131 de orixe porcina, as cales están escasamente documentadas. Realizamos WGS, analizamos o CG e o plasmidoma de sete cepas procedentes de leitóns con diarrea e catro cepas procedentes de carne de porco. As 11 cepas pertenceron ao serotipo O25b:H4 e ao clado B de ST131 presentando o alelo *fim*H22 ou mutacións alélicas derivadas do mesmo. As 11 cepas codificaron FV que as clasificaron como equivalentes a patóxenos aviarias, uropatóxenos e patóxenos extraintestinales de *E. coli* (equivalentes a APEC, UPEC e ExPEC) e constituíron o virotipo D. O CG realizouse incluíndo os xenomas de todas as cepas de orixe porcina ademais de 73 cepas ST131 de referencia con distintas orixes. Dentro do clado B, o CG se segregou dando lugar a nove subgrupos, o que nos permitiu describir cinco novos subclados (B6, B6- like, B7, B8, B9). Observamos unha asociación entre as cepas pertencentes ao subclade B6, PST43, virotipo D2 e con orixe alimentaria, e unha asociación entre cepas pertencentes ao subclade B7, PST9, virotipo D5 e procedentes de leitóns con diarrea (P=0,007). A distancia entre os xenomas de orixe humana e porcina dos subclados B6 e B7 tivo unha media de 20 SNP/Mb e 15 SNP/Mb, respectivamente. Os plásmidos [F2:A-:B1]-IncF, tipo COLE1 e IncX foron os máis prevalecentes. Ademais, os plásmidos IncF codificaron unha rexión ColV frecuente entre cepas APEC. Os xenes de resistencia antimicrobiana achados con maior frecuencia foron aqueles que confiren resistencia á penicilina, tetraciclina, quinolonas e colistina. O xene *mcr*-1.1 detectouse en 5 das 11 cepas de orixe porcina, integrado no cromosoma dunha cepa e vehiculado por plásmidos no resto (dous MOB_{H11}/IncHI2-ST4, un MOB_{P3}/IncX4 e un MOB_{F12}/IncF [F2:A-:B1] supostamente cointegrado cun IncHI2). As contornas do casete *mcr*-1 mostraron variabilidade. Con todo, observamos estruturas conservadas dentro de una mesma familia de plásmidos. En conclusión, a análise do CG definiu cinco novos subclados. As cepas ST131 de orixe porcina pertencen aos novos subclados B6 e B7. Ademais, as cepas de orixe clínica humanas e as cepas de orixe porcina están estreitamente relacionadas. As 11 cepas ST131 porcinas albergaron unha ampla variedade de plásmidos, xenes de virulencia e resistencia. Ademais, os plásmidos epidémicos IncX4 e IncHI2 son responsables da adquisición do xene *mcr*-1.1. Establecemos a hipótese de que o plásmido APEC-IncF adquiriu o xene *mcr*-1.1 mediante a cointegración dun plásmido IncHI2, o cal é preocupante debido á combinación de atributos de virulencia e resistencia nun só elemento xenético móbil.

Dos resultados obtidos, concluímos:

1. Aproximadamente o 10% das infeccións extraintestinales causadas por *E. coli* acontecidas en 2016 nos dous hospitais estudados foron causadas por cepas do clon ST131, e aproximadamente o 60% destas infeccións foron causadas por cepas pertencentes a só 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 e ST141).

2. A maioría das cepas ST12, ST73, ST95 e ST141 causantes de infeccións extraintestinales en seres humanos foron susceptibles á maioría dos antibióticos, o que indica que a MDR non foi a razón do seu éxito. Os resultados do presente estudo apoian a idea de que o seu éxito débese principalmente a que albergan unha gran cantidade de xenes que codifican FV.
3. Entre as cepas MDR causantes de infeccións extraintestinales en seres humanos, predominan catro clons: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 e D-CH35-27-ST69.
4. A pesar da enorme diversidade xenética observada na nosa colección de cepas de *E. coli* produtoras de BLEE (71 clons presentes en 188 cepas BLEE), pódese concluír que a maioría das cepas pertencen a só tres complexos clonales (CC10, CC23 e CC131) e que ST131 subclade C2 asociado á produción de CTX- M-15 segue sendo a liñaxe de *E. coli* máis prevalecente entre as cepas de *E. coli* produtoras de BLEE identificadas nos hospitais español e francés estudados.
5. Entre as nosas 394 cepas de *E. coli* procedentes de infeccións extraintestinales, <10% son capaces de producir niveis altos de biopelícula de forma rápida e persistente durante 5 h. Estas cepas pertencen a uns poucos clons descritos previamente en varios estudos como colonizadores intestinais dominantes en mamíferos e aves, e inclúen ao clon B2-CH40-22-ST131, é dicir, o antepasado do globalmente estendido clon B2-CH40-30-ST131, que é o clon dominante entre as cepas débiles produtores de biopelículas.
6. Na análise do CG de cepas de *E. coli* produtoras de BLEE destaca a diversidade xenómica. Incluíndo representantes dos filogrupos A, B1, B2, C, E e F. Describimos dous novos clados no complexo clonal ST131 representados por xenomas que albergan os alelos *fim*H5 e *fim*H35. O clado C caracterízase por mostrar unha menor distancia de SNP no seu xenoma esencial en comparación ás súas clados irmáns A e B.
7. Existe unha correlación entre os perfís de xenes que codifican FV e as STs.
8. Existe unha impresionante diversidade de plásmidos entre os xenomas de *E. coli* produtores de BLEE. Con todo, as familias de relaxasas MOB_F, MOB_P e MOB_Q representan o 92% do plasmidoma e Col-like e IncF son os grupos de incompatibilidade máis frecuentes. Ademais, a maioría dos xenes que codifican factores asociados á virulencia e á resistencia son vehiculados por plásmidos pertencentes á familia IncF.
9. Na liñaxe ST131 a diversidade de grupos de incompatibilidade de plásmidos é maior no subclado C2 que no clado A e o subclado C1-M27. O clado A está relacionado coa presenza de ICEs e plásmidos [F29:A-:B10]-IncF.
10. O tracto intestinal dos cans sans constitúe un importante reservorio de cepas ExPEC e/ou UPEC e, en menor medida, de cepas de *E. coli* MDR.
11. Entre as cepas de orixe canina que presentaron o status ExPEC e/ou UPEC, o clon B2-CH103-9-ST372 foi dominante.
12. A análise de WGS suxire que as cepas de orixe canina do clon B2-CH103-9-ST372, pertencentes ao grupo 1 (serotipos O4:H31, O21:H31 e O83:H31) e ao grupo 6 (serotipo O6:H31) poderían causar infeccións extraintestinales tanto en seres humanos como en cans, mentres que as cepas deste clon pertencentes ao grupo 2 e os serotipos O18:H31 e O45:H31 poderían causar só infeccións en seres humanos.

13. O clado B de ST131 mostrou unha enorme diversidade xenética e definimos cinco novos subclados (B6, B6-like, B7, B8 e B9). A maioría das cepas ST131 de orixe porcina pertencen aos novos subclados B6 e B7.
14. A maioría das cepas de ST131 de orixe porcina son MDR (91%) e albergan moitos xenes de resistencia aos antibióticos. A resistencia á colistina introduciuse a través de elementos xenéticos móbiles e logra estabilizarse no cromosoma. Os plásmidos epidémicos IncX4 e IncHI2 que albergan o xene *mcr*-1 son responsables da resistencia adquirida á colistina codificada polo xene *mcr*-1.1. A contorna circundante do casete *mcr*-1 é variable, pero dentro da mesma familia de plásmidos, a inserción e a estabilización conserva estruturas comúns.

Palabras chave: *Escherichia coli*, *E. coli*, estrutura clonal, formación de biofilm, xenes de virulencia, resistencia a antibióticos; β-lactamases, CTX-M, BLEE, ExPEC, UPEC, clons de alto risco, ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131, ST141, ST372, ST1193, cans, porcino, colistina, *mcr*-1.1, WGS, xenoma esencial, plasmidoma.

RÉSUMÉ

Escherichia coli est la principale cause d'infections des voies urinaires (IVU) et de la circulation sanguine chez l'homme et l'animal. La plupart de ces infections sont dues à des souches d'agents pathogènes connus comme *E. coli* pathogène extraintestinales (ExPEC) ou *E. coli* uropathogènes (UPEC). Nombreux gènes de virulence ont été associés à des souches qui provoquent des infections extraintestinales, telles que les adhésines, les toxines, les sidérophores et les antigènes capsulaires, qui leur permettent de coloniser les surfaces de l'hőte, de capturer le fer disponible, d'endommager les tissus de l'hőte et de contourner les systèmes de défense de l'hőte. Le traitement de ces infections a été sérieusement compliqué par l'émergence de souches multirésistantes (MDR) et en particulier par la rapide propagation de souches d'*E. coli* productrices de β-lactamases à spectre étendu (ECBLSE).

La présente thèse de doctorat comprend cinq publications (Flament-Simon SC *et al.* 2020 Antibiotics 9 (4) 161 doi.org/10.3390/antibiotics9040161; Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10: 125 doi: 10.3389/fcimb .2020.00125; Flament-Simon SC *et al.* 2019 Front Microbiol 10: 1183 doi: 10.3389/fmicb.2019.01183; Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms 8111712; Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11: 387 doi: 10.3389/fmicb.2020.00387) et des données non publiées avec les objectifs suivants:

I. Évaluer la prévalence de la séquence type 131 (ST131) et de ses sous-clones parmi les souches d'*E. coli* provoquant des infections extraintestinales chez l'homme isolées de cas consécutifs en 2016 et parmi des souches ECBLSE isolées en 2015, dans deux hőpitaux (Hőpital universitaire Lucus Augusti et AP-HP Beaujon), en accordant une attention particulière à l'apparition de nouveaux clones hautement virulents et/ou MDR.
II. Acquérir des connaissances sur les caractéristiques phénotypiques et génotypiques des souches ST131 et d'autres clones à haut risque isolés chez l'homme et l'animal, déterminer les phylogroupes, les clonotypes, les séquences type (ST), les sérotypes O:H, les gènes codant pour les Facteurs de Virulence (FV), la résistance aux antibiotiques, en plus du séquençage total du génome (*whole genome sequencing*) (WGS) d'un sous-ensemble de souches pour effectuer l'analyse de leur génome essentiel (*core genome*) (CG) et de leur plasmidome.
III. Acquérir des connaissances sur la formation précoce de biofilm (*early biofilm formation*) (EBF) sur les souches d'*E. coli* impliquées dans des infections extra-intestinales et comparer la capacité de formation de biofilm avec les caractéristiques génotypiques des souches.
IV. Étudier les animaux d'élevage (porcs) et les animaux de compagnie (chiens) comme réservoirs des souches ST131 et d'autres clones à haut risque pour les humains et rechercher l'éventuel transfert de clones.

Au total, nous avons mené six études d'épidémiologie moléculaire, ainsi qu'une analyse génomique des souches d'*E. coli*. Au cours des quatre premières études, nous avons analysé un total de 394 souches productrices d'infections extraintestinales prélevées dans deux hőpitaux (hőpital universitaire Lucus Augusti à Lugo et hőpital Beaujon à Paris), tandis que dans les deux autres études nous avons analysé des souches d'origine

canine et porcine. La caractérisation phénotypique comprend: (I) le sérotypage, (II) les tests de sensibilité aux antibiotiques et (III) la détermination de production de biofilm. La caractérisation moléculaire comprend: (I) la description des profils de virulence par l'identification des gènes codant pour les FV au moyen de la réaction en chaîne par polymérase (PCR) suivie de la définition des profil de virulence liés aux pathotypes, (II) la description des profils de résistance aux antibiotiques par PCR suivi d'un séquençage d'amplicon, (III) la détermination des groupes phylogénétiques d'*E. coli* par PCR, (IV) la détermination des clonotypes (*fum*C-*fim*H) et des séquences type (ST) par PCR suivie d'un séquençage d'amplicon, (VI) l'identification des clades et sous-clades d'*E. coli* ST131 par PCR, (VII) l'identification du gène *fim*B par PCR suivie d'un séquençage d'amplicon pour un sous-ensemble des souches de la collection de bactéries d'origine humain et finalement, (VIII) nous avons réalisé le WGS d'un total de 109 souches d'*E. coli* appartenant à trois études différentes, permettant d'effectuer l'analyse de leur GC et de leur plasmidome. L'analyse statistique comprend le test de normalité de Shapiro-Wilk, le test non paramétrique de Mann-Whitney-Wilcoxon, le test exact de Fisher et le test d'analyse de variance (ANOVA). Les valeurs $P<0,05$ ont été considérées comme statistiquement significatives. La corrélation entre les variables quantitatives a été évaluée en utilisant le coefficient de corrélation de Pearson, et l'évaluation de la reproductibilité a été réalisée en utilisant la méthode statistique de Bland et Altman, (1986).

Au cours de la première étude, nous avons réalisé l'analyse moléculaire d'un total de 196 souches d'*E. coli* dérivées de cas consécutifs d'infections extraintestinales isolés tout au long de 2016 en Espagne (100 souches de l'hŏpital universitaire Lucus Augusti de Lugo) et en France (96 souches de l'Hŏpital Beaujon de Clichy). Nous avons déterminé les phylogroupes, les clonotypes, les séquences type (ST), les sérotypes O:H, les gènes codant pour des FV et les gènes codant pour la résistance aux antibiotiques. Nous avons observé qu'environ 10% des infections étaient causées par des souches du clone ST131 dans les deux hŏpitaux et qu'environ 60% de ces infections étaient causées par des souches appartenant à seulement 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 et ST141). Les souches ST88 étaient fréquentes, notamment en Espagne, tandis que les souches ST141 étaient significativement prédominantes en France. Nous avons identifié 23 souches ST131 qui représentaient quatre clonotypes: CH40-30, CH40-41, CH40-22 et CH40-298. Seules 13 souches (6,6%) étaient porteuses de β-lactamases à spectre étendu (BLSE). Cependant, 37,2% des souches étaient MDR. Environ 40% des souches MDR appartenaient à seulement quatre des clones dominants (B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 et D-CH35-27-ST69). Parmi les souches MDR restantes, deux appartenaient au clone B2-CH14-64-ST1193, c'est-à-dire le nouveau clone MDR émergeant dans le monde. De plus, nous avons identifié une souche hybride ExPEC-enteroaggregative appartenant au clone A-CH11-54-ST10.

Lors de la deuxième étude, nous avons effectué l'analyse moléculaire d'un total de 188 souches d'*E. coli* productrices de BLSE collectées tout au long de 2015 à l'hŏpital universitaire Lucus Augusti et à l'hŏpital AP-HP Beaujon, afin d'analyser les clades et sous-clades du clone ST131, d'autres STs et caractériser les types d'enzymes BLSE produites, les sérotypes et les gènes codant pour des FV. ST131 a été détecté dans 33 (39,1%) et 46 (47,9%) des isolats de l'hŏpital universitaire Lucus Augusti et de l'hŏpital Beaujon, respectivement. Les 109 souches restantes constituaient 57 STs différentes, où les STs suivantes comprenaient au moins trois représentants: ST10 (8 souches), ST23 (3), ST38 (4), ST58 (3), ST88 (5), ST95 (4), ST167 (3), ST354 (5), ST361 (3), ST410 (6), ST648 (4), ST744 (3) et ST1615 (6). ST354, ST410 et ST1615 étaient significativement

(P <0,05) plus fréquents chez Lucus Augusti (5,4%, 6,5% et 6,5%) qu'à Beaujon (0% pour les trois STs). De plus, nous décrivons deux souches appartenant au clone émergent B2-CH14-64-ST1193. CTX-M-15 a été l'enzyme BLSE la plus fréquemment détectée dans les deux hŏpitaux (44,6% à Lucus Augusti et 50,0% à Beaujon). CTX-M-14 a été significativement (P=0,0003) plus fréquente à Lucus Augusti (31,5%) que chez Beaujon (10,4%), tandis que CTX-M-1 (20,8% *versus* 7, 6%; P=0,008) et CTX-M-27 (15,6% *versus* 6,5%; P=0,0389) étaient plus fréquentes à Beaujon qu'à Lucus Augusti. Les souches ST131 présentaient une moyenne de gènes codant pour des FV plus élevé (moyenne 13,367) par rapport aux souches non-ST131 (moyenne 7,661) (P<0,001). Parmi les 79 souches ST131, la majorité (52; 65,8%) appartenait au sous-clade C2 (également appelé sous-clone H30Rx), suivi des souches appartenant au sous-clade C1 (groupe C1-M27: 16 souches, 20,3%; non-C1 -Groupe M27: 6 souches, 7,6%) et au clade A (4 souches; 5,1%). De plus, les souches du sous-clade C2 présentaient une moyenne de gènes codant pour des FV plus élevé (moyenne de 14 250) que les souches appartenant au groupe C1-M27 (moyenne de 10 875) (P <0,001).

Durant la troisième étude, nous avons analysé la formation de biofilm pour les souches d'*E. coli* qui causent des infections extraintestinales chez l'homme. La formation de biofilm sur *E. coli* a été évaluée principalement pour des groupes spécifiques d'*E. coli* pathogènes. Nous évaluons ici la formation précoce de biofilm (EBF), c'est-à-dire le stade d'adhésion, à l'aide du BioFilm Ring Test® pour 394 isolats d'origine clinique d'*E. coli* (EC) [196 cas consécutifs (ECC) isolés en 2016 et 198 E *coli* producteur de BLSE (ECBLSE) isolés en 2015]. Après quoi, la capacité de formation du biofilm a été comparée aux phylogroupes, clonotypes (CH) et séquences types (ST) utilisés pour définir les clones, les facteurs de virulence (FV) et FimB. Selon les niveaux de production de biofilm à 2h, 3h et 5h, et la cinétique EBF pendant 5h, les souches appartenant à ECC et ECBLSE ont été séparées en trois groupes EBF: forts producteurs (G1), producteurs modérés (G2) et faibles producteurs (G3). A 2h, les forts producteurs étaient plus fréquents chez ECC (n=28; 14,3%) que chez ECBLSE (n=8; 4%) (P=0,0004). Les souches ECC et ECBLSE présentaient une cinétique individuelle de EBF similaire par groupe, par conséquent, la comparaison des caractéristiques des souches appartenant à chaque groupe des collections ECC et ECBLSE a été appliquée ensemble après 2h d'incubation, étant les 2h le temps de lecture le plus représentative de la ségrégation des souches ECC et ECBLSE en trois groupes. Le phylogroupe B2 représenté par 51,3% des 394 souches était plus fréquent chez G1 (77,8%) que parmi G3 (47,6%) (P=0,0006). Les 394 souches constituaient 153 clones dont 31 comprenaient au moins trois représentants. Les clones B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5/14-ST141 et E-CH100-96-ST362 étaient associés à G1 (P <0,03) représentant 41,7% des souches G1. Le clone B2-CH40-30-ST131 était associé à G3 (P <0,0001) représentant 25,5% des souches G3. La moyenne des FV était plus élevée chez G1 que chez G3 (P <0,001). La variante protéique FimB-P2 a été associée à G1 (P=0,0011) et la variante protéique FimB-P1 à G3 (P=0,0023). Les clones, certains FV et FimB étaient associés à l'EBF, le groupe clonal étant l'explicatif de 72% de la variabilité de l'EBF. Sur les 394 souches analysées, moins de 10% ont pu produire des niveaux élevés de biofilm rapidement et de manière persistante pendant 5 h. Ces souches appartenaient à un nombre limité de clones, décrits dans plusieurs études antérieures comme colonisateurs intestinaux dominants chez les mammifères et les oiseaux, dont le clone B2-CH40-22-ST131, c'est-à-dire l'ancêtre du clone globalement étendu B2-CH40-30-ST131 et dominant dans le groupe des souches faibles productrices de biofilm (G3).

Au cours de la quatrième étude, nous cherchons à approfondir les connaissances sur le complexe clonal ST131 produisant des BLSE d'*E. coli* (groupe STC131 composé de 30 génomes) et d'autres lignées d'*E. coli* émergentes produisant des BLSE (groupe non-STC131 composé de 45 génomes), par l'analyse du séquençage complet du génome d'un total de 75 souches BLSE d'*E. coli* (51 génomes de l'hŏpital universitaire Lucus Augusti et 24 génomes de l'hŏpital de Beaujon). Ces génomes proviennent de différents échantillons cliniques: 57 isolats d'urine, 11 isolats de sang et 7 isolats de différentes sources, dont un de bile, un d'ascite, un d'os, deux de plaies et deux d'infections digestives. Nous avons effectué une recherche phylogénétique comprenant l'analyse du GC et du plasmidome. Grăce à l'analyse du GC, nous avons observé une corrélation avec les groupes phylogénétiques des génomes et décrit deux nouveaux clades appartenant au complexe clonal ST131 représenté par des souches avec les allèles *fim*H5 et *fim*H35. Nous avons étudié *in silico* la présence de 164 gènes codant pour des FV et 13 mutations ponctuelles (MP) liées à l'expression de la résistance aux antibiotiques sur le chromosome des 75 génomes BLSE d'*E. coli*. Le groupe des génomes STC131 présentaient une moyenne plus élevé de gènes codant pour des FV que le groupe des génomes non-STC131 (62 *versus* 47) et une moyenne plus élevée de MP que le groupe non-STC131 (5 *versus* 3). En outre, nous avons observé une association claire entre le statut ExPEC, le statut UPEC et le groupe génomique STC131 ($P < 0,0001$). Les 75 génomes BLSE abritaient 21 présumés éléments conjugatifs intégratifs (ICE) et un total de 228 plasmides, dont 86 (32%) trouvés dans le groupe STC131 et 142 (62%) trouvés dans le groupe non-STC131. 92% des plasmides appartenaient aux familles de relaxases MOB_F, MOB_P et MOB_Q. De plus, nous décrivons une grande diversité de groupes d'incompatibilité (Inc.), cette diversité étant plus grande dans le groupe non-STC131 par rapport au groupe STC131 (12 Inc. incluant Col-like, F, Q, I, X, Y, P , 13, HI2, N, R et B/O/K/Z *versus* 6 Inc. y compris Col-like, F, X, 13, N et B/O/K/Z, respectivement). Nous avons observé que les plasmides de type Col et IncF étaient dominants dans les deux groupes de génomes. Les plasmides IncF présentaient une composition variable, étant porteurs de presque tous les gènes de virulence et de résistance trouvés dans le plasmidome des 75 génomes BLSE analysés. De plus, nous avons localisé les gènes codant pour les enzymes BLSE dans 67 des 75 génomes, dont 21 étaient intégrés dans le chromosome (31%) et 46 étaient portés par des plasmides (69%). Nous avons également observé une corrélation entre la présence de types spécifiques d'IncF et les différents clades de ST131. Nous avons trouvé deux gènes codant pour $bla_{CTX\text{-}M\text{-}27}$ portés par des plasmides [F1:A2:B20]-IncF et cinq gènes codants pour $bla_{CTX\text{-}M\text{-}15}$ portés par des plasmides [F2:A1:B-]-IncF associés au sous-clades C1 et C2, respectivement. Tout cela prŏne l'existence d'une association plasmide-clade. Concernant le clade A, nous décrivons la présence de plasmides [F29:A-:B10]-IncF et la présence d'ICE comme caractéristiques distinctives des génomes du clade C.

Lors de la cinquième étude, du point de vue «une seule santé» et de la préoccupation mondiale concernant l'augmentation de la résistance aux antimicrobiens, nous avons étudié 197 souches d'*E. coli* isolées chez des chiens en bonne santé en Espagne entre 2013 et 2017 pour etablir les quelles étaient ExPEC, UPEC et MDR. Un total de 91 (46,2%) souches ont été classées au niveau moléculaire en ExPEC et/ou UPEC, constituaient 50 clones, parmi lesquels quatre clones dominaient (B2-CH14-180-ST127, B2-CH52-14-ST141, B2-CH103-9-ST372 et F-CH4-58-ST648) et 15 avaient déjà été identifiées parmi les souches ayant causé des infections extraintestinales chez les patients espagnol et français en 2015 et 2016. Un total de 28 (14,2%) souches ont été classées comme MDR, associés aux phylogroupes B1, D et E, représentés par 24 clones, dont huit avaient été

préalablement identifiés parmi des isolats cliniques humains. Nous avons sélectionné 23 souches ST372, 21 souches de chiens sains et deux souches d'isolats cliniques humains pour WGS, après quoi nous avons construit un arbre phylogénétique basé sur (*single nucleotide polymorphisms*) SNPs avec ces 23 génomes et 174 autres génomes obtenus à partir de bases de données publiques (128 souches d'origine canine et 46 souches d'origine humain). Ces 197 génomes ont été séparés en six groupes. Le groupe 1 comprenait 74,6% des génomes, constitués principalement de génomes de souches d'origine canine (P<0,00001). Les groupes 4 et 6 comprenaient également des génomes de souches d'origine canine, tandis que les groupes 2, 3 et 5 étaient significativement associés aux génomes de souches d'origine humaine. Nous avons trouvé plusieurs clones avec des sérotypes communs parmi ceux trouvés chez les chiens et les humains, suggérant un potentiel transfert bidirectionnel de clones, soutenant la perspective d'un monde avec «une seule santé».

Durant la sixième étude, nous avons analysé des souches appartenant au clone ST131 d'origine porcine, qui sont rarement documentées. Nous avons réalisé le WGS, analysé le GC et le plasmidome de sept souches de porcelets souffrant de diarrhée et de quatre souches isolés de viande de porc. Les 11 souches appartenaient au sérotype O25b:H4 et au clade B de ST131, présentant l'allèle *fim*H22 ou des mutations alléliques dérivées de celui-ci. Les 11 souches codées des FV qui les classaient comme équivalentes aux pathogènes aviaires, uropathogènes et pathogènes extraintestinaux d'*E. coli* (équivalent à APEC, UPEC et ExPEC) et constituaient le virotype D. L'analyse du GC a été réalisée en incluant les génomes de toutes les souches d'origine porcins ainsi que 73 souches ST131 de référence d'origines différentes. Au sein du clade B, le CG a été ségrégué donnant naissance à neuf sous-groupes, ce qui nous a permis de décrire cinq nouveaux sous-clades (B6, B6-like, B7, B8, B9). Nous avons observé une association entre les souches appartenant au sous-clade B6, PST43, virotype D2 et d'origine alimentaire, et une association entre les souches appartenant au sous-clade B7, PST9, virotype D5 et isolés de porcelets souffrant de diarrhée (P=0,007). La distance moyenne entre les génomes d'origine humaine et porcine des sous-clades B6 et B7 était de 20 SNP/Mb et 15 SNP/Mb, respectivement. Les plasmides [F2: A-:B1]-IncF, de type ColE1 et IncX étaient les plus fréquent. Par ailleur, les plasmides IncF codaient pour une région ColV commune parmi les souches APEC. Les gènes de résistance aux antibiotiques les plus fréquemment trouvés étaient ceux qui confèrent une résistance à la pénicilline, à la tétracycline, aux quinolones et à la colistine. Le gène *mcr*-1.1 a été détecté dans 5 des 11 souches d'origine porcine, intégré dans le chromosome d'une souche et porté par des plasmides pour le reste (deux MOB_{H11}/IncHI2-ST4, un MOB_{P3}/IncX4 et un MOB_{F12}/IncF [F2:A-:B1] vraisemblablement co-intégré avec un IncHI2). Les environnements du gène *mcr*-1 ont montré une grande variabilité. Cependant, nous observons des structures conservées au sein d'une même famille de plasmides. En conclusion, l'analyse du GC a défini cinq nouveaux sous-clades. Les souches ST131 d'origine porcine appartiennent aux nouvelles sous-clades B6 et B7. De plus, les souches cliniques humaines et les souches porcines sont étroitement liées. Les 11 souches porcines ST131 abritaient une grande variété de plasmides, de gènes de virulence et de résistance. De plus, les plasmides épidémiques IncX4 et IncHI2 sont responsables de l'acquisition du gène *mcr*-1.1. Nous émettons l'hypothèse que le plasmide APEC-IncF a acquis le gène *mcr*-1.1 par co-intégration d'un plasmide IncHI2, ce qui est inquiétant en raison de la combinaison d'attributs de virulence et de résistance dans un seul élément génétique mobile.

A partir des résultats obtenus, nous concluons:

1. Environ 10% des infections extraintestinales causées par *E. coli* survenues en 2016 dans les deux hőpitaux étudiés étaient causées par des souches du clone ST131, et environ 60% de ces infections étaient causées par des souches appartenant à seulement 10 STs (ST10 , ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 et ST141).
2. La plupart des souches ST12, ST73, ST95 et ST141 causant des infections extraintestinales chez l'homme étaient sensibles à la plupart des antibiotiques, ce qui indique que la MDR n'était pas la raison de leur succès. Les résultats de la présente étude soutiennent l'idée que leur succès est principalement dű au fait qu'ils hébergent un grand nombre de gènes codant pour des FV.
3. Parmi les souches MDR causant des infections extraintestinales chez l'homme, quatre clones prédominent: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 et D-CH35-27-ST69.
4. Malgré l'énorme diversité génétique observée dans notre collection de souches d'*E. coli* productrices de BLSE (71 clones présents dans 188 souches de BLSE), on peut conclure que la plupart des souches appartiennent à seulement trois complexes clonaux (CC10, CC23 et CC131) et que le sous-clade C2 de ST131 associée à la production de CTX-M-15 reste la lignée d'*E. coli* la plus répandue parmi les souches d'*E. coli* productrices de BLSE identifiées dans les hőpitaux espagnols et français étudiés.
5. Parmi nos 394 souches d'*E. coli* provenant d'infections extraintestinales, <10% sont capables de produire rapidement et de manière persistante des niveaux élevés de biofilm pendant 5 h. Ces souches appartiennent à quelques clones précédemment décrits dans diverses études comme colonisateurs intestinaux dominants chez les mammifères et les oiseaux, et comprennent le clone B2-CH40-22-ST131, c'est-à-dire l'ancętre du clone globalement étendu B2-CH40-30-ST131, qui est le clone dominant parmi les souches faibles productrices de biofilm.
6. Dans l'analyse du GC des souches d'*E. coli* productrices de BLSE, la diversité génomique se démarque. Y compris des représentants des phylogroupes A, B1, B2, C, E et F. Nous décrivons deux nouveaux clades dans le complexe clonal ST131 représenté par des génomes abritant les allèles *fim*H5 et *fim*H35. Le clade C se caractérise par une distance mineur de SNPs au sein du génome essentiel par rapport à ses clades frères A et B.
7. Il existe une corrélation entre les profils des gènes codant des FV et les STs.
8. Il existe une diversité impressionnante de plasmides parmi les génomes d'*E. coli* producteurs de BLSE. Cependant, les familles de relaxases MOB_F, MOB_P et MOB_Q représentent 92% du plasmidome et Col-like et IncF sont les groupes d'incompatibilité les plus fréquents. De plus, la plupart des gènes codant pour des facteurs associés à la virulence et à la résistance sont portés par des plasmides appartenant à la famille IncF.
9. Dans la lignée ST131, la diversité des groupes d'incompatibilité plasmidique est plus grande pour le sous-clade C2 que pour le clade A et le groupe C1-M27 du sous-clade C1. Le clade A est lié à la présence d'ICEs et de plasmides [F29:A-:B10]-IncF.
10. Le tractus intestinal des chiens sains constitue un réservoir important de souches ExPEC et/ou UPEC et, dans une moindre mesure, de souches d'*E. coli* MDR.

11. Parmi les souches d'origine canine présentant un statut ExPEC et/ou UPEC, le clone B2-CH103-9-ST372 était dominant.
12. L'analyse du WGS suggère que les souches d'origine canine du clone B2-CH103-9-ST372, appartenant au groupe 1 (sérotypes O4:H31, O21:H31 et O83:H31) et au groupe 6 (sérotype O6:H31) pourraient provoquer des infections extraintestinales tant chez l'homme que chez le chien, alors que les souches de ce clone appartenant au groupe 2 et les sérotypes O18:H31 et O45:H31 ne pourraient provoquer d'infections que chez l'homme.
13. Le clade B de ST131 a montré une énorme diversité génétique et nous avons défini cinq nouveaux sous-clades (B6, B6-like, B7, B8 et B9). La plupart des souches ST131 d'origine porcine appartiennent aux nouvelles sous-clades B6 et B7.
14. La plupart des souches ST131 d'origine porcine sont MDR (91%) et contiennent de nombreux gènes de résistance aux antibiotiques. La résistance à la colistine a été introduite par des éléments génétiques mobiles et parvient à se stabiliser sur le chromosome. Les plasmides épidémiques IncX4 et IncHI2 hébergeant le gène *mcr*-1 et sont responsables de la résistance acquise à la colistine codée par le gène *mcr*-1.1. L'environnement du gène *mcr*-1 est variable, mais au sein d'une même famille de plasmides, l'insertion et la stabilisation conservent des structures communes.

Mots clés: *Escherichia coli*; *E. coli*, structure clonale, formation de biofilm, gènes de virulence, résistance aux antibiotiques; β-lactamases, CTX-M, BLSE, ExPEC, UPEC, clones à haut risque, ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131, ST141, ST372, ST1193, chiens, porcs, colistine, *mcr*-1.1, WGS, génome essentiel, plasmidome.

ABBREVIATIONS

AAC Aminoglycoside acyltransferase
ACT Artemis comparison tool
adk Adenylate Kinase gene
aEPEC Atypical enteropathogenic *E. coli*
AIEC Disease-associated adherent-invasive *E. coli*
AM/AMP Ampicillin
AMC Amoxicillin-clavulanic acid
AME Aminoglycoside modifying enzyme
AMP/SAM Ampicillin-sulbactam
AN Amikacin
ANT Aminoglycoside nucleotidyltransferase
APEC Avian pathogenic *E. coli*
APH Aminoglycoside phosphotransferase
ARG Antibiotic resistance gene
ARG-ANNOT Antibiotic resistance gene-annotation
AST Antimicrobial susceptibility test
ATM Aztreonam
ATP Adenosine triphosphate
BEAs Biofilm eradication agents
BFI Biofilm formation index
BHI Brain heart infusion agar
BLAST Basic Local Alignment Search Tool
bp Base pairs
BRIG BLAST Ring Image Generator
C Chloramphenicol
CARD Comprehensive antibiotic research database
CAZ Ceftazidime
CC Clonal complexes
CDS Coding Sequence
CEC *E. coli* consecutively isolated
CG Core genome
CGE Center for Genomic Epidemiology
cgMLST Core genome MLST
CH Clonotype (*fum*C-*fim*H alleles)
CHL Chloramphenicol
Chr Chromosome
CIP Ciprofloxacin
CLSI Clinical and Laboratory Standards Institute
COL/CL Colistin
CTX Cefotaxime
CTX-M Cefotaximase of Munich
CXM Cefuroxime
CZ Ceftazidime
D Doxycycline
DAEC Diffusely adherent *E. coli*
DEC Diarrheagenic *E. coli*
DHPS Dihydropteorate synthetases
DNA Deoxyribonucleic acid

DPMT PCR-based degenerate primer MOB typing
EAEC Enteroaggregative *E. coli*
EBF Early biofilm formation
EHEC Enterohaemorrhagic *E. coli*
EIEC Enteroinvasive *E. coli*
EPEC Enteropathogenic *E. coli*
ESBL Extended-spectrum betalactamase
ESBLEC Extended-spectrum β-lactamase-producing *Escherichia coli*
ETEC Enterotoxigenic *E. coli*
EUCAST European Union Committee on Antimicrobial Susceptibility Testing
ExPEC Extraintestinal pathogenic *E. coli*
FAB FII, FIA and FIB IncF replicons
FAO Food and Agriculture Organisation
FBP Foodborne pathogen
FD Nitrofurantoin
FED Cefepime
fimH Type 1 fimbrial adhesin gene
FMT Fecal microbiota transplant
FOS Fosfomycin
FOX Cefoxitin
fumC Fumarate hydratase gene
GEN/GM Gentamicin
GES Guyana extended-spectrum-lactamase
GIM German imipenemase
GI Genomic island
gyrB DNA gyrase gene
HGT Horizontal gene transfer
HNM Non-motile strains
HPI High pathogenicity island
HUH His-hydrophobe-His endonuclease superfamily
I Intermediate
IBCs Intracellular bacterial communities
icd Isocitrate/isopropylmalate dehydrogenase gene
ICE Integrative conjugative element
IM Inner membrane
IMI Imipenemase
IMP Imipenem
IMP Imipenemase metallo-betalactamase
In Integrons
Inc. Incompatibility groups
InPEC Intestinal pathogenic *E. coli*
IR Inverted repeat
IS Insertion sequence
kb Kilobases
KPC Klebsiella pneumoniae carbapenemase
LPS Lipopolysaccharide
LREC from Spanish "Laboratorio de Referencia de *Escherichia coli*"
MATE Multidrug and toxic compound extrusion family

MBL Metallo-betalactamase
mcr Mobile colistin resistance gene
mdh Malate dehydrogenase gene
MDR Multidrug-resistant
MDT Secondary multidrug transporter
MFS Major facilitator superfamily
MGE Mobile genetic element
MI Minocycline
MIC Minimum inhibitory concentration
MLEE Multi-locus enzyme electrophoresis
MLST Multi-locus sequence typing
MOB Relaxase protein
Mpf Mating pair formation proteins
na Not-applicable
NAL Nalidixic acid
NCBI National Center for Biotechnology Information
nd Not-detected
NDM New Delhi metallo-betalactamase
NGS Next-generation sequencing
NIH National Library of Medicine
NMEC Neonatal meningitis-associated *E. coli*
NN Tobramycin
OIE World Organisation for Animal Health
OM Outer membrane
ONT Non-typeable O-antigen
ORF Open reading frame
OXA-48 Oxacillinase-48-types betalactamase
p Plasmid
PAI Pathogenicity island
pAmpC Plasmidic aquired cephalosporinase
PBRT PCR-based replicon typing
PCR Polymerase chain reaction
PDR Pan-drug-resistant
PFGE Pulse-field gel electrophoresis
PI/TZP Piperacillin-tazobactam
PLACNET Plasmid constellation network
PM Point mutations
pMLST PCR-based replicon subtyping
PMQR Plasmid-mediated quinolone resistance
PST Sequence type according to the Pasteur scheme
purA Adenylosuccinate dehydrogenase gene
QIRs Quiascent intracellular reservoirs
R Resistant
rmtB Plasmid-mediated 16S rRNA methylases
RND Resistance-nodulation-cell division family
RST Replicon typing scheme
rUTI Recurrent UTI infection
S Susceptible

SEPEC Sepsis-causing *E. coli*
SGS Second-generation sequencing or short-read sequencing technologies
SIM Seoul imipenemase
SME Serratia marcescens enzyme
SMR Small multidrug resistance family
SMRT Single-molecule real-time
SNP Single nucleotide polymorphism
SPM Sao Paulo metallo-betalactamase
spp Specie
SS Saline solution
ST Sequence type according to Achtman scheme
STc Sequence type complexes
STEC Shiga toxin-producing *E. coli*
ST Sequence type
SXT Trimethoprim-Sulfamethoxazole
tEPEC Typical and enteropathogenic *E. coli*
TGS Third-generation sequencing or long-read sequencing appeared
TI Ticarcillin
TMP/SMX Trimethoprim-sulfamethoxazole
Tn Transposons
TOB Tobramycin
tra Transfer region genes
TSA Tryptone soya agar
UPEC Uropathogenic *E. coli*
USC Universidade de Santiago de Compostela
UTI Urinary tract infections
VF Virulence factor
VFDB Virulence factor database
VIM Verona integron-encoded metallo-betalactamase
vs. Versus
wgMLST Whole genome MLST
WGS Whole-genome sequencing
WHO World Health Organisation
XDR Extensively drug-resistant

1. INTRODUCTION

1.1. A BRIEF INTRODUCTION TO MICROBIOLOGY

Microbiology is the study of microorganisms including bacteria (bacteriology). Bacteria are believed to be the oldest living organisms on Earth. Their traces date back 3 billion years (Hacker and Carniel, 2001). The Eubacteria (*Bacteria sensus stricto*) and the *Archaea* (ancient bacteria) domains represent two of the three greatest branches of cellular life, both constituted by unicellular prokaryotes. The third branch includes plants, animals, and fungi kingdoms, pluricellular eukaryotes. The Eubacteria domain is divided into 40 kingdoms and the *Archaea* domain is divided in two kingdoms (Dykhuizen, 2005). Therefore, bacteria are very diverse and include a wide spectrum of different shape and size species. The three most basic bacteria shapes are known as coccus (spherical), bacillus (rod-shaped) and spiral (twisted). However, there are also pleomorphic bacteria that can adopt several shapes. In regards to size, the smallest bacterium which has been discovered is named *Nanobacterium* and it has 100 nm in length, while the largest discovered bacterium named *Thiomargarita namibiensis* has a length of 0.75 mm. Besides, bacteria can be found everywhere, within animals, plants, soil, water and air (Dykhuizen, 2005).

From a human point of view bacteria has two sides, as they can appear to be our "enemies" or "allies". On one hand, many diseases have their origin in bacteria and they are responsible for some of the most devastating pandemics in human history, such as plague (*Yersinia pestis*) or cholera (*Vibrio cholerae*) (Morabia, 2020). On the other hand, many other types of bacteria are great contributors in environmental protection and/or in the human industry, for instance, food production. Besides, bacteria are part of animals' microbiota, including humans, which include up to 500 species of bacteria. The microbiota performs a wide range of essential and beneficial functions, such as controlling mucosal immunity, breaking down nutrients and preventing pathogen colonization. Moreover, according to medical research, bacteria can reduce intestinal decay and slow down aging (Kundu *et al.*, 2017; Levy *et al.*, 2017).

Biological evolution drives from the plasticity of the genomes, the rate of phenotype generation and the selective pressure exerted by the environment (Arber, 2000). In relation to the evolution of bacteria, human antimicrobial therapy is exercising an artificial selective pressure that is contributing to the selection of antimicrobial-resistant mutants (Dykhuizen, 2005). In eukaryotes, genetic variability results from sexual reproduction. Instead, prokaryotes reproduce asexually and have other mechanisms contributing to genetic variabilities, such as point mutations, recombinations and horizontal gene transfer (HGT). HGT allows bacteria to transfer, replace, or acquire pieces of DNA (or whole new genes) into their genomes. Furthermore, bacterial genetic exchange can be done between species or even genera. Therefore, bacteria evolve in impressively short periods of time (Ochman *et al.*, 2000).

Antoine van Leeuwenhoek (1632-1723), Louis Pasteur (1822-1895), and Robert Koch (1843-1910) are known as the first contributors to microbiology. Since then, the microbiology science has become vast and multidisciplinary, evolving in parallel with the development of new technologies. Nowadays the bacterial genome can be entirety sequenced and numerous whole-organism (or "omics") methods can be applied for bacteriology research (van Dijk *et al.*, 2014; Lo *et al.*, 2017).

Despite the continuous development of science, outbreaks are nearly constant and pandemics still occur, of which the current COVID-19 is proof. Research against constantly mutating microorganisms is far from over.

1.2. *ESCHERICHIA COLI* AND GENOMICS

The protagonist of this thesis, *Escherichia coli,* is a non-sporulating facultative anaerobe Gram-negative bacterium which belongs to the *Enterobacteriaceae* family. *E. coli* was discovered by Theodor Escherich in 1885 (Escherich, 1885) and it currently is one of the most studied bacteria worldwide. Interestingly, this bacterium has huge genome plasticity representing a versatile and diverse specie (Dobrindt *et al.*, 2003; Beghain *et al.*, 2018; Denamur *et al.*, 2020). Whilst *E. coli* is part of the gastrointestinal microbiota of warm-blooded animals and reptiles (Gordon and Cowling, 2003) (non-pathogenic/commensal strains), some *E. coli* strains can also be the etiological agent of a diverse range of clinical diseases that affect all age groups in humans and animals (Blanco *et al.*, 2002), causing more than 2 millions deaths per year in humans (Kosek *et al.*, 2003; Russo, 2003). Also, environmental strains have been estimated to represent half of *E. coli* population (Savageau, 1983). In addition, among resistant bacteria, *E. coli* is the most common Gram-negative bacterial pathogen (Poirel *et al.*, 2018). In consequence, the ability to distinguish types beyond the species is essential and helpful in many circumstances in *E. coli*.

Nowadays, it is known that the genus *Escherichia* is composed of *E. albertii*, *E. fergusonii*, five cryptic *Escherichia* linages (clades I to V) (Walk *et al.*, 2009; Clermont *et al.*, 2011; Luo *et al.*, 2011) and *E. coli sensu stricto* (Beghain *et al.*, 2018). *E. coli* cryptic lineages are highly divergent at the nucleotide level but phenotypically similar to *E. coli sensu stricto* and conventional phenotypic test such as API20 *Enterobacteriaceae* (BioMérieux) or MALDI-TOF mass spectrometry fail to distinguish them (Walk *et al.*, 2009). In this thesis we will refer to *E. coli sensus stricto* or classical *E. coli*.

Due of the common rearrangements and recombinational events that occur in bacterial genomes, bacterial progeny (clones) may not be identical (Ochman *et al.*, 2000; Spratt, 2004). In bacteriology, clonal lineages refer to closely related strains which have diverged from a common ancestor and are indistinguishable or highly similar to each other, using a particular molecular typing procedure (Mathers *et al.*, 2015).

Genomics is known to be the study of the structure and function of the genome of an organism. For its analysis, DNA sequencing methods, recombinant DNA and bioinformatics tools are employed (Johnson and Russo, 2018). The first *E. coli* genome which was completely sequenced was *E. coli* K-12 strain MG1655 in 1997 (Blattner, 1997). To date, thousands of *E. coli* genomes at different levels of completion are available in public databases. As of August 2020, the National Center for Biotechnology Information (NCBI) database accounted for 20,587 genome assembly and annotation reports of *E. coli* belonging from a wide range of sources.

The number of genes present in the *E. coli* genome varies from approximatively 4,000 to 6,000 genes (which corresponds to 4.2-6.0 Mbp), whereas a set of around 2000 genes provides the genetic information required for essential cellular processes which are present in all members of a specie (core genome) (Bergthorsson and Ochman, 1998). There is also a flexible set of strain-specific genes which provide genetic information that enables the harboring strains to adapt to particular environmental conditions (accessory genome). The accessory genome includes all the genes encoding bacteriophage elements, virulence-associated determinants and acquired resistance to antimicrobials.

The vast amount of sequence information generated in the era of "genomics" strongly supports horizontal gene transfer (HGT) as a cornerstone in the adaptive evolution of prokaryotes, including *E. coli*. As more *E. coli* strains are sequenced, the core genome shrinks while the accessory genome increase (Dobrindt, 2005; Brockhurst *et al.*, 2019;

Touchon *et al.*, 2020). In a recent study, it was shown that the pan genome (the total number of genes) of 4,000 *E. coli* strains accounted for more than 26,000 genes (Decano and Downing, 2019).

1.3. PATHOGENIC *E. COLI* AND PATHOGENOMIC

The term "colibacillosis" broadly refers to a disease caused by *E. coli* (Poirel *et al.*, 2018). Historically, pathogenic *E. coli* strains were separated into disease-associated groups (pathotypes) (recently reviewed in Denamur *et al.*, 2020).

Pathogenic *E. coli* are classified into two categories: (I) intestinal pathogenic *E. coli* or InPEC when they are the causative agent of intestinal infections and (II) extraintestinal pathogenic *E. coli* or ExPEC when they are the causative agent of extraintestinal infections once outside the gastrointestinal tract (reservoir) of the host (Russo and Johnson, 2000). ExPEC within the intestine rarely causes any complications and exists in a beneficial symbiotic relationship with the intestinal microbiota (Wiles *et al.*, 2008).

There are seven well-studied InPEC pathotypes classified by pathogenicity mechanisms attributed to their ability to cause gastrointestinal diseases like diarrhea (Nataro and Kaper, 1998; Kaper *et al.*, 2004), including Shiga toxin-producing *E. coli* (STEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), diffusely adherent *E. coli* (DAEC), enteroinvasive *E. coli* (EIEC), and the Crohn's disease-associated adherent-invasive *E. coli* (AIEC) (Darfeuille-Michaud *et al.*, 1998). The Enterohaemorrhagic *E. coli* (EHEC) is a type of STEC that can cause severe enteric diseases, such as hemolytic uraemic syndrome or hemorrhagic colitis leading to acute renal failure and often death (Kaper *et al.*, 2004).

ExPEC represents a distinctive group of *E. coli* strains that posseses an enhanced ability to overcome the host's defenses and cause a wide spectrum of syndromes, ranging from urinary tract infections (UTI) to life-threatening bloodstream infections (e.g. skin and soft tissue infections, biliary infections, infective pyelonephritis, pelvic inflammatory disease, bone infections, neonatal meningitis and avian colisepticemia). There are four well-studied ExPEC pathotypes, including avian pathogenic *E. coli* (APEC), uropathogenic *E. coli* (UPEC), neonatal meningitis-associated *E. coli* (NMEC) and sepsis-causing *E. coli* (SEPEC) (Dale and Woodford, 2015). However, the attribution to one specific pathotype does not exclude the strain to be able to cause infection at another anatomical site (Ron, 2006; Dale and Woodford, 2015). During recent years, complex hybrid pathotypes have also been described as, for example, a shiga toxin (STEC) EAEC carrier which caused illness in more than 3 thousand people as well as 54 deaths during a large outbreak in 2011 (Frank *et al.*, 2011).

Pathogenomics investigates processes involved in bacterial pathogenesis on the genome level. Distinctive genetic features can be linked to different lifestyles in *E. coli* (Dobrindt, 2005; Tenaillon *et al.*, 2010). The differentiation of (pathogenic) variants facilitate risk assessment, diagnostics and the prevention of *E. coli* infections. To date, molecular studies have yielded numerous important insights on this matter (Dobrindt, 2005; Johnson and Russo, 2018).

A genomic comparison of the commensal strain MG1655 (*E. coli* K12), strain EDL933 (EHEC) and strain CFT073 (UPEC) (Table 1) revealed that less than 40% (2996 genes) of their combined set of proteins were common to all three strains (Welch *et al.*, 2002). Furthermore, a comparison between MG1655 and EDL933 genomes showed 0.53 Mb and 1.4 Mb strain-specific DNA sequences, respectively. The comparison with the

CFT073 genome reveals 1.3 Mb without a counterpart in either of the other genomes (Dobrindt, 2005).

Table 1. Characteristics of publicly available complete *E. coli* genome sequences

Strain	Pathotype	Size (bp)	G+C content (%)	No. of predicted ORFs	Accession number	References
MG1655	K-12	4,639,221	50.8	4,294	U00096	(Blattner, 1997)
CFT073	UPEC	5,231,428	50.5	5,533	AE014075	(Welch *et al.*, 2002)
EDL933	EHEC	5,528,445	50.5	5,361	AE005174	(Perna *et al.*, 2001)

Adapted from (Dobrindt, 2005)

1.3.1. EPIDEMIOLOGY AND *E. COLI* ATTRIBUTES

Epidemiology studies diseases within populations, whilst molecular epidemiology emphasises observations on a molecular level in order to identify bacterial traits and host characteristics that may help to predict the occurrence, severity and/or clinical manifestations of a particular disease. This field of research also makes a contribution to identifying relevant reservoirs and transmission pathways, which are helpful in developing preventive strategies. As a limitation, molecular epidemiological studies require complementary experimental assessment to be conclusive, because gene expression has to be taken under consideration to make phenotype-genotype associations (Honardoost *et al.*, 2018; Johnson and Russo, 2018).

Firstly, *E. coli* is a zoonotic pathogen, which means that it can be transmitted, in both directions, between animals and humans, and cause various diseases (Manges and Johnson, 2012). ExPEC strains are important pathogens to both humans and certain animals. In this thesis, we will focus on the study of ExPEC strains and their hosts (also referred to as reservoirs).

1.3.1.1. ExPEC related to human diseases

ExPEC strains are the leading causative agent of human extraintestinal infections globally (Manges *et al.*, 2019). In numerous cases, they are opportunistic pathogens causing ExPEC-related illeness in patients with a compromised immune system (Russo and Johnson, 2000; Foxman, 2003; Nielubowicz and Mobley, 2010; Singer, 2015; Johnson *et al.*, 2016a; Biran and Ron, 2018).

In particular, UPEC strains are the infective agent of urinary tract infections (UTIs) and pyelonephritis in up to 80% of cases (Dale and Woodford, 2015), while NMEC strains are one of the most common causes of neonatal bacterial meningitis accounting to 30% of cases (Chu *et al.*, 2014). In addition, NMEC infections are associated with rates of 10 to 40% mortality and up to 50% of neurological sequelae in survivors (Mulder and Zanen, 1984; Mulder *et al.*, 1984; Klinger *et al.*, 2000; Stoll *et al.*, 2011). Furthermore, ExPEC are the leading cause of abscesses as well as a major cause of septicaemia.

1.3.1.2.ExPEC related to animal diseases

ExPEC strains are also the causative agent of diseases in food-producing animals and are associated with important economic losses in the food industry worldwide. In poultry, APEC strains cause infections by inhalation of fecal dust which spread, causing septicemia, pericarditis, and high rates of mortality (Antŷo *et al.*, 2008). In dairy cattle, *E. coli* is responsible for more than 80% of cases of acute mastitis, a common inflammatory response of the mammary gland that significantly decrease milk production (Shpigel *et al.*, 2008; Poirel *et al.*, 2018). In pig farms, outbreak of acute sepsis, pneumonia and meninguitis are also reported (Tan *et al.*, 2012; Ma *et al.*, 2020). Besides,

E. coli is also responsible for infections of the lower and upper urinary tract in companion animals (Pomba *et al.*, 2016; Bouillon *et al.*, 2018; Hutton *et al.*, 2018).

Secondly, *E. coli* as a common inhabitant of the gastrointestinal tract, it is subject to frequent encounters with antibiotics leading to developing resistance against the antibiotics that are frequently consumed by the host (Looft *et al.*, 2012). Thus, this bacterial group can be regarded as a sentinel for resistance to antibiotics (SCENIHR, 2010). Besides, *E. coli* is discharged to the environment through feces and wastewater treatment plants (Anastasi *et al.*, 2012) and for this reason, it is considered a major source of contamination and hazard for infection under a "one health" perspective (Pomba *et al.*, 2016; Poirel *et al.*, 2018).

Thirdly, host to host direct transmission through close contact also represents a major route of dispersion of *E. coli* strains, including human close contact with companion animals (Pomba *et al.*, 2016).

Lastly, it is known that *E. coli* can also be transmited through food and for this reason some food-producing animals which are reservoirs of ExPEC strains are considered major hazards. The linkage between ExPEC strains in food products and human outbreaks has lead to the definition of foodborne ExPEC infection.

1.4. TYPING METHODS IN *E. COLI*

Apart from pathotype, individual strains of *E. coli* can be classified using a variety of molecular and non-molecular techniques. The implementation of molecular methodologies has allowed a more detailed analysis of *E. coli* lineages (Clermont *et al.*, 2011; Beghain *et al.*, 2018). With the application of-multi-locus enzyme electrophoresis (MLEE), a detailed phylogenetical study of *E. coli* variants started. Milkman associated intragenic recombination of certain segments of the genome with bacterial lineages and proposed to use it to decipher the species phylogeny (Milkman and Stoltzfus, 1988). Ultimately, techniques such as pulse-field gel electrophoresis (PFGE) and whole-genome sequencing (WGS) provide the greatest resolution for this purpose and are both valuable in the context of surveillance and outbreak analysis. However, PFGE is a labor-intensive technique with a complex interpretation which complicates data exchange amongst laboratories. For this reason, PFGE has rapidly become an out-dated method since DNA sequencing has become more economical and widely available (Johnson and Russo, 2018).

1.4.1. SEROTYPE DETERMINATION

Serotypes are determined by the identification of three types of somatic (O), capsular (K) and flagellar (H) antigens of *E. coli* (serotyping). From which the K antigen is considerably more complicated to determine. Thus, the determination of O-groups (serogroup) and H-groups became a gold standard (Orskov and Orskov, 1992). There are 182 different O-groups and 53 H-types in *E. coli* and thousands of different combinations (O:H) have been described (Nataro and Kaper, 1998).

Serotyping was the first typing method available to delineate variants within *E. coli* (Kauffmann and Ørskov, 1956; Orskov *et al.*, 1977) because these biomarkers are often associated with specific pathotypes (Kaper *et al.*, 2004).

The O-antigen refers to a unique polysaccharide of the outer membrane of *E. coli*. This antigen is chromosomally encoded in a cluster designated as the O-antigen gene cluster (O-AGC) which includes *wzx, wzy, wzm*, and *wzt* genes (Stenutz *et al.*, 2006; Fratamico *et al.*, 2016). The H-antigen refers to the flagellar protein that is codified by up to 40 flagellin-associated genes (Macnab, 1996), most of them encoded in the *fliC* locus

(Machado *et al.*, 2000; Wang *et al.*, 2003; Fratamico *et al.*, 2016). Currently, serotyping can be performed by conventional phenotypic methods as well as by molecular methods (see below).

- Phenotypic technique: the conventional serotyping method uses antisera with specific antibodies to test the surface antigens of *E. coli*, which are detected by agglutination reactions (serology).
- Genotypic techniques: PCR-based or *in silico* determination through WGS. Molecular methods for H-typing target *fliC* genes and other less frequent flagellin-associated genes (*flkA, fllA*, *flmA or flnA*) and for O-typing target *wzx, wzy, wzm*, and *wzt* genes (Joensen *et al.*, 2015).

The phenotypic determination of *E. coli* antigens has several drawbacks. In fact, many *E. coli* strains could remain non-typeable because their O- or H-antigens are not expressed *in vitro* or because they autoagglutinate during the procedure. Also, this method requires specific antiseras which are only available in specialized laboratories and might vary in quality showing cross-reactivity with different serogroups. Additionally, conventional serotyping is a relatively costly and labor-intensive technique. For these reasons, *in silico* serotyping is likely to replace traditional serotyping in the future (Fratamico *et al.*, 2016). Nonetheless, for the identification of potential vaccine candidates it is necessary to identify the most unique genes from pathogenic microorganisms that are most effective in stimulating immunity to produce antibodies (Moriel *et al.*, 2016). In this regard, the phenotypic determination of *E. coli* antigens provides relevant information that cannot be retrieved through genotypic methods.

1.4.2. PHYLOGENETICAL TYPING

In general, *E. coli* strains with common phylogenetical origins show similar traits, including serotype and pathotype (Picard *et al.*, 1999; Gordon and Cowling, 2003; Méric *et al.*, 2013; Clermont *et al.*, 2019). Thus, the attribution of a given strain to a clonal or phylogenetic group is meaningful and it is usually performed in epidemiological studies.

E. coli phylogroups and clonal groups include closely related strains classed under different criteria (see below).

1.4.2.1 Phylogroup determination

Phylogroups are defined as a large phylogenetic entity within the species (Denamur *et al.*, 2020). Thanks to technological advances, eight phylogroups for *E. coli* (*sensu stricto*) have been recognised to date (A, B1, B2, C, D, E, F, and G). Besides, an additional group named H has also been described (Lu *et al.*, 2016b).

The phylogenetical studies performed by Beghain *et al.*, (2018) and Clermont *et al.* (2019), show that phylogroups B2, F, D and G are the more ancient lineages conforming one cluster which had further diverge with the emergence of phylogroup E, followed by phylogroups C, B1 and A which constituted an other cluster. Interestingly, phylogroups B2, F and D encompass the majority of extraintestinal pathogenic strains, whereas commensal *E. coli* strains as well as strains associated with enteric diseases are generally derived from the most recently diverged phylogroups (Picard *et al.*, 1999; Escobar-Páramo *et al.*, 2004).

Clermont and colleagues have provided several PCR assays to easily perform phylotyping through the determination of "marker" genes (*arpA*-1, *arpA*-gE, *trpA*-gC, *chuA*, *yjaA*, *Tsp*E4.C2, *ybgD* and *cfaB*) (Clermont *et al.*, 2000, 2013, 2019). This approach has been extended to *in silico* typing with WGS.

1.4.2.2 Clonal group determination

Thanks to the conserved nature of the core genome in *E. coli*, it is possible to establish the taxonomic distances between strains based on the comparison of highly conserved chromosomal genes (housekeeping genes) (Chaudhuri and Henderson, 2012). The multi-locus sequence typing (MLST), using either the Warwick University (Wirth *et al.*, 2006) or the Pasteur Institue (Jaureguy *et al.*, 2008) schemes in addition to the two-locus CH typing scheme (Weissman *et al.*, 2012) are the most extended nucleotide sequence-based techniques used for this purpose in *E. coli*.

In MLST schemes, an arbitrary number is assigned to each unique loci sequence (seven loci in the Warwick University scheme and eight loci in the Pasteur Institute scheme). MLST allows classifying strains into distinct sequence types (STs), which are defined as isolates with identical allelic profiles (Maiden *et al.*, 1998) and into broader clonal complex (CC) (clonal population with strong association of alleles) (Maiden *et al.*, 2013).

The CH scheme allows predicting the MLST-based profile with up to 95% accuracy (Weissman *et al.*, 2012) by the determination of the *fum*C and *fim*H alleles (two locus). Although CH typing will not replace traditional MLST for more detailed phylogenetic studies, CH typing works as a rapid screening test. Besides, clonotypes split large STs into a smaller number of clonal subgroups (clonotypes) allowing to decipher more clonal diversity (Dale and Woodford, 2015).

Moreover, WGS allows to perform a quantitative analysis of strain divergences by studying single nucleotide polymorphism (SNPs) in the core genome of bacterial population and enables *in silico* typing of strains for various genotyping schemes including MLST and CH. Besides, the MLST approach has recently been extended to core genome MLST (cgMLST) and whole genome MLST (wgMLST) with the purpose of obtaining a higher level of resolution in terms of taxonomic distances (Maiden *et al.*, 2013).

1.5. VIRULENCE PROFILES IN *E. COLI* AND VIRULOME

In 1988, Johnson and colleagues described “Virulence Factor (VF)” as any bacterial function that facilitates infection (Johnson *et al.*, 1988).

Molecular epidemiology studies have proved that some molecular markers exhibit a greater prevalence within *E. coli* isolates from diverse extraintestinal infection compared to those isolated from feces of healthy carriers (Johnson *et al.*, 2003; Moreno *et al.*, 2008; Kudinha *et al.*, 2013; Nielsen *et al.*, 2014). However, the epidemiological association with virulence dose not necessarily reflect causality (Schreiber *et al.*, 2017; Johnson and Russo, 2018).

Three hypotheses of disease causation have been proposed for ExPEC isolates. Firstly, it has been suggested that most urinary tract infections (UTI) are opportunistic infections caused by bacteria that predominate in the faecal flora (the prevalence hypothesis). Secondly, it has been proposed that *E. coli* requires some specific virulent traits to become pathogenic and represent a distinctive subset of the fecal *E. coli* population (special pathogenicity hypothesis) (Whittam *et al.*, 1989; Johnson *et al.*, 2003, 2005b, 2005c; Moreno *et al.*, 2008; Thompson *et al.*, 2011; Nielsen *et al.*, 2014). However, neither of those previous definitions are truly adequate, nor are the definitions mutually exclusive (Manges *et al.*, 2019). In third place, there is now growing evidence that a commensal *E. coli* will develop into a pathogen through the selection of VF-roles which are related with the ecology of the bacteria more than with their direct contribution

to pathogenesis (the coincidental evolution hypothesis) (Grozdanov *et al.*, 2004; Dobrindt, 2005; Tenaillon *et al.*, 2010; Denamur *et al.*, 2020). Paradoxically, certain data suggests that virulence may be a product of commensalism, as it has been shown that a considerable fraction of virulence-associated genes in ExPEC are also present in many commensal *E. coli* isolates, increasing their adaptability and competitiveness (Nowrouzian *et al.*, 2001, 2003; Le Gall *et al.*, 2007; Duprilot *et al.*, 2019). Besides, colonising sites outside the gastrointestinal tract are unlikely to provide any selective advantage in terms of transmissibility (Johnson and Russo, 2018; Denamur *et al.*, 2020).

In ExPEC a wide set VFs have been described (virulome). It is apparent there is a considerable variability of VF profiles within phylogenetic groups, clonal groups, and even among subclones and that no single virulence factor is sufficient to cause disease (Lane *et al.*, 2007; Hannan *et al.*, 2012; Sarkar *et al.*, 2014; Conover *et al.*, 2016).

Additionally, certain VFs commonly occur together among clinical isolates suggesting genetic linkage (Johnson and Stell, 2000; Kanamaru *et al.*, 2003; Dahbi *et al.*, 2014). For example, in NMEC the invasion of brain endothelium protein (*ibeA* gene) and the cytotoxic necrotising factor 1 (*cnf-1* gene). Thus, even if known VFs are not the actual determinants of virulence, they are useful markers and potential clinical predictors.

VFs are usually grouped by functional categories, including those related to (I) adherence and colonisation, (II) nutritional and metabolism flexibility (III) evasion of the host immune system (IV) persistence and (V) toxin production. Additionally, it has been observed that clinical isolates often contain multiple VFs from a particular functional category and it has been hypothesised that a synergistic interaction among VFs may occur (Holden and Gally, 2004). Conversely, some virulent strains lack known representatives of one or more of these functional categories and it has been hypothesised that this apparent deficit may be compensated by other functional categories. It is also possible that these strains contain VFs from other unrecognised categories (Johnson *et al.*, 2001, 2002). It is hoped that genome comparison between closely related pathogenic and non-pathogenic variants will reveal new potential targets for preventive interventions and development of vaccines.

1.5.1. Adherence and colonisation

The attachment of bacteria to mucosal surfaces is the initial event in the pathogenesis of most infectious diseases. Moreover, the adherence ability of pathogenic *E. coli* is frequently related to host and even tissue specificity (adhesins-receptor) promoting colonisation (Beachey, 1981; Kline *et al.*, 2009). Different surface appendages have been associated with the initial stage of infection in ExPEC, including (I) fimbral adhesins (II) non-fimbral adhesins and (III) flagella.

1.5.1.1. Fimbrial adhesins

Fimbriae or fimbrias are long proteinaceous organelles that extend from the surface of many bacteria and mediate in diverse functions, including adherence and biofilm formation. Fimbrial adhesins are often located at the tip of the fimbrial organelle and recognise specific ligands of the host epithel (Kline *et al.*, 2009). In Gram-negative bacteria, fimbrias are usually assembled via different protein translocation systems, such as the chaperone–usher (CU) pathway, type IV secretion pathway and the extracellular nucleation precipitation pathway. Among which the CU pathway is the most abundant and therefore constitute for example type 1 fimbriae, P fimbriae, Dr/Afa family fimbriae and S/F1C fimbriae (Busch and Waksman, 2012). ExpEC isolates often express more

than one type of fimbriae, which may increase the probability of a successful survival under changing environmental conditions (Holden and Gally, 2004).

- Type 1 fimbriae is encoded in the bacterial chromosome by the *fim* operon that codifies for several subunits named FimA, FimC, FimD, FimF, FimH and FimG (Thankavel *et al.*, 1997), where FimH represent the actual adhesin (monoadhesin fimbriae) (Mulvey, 2002; Kline *et al.*, 2009; Busch and Waksman, 2012).
- P-fimbriae or pyelonephritis-associated pili is encoded by the *pap* operon that codifies for several subunits named PapA, PapC, PapE, PapF and PapK, where PapG represent the actual adhesin (monoadhesin fimbriae) (Mulvey, 2002; Kline *et al.*, 2009; Busch and Waksman, 2012).
- Dr/Afa family fimbriae constitute polyadhesins fimbrias encode by the *afa*, *dra*, and *daa* operons which have a similar genetic organization (Servin, 2005, 2014). Dr family fimbriae can bind red blood cells (RBCs) thanks to receptor specificity to the Dr blood group antigen (Nowicki *et al.*, 1988, 2001).
- S/F1C fimbriae. The S fimbriae is encoded by the *sfa* operon that codifies for several subunits named SfaA, SfaG, SfaH and SfaS, where the SfaS is actual adhesin. The F1C fimbriae is an analog of the S fimbriae, which is encoded by *foc* genes (Mulvey, 2002).
- Curli fibrils are adhesive amyloids organelles assembled via a nucleation-dependent process and which biosynthesis is encoded by the *csg* operon that codifies for the curli fibers proteins CsgA and CsgB (Kline *et al.*, 2009).

1.5.1.2. Non-fimbrial adhesins

This group is constituted of a wide number of heterogeneous structures with adhesive properties such as autotransporter (AT) proteins, which contribute to the adherence and colonisation of ExPEC, like for example the type 1 secretion A protein (TosA), the factor adherence *E. coli* protein (FdeC), the IrgA homologue adhesin (Iha), the temperature-sensitive hemagglutinin (Tsh), the outer membrane protein antigen 43 (Ag-43) and the surface-located trimeric AT proteins (UpaH, UpaC and UpaG) (Tarr *et al.*, 2000; Johnson *et al.*, 2005a; Ulett *et al.*, 2007; Allsopp *et al.*, 2010, 2012; Kobayashi *et al.*, 2010; Wells *et al.*, 2010; Vigil *et al.*, 2011; Nesta *et al.*, 2012; Xicohtencatl-Cortes *et al.*, 2019)**.**

1.5.1.3.Flagella

The flagella is an organelle made from protein monomers called flagellins (FliC proteins) and it is composed of a basal body, a hook, a motor, and a filament (Macnab, 1996; Machado *et al.*, 2000; Wang *et al.*, 2003). This structure confers adhesive and invasive abilities, motility and plays different roles during biofilm formation (Pratt and Kolter, 1998).

1.5.2. NUTRIENTS AVAILABILITY AND METABOLIC PATHWAYS

Within the intestine, most nutrients are in sufficient quantities and the carbon source is considered as the main nutritional limitation for *E. coli* (Conway and Cohen, 2015). In contrast, outside the intestine, there is a high range of nutritionally-divers environments to overcome (Mann *et al.*, 2017). For this purpose, in ExPEC isolates VFs have been described and established into three categories: (I) the regulation of metabolic pathways (II), secretion systems machinery and (III) iron acquisition.

1.5.2.1. Regulation of metabolic pathways

The two-component signaling system (TCS) is the main signal transduction pathway to sense and respond to environmental stimuli in bacteria, including quorum sensing signals (Stock *et al.*, 2000). The TCS is composed of a membrane-bound sensor histidine kinase (HK) and a cytoplasmatic response regulator of gene expression (RR). It has been proposed that the TCS has a role in mediating the effective regulation of some metabolic pathways in ExPEC. Some TCSs that have been associated with extraintestinal pathogenesis are the BarA/UvrY, EvgS/EvgA, KguS/KguR and PhoQ/phoP systems (Tomenius *et al.*, 2006; Eguchi *et al.*, 2011; Cai *et al.*, 2013).

1.5.2.2. Secretion systems machinery

The secretion systems machinery contributes to pathogenesis by facilitating the transport of molecules across bacterial membranes (physical barriers), for example siderophores, exopolysaccharides and toxins (Fox *et al.*, 2020).

1.5.2.3. Iron acquisition systems

Iron is a critical nutritional requirement for the survival and growth of bacteria. Thus, ExPEC isolates typically contain genes codifying for iron sequestration. It is therefore essential for septicemic pathogens to be able to acquire the iron that is bound to blood proteins (such as ferritin, hemosiderin) (Robinson *et al.*, 2018).

Siderophores are molecules synthesized to bind ferric ions (Fe^{3+}) that belong to a high-affinity iron acquisition system energized by the Ton complex. The Ton complex is a cytoplasmatic membrane-localised complex that interacts with the bacterial outer membrane specific siderophores-receptors to facilitate iron uptake. The biosynthesis of siderophores in *E. coli* requires to be activated through an EntD transferase (PPTase) (Martin *et al.*, 2013). In ExPEC, several types of siderophores have been described, with different chemical structures and affinity for iron, including (I) enterobactin, (II) salmochelins, (III) aerobactin and (IV) yersiniabactin (Robinson *et al.*, 2018).

- Enterobactin is encoded within the core genome of *E. coli*. The initial biosynthesis is codified by *ent* genes and utilisation is codified by *fep* and *fes* genes.
- Salmochelin was originally identified in *Salmonella* species and it is coded by a non-conserved *iro* gene cluster. The initial biosynthesis is codified by *iroB, iroC* and *iroE* and utilisation is codified by *iroN* and *iroD*.
- Aerobactin was originally identified in *Aerobacter aerogenes* and it is coded by a non-conserved set of genes. The initial biosynthesis is codified by the operon *iucABCD* and utilisation is codified by *iutA* (outer membrane receptor).
- Yersiniabactin (Ybt) was originally identified in *Yersinia pestis*. The Ybt system is encoded in a multioperon pathogenicity island named as the *Yersinia* high pathogenicity island (HPI). The initial biosynthesis is codified by *ybtASDUET, irp1 and irp2* and utilisation is codified by *fyuA* and *ybtPQ*.

Recent data suggests a complementary role for Ybt as a metallophore that could conform stable complexes with copper and other metals, such as cobalt and nickel (Koh *et al.*, 2015).

Other iron uptake systems are the Sit, Chu, Hma, Fec and Fhu systems (Nagy *et al.*, 2001; Rodriguez-Siek *et al.*, 2005; Hagan and Mobley, 2009; Chu and Vogel, 2011). Besides, ExPEC may also produce toxins such as alfa-haemolysin, which causes cell lysis and the subsequent release of micronutrients, like heme, which is the most abundant iron complex in the host (Wiles and Mulvey, 2013).

1.5.3. EVASION OF THE HOST IMMUNE SYSTEM

Successful pathogens have developed a wide range of anti-immune strategies that play critical roles in their abilities to cause disease (Finlay and McFadden, 2006). In this regard, many VFs have been described in ExPEC isolates, such as protectins. For example, the increased serum survival (Iss), the production of conjugal transfer surface exclusion protein (TraT) and the presence of an outer membrane protein A (OmpA) (Johnson *et al.*, 2000). Additionally, surface molecules such as capsules, lipopolysaccharides (LPS) and outer-membrane vesicles also play a significant role avoiding the host immune responses (Phan *et al.*, 2013).

1.5.3.1.Lipopolysaccharides (LPS) and capsules

The external side of the outer-membrane in ExPEC is composed of various lipopolysaccharide (O antigen) and it can also be covered with a capsular layer (K antigen). These external structures can inhibit the complement-mediated killing and the phagocytic action of the host immune system (Johnson, 1991; Russo *et al.*, 2009; Goller and Seed, 2010; Phan *et al.*, 2013).

1.5.3.2.Outer-membrane vesicles

Outer-membrane vesicles contribute to vehiculate and protect (structural protection) many toxins, adhesins, enzymes and other non-protein antigens while they are released from the bacterial cell to the external environment (Ellis and Kuehn, 2010).

1.5.4. PERSISTENCE

In order to maintain the infection, ExPEC isolates have developed several strategies such as: (I) biofilm production (further explained in more detail, see biofilm formation), (II) the constitution of intracellular bacterial communities (IBCs) and (III) the constitution of quiascent intracellular reservoirs (QIRs).

Both IBCs and QIRs contribute to the virulence of UPEC. Briefly, IBCs are clonal collections of bacterial cells encased within a biofilm-like matrix and/or the cytoplasm of superficial cells of the bladder. These IBCs have been observed in the murine model, human cell-culture model and shed into human urine (Anderson, 2003; Justice *et al.*, 2004; Rosen *et al.*, 2007; Andersen *et al.*, 2012). QIRs are small communities consisting of 4 to 10 bacterial cells that are nonreplicating and can be present in both superficial epithelial and transitional bladder cells. QIRs can persist for 12 weeks (Mysorekar and Hultgren, 2006). Beyond being protected from antibiotics (Blango *et al.*, 2014), such reservoirs have been associated with recurrent infections (Wiles *et al.*, 2008; Goller and Seed, 2010; Hannan *et al.*, 2012).

1.5.5. TOXIN PRODUCTION

Toxins are molecules which may cause defects in the function or shape of the host cells, stop the cell cycle, or even originate cell lysis. In ExPEC, the production of several toxins has been described - this includes the alfa-haemolysin (Hly), the cytotoxic

necrotizing factor 1 (CNF1), the cytolethal distending toxin (CDT), the vacuolating autotransporter cytotoxin (Vat) and the secreted autotransporter toxin (Sat) (Wiles *et al.*, 2008; Denamur *et al.*, 2020).

1.6. UROPATHOGENIC *E. COLI* (UPEC)

Urinary tract infection (UTI) is diagnosed when a certain number of bacteria are retrieved from urine (generally $>10^5$/ml) (Smelov *et al.*, 2016). UTIs have numerous outcomes, including asymptomatic bacteriuria, acute, chronic, and recurrent UTI infection (rUTI). Besides, UTIs can be clinically classified as uncomplicated or complicated. The majority of uncomplicated UTIs are manifested as infections of the lower urinary tract, which has two levels: the infection and inflammation of the urethra (urethritis) or the urinary bladder (cystitis). In complicated UTIs, the infection ascends thought ureters to the upper urinary tract resulting in kidney infection (pyelonephritis) and even potentially bloodstream infection (septicemia). Fortunately, UTIs rarely cause any serious damage and they can be naturally cleared by the host immune system. Nonetheless, some complicated cases required prolonged therapy and are associated with serious complications and treatment failure (Micali *et al.*, 2014; Flores-Mireles *et al.*, 2015; Smelov *et al.*, 2016; Zacchè and Giarenis, 2016; Terlizzi *et al.*, 2017).

In humans, a prevalence of about 150 million UTIs per year worldwide has been estimated (Harding and Ronald, 1994). Also, it has been stated that almost half of women and 12% of men will experience at least one UTI during their lifetime. Furthermore, it is estimated that one-third of women will be prescribed antibiotics to treat a UTI before the age of 24 and will experience a rUTI despite antibiotic treatments (Foxman *et al.*, 2000; Foxman, 2003, 2014; Micali *et al.*, 2014; Tamadonfar *et al.*, 2019).

There are different pathogens implicated in UTIs, including Gram-negative bacilli like *Klebsiella* or Gram-positive cocci such as *Enterococcus* or *Staphylococcus*. However, it has been estimated that uropathogenic *E. coli* (UPEC) accounts for up to 75% of all cases. Furthermore, UPEC infections have an associated risk of around 30% of recurrent infection within 6 months (Foxman *et al.*, 2000; Ronald, 2003; Singer, 2015).

It has been recognised that fecal microbiota is a primary source of *E. coli* causing UTIs via the fecal-urethral pathway (Yamamoto *et al.*, 1997; Johnson *et al.*, 2016a). The fecal-urethral infection model derivates from the ascending infection model developed by Tomas Stamey in the late 1970s. This model suggests that UTIs are caused by strains which are present in the host's own microbiota. Following a progressive ascending colonisation process through the urethra up to the bladder.

UPEC pathogenesis cycle starts with the adhesion to the bladder and further progress with the colonisation of the bladder lumen. Some UPEC may reach interstitial spaces and/or even intracellular sites and constitute intracellular bacterial communities (IBCs). As a host response to infection, the urothelium typically exfoliates and exposes the underlying transitional epithelium that UPEC can further invade, conforming small quiescent intracellular reservoirs (QIRs) (Terlizzi *et al.*, 2017).

Favoring the prevalence hypothesis, a correlation with the host fecal predominant *E. coli* strains and the urine-infective *E. coli* strain has been observed (Yamamoto *et al.*, 1997). Concerning the special pathogenicity hypothesis, several studies have demonstrated that UPEC exhibits more VFs than the host concurrent fecal clones (Johnson *et al.*, 2003, 2005c, 2005b; Moreno *et al.*, 2008; Bahadori *et al.*, 2019). In a recent study, Bahadori and colleagues found that in around 70% of patients the urine-

infective *E. coli* strain was distinct from any fecal strain in the same patient (Bahadori *et al.*, 2019).

There is evidence corroborating host-to-host transmission (whether direct or indirect) and different ways of transmission have been addressed including sexual activities, vaginal microbiota, contaminated food, or even close animal contact (Singer, 2015). Regarding rUTI, studies reported recurrences caused by the same strain which were identified in the previous infection (Ulla-Maija Kãrkkãinen, 2000; Jantunen *et al.*, 2002; Johnson *et al.*, 2016a). Those recurrences could suggest (I) an underlying host predisposition and/or (II) the existence of a persisting endogenous focus, and/or (III) the reintroduction of the strain from a persisting external reservoir. In either situation, broader knowledge and investigation of the route of transmission and spread of ExPEC strains may help to control and prevent further infections (Singer, 2015).

1.7. RESISTANCE PROFILE IN *E. COLI* AND RESISTOME

In the following section, we revisit the theoretical framework around antibiotic therapy and bacterial resistance.

1.7.1. ANTIBIOTICS AND THEIR USE

Antibiotics are a diverse group of antimicrobial substances which target specific bacterial structures or processes that interfere with bacterial growth and reproduction (bacteriostatic effect) and/or cause bacterial death (bactericidal effect) (Pal *et al.*, 2017).

The introduction of antibiotics, through the discovery of penicillin by Alexander Fleming in 1928, revolutionised the treatment of bacterial diseases and was a milestone for the development of modern medicine. Thanks to antibiotics, many lethal diseases such as scarlet fever, purulent pharyngitis, diphtheria, or gonorrhea were effectively inhibited and human life expectancy was substantially extended (Laxminarayan *et al.*, 2016; Durand *et al.*, 2019).

Nowadays, antibiotics are extensively used for different proposes in human society (i.e. human medicine, veterinary medicine, aquaculture and agriculture) (Cabello, 2006; Van Boeckel *et al.*, 2014; Meek *et al.*, 2015). In fact, it has been estimated that the use of antimicrobials in food-production animals reached approximately 131,109 tons in 2013 and will continue to increase (Van Boeckel *et al.*, 2017). In medicine, antibiotics are used to treat and prevent infections (prophylaxis). For example, during medical procedures such as surgery or organ transplants, and they are crucial for patients with compromised immune function (Holmes *et al.*, 2016; Laxminarayan *et al.*, 2016). However, since the "golden age" of the development of antibiotics in 1960s, very few new antibiotics have been discovered. The main antibiotic classes and their mechanism of action are summarised in Table 2.

The use of antibiotics by humans began less than 70 years ago, but bacteria have been evolving with natural antimicrobial compounds for billions of years (Martínez, 2012, 2014). The human overuse and misuse of these compounds is exerting a selective pressure among bacteria which has favoured and accelerated the development of bacterial resistance mechanisms. Subsequently, once treatable bacteria are now either untreatable or require to use last line of antibiotics (Ventola, 2015). Therefore, Multidrug-resistant (MDR), extensively drug-resistant (XDR) and pan-drug-resistant (PDR) strains of *E. coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii* and *Pseudomonas aeruginosa* are now reported worldwide (Boucher *et al.*, 2009; Shlaes *et al.*, 2013).

The World Health Organization (WHO) has warned that a postantibiotic era could be approaching (WHO, 2014) and has recommended to keep in reserve some antibiotics, by

introducing the “Access”, “watch” and “reserve” (AWaRe) classification of antibiotics (WHO, 2017). In 2019, there was an update on the AWaRe classification Database (see AWaRe web portal, in https://adoptaware.org/). The “access” category includes antibiotics with activity against a wide range of common susceptible pathogens, which can be used for empiric treatments (48 antibiotics). The “watch” category includes antibiotics with higher resistance potential that should be used with precaution (110 antibiotics). Lastly, the “reserve” category includes antibiotics which should be used only if other options have failed (last-resort antibiotics) (22 antibiotics). Besides, in 2004 the world organisation for animal health (OIE) has also developed a list of critically important antimicrobial agents in veterinary medicine which has also been updeated (OIE, 2019).

Table 2. Antibiotics classes, agents and mechanism of action.

Antibiotic Class	Representative antibiotic agents	Antimicrobial action
Beta-lactams	Penicillins[1], cephalosporins (1st, 2nd, 3rd and 4th generations), monobactams (aztreonam), carbapenems (imipenem, meropenem), cefamycin (cefoxitin)	Inhibit bacteria cell wall biosynthesis (bactericidal)
Aminoglycosides	Streptomycin, neomycin, kanamycin, gentamicin and tobramycin	Inhibit the synthesis of proteins by bacteria (bactericidal)
Quinolones and Fluoroquinolones	Nalidixic acid, delafloxacin, norfloxacin, Flumequine, enrofloxacin, ciprofloxacin, levofloxacin and moxifloxacin	Interfere with bacteria DNA replication and transcription (bactericidal)
Glycopeptides	Vancomycin, dalbavancin, oritavancin and teicoplanina	Inhibit bacteria cell wall biosynthesis (bactericidal)
Lipopeptides	Daptomycin and surfactin	Disrupt multiple cell membrane functions. Inhibit the synthesis of proteins, RNA and DNA by bacteria (bactericidal)
Streptogramins	Streptogramin A+B, virginiamycin A+B and Pristinamycin	Inhibit the synthesis of proteins by bacteria (bactericidal).
Polypeptide antibiotics	Polymyxin B, polymyxin E (colistin) and bacitracin	Inhibit bacteria cell wall biosynthesis (bactericidal)
Epoxide antibiotics	Fosfomycin	Inhibit bacteria cell wall biosynthesis (bactericidal)
Nitrofurans	Nitrofurantoin	Inhibit enzymatic pathways (bactericidal)
Nitroimidazoles	Metronidazole and ornidazole	Interfere with bacteria DNA replication and transcription (bactericidal)
Ansamycins	Geldanamycin, naphthomycin and Rifamycin	Inhibit the synthesis of RNA by bacteria (bactericidal)
Macrolides	Erythromycin, josamycin, midecamycin, spiramycin, fidaxomicin, azithromycin and clarithromycin	Inhibit the synthesis of proteins by bacteria (bacteriostatic agent, occasionally bactericidal at high dose levels)
Lincosamides	Lincomycin and clindamycin	Inhibit the synthesis of proteins by bacteria (bacteriostatic agent, bactericidal agent at high dose levels)
Tetracyclines	Tetracycline, chlortetracycline, oxytetracycline, eravacycline, minocycline, doxycycline, limecycline and glycylcyclines (3rd generation tetracyclines: tigecycline)	Inhibit the synthesis of proteins by bacteria (bacteriostatic)
Phenicols	Chloramphenicol	Inhibit the synthesis of proteins by bacteria (bacteriostatic)
Aminocyclitols	Spectinomycin	Inhibit the synthesis of proteins by bacteria (bacteriostatic, occasionally bactericidal at high dose levels)
Sulfonamides	Prontosil, sulfanilamide, sulfadiazine, sulfamethoxazole	Inhibit metabolic pathways of bacteria (bacteriostatic)
Trimethoprim	Trimethoprim	Inhibit metabolic pathways of bacteria (bacteriostatic)
Oxazolidinones	Linezolid, posizolid, cycloserine, Tedizolid	Inhibit the synthesis of proteins by bacteria (bacteriostatic)
Pleuromutilins	Pleuromutilin and Tiamulin	Inhibit the synthesis of proteins by bacteria (bacteriostatic)
Pseudomonic acid antibiotics (Carboxylic acid)	Mupirocin	Inhibit the synthesis of proteins by bacteria (bacteriostatic, bactericidal agent at high dose levels)
Steroid antibiotics	Fusidic acid	Inhibit the synthesis of proteins by bacteria (bacteriostatic)
Ketolides	Telithromycin	Inhibit the synthesis of proteins by bacteria (bacteriostatic, bactericidal agent at high dose levels)

[1]Frequently used in combination with clavulanic acid, sulbactam or tazobactam (beta-lactamase inhibitors) (Blondeau, 2009; Fernández-Rodríguez and Prieto-Lago, 2017b; Durand *et al.*, 2019).

An important task for medical microbiology is to perform antimicrobial susceptibility tests (ASTs) to guide antibiotic therapy. ASTs provide the phenotypic resistance profile of bacteria and a broad variety of methods have been developed for this purpose (phenotypic testing has been reviewed in detail by Jorgensen and Ferraro, (2009). Fundamentally, AST assesses how well bacteria grow in presence of a tested antibiotic and the results can be expressed either as a concentration (minimum inhibitory concentration or MIC) or zone diameter (in millimeters). These results allow classifying the bacterial strain phenotype to a given antibiotic in one of three categories: susceptible, resistant and intermediate.

According to the international standard, (ISO 20776-1) these terms are defined as follows: (I) "susceptible" when a bacterial strain is inhibited *in vitro* by a concentration of drug that is associated with a high likelihood of therapeutic success, (II) "resistant" when a bacterial strain is inhibited *in vitro* by a concentration of drug that is associated with a high likelihood of therapeutic failure (II) "intermediate" when a bacterial strain is inhibited *in vitro* by a concentration of drug that is associated with an uncertain therapeutic effect - this last category suggests that a higher-than-normal drug dosage may be needed to ensure efficacy.

As a whole, data of this type can be used to form a picture of the degree of resistance to each drug in the population at large and classify bacteria in one of the following three categories: (I) "usually sensible" when less than 10% of the bacteria population showed resistance (II) "not usually sensible" when 10 to 15 % of the bacteria population showed resistance and such antibiotic is not recommended for "empirical" treatments, and (III) "usually resistant" when more than 50% of the bacteria population was not sensible and such antibiotic is not recommended for treatment (Rodloff *et al.*, 2008; Blondeau, 2009; Martínez, 2014; Fernández-Rodríguez and Prieto-Lago, 2017a).

For the success of antimicrobial therapy, it is essential to correctly assign clinical isolates to the interpretive categories. Thus, standardizing testing techniques and establishing breakpoints (Hombach *et al.*, 2013). In this regard, the international organisations Clinical and Laboratory Standards Institute (CLSI) and the European Union Committee on Antimicrobial Susceptibility Testing (EUCAST) provide standards with the aim of converting MIC and zones of inhibition measurements into categorical resistance determinations (CLSI, 2017; EUCAST, 2020).

1.7.2. BACTERIAL RESISTANCE IN *E. COLI*

Resistance can be defined as the ability of a microorganism to multiply in the presence of a toxic compound and is applicable to bacteriostatic and bactericidal antibiotics, whereas tolerance determines the absence of growth yet with the existence of bacterial survival in the presence of a bactericidal antibiotic (Lewis, 2007, 2008).

Bacteria can display intrinsic resistance or develop acquired resistance against antibiotics (see below) and the resistome includes all existing resistance genes (known or unknown) of a given bacteria (Perry *et al.*, 2014).

Intrinsic resistances occur in all taxonomically related bacteria, when the physiological or biochemical makeup of that species naturally makes these groups of bacteria resistant to a specific drug. Intrinsic resistances are chromosomally encoded and can only be transmitted vertically (descendence). Acquired resistance is only present in a certain lineage of bacteria derived from a susceptible parent due to the presence of a resistance gene that is not natively present. Acquired resistances are frequently encoded outside the chromosome in mobile genetic elements (MGEs) and can be transferred horizontally between bacteria.

Moreover, is not uncommon for a single bacterium to express resistance to different compounds simultaneously. The MDR, XDR and PDR denominations describe the most worrying patterns of resistance. MDR was defined as "acquired non-susceptibility to at least one agent in three or more antimicrobial categories", XDR was noted as "non-susceptibility to at least one agent in all but two or fewer antimicrobial categories" and PDR was defined as "non-susceptibility to all agents in all antimicrobial categories" (Magiorakos *et al.*, 2012). These phenomenons are mainly caused by co- or cross-resistance mechanisms.

Co-resistance occurs when a bacterium due to the acquisition of two or more different resistance genes acquires resistance to more than one family of antibiotics. In contrast, cross-resistance means that the acquisition of one specific mechanism of resistance results in resistance to more than one family of antibiotics. These phenomenons can maintain and promote antibiotic resistance in bacterial populations even in the absence of one of the affected compounds (co-selection) (Pal *et al.*, 2017).

There are six main mechanisms of bacterial resistance: (I) enzymatic alteration of the antibiotic; (II) alteration or protection of the target site; (III) restriction of access to the target, (IV) overproduction of the target (V) reduced permeability and (VI) active efflux (Martínez, 2014; Munita and Arias, 2016).

Over the past few decades, antimicrobial resistance among *E. coli* clinical isolates has increased significantly. In what follows we will outline the effluxome activity and acquired antibiotic resistance genes (ARGs) reported in *E. coli*.

1.7.2.1. The *E. coli* effluxome

The "effluxome" of *E. coli* includes a whole set of transporters which function in a coordinated mode in order to maintain appropriate homeostasis under changing environmental conditions (Schuldiner, 2018). Furthermore, it contributes to reducing the intracellular antibiotic concentration by decreasing passive uptake (bacterial porines) or increasing active efflux (bacterial efflux systems) through bacterial membranes in a process termed efflux (Figure 1).

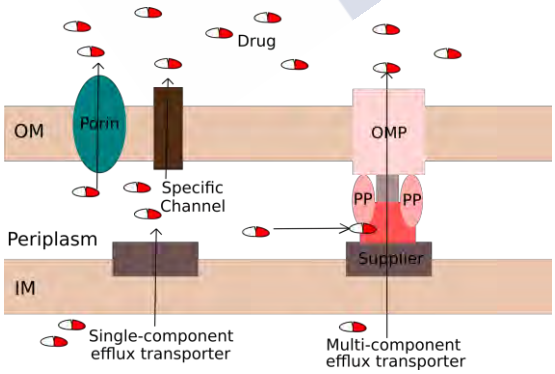


Figure 1. *E. coli* effluxome chematic representation. The influx of drugs (shown as pills) through the OM (outer membrane) and IM (inner membrane) occurs in one or more of the following three pathways: porin channels, specific protein channels and the LPS-containing asymmetric lipid bilayer region. Further, these drugs can by extruded out by single-component efflux transporters (e.g., Tet pumps) or multi-component efflux transporters (e.g., AcrAB-TolC tripartite efflux systems that typically contain a supplier, an OM channel protein [OMP], and an accessory periplasmatic adaptor protein [PP]) (adapted from Li *et al.* (2015)).

What is more, recent studies have pointed out efflux as a first-line defense mechanism which allows bacteria to fix specific beneficial mutation favorising the development and spread of resistant mutants. Also, several studies support that efflux pump overproduction mechanism in *E. coli* isolates is becoming more widespread over time (Pagès *et al.*, 2008; Fernandez and Hancock, 2012; Nikaido and Pagès, 2012; Schuldiner, 2018).

1.7.2.1.1. Bacterial Porines

Porin proteins are channels which allow the passive diffusion of hydrophilic small molecules through the outer membrane (OM) permeability barrier. Therefore, porins are one major route of entry for hydrophilic antibiotics within the bacterial cell. For this reason, mutations affecting the expression and/or function of porins have a direct impact on the susceptibility of bacteria to antibiotics.

The loss of a particular porin has generally a minor to moderate effect on the overall resistance of the microorganism, leading to a low-level resistance (not phenotypically significant). However, the accumulation of mutational events affecting various resistance mechanisms can gradually confer a high-level resistance (phenotypically significant). Several types of porins have been described and can be broadly classified as general porins, such as OmpF and PhoE in *E. coli* and specific porins for those which enable the uptake of specific substrates (Pagès *et al.*, 2008; Fernandez and Hancock, 2012).

1.7.2.1.2. Bacterial efflux systems

Bacterial efflux systems or pumps are energy-dependent systems which facilitate the transport of toxic compounds out of the bacterial cell.

Efflux systems can be divided in two main classes: (I) the ATP-binding cassette (ABC) transporters and (II) the secondary multidrug transporters (MDTs). The major difference between those classes is the source of energy required for transport. ABC-type systems use the energy derived from ATP hydrolysis, whereas MDTs utilize proton motive force. MDTs comprise the majority of clinically relevant efflux systems and encompass (I) the major facilitator superfamily (MFS), (II) the small multidrug resistance (SMR) family, (III) the multidrug and toxic compound extrusion (MATE) family, and (IV) the resistance-nodulation-cell division (RND) family (Fernandez and Hancock, 2012; Schuldiner, 2018).

Some members of the ABC transporters (e.g., MacB), the RND family (e.g., AcrAB–TolC complex), the MATE family (e.g., MdtK), and even the MFS (e.g., EmrB) are structured as tripartite complexes. The suppliers (first component) are located in the inner membrane (IM) (cytoplasmic membrane) and remove toxic compounds from the cytoplasm to the periplasmic space. Suppliers interact with the periplasmic adaptor protein (second component) and the outer membrane (OM) channel (third component) to finally excrete the desired substance to the external medium (Fernandez and Hancock, 2012; Schuldiner, 2018).

Most efflux systems are chromosomally encoded and can transport multiple different substrates despite certain examples of specific efflux pumps. In general, specific efflux pumps are encoded in mobile genetic elements (MGEs)(Butaye, 2003).

1.7.2.2. Acquired resistance to β-lactams in *E. coli*

The beta-lactam family includes a broadly used group of antibiotic agents due to their wide spectrum of action and low toxicity.

Bacterial resistance to beta-lactams is mainly due to the production of enzymes (betalactamases). Some betalactamases are intrinsic to *E. coli* (i.e., penicillin binding proteins and class C betalactamases) and are chromosomally encoded, although these

enzymes generally express a low level of resistance and are not phenotypically significant. Most betalactamases are harbored in mobile genetic elements (MGEs) and can be transferred horizontally between bacteria, confering high levels of drug resistance (Munita and Arias, 2016).

Currently, hundreds of betalactamases have been described and two classification schemes have been globally accepted based on either their structure or by their functional properties. Firstly, the Ambler or structural classification is based on the enzyme amino acid identities and active site. It separates betalactamases into 4 classes: A, C and D serine betalactamases and B metallo-betalactamases (Ambler, 1980). On another note, the Bush-Jacoby or functional classification separates betalactamases into four categories (each with several subgroups) according to their biochemical function and correlated the properties of the enzyme with the expressed resistance (Bush and Jacoby, 2010; Bush, 2013).

1.7.2.2.1. Extended-spectrum β-lactamases (ESBLs)

Extended-spectrum betalactamases (ESBLs) are a group of enzymes which cause resistance to the oxyiminocephalosporins (i.e., cefotaxime, ceftazidime, ceftriaxone, cefuroxime, cefepime) and the monobactams (i.e., aztreonam), but not the cephamycins (i.e., cefoxitin) or the carbapenems (i.e., imipenem, ertapenem). They are inhibited by betalactamase inhibitors (i.e., clavulanic acid, sulbactam, and tazobactam) (Munita and Arias, 2016). Class A betalactamases are the most common sources of acquired ESBLs in *E. coli*, and include SHV, TEM and CTX-M families.

The cefotaximase of Munich (CTX-M) is a family of enzymes which has become the most prevalent ESBL driver worldwide (CTX-M enzymes origin and diffusion has been recently reviewed in detail by Bevan *et al.* (2017). This group has been structurally classified into several subfamilies, among which subfamilies CTX-M-1 (e.g., CTX-M-15 variant) and CTX-M-9 (e.g., CTX-M-14 and CTX-M-27 variants) are the most frequently reported worldwide.

Extended-spectrum β-lactamase-producing *Escherichia coli* (ESBLEC) is an important cause of urinary tract and bloodstream infections, as well as other types of human extraintestinal infections (Arnaud *et al.*, 2015; Mamani *et al.*, 2019). The main sequence type (ST) responsible for the global increase in ESBLEC is, without a doubt, ST131 (Nicolas-Chanoine *et al.*, 2007b, 2014). Subclade C2 (also known as subclone *H*30Rx) associated with the production of CTX-M-15 seems to be the most expanded and successful ST131 sublineage (Banerjee *et al.*, 2013b; Price *et al.*, 2013; Dahbi *et al.*, 2014; Peirano *et al.*, 2014b; Sauget *et al.*, 2016). However, cluster C1-M27 that belongs to subclade C1 and produces CTX-M-27 has recently expanded, first in Japan (Matsumura *et al.*, 2016, 2017b), then in other countries (Thailand, Australia, Canada, USA, France, Italy, Germany, The Netherlands and Spain) (Blanc *et al.*, 2014; Birgy *et al.*, 2016; Bevan *et al.*, 2017; Merino *et al.*, 2018; Peirano and Pitout, 2019). The other STs frequently found among ESBLEC are ST10, ST38, ST405, ST410, ST648 and ST1193 (Brisse *et al.*, 2012; Naseer *et al.*, 2012; Peirano *et al.*, 2012; Izdebski *et al.*, 2013; Peirano and Pitout, 2019). Besides there are notable differences with regard to the ESBLEC epidemiology according to the countries, including, sometimes, the emergence of new lineages. Thus, a new clade of ST131 associated with the production of the CTX-M-101 enzyme and the *fimH*27 allele has recently been identified in *E. coli* responsible for bacteremia in Denmark (Roer *et al.*, 2017).

The horizontal gene transfer (HGT) of a wide variety of mobile genetic elements (MGEs) has been associated with the dissemination of betalactamases. Likewise, the

recurrent rearrangements in and around ESBL genes have made their surrounding genetic environments very diverse (Dhanji *et al.*, 2011; Munita and Arias, 2016; Bevan *et al.*, 2017).

1.7.2.2.2. Plasmidic aquired cephalosporinases (pAmpC)

Plasmidic aquired cephalosporinases (pAmpC) are plasmid-borne enzymes that cause resistance to narrow-, broad-, and expanded-spectrum cephalosporins and cephamycins and which are not inhibited by betalactamase inhibitors (Munita and Arias, 2016). This group belongs to class C betalactamases.

During the last decade, a wide variety of plasmid-borne *ampC* genes have been reported (Jacoby, 2009), including CMY-, DHA-, and ACC-type betalactamases, amongst which the CMY-type is the most prevalent (Philippon *et al.*, 2002).

1.7.2.2.3. Carbapenemases

Carbapenemases are a group of enzymes which have the ability to hydrolyze almost all beta-lactam antibiotics, including carbapenems (last line of beta-lactams). Carbapensemases include members belonging to classes A, B, and D betalactamases which confer a wide and versatile spectrum of resistance to beta-lactams.

Class A includes *Serratia marcescens* enzymes (SME), imipenemases (IMI), New Delhi metallo-betalactamases (NDM), Guyana extended-spectrum-lactamases (GES) and the *Klebsiella pneumoniae* carbapenemases (KPC) families; class D includes oxacillinase-48-types betalactamases (OXA-48); and class B which are metallo-betalactamases (MBLs) includes Imipenemase metallo-betalactamases (IMP), Verona integron-encoded metallo-betalactamases (VIM), Sao Paulo metallo-betalactamases (SPM), German imipenemases (GIM) and Seoul imipenemases (SIM) families (Munita and Arias, 2016).

These enzymes were first discovered among nonpathogenic bacteria and were commonly chromosomally encoded. Nowadays, plasmid-encoding carbapenemases have been reported in several studies, involving KPC, NDM and OXA-58 enzymes that represent major threats for human medicine (Nordmann *et al.*, 2011). Fortunately, carbapenemase-producers are not as widespread as ESBL-producers within *E. coli* yet, but they are increasing (Peirano *et al.*, 2014a).

1.7.2.3. Acquired resistance to polimyxins and other families of antibiotics in *E. coli*

Due to the extensive use of antibiotics, the therapeutic options against enterobacteria have changed over the last years and resistance to almost all classes of antibiotics has been reported in *E. coli*. Worryingly, resistance to last-resort antibiotics such as carbapenems and polimyxins has already been reported in the pandemic *E. coli* clone ST131 (Trobos *et al.*, 2009; Schink *et al.*, 2013; Hasman *et al.*, 2015; Sonnevend *et al.*, 2016; Ewers *et al.*, 2016; Kuo *et al.*, 2016; de Toro *et al.*, 2017; Ortiz de la Tabla *et al.*, 2017; García-Meniño *et al.*, 2018; Liu *et al.*, 2018b; Ellaby *et al.*, 2019; Hojabri *et al.*, 2019; Reid *et al.*, 2019). Colistin (polymyxin E) is associated with nephro and neurotoxicity. However, it has been widely used in veterinary medicine. Livestock, and particularly porcine farming, has been singled out as reservoir for colistin resistance (García-Meniño *et al.*, 2019) and foodborne pathogens (FBPs) that could have devastating health and economic consequences (Sekse *et al.*, 2017).

Intrinsic resistance to colistin has been related with two chromosomally encoded systems, the PhoPQ component and the *pmr*CAB operon (Olaitan *et al.*, 2014; Poirel *et al.*, 2018). However, the newly discovered plasmid-borne mobile colistin resistance (*mcr*) gene is responsible for a transferable mechanism of resistance (Liu *et al.*, 2016d). Since

the description of the *mcr*-1 gene by Liu *et al.* (2016b), several amino acid variants have been described, encoded in *mcr*-2 to *mcr*-9 genes (AbuOun *et al.*, 2017; Borowiak *et al.*, 2017; Carattoli *et al.*, 2017; Yin *et al.*, 2017; Wang *et al.*, 2018b; Yang *et al.*, 2018; Carroll *et al.*, 2019). Besides, several sub-variants for some of them have been reported, i.e. thirteen nucleotide and protein variants in the *mcr*-1-family (designated *mcr*-1.1 to *mcr*-1.13) (Partridge *et al.*, 2018a).

Colistin plasmid-borne resistance has widely spread geographically (Kempf *et al.*, 2016). It has been described in numerous genera of *Enterobacteriaceae* including *Escherichia, Moraxella, Klebsiella, Salmonella, Enterobacter, Cronobacter, Shigella, Kluyvera, Citrobacter* and *Raoultella* (Campos *et al.*, 2016; Zhao and Zong, 2016; Liu *et al.*, 2016d, 2016a; Olaitan *et al.*, 2016; Pham Thanh *et al.*, 2016; Stoesser *et al.*, 2016a; Zeng *et al.*, 2016; AbuOun *et al.*, 2017; Li *et al.*, 2017c; Luo *et al.*, 2017). Although *mcr* genes have been found in a large diversity of clones, Matamoros *et al.* (2017) and García-Meniño *et al.* (2019) established the ST10 and ST155 *E. coli* clones as potential reservoirs of the *mcr*-1 gene.

The *mcr* gene has been detected in many plasmid types, including IncI2, IncHI2, IncP, IncX4, IncY, IncF and ColE10-like ones (Madec and Haenni, 2018) from different origins (Sun *et al.*, 2018). Nevertheless, Matamoros *et al.*, (2017) and García-Meniño *et al.* (2019) found that the majority of the *mcr*-carrier plasmids belonged mainly to four plasmid incompatibility groups: IncX4, IncI2, IncHI2 and ColE10-like. In contrast, it remains rare to find colistin resistance genes chromosomally encoded, in accordance to Li *et al.* (2018) studies, where the prevalence of chromosomal *mcr* carrier isolates was estimated in 4% of the analysed isolates.

Broadly, the *mcr-1* cassette is described as an approximately 2,600 base pair fragment containing the *mcr*-1 gene followed by a phosphoesterase (Poirel *et al.*, 2016). It has been proposed that the IS*Apl1* insertion sequence (IS) mediates the transmission of *mcr*-1 by forming circular intermediates, which can translocate (Tegetmeyer *et al.*, 2008; Chandler and Siguier, 2013; Snesrud *et al.*, 2016; Zurfluh *et al.*, 2016a, 2016b; Li *et al.*, 2017a).

Table 3 summarises ARGs described in *E. coli* which confer resistance to families of antibiotics distinct from beta-lactams. Among them it can be noted the recently described plasmid-borne *mcr* gene, which confers resistance to colistin (Liu *et al.*, 2016d).

Table 3. Resistant determinants described in *E. coli* (excluding b-lactam resitant determinants).

Drug class	Efflux mechanisms	Alterations in drug targets	Other acquired mechanisms
Quinolones and fluoroquinolones	RND-type efflux pump (e.g., AcrAB-TolC) and QepA and OqxAB active efflux pumps[1]	Mutations in DNA gyrase (GyrA and GyrB subunits) and topoisomerase IV (ParC and ParE subunits)[2]	Plasmid-mediated quinolone resistance (PMQR): Qnr-like proteins (QnrA, QnrB, QnrC,QnrD, and QnrS) and AAC(6′)-Ib-cr acetyltransferase[3].
Aminoglycosides	RND-type efflux pump (e.g., AcrAD-TolC).	Mutations on the 16S RNA and/or the S5 and S12 ribosomal proteins.	Plasmid-mediated 16S rRNA methylases (rmtB)[4] and aminoglycoside modifying enzymes (AMEs)[5]: aminoglycoside acyltransferases (AACs), aminoglycoside nucleotidyltransferases (ANTs), and aminoglycoside phosphotransferases (APHs).
Epoxide antibiotics (fosfomycin)		Mutations in the glpT and uhpA/T genes encoding proteins related to the fosfomycin uptake system.	Fosfomycin-modifying enzymes such as the metalloenzymes FosA, FosB, and FosX or the kinases FomA and FomB.
Tetracyclines	MF-type efflux pump (e.g., Tet efflux proteins [tet(A), tet(B), tet(C), tet(D), tet(E), tet(G), tet(J), tet(L)]).		Tetracycline resistance genes encoding ribosome protective proteins [tet(M) and tet(W)], and an oxidoreductase that inactivates tetracyclines [tet(X)].
Phenicols	MF-type efflux pump (e.g., active efflux of non-fluorinated phenicols (*cml*A genes) or fluorinated and non-fluorinated phenicols (*floR* genes)).	Target site methylation by rRNA methylases (*cfr* gene[6])	Enzymatic inactivation of non-fluorinated phenicols by chloramphenicol acetyltransferases (*cat* genes).
Sulfonamides		Mutations in the chromosomal dihydropteroate synthase (DHPS).	Acquisition of dihydropteorate synthetases (DHPS) (*sul*1, *sul*2, or *sul*3 genes).
Trimethoprim		Promoter mutation and overproduction of chromosomal dihydrofolate reductase (DHFR)	Acquisition of dihydrofolate reductases (*dfr*A and *dfr*B genes).
Macrolides	MF-type efflux pump (e.g., macrolide specific efflux pump encoded by *mef(*A)) and RND-type efflux pump (e.g., AcrAB-TolC)	Modification of the macrolide target (23S rRNA) by methylases (*erm*(B) gene).	Acquisition of phosphotransferases (*mph*(A), *mph*(B) genes) and esterases (*ere*(A), *ere*(B) genes).
Polypeptide antibiotics (Polymyxins)		Modification of the outer membrane through mutations in PmrA and/or PmrB or in the PhoPQ two-component system or its regulator MgrB.	Acquisition of MCR phosphoethanolamine transferases (*mcr* gene)

[1]Plasmid-born; [2]Single mutations in the gene *gyr*A may confer resistance to quinolones, but for resistance to fluoroquinolones, further mutations within *gyr*A and/or *par*C are needed; [3]Also confer resistance to aminoglycosides; [4]Affect the ability of aminoglycosides to bind to the ribosome via methylation of the aminoglycoside-binding site; [5]For each enzyme, several variants have been reported; [6]Confers resistance to five classes of antimicrobial agents, including fluorinated and nonfluorinated phenicols (Munita and Arias, 2016; Poirel *et al.*, 2018)

1.8. BIOFILM FORMATION

Since the first definition of biofilms provided by Costerton *et al.* (1978) 30 years ago, it is well established that the majority of bacteria found in nature exists attached to surfaces within the structured biofilm ecosystem (Costerton *et al.*, 1978; Hall-Stoodley *et al.*, 2004).

Biofilms are complex three-dimensional communities of microorganisms which are adhered to a surface and housed in a protective extracellular polymeric substance. The production of biofilm follows three steps: (I) early development, (II) maturation and (III) dispersion. The early development step starts with the initial adhesion of planktonic bacteria (free-floating) to a surface and followed by the matrix formation. The maturation step occurs while the biofilm reaches its maximum cell density. Lastly, during the dispersion step microcolonies of bacteria are released to colonise new surfaces (for a comprehensive review, see Lebeaux *et al.*, 2014).

In biofilms bacteria communicate (quorum Sensing) to display unified responses that benefit the entire community (Dickschat, 2010). For example, by controlling the population density (access to nutrients) or by the expression of co-ordinated virulence phenotypes. Also, biofilms contribute to the bacteria evasion of the host immune response and induce a decrease of antibiotic susceptibility (tolerance) (Lewis, 2007, 2008). Bacterial biofilms are known for their resistance to antibiotics, disinfectants, and components of the innate and adaptative inflammatory defense system of the body (Høiby *et al.*, 2011). Accordingly, biofilm-growing bacteria cause chronic infections, persisting inflammation, tissue damage, and foreign body infections (Høiby *et al.*, 2011). Thus, it was shown that persistence of staphylococcal infections related to foreign bodies is due to biofilm formation (Sabaté Brescó *et al.*, 2017). Likewise, chronic *Pseudomonas aeruginosa* lung infections in cystic fibrosis patients are caused by biofilm-growing mucoid isolates (Høiby *et al.*, 2010). It has been estimated that in humans around 80% of all bacterial infections involve biofilms (Davies, 2003). Besides, biofilms have also been related with a great part of food-borne outbreaks due to its ability to growth and persist on food processing facilities (Piercey *et al.*, 2016).

Antibiotic tolerance within biofilms is multifactorial and reversible (it disappears when the biofilm is disrupted). It has been proved that the persister cells are not resistant mutants (Lebeaux *et al.*, 2014). Nonetheless, biofilm communities can be exposed to subinhibitory concentrations of antibiotics. Factor which is known to increase the likelihood of selecting antibiotic-resistant mutants. Besides, biofilms provides a favourable environment for horizontal gene transfer (HGT) .Thus, contributing to the exchange of antibiotic resistance genes (Ghigo, 2001; Madsen *et al.*, 2012).

Concerning *E. coli*, which can exist as a harmless commensal in the mammalian digestive tract and as a pathogen causing significant morbidity and mortality worldwide, its ability to form biofilm has been extensively studied from non-pathogenic *E. coli* K12 strains (Beloin *et al.*, 2008). With regard to pathogenic *E. coli*, biofilm involvement in pathogenesis has been well defined in diarrheagenic *E. coli*, notably enteroaggregative *E. coli* (Sherlock *et al.*, 2004; Sheikh *et al.*, 2008; Schiebel *et al.*, 2017) and in adherent-invasive *E. coli* that have been implicated in the origin and perpetuation of Crohn's disease (Martinez-Medina *et al.*, 2009b). Production of biofilm by extraintestinal pathogenic *E. coli* (ExPEC) was mostly assessed in uropathogenic *E. coli* (Watts *et al.*, 2010; Ponnusamy *et al.*, 2012; Agarwal *et al.*, 2013; Tapiainen *et al.*, 2014). All these studies highlighted various biofilm formation phenotypes among uropathogenic *E. coli* including isolates obtained from patients with an indwelling catheter. Assessment of

biofilm formation focusing on the pandemic extended-spectrum β-lactamase (ESBL)-producing and multidrug-resistant ST131 *E. coli* clone (Nicolas-Chanoine *et al.*, 2014) showed highly contrasting results: very low levels (Novais *et al.*, 2012) and moderate levels (Hussain *et al.*, 2014) of biofilm formation. Sarkar *et al.* (2016) studying ST131 isolates expressing or non-expressing type 1 fimbriae showed that biofilm growth depended on type 1 fimbriae expression and assay conditions. In another recent study it is shown that ST131 isolates may express different phenotypes according to their *fimH* allele encoding adhesins of type 1 fimbriae, namely, significantly higher levels of early biofilm production by isolates of ST131 H22 subclone than those of ST131 H30 subclone (Nicolas-Chanoine *et al.*, 2017).

1.9. MOBILE GENETIC ELEMENTS (MGES) AND EVOLUTION OF PATHOGENIC *E. COLI*

Mobile genetic elements (MGEs) are segments of DNA which encode enzymes and other proteins that mediate the movement of other DNA pieces and can be divided into two groups as follows: (I) MGEs with the ability to move DNA segments within the genome of a single bacterium (i.e., intracellular mobility) and (II) MGEs that allow movement of DNA between bacteria (i.e., intercellular mobility) (Partridge *et al.*, 2018b; Peirano and Pitout, 2019).

MGEs contribute to chromosome rearrangements and plasmid integration of genes. Thus, facilitating the exchange of genes between bacteria. All the capture, transposition and recombination of DNA through MGEs have a profound impact on the ongoing evolution of bacterial pathogens (Hacker and Kaper, 2000; Reid *et al.*, 2000; Dobrindt *et al.*, 2002, 2004). Notably, resistance and virulence-associated genes are frequently located on MGEs such as prophages, plasmids and genomic islands (GIs) (Dobrindt, 2005).

1.9.1. INTRACELLULAR MOBILITY

MGEs contributing to intracellular mobility include (I) insertion sequences (IS), (II) transposons (Tn) and (III) integrons (In). These types of MGEs can be found in different locations as well as in multiple copies within the bacterial genome (Orlek *et al.*, 2017).

Insertion sequences (IS) and transposons (Tn) are DNA segments which are able to move almost randomly to new locations through a 'cut and paste' mechanism. On the other hand, integrons (In) are genetic elements that can capture and allow the expression of exogenous genes using a site-specific recombination mechanism (Partridge *et al.*, 2018b).

- Insertion sequences (IS) elements are the simplest types of MGEs and they typically contain a single open reading frame (ORF) coding for its own transposition (*tnp* gene). The major part are flanked by identical short terminal inverted-repeat (IR) sequences in opposite orientation (its binding sites) (Siguier *et al.*, 2014).
- Transposons (Tn) are larger than IS elements and besides the function required for its transposition, they usually carry a resistance gene (or another marker). We refer to a composite transposon when a pair of closely related IS elements are present at each end of a transposon. Afterwards, the merged elements move together leading to the transposition of the composite region (Partridge *et al.*, 2018b).

- Integrons's (In) basic structure includes a site-specific integrase (*int*I), an integron-associated attachment site (*att*I), an integron-carried promoter (Pc) and a gene cassette with a downstream attachment site (*att*C). The integron promotor allows the expression of the cassette-associated genes (usually coding for resistance determinants). Integrons are divided into different classes (I–IV) according to the sequence of their integrase. The ones most frequently detected among clinical isolates belong to class I (Gillings, 2014).

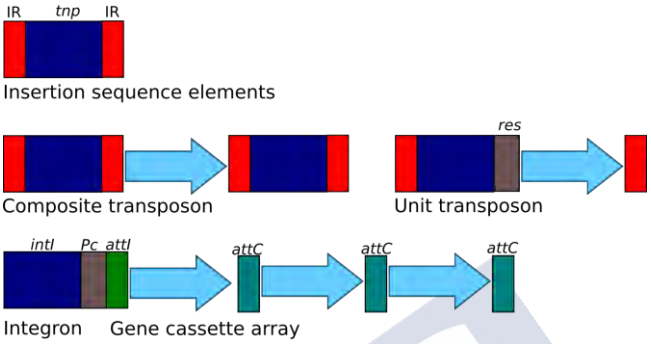


Figure 2. schematic representation of insertion sequence elements, composite transposons and unit transposons. Insertion sequences are flanked by inverted repeats (IRs) that are short identical sequences in opposite orientation. IR inverted repeat, *res* gene encoding for resolvase, *tnp* gene encoding for transposase, *intI* gene encoding for integrase, Pc integron-carried promoter, *att*I integron-associated attachment site, *att*C downstream attachment site.

1.9.2. INTERCELLULAR MOBILITY

The intercellular bacterial mechanisms of genetic exchange include (I) conjugation/mobilization (mediated by plasmids and integrative conjugative elements), (II) transduction (mediated by bacteriophages), and (III) transformation (uptake of extracellular DNA) (for a comprehensive review, see Thomas and Nielsen, 2005).

1.9.2.1. Plasmids

The bacterial genome is composed of a circular chromosome and plasmids. Plasmids are extrachromosomal elements that can self-replicate and harbor genetic information which is not essential for the survival of bacteria. This implies that although several plasmids can coexist in a single bacterium, they can also be naturally lost (Partridge *et al.*, 2018b).

The plasmid structure includes a backbone harbouring a set of housekeeping genes coding for essential plasmid functions that can be classified into three modules (see below), as well as a variable assortment of accessory genome (Orlek *et al.*, 2017; Partridge *et al.*, 2018b).

The accessory genome represents a wide group of adaptive genes (usually resistance and virulence codifying genes) which might benefit bacteria in a particular environment. The accessory genome may also harbor other MGEs such as insertion sequences, transposons and/or integrons (Partridge *et al.*, 2018b).

1.9.2.1.1. Replication and copy number control module

Plasmids encode the initiation of their own replication. Three strategies have been discerned in circular plasmids: (I) Theta-mode (II) Rolling circle and (III) strand displacement. However, most plasmids require to exploit part of the host replication

machinery (helicases, polymerase, etc.) in order to replicate (del Solar *et al.*, 1998; Partridge *et al.*, 2018b).

Plasmid replication usually initiates in the *ori*V (origin of vegetative replication) region that is recognised by replication proteins (*rep* genes). Furthermore, multireplicon plasmids (plasmids carrying multiple replication regions) are quite common, suggesting that the fusion between plasmids frequently occurs (cointegrated plasmids). In some cases, those extra-replicons decrease bacteria fitness and must be eliminated, although they can also be advantageous by increasing the plasmid host range (more flexible replicative machinery) (Orlek *et al.*, 2017; Partridge *et al.*, 2018b).

Once replicated, plasmids must also regulate their number of copies per cell to ensure that they do not excessively burden the host (too high number of copies) or become lost during cell division (too low number of copies). For this purpose, two basic strategies have been discerned: (I) the antisense (countertranscript) RNA systems and (II) the blockage of *ori* sites. As a general rule, small plasmids usually show a high number of copies per cell (>15 plasmids/cell) whilst large plasmids show a low number of copies due to fitness cost (Sengupta and Austin, 2011; Orlek *et al.*, 2017; Partridge *et al.*, 2018b).

It is impossible for plasmids with the same replication mechanism to coexist in the same cell, a phenomenon termed as incompatibility. This is usually due to cross-talk between the replication initiation systems of the two plasmids that "confuses" the copy number control (the two different plasmids are perceived as the same), leading to segregational instability (Couturier *et al.*, 1988; Novick, 1989). This incompatibility trait has provided the basis for the initial classification of plasmids (see plasmid classification).

1.9.2.1.2. Vertical inheritance module

In order to guarantee their maintenance, plasmids must ensure to be distributed between daughter cells when the cell division takes place (segregational stability). In this regard, the vertical module includes multimer resolution (*res* genes), partitioning (*par* genes) and post-segregational killing toxin-antitoxin (TA) systems (Sengupta and Austin, 2011; Orlek *et al.*, 2017; Partridge *et al.*, 2018b).

Low-copy number plasmids require a sophisticated segregation system to ensure inheritance within a bacterial lineage, for example the ParRMC system in R plasmids and SopABC system in F plasmids. In contrast, high-copy number plasmids can be inherited through a random segregation process (Partridge *et al.*, 2018b).

Some plasmids can only be efficiently maintained in closely related bacterial taxa and they tend to be restricted to certain species or clones within species, (narrow-host-range plasmids) whilst others can replicate and maintain themselves within diverse genera of bacteria (broad-host-range plasmids).

1.9.2.1.3. Horizontal transfer module

Plasmids can be transferred between bacteria in a process termed as conjugation or mobilisation. The conjugative machinery sets up a specialised transfer pore and induces the transport of DNA from the donor to the recipient cell.

This module varies within conjugative (self-transmissible) and non-conjugative plasmids (mobilizables). Self-transmissible plasmids encode the complete protein machinery for conjugal transfer, while mobilizable plasmids contain a minimal gene set and can only be transmitted in presence of another "helper plasmid". Thus, this module is related with the plasmid size: conjugative plasmids tend to be large, (30 to 500 kb) whilst mobilizable plasmids (3 to 35 kb) and non-mobilizable infective plasmids (1 to 5 kb) are commonly small (Ramsay and Firth, 2017; Partridge *et al.*, 2018b).

The transfer region (*tra* genes) of conjugative plasmids includes DNA transfer replication (DTR) proteins (Mob; classified into 9 families) and mating pair formation proteins (Mpf; classified into 8 types)(Guglielmini *et al.*, 2014). The Mob protein is a crucial relaxosome component for the initiation of DNA transfer in both conjugative and mobilizable plasmids. Additionally, this module encodes entry (surface) exclusion proteins which prevent the host from acting as recipient for the same or related plasmids (Garcillán-Barcia and de la Cruz, 2008).

During conjugation, the Mob protein (relaxase) binds the origin of transfer (*ori*T) site and nicks DNA in a single-stranded substrate suitable for transfer (Hayes, 2001). Mob proteins belong to nine families included in four protein superfamilies (HUH endonucleases, tyrosine recombinases, Rep_trans RCR initiators, and PD-(D/E) XK restriction endonucleases). Six major families of relaxases can be found within Enterobacteria: MOB_F, MOB_H, MOB_Q, MOB_C, MOB_P, and MOB_V (Francia *et al.*, 2004; Garcillán-Barcia *et al.*, 2009, 2011; Alvarado *et al.*, 2012) which belong to the His-hydrophobe-His (HUH) endonuclease superfamily (Chandler *et al.*, 2013). These proteins provide the basis for the classification of self-transmissible and mobilizable plasmids (see plasmid classification).

1.9.2.2. Prophages

Prophages are a fragment of DNA incorporated into the bacterial genome through bacteriophages, which are viruses-like structures which infect and replicate into bacterial cells. They are extremely diverse in size, morphology and genomic organisation. The spread of genes (frequently virulence-associated) by phages has been proved to be very frequent within bacteria (Novick *et al.*, 2010; Kasman and Porter, 2020).

1.9.2.3. Genomic islands (GIs)

A genomic island (GI) is a distinct region of the bacterial chromosome which varies in size (they may be composed of several hundred genes) and composition. Generally, GIs encompass elements with mobility functions (referred to as integrative conjugative elements), resistance determinants (referred to as resistance islands) or virulence factors (referred to as pathogenicity islands) and in many cases are flanked by direct repeats.

Although there is still limited knowledge about the origin of GIs, it has been speculated that they derive from the integration of plasmids or phages into the chromosome (Dobrindt *et al.*, 2004).

1.9.2.3.1. Integrative Conjugative Elements (ICEs)

An integrative Conjugative Element (ICE) is a self-transmissible fragment of DNA that is integrated into the bacterial chromosome and replicated as part of it, but it can also be excised from the chromosome and transferred horizontally in a plasmid-like way (Burrus *et al.*, 2002; Carraro and Burrus, 2015). Structurally, ICE typically consists of a backbone (containing phage-like integration/excision functions, plasmid-like conjugation/ maintenance components and a regulation module) as well as accessory genes that confer beneficial traits for the host (for a comprehensive review please refer to Delavat *et al.*, 2017).

1.9.2.3.2. Pathogenicity islands (PAIs)

Pathogenicity islands (PAIs) are fragments of DNA encoding various virulence-associated genes (Dobrindt *et al.*, 2004). Structurally, PAIs are generally large in size (>10 kb) and they have a G + C content which differs from the host bacterial core genome. These elements are frequently flanked by repeated sequences (attachment sites *att*L and *att*R) and carry many fragments of other mobile and accessory genetic elements, such as

bacteriophages, plasmids and insertion sequence. In contrast to the other MGEs, most PAIs lack an origin of replication and can neither replicate nor mobilise by themselves.

It has been suggested that PAI mobility is manly mediated by PAI-encoded integrases through a site-specific recombination process and that horizontal transfer is mediated by other MGEs (Middendorf *et al.*, 2004; Hochhut *et al.*, 2006). Genomic research has shown that pathogenicity islands (PAIs) have played a major role in the transformation of avirulent into virulent bacteria. As a typical example through the adcquisition of the so-called "high pathogenicity island" (HPI), initially described in *yersiniae* (Schubert *et al.*, 1998; Desvaux *et al.*, 2020).

PAIs are usually inserted close to a transfer-messenger RNA (tmRNA) locus and within *E. coli* 15 tDNA genes have been described as hotspots of integration, from which *asnT*, *aspV*, *leuX*, *metV*, *pheV* and *thrW* are the most frequently targed genes (PAI markers) (Desvaux *et al.*, 2020).

Almost identical PAIs have been reported within *E. coli* pathotypes. However, great variability exists with regard to their composition, structural organisation and chromosomal localisation (Hacker and Kaper, 2000; Hacker and Carniel, 2001; Dobrindt *et al.*, 2002, 2004; Welch *et al.*, 2002; Dobrindt, 2005). Samei A., *et al.* 2016 and Naderi G., *et al.* 2016 analysed the prevalence of PAI markers among pathogenic and commensal *E. coli* isolates and concluded that PAI markers were more prevalent among UPEC and DEC strains than commensal strains (98.7% and 94.8%, respectively vs. 88% in commensals). They also found that the average of PAIs per isolate was higher in UPEC and DEC isolates than in commensals (Naderi *et al.*, 2016; Samei *et al.*, 2016). Thus, PAIs can be regarded as a feature that distinguishes pathogens from commensal in *E. coli*. Besides, most VFs of ExPEC are encoded by PAIs (for a comprehensive review see Desvaux *et al.*, 2020).

1.10. PLASMID CLASSIFICATION AND TYPING

Two central schemes for plasmid typing are widely used: (1) the replicon typing scheme which relies on the classical plasmid classification based on the 28 defined Incompatibility groups (Inc) defined in *Enterobacteriaceae* (Carattoli *et al.*, 2005; Rozwandowicz *et al.*, 2018) and (2) the relaxase (MOB) typing scheme based on differences in conjugation systems which relies on Mob proteins (Francia *et al.*, 2004; Alvarado *et al.*, 2012).

In 1988, Couturier and colleagues cloned replicons representing *Enterobacteriaceae* Incompatibility groups, after which plasmids started to be classified according to Southern blot hybridizations (Couturier *et al.*, 1988). However, this method is laborious and can be hardly performed on a large number of strains. Nowadays, the PCR-based replicon typing (PBRT) method allows a rapid determination of various genetic loci (including rep genes and replication regulatory sequences) for plasmid replicon typing. The PCR-based replicon subtyping (pMLST) method allows the assignment of a pMLST subtype for six common replicon types in *Enterobacteriaceae* (IncF, HI1, HI2, I1, N, A/C) by the determination of 2 to 6 core loci (depending on the broader replicon type).

The PCR-based degenerate primer MOB typing (DPMT) method allows to classify γ-Proteobacterial plasmids into five of the main relaxase (MOB) types, based on the determination of the N-terminal relaxase sequences (Alvarado *et al.*, 2012).

Replicon and relaxase typing are complementary and when performed together they can improve plasmid analysis. Additionally, in both silico replicon and relaxase (MOB) typing can also be performed using friendly web-tools (Table 4).

The correspondence between typing schemes is summarised in Table 4. The replicon typing scheme uses traditional nomenclature, allows pMLST subtyping and classifies approximatively 75% of *Enterobacteriaceae* plasmids and 50% of plasmids of Gram-positive bacteria. Nevertheless, as a drawback, some plasmids may contain more than one replication protein. Instead, there is only one relaxase per plasmid. Relaxases classifies approximatively 50% of *Enterobacteriaceae* plasmids and a less percentage of Gram-positive bacteria. The main drawback is that it only allows the classification of transmissible (mobilizable) plasmids. In comparison, MOB typing has higher sensitivity and lower resolution than the replicon typing method (Orlek *et al.*, 2017).

The phylogenetic establishment of plasmids remains challenging due to the plastic nature of plasmid genomes (Boyd *et al.*, 1996). Also, current typing schemes are not able to classify all plasmids (Orlek *et al.*, 2017). Although it is still evolving, WGS has brought forth many new opportunities in this regard. Some common *in silico* whole genome sequence (WGS) based tools that can be used for plasmid analysis are summarised in Table 5.

Table 4. Summary of plasmid features and correspondence between MOB and Replicon typing.

Inc. type (Incopatibility group)	Replicon (replication machinery)	MOB type (relaxase proteins)	Size (kb)	Copy number	Transferability	Host range
IncQ (IncQ1, IncQ2 and p11745 or pIGWZ12-like)	Q	MOB_{Q11} (IncQ1); MOB_{Q12} (p11745-like); MOB_{P14} (IncQ2) and MOB_{Qu} (pIGWZ12-like)	8-14	medium (4-12 copies/cell)	mobilizable	wide host range
Col-plasmids (Type I-GA and Type II-GB)	Col	$MOB_{P5/HEN}$	2-10 (Type I, ColE1-like); ~40 (Type II)	high (Type I); low (Type II); (1-20 copies/cell)	mobilizable	unknown
IncU/G (IncP-6)	U/G	MOB_{P3} and MOB_{P4} ($MOB_{P14\text{-}IncP\text{-}6}$)	29-60	low	Conjugative / mobilizable	wide host range
IncL/M	L/M	MOB_{P13}	50-80	low	Conjugative	broad
IncN (IncN1 and IncN2)	N	MOB_{F11}	30-70	low	Conjugative	broad
IncX (IncX1 to IncX8)	X	MOB_{P3} and MOB_{P4}	30-50	low (~1-3 copies/cell)	Conjugative	narrow
IncW	W	MOB_{F11}	up to 40	low	Conjugative	broad
IncR	R	unknown	40-160	low	mobilizable	broad (not clear)
IncY (Group of prophages)	Y	unknown	90-100	low	conjugative	broad (Enterobacter-iaceae)
IncA/C (IncA/C1-IncA and IncA/C2-IncC, IncP-3)	A/C	MOB_H (IncA-MOB_{H121} and IncC-MOB_{F12})	18-230	low	Conjugative	narrow
IncF[1]	FII/FIA/FIB	MOB_{F12}	45-200	low	Conjugative	broad (Enterobacter-iaceae)
IncI-complex (IncI/K/B/O/Z)	I1/Iγ/B/O/K/Z	MOB_{P12}	50-250	low	Conjugative	narrow
IncP (currently many subgroups are described)	P	MOB_{P11} (IncP1-complex)	70-280	low	Conjugative	broad
IncH (IncHI1 to HI5 and HI-like)	HI1A/HI1B/HI2	MOB_{H11} (IncHI1 and IncHI2)	75-400	low	Conjugative	wide host range
IncT	T	MOB_H	~217	low	Conjugative	narrow (not clear)

[1]IncF plasmids are usually multireplicons carriers with a highly mosaic structure. The replicon typing scheme (RST) is based on diversity in the replicon regions and applies a FAB numbers formula for IncF subtyping which is obtained by combinations of three replicons (FII:FIA:FIB, e.g., F1:A2:B2) (Carattoli *et al.*, 2005; Villa *et al.*, 2010; Alvarado *et al.*, 2012; Yano *et al.*, 2013; Rozwandowicz *et al.*, 2018)

Table 5. Summary of common *in silico* whole genome sequence (WGS) based tools that can be used for plasmid analysis.

Tool (s)	Goals and comments
SRST2	Mapping-based tool to detect loci of interest from reads by mapping to a reference database.
KmerResistance	Detect resistance genes by examining co-occurrence of k-mers between the query and a reference database.
Artemis comparison tool (ACT), BLAST Ring Image Generator (BRIG) and Easyfig.	Comparative genomics tools that allow to visualize homologies and gene content similarities between contigs and references using BLAST[1].
PlasmidFinder and plasmid MLST (pMLST)	Detect replicon type/subtype from contigs based on their similarity with reference database.
ResFinder	Detect resistance genes from contigs based on their similarity with a reference database using BLAST[1].
Comprehensive antibiotic research database (CARD)	Detect resistance genes from contigs based on their similarity with CARD database using BLAST[1].
Antibiotic resistance gene-annotation (ARG-ANNOT)	Detect resistance genes from contigs based on their similarity with ARG-ANNOT database using BLAST[1].
ISMapper	Mapping-based tool to localize insertion sequences. Localize specific genes from a contig assembly.
cBar	Distinguish plasmid from chromosomal sequences based on pentamer frequencies.
PLACNET	Resolve plasmid structures from ambiguous assembly graphs according to the homology of contigs to reference sequences. The identification of plasmids is base on multiple traits (coverage, size, replication initiator proteins and relaxases). Semi-automatic tool (biological knowledge required).
Recycler	Resolve plasmids or circular phages structures from ambiguous assembly graphs by assuming that different genetic units have distinct copy numbers. Fully automatic tool.
plasmidSPAdes	Resolve plasmid structures from ambiguous assembly graphs based on the estimated chromosomal coverage. Fully automatic tool.
Bandage	Assembly graph visualization and annotation tool that resolves plasmid structures from ambiguous assembly graphs. Facilitates interaction with de Bruijn graphs made by *de novo* assemblers such as Velvet (Zerbino and Birney, 2008), SPAdes (Bankevich *et al.*, 2012) and Trinity (Grabherr *et al.*, 2011).
Mlplasmids	Predict plasmids and chromosome sequences regardless of contig coverage and can be used as a basis for plasmid classification by other tools.
MOB-suite	Automated annotation of MOB Relaxases.
MOBscan	Automated annotation of MOB Relaxases.
PlasFlow	Identification of bacterial plasmid sequences based on genomic signatures.

[1]Basic local alignment serch tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Carver *et al.*, 2005; Zhou and Xu, 2010; Alikhan *et al.*, 2011; Sullivan *et al.*, 2011a; Zankari *et al.*, 2012b; McArthur *et al.*, 2013; Lanza *et al.*, 2014; Carattoli *et al.*, 2014; Gupta *et al.*, 2014; Inouye *et al.*, 2014; Wick *et al.*, 2015; Hawkey *et al.*, 2015; Antipov *et al.*, 2016; Rozov *et al.*, 2016; Clausen *et al.*, 2016; Arredondo-Alonso *et al.*, 2017; Krawczyk *et al.*, 2018; Robertson and Nash, 2018; Garcillán-Barcia *et al.*, 2020)..

1.11. CLONES AND EVOLUTIONARY SUCCESS OF *E. COLI*

The term clone refers to bacteria isolates which share similar traits, indicating that they belong to the same lineage and have a common ancestor, thus this knowledge allows to link isolates from different geographical areas and hosts.

Notable differences in the prevalence of *E. coli* linages are observed among different hosts, regarding to the host specie (animals and human populations) and the host environment (Tenaillon *et al.*, 2010). Besides, intra-host diversity has been reported in several studies and attributed to migration of strains. In this regard, strain sharing occurs more frequently amongst household members, including pets (Johnson *et al.*, 2008c; Pomba *et al.*, 2016).

Phylogenetical groups prevalence differs between animals and humans. In a recent study through a compiled analysis of 1,117 human isolates and 1,154 animal isolates, it was observed that the B1 phylogenetical group occupated the third place of prevalence in humans, while in animals occupated the first place. After which, in both humans and animals the prevalence of phylogenetic groups varies from higher to lower in the following order: A, B2 and D (Tenaillon *et al.*, 2010).

Furthermore, it has become evident that certain ExPEC lineages or clonal groups are responsible for a large fraction of human and animal extraintestinal infections. Concerningly, these lineages are becoming increasingly multi-drug resistant (Smith *et al.*, 2007; Manges and Johnson, 2012; Denamur *et al.*, 2020). This association has led to the definition of "high-risk clones".

Baquero *et al.* (2013) defined as "multi-drug-resistant (MDR) high-risk clone" those having the following characteristics: (I) international distribution, (II) possess various antimicrobial resistance determinants, (III) successful colonisation and persistance in hosts for long time intervals (>6 months), (IV) effective transmission among hosts, (V) extended pathogenicity and fitness and (VI) cause severe and/or recurrent infections. Examples of MDR high-risk clones among *E. coli* (termed under sequence types according to the Warwick University scheme) are ST10 clonal complex, ST38, ST69, ST88, ST131, ST155, ST167, ST315, ST393, ST405, ST410 and ST648 (Manges and Johnson, 2012; Riley, 2014; Wang *et al.*, 2016; Bubpamala *et al.*, 2018; van Hoek *et al.*, 2018; Denamur *et al.*, 2020). Currently, some other clones are emerging among ESBL-producing *E. coli* like ST1193 (Valenza *et al.*, 2019). It has been suggested that the association with certain epidemic plasmids harboring numerous ARGs is an important contributor to the high-risk clones success (Mathers *et al.*, 2015). Nonetheless, antibiotic susceptible lineages have also emerged globally among ExPEC infections, e.g. ST12, ST73, ST95, ST127 and ST141 (Gibreel *et al.*, 2012; Denamur *et al.*, 2020). Although the reason why these clones are prevailing over the others remains unclear, antimicrobial resistance and virulence factors are suspected contributors (Dobrindt, 2005; Stoesser *et al.*, 2016b).

Therefore, the fact of high-risk clones contributing to the spread of global multidrug resistance among Enterobacteria is undeniable. High-risk clones increase the prevalence of antibiotic resistance vertically due to co-evolution with certain ARG-carrier components (e.g. integrons, transposons and plasmids) (Mathers *et al.*, 2015). Additionally, eminent clones favor the transmission of antibiotic resistance horizontally, due to their ability to survive for long periods of time and reproduce efficiently among bacterial populations.

1.11.1. THE EXPEC CLONES

The majority of ExPEC isolates belong to B2 phylogenetic group. However, there is enormous diversity among *E. coli* isolates causing human and animal extraintestinal infections, but, epidemiological studies indicate that certain O:H serotypes, sequence types (STs) and clonotypes are more predominant and especially successful (Riley, 2014; Kallonen *et al.*, 2017; LeCuyer *et al.*, 2018; Zogg *et al.*, 2018; Manges *et al.*, 2019; Flament-Simon *et al.*, 2020b; Valat *et al.*, 2020).

Twenty major STs (in order of highest to lowest prevalence: ST131, ST69, ST10, ST405, ST38, ST95, ST648, ST73, ST410, ST393, ST354, ST12, ST127, ST167, ST58, ST88, ST617, ST23, ST117 and ST1193) accounted for 85% of the *E. coli* isolates causing human extraintestinal infections from 217 meta-analyzed studies (1995 and 2018), systematically reviewed by Manges *et al.* (Manges *et al.*, 2019). However, most of the studies have been carried out on MDR and ESBL-producing isolates, but very few have been focused on any type of *E. coli* causing extraintestinal infections and, furthermore, on their clonal structure. Therefore, there is probably an overestimation of some STs and an underestimation of others.

Seven clonal complexes (CC) or sequence type complexes (STc) harbour the most reported ExPEC clones isolated from humans in pathogenic conditions whitch are STc12, STc14, STc69 (also termed as clonal group A), STc73, STc95, STc127 and STc131 and (Denamur *et al.*, 2020). Table 6 summarised the principal features of those eminent ExPEC STc.

Table 6. Principal features of eminent ExPEC sequence type complexes (STc).

Sequence type complex (STc)	Phylogroup	Serotypes	Resistance profile
STc12	B2	Mainly associated with O4:H1/H5	Emerging ESBL
STc14 (including ST550, ST14 and ST1193)	B2	Mainly associated with O75:H5	ESBL producers
STc73	B2	Mainly associated with O6:H1	Mostly susceptible
STc69	D	Mainly associated with O11, O15, O17, O73, O77, O117 and H18	ESBL producers
STc95	B2	Mainly associated with O1, O2, O18, O45 and H4/H7	Mostly susceptible
STc127	B2	Mainly associated with O6:H31	Mostly susceptible
STc131	B2	Mainly associated with O16:H5 and O25:H4	ESBL producers

Sources: (Adams-Sapper *et al.*, 2013; Danzeisen *et al.*, 2013; Basmaci *et al.*, 2015; Gordon *et al.*, 2017a; Bourne *et al.*, 2019; La Combe *et al.*, 2019; Bogema *et al.*, 2020; Denamur *et al.*, 2020; Kidsley *et al.*, 2020a).

Three recent studies showed the dominance of some STs in dogs in Australia, the United States, and France, such as ST372, assessed to be specifically associated with dogs, and ST12, ST73, ST127 and ST141, assessed to be specifically associated with humans (LeCuyer *et al.*, 2018; Kidsley *et al.*, 2020b; Valat *et al.*, 2020). On the other hand, within-household sharing of ExPEC ST73 and ST95 strains, those with same serotypes and VF-encoding genes have been documented in the United States among humans and dogs (Johnson *et al.*, 2008a). Furthermore, in Australia and the United States, human and canine *E. coli* ST127, ST131 and ST1193 that exhibited identical virulence genotypes and highly similar PFGE profiles have been identified (Johnson *et al.*, 2008b; Platell *et al.*, 2011, 2012). These findings suggest that some *E. coli* infections may sometimes be a zoonosis in either direction (human to pet or pet to human).

It seems that porcine ExPEC diseases are increasing frequency in China (Tan *et al.*, 2012; Zhu *et al.*, 2017; Ma *et al.*, 2020). In contrast to most humans and companion animal ExPEC clones, swines ExPEC isolates frequently belong to phylogroups A and B1 (Ding *et al.*, 2012). In a recent study, nine main clonal complexes were described among the B2 and D ExPEC isolates from swine: CC12, CC648, CC354, CC141, CC405, CC457, CC1011, CC117 and CC95, witch based on their virulence profiles were classified into four clusters (named I to IV) (Ma *et al.*, 2020). Cluster IV harboured strains belonging from B2 phylogroup and with the highest average of virulence-associated genes, and included three clonal complexes: CC12 (including ST12 and ST961 strains), CC95 (including ST95 strains) and CC141 (including ST141 strains) (Zhu *et al.*, 2017). In poultry, STc88 belonging to phylogroups C (including ST410 strains) and STc95 are predominant among APEC strains (Moulin-Schouleur *et al.*, 2007).

1.11.2. THE HIGH-RISK CLONE SEQUENCE TYPE 131 (ST131) OF *E. COLI*

E. coli ST131 is one of the most prevalent ExPEC clones. It is also responsible for the worldwide dissemination of MDR isolates within the *E. coli* group, mostly due to the acquisition of *bla*$_{\text{CTX-M-15}}$ carriers IncF epidemic plasmids (Nicolas-Chanoine *et al.*, 2007b, 2014).

Phylogenetical studies have revealed that *E. coli* ST131 belongs to four clades (A, B, C and I) and many different subclades, three within clade C (C1, C2, and C3) and six within clade B (from B1 to B5, plus intermediate B0). Recently, whole genome sequencing studies have shown that clade C has evolved from clade B. Clades are designated according to their *fimH* allele (type 1 fimbriae adhesin gene), while subclades are associated with antibiotic resistance patterns. In summary, clade B (*fim*H22) is a fluoroquinolone (FQ) and cephalosporin susceptible ancestor which evolved in clade C (*fim*H30) that was resistant to FQ and the predecessor of two sister subclades C1 (often carrier of the *bla*$_{\text{CTX–M–14}}$ enzyme) and C2 (often carrier of the *bla*$_{\text{CTX–M–15}}$ enzyme). Subsequently, two additional ST131 sublineages referred to C1-M27 (positive for the *bla*$_{\text{CTX–M–27}}$ enzyme) and C1-nM27 have been described. Clade I constitutes a rare variant that had also emerged from clade B but presented the *fim*H54 allele.

Even though ST131 clone is prevalent in human samples, it remains less frequent from animals and it is especially rare from porcine source (Reid *et al.*, 2019). However, it has already been identified in many animal species including Antarctic pinnipeds (Mora *et al.*, 2018). Indeed, in ST131 strains have been reported from companion animals as well as food products from animal origin like poultry meat (Ewers *et al.*, 2010; Liu *et al.*, 2018a; Zogg *et al.*, 2018). The majority of animal isolates belong to clade B, whereas those which cause extraintestinal infections in humans predominate those of clade C and particularly to subclades C1-nM27, C1-M27 and C2 (H30Rx). This could be one reason for the underrepresentation of ST131 clade B isolates in the literature and sequence databases (Reid *et al.*, 2019).

Some studies suggested a negative association between resistance genes and virulence-associated genes. This is not the case with ST131 clone. In a study performed by Colpan and colleagues, it was shown that fluoroquinolones resistant linages of ST131 that acquired *bla*$_{\text{CTX-M-15}}$ had also tended to further accumulate additional virulence-associated factors (Colpan *et al.*, 2013). In this regard, it has been postulated that the *gyr*A-*par*C mutations (i.e., *gyr*A1AB and *par*C1aAB) selected by fluoroquinolones could have reduced the fitness costs associated with the acquisitions of the IncF epidemic plasmids acting as compensatory mutations. It has also been noted that the inherent combination of different virulence factors, biofilm production, and high metabolic

potential of ST131 clone has contributed to its fitness. Paradoxically, a recent study had proved that the most worldwide disseminated ST131 clade (clade C) in human infections displays a more attenuated virulence than its ancestor clade (clade B). Two explanations were proposed: (I) that the acquired FQ resistance came at a direct cost, but was compensated by allowing calde C to spread and (II) that a lower virulence may confer an evolutionary advantage by balancing transmission and host mortality. Further studies are needed to clarify these matters (Duprilot *et al.*, 2019).

The antimicrobial resistance of human ExPEC and UPEC isolates has increased dramatically due to the emergence of the pandemic clone ST131 and more especially to subclade C2 (also known as subclone *H*30Rx) (Nicolas-Chanoine *et al.*, 2007a, 2014; Price *et al.*, 2013; de Toro *et al.*, 2017; Matsumura *et al.*, 2017b; Jamborova *et al.*, 2018; Flament-Simon *et al.*, 2020a). This subclone has also been occasionally isolated from dogs in several countries (Kawamura *et al.*, 2017; Belas *et al.*, 2018; Maeyama *et al.*, 2018; Bortolami *et al.*, 2019; Melo *et al.*, 2019). The emergence of multidrug resistance (MDR) among *E. coli* causing infections in dogs is of great concern and increases the risk of treatment failure (Johnson *et al.*, 2009; Harada *et al.*, 2012; Wagner *et al.*, 2014; Bogaerts *et al.*, 2015; Falgenhauer *et al.*, 2016; Liu *et al.*, 2016b, 2016c; Zhang *et al.*, 2018; Marques *et al.*, 2018; Chen *et al.*, 2019b; Pepin-Puget *et al.*, 2020). Additionally, exposure to dogs and/or dog faeces has been identified as a risk factor for the development of drug-resistant *E. coli* UTI in women (Ukah *et al.*, 2018).

1.12. ADVANCES IN MICROBIOLOGY RESEARCH: NEXT GENERATION SEQUENCING (NGS) TECHNOLOGIES

In 1977, less than 50 years ago, the first DNA sequencing methods were developed (Maxam and Gilbert chemical chain termination method and Sanger sequencing) allowing to decipher complete genome sequences for the first time. Second-generation sequencing (SGS) or short-read sequencing technologies came after, including the methodologies developed by Roche Diagnostics, Illumina and Life Technologies companies. These were based on the polymerase chain reaction (PCR) and allowed to sequence the whole genome of bacteria. Afterwards, third-generation sequencing (TGS) or long-read sequencing appeared (no-PCR based technologies). In 2011, Pacific Biosciences (PacBio) developed the "single-molecule real-time" (SMRT) sequencing methodology A few years after, in 2014, Oxford Nanopore Technologies launched the "nanopore sequencing" methodology.

The so-called next-generation sequencing (NGS) includes both SGS and TGS. This technology is now revolutionising the field of genomic research as it enables to explore genomes with an unprecedented resolution. For instance, contributing to distinguish pathogenic *E. coli* or tracing back outbreaks. Besides, WGS can be implemented for bacterial-resistance surveillance programs worldwide. However, NGS has still some challenges to overcome to be fully implemented worldwide (reviewed in Quainoo *et al.*, 2017):

- To be cost-effective.
- To be standardised to simplify interpretation and comparison of the data worldwide.
- To overcome sequencing limitations related to read length, read depth, and error rate.
- To improve bioinformatics analysis algorithms and develop friendly analysis tools for clinical microbiologists with few knowledges on bioinformatics.

In brief, there are two WGS analysis methods: (I) mapping the sequenced genome with another previously reconstructed genome (or loci) using alignment tools, and/or (II) without assuming a reference genome (*de novo* assembly). The use of coregenome sequences and references has proved to be adequate to infer phylogenies. Nonetheless, by definition coregenome phylogenies ignore the accessory genome (including plasmids, virulence genes, bacteriophages, and resistance genes). In fact, *de novo* assembly allows identifying structural variations and genomic rearrangements.

1.12.1. THE READ LENGTHS LIMITATIONS

Long repetitive elements in bacterial genome represent a particular obstacle (Arredondo-Alonso *et al.*, 2017; Orlek *et al.*, 2017). With SGS technology, the sequence length (35 to 700 bp) is shorter than with traditional Sanger sequencing platforms. Thus, genome assembly is complicated, and most assembly programs collapse and fail to assemble reads covering long repetitive elements. As a consequence, regions between repeats are usually found in separate contigs. TGS generates long-read sequences (10 kb to over 100 kb). However, this technology remains considerably more expensive and has higher error rates at its base level (Lu *et al.*, 2016a; Orlek *et al.*, 2017; Boolchandani *et al.*, 2019).

The read lengths limitations of SGS does not interfere with outbreak analysis; as phylogenetic trees are not reconstructed based on repetitive elements. In contrast, resistance regions in plasmids or genomic islands contain multiple copies of the same mobile element which constitute long repeats.

In summary, genomic tools allow a compressive characterization of FBPs and the identification of clonal groups of bacteria that represent public health hazards (Kovac *et al.*, 2017). However, epidemiological surveillance of epidemic plasmids related to antimicrobial resistances or/and virulence genes spreading in the bacterial population is still complicated due to the intrinsic plasticity of plasmids (Orlek *et al.*, 2017). Future developments in real-time and long-read sequencing are predicted to overcome most of the technology limitations which we are facing today (Eid *et al.*, 2009; Quainoo *et al.*, 2017).

1.12.2. GAPS IN THIS GENOME COMPENDIUM

The major clonal complexes and outbreaks-associated strains are largely represented in public genome databases. In contrast, there is a limited representation of commensal strains and other divergent lineages (Moriel *et al.*, 2016). Besides, many strains lack information regarding their origin and clinical association as well as there is a surprisingly great proportion of unknown ORFs (Lo *et al.*, 2017).

1.12.3. PREDICTION OF PHENOTYPES

The expression of certain genes is often context-dependent (Boolchandani *et al.*, 2019). Consequently, phenotypic traits predicted from resistance and virulence-associated genes remain uncertain.

Several studies showed up to 95% concordance between resistance phenotypic testing and genomic resistant profiles (Stoesser *et al.*, 2013; Zankari *et al.*, 2013; Tyson *et al.*, 2015; McDermott *et al.*, 2016). Nevertheless, certain complex resistance mechanisms can remain undetected, such as carbapenem resistance arising from the combination of extended-spectrum betalactamases and efflux pumps or porin impermeability (Baroud *et al.*, 2013), as well as novel antibiotic resistance genes or mutations (Ledeboer and Hodinka, 2011; Fleece *et al.*, 2018).

To date, WGS cannot entirely replace phenotypic experiments, but they can greatly contribute to gain knowledge of bacterial pathogenesis.

1.13. THERAPEUTIC APPROACH AND DEVELOPMENT OF ALTERNATIVE TREATMENTS

Against bacterial infections, two approaches have been addressed: (I) direct strategies that target the bacteria viability (e.g. antibiotics) and (II) indirect strategies that enhance the host immune responses (e.g. vaccines).

The research of anti-ExPEC targets can be particularly challenging for various reasons. To begin with, due to the heterogeneity of ExPEC strains. Secondly, as the expression of several genes is involved in ExPEC pathogenesis having overlapping activities. Thus, inhibiting one of them would probably be insufficient to prevent the infection (Cash, 2014). Lastly, because ExPEC strains are common inhabitants of the gastrointestinal tract in the host and therefore it is important to consider potential side effects on the balance of the intestinal microbiome (Brumbaugh and Mobley, 2012; O'Brien *et al.*, 2016).

Tigecycline (a semisynthetic tetracycline) is among the newest availables antibiotics against Gram-negative pathogens, as very few new antibiotics have been developed during the last decades. Fortunately, other alternatives are being investigated: (I) the use of nutritional supplementation therapies or probiotics and fecal microbiota transplant (FMT) to restore the microbiome (Saha *et al.*, 2019; Saïdani *et al.*, 2019), (II) the development of pilicides, curlicides and mannosides to prevent bacterial invasion, and biofilm eradication agents (BEAs) to inhibit biofilm formation (Han *et al.*, 2010, 2012; Terlizzi *et al.*, 2017) and (III) the development of anti-plasmid strategies to remove harmful plasmids, and phage therapy to kill bacteria (Schmidt, 2019; Ferry *et al.*, 2020). Regarding vaccines, some new potential candidates have been identified using the reverse-vaccinology approach (Nesta *et al.*, 2012; Moriel *et al.*, 2016; Terlizzi *et al.*, 2017).

1.14. POLICIES AGAINST THE EMERGENCE OF ANTIBIOTIC RESISTANCE

Currently many pathogens have started to become resistant to the second- and third-line drugs, including last-resort antibiotics (Gupta *et al.*, 2011; Johnson and Woodford, 2013; Ventola, 2015; Liu *et al.*, 2016d). Therefore, antibiotics such as colistin and tigecycline are in some cases the last-resort against infections caused by MDR, XDR and PDR Gram-negative pathogens (Holmes *et al.*, 2016; Laxminarayan *et al.*, 2016; Yang *et al.*, 2016; Yao *et al.*, 2016).

In particular, *E. coli* has become one of the most common MDR organisms, with ESBL-producing *E. coli* (ESBLEC) being dramatically extended world-wide (Nicolas-Chanoine *et al.*, 2007b; de Toro *et al.*, 2017; Mamani *et al.*, 2019). The hypothesis regarding the global circulation of antimicrobial-resistant bacteria and ARGs between animals, humans and the environment is now strongly supported by several studies which have found ESBLEC strains carried by hosts from natural ecosystems, including vultures (*Neophron percnopterus var. majorensis*), free-tailed bats (*Tadarida teniotis*) and many other wild animals (Radhouani *et al.*, 2014; Cristóvýo *et al.*, 2017; Garcẽs *et al.*, 2019; Carvalho *et al.*, 2020).

The description of MCR-producing *E. coli* strains appears to be concerning, as the co-occurrence of MCR, ESBL and carbapenemases production may result in untreatable infections. The co-occurrence of MCR in both ESBL-producing and carbapenemases

producing *E. coli* isolates have been reported in several studies (Haenni *et al.*, 2016; McGann *et al.*, 2016; Hernández *et al.*, 2017; Jin *et al.*, 2018). Likewise, the co-occurrence of MCR, ESBL and Carbapenemase genes in *E. coli* strains form animal origin has also been reported (e.g., one *E. coli* ST167 strain isolated from a chicken wing coproduced MCR-1, NDM-9, CTX-M-65 among other resistant determinants (Yao *et al.*, 2016); and one *E. coli* ST648 and one *E. coli* ST156 strains isolated from a single duck co-produced MCR-1, NDM-5 and CTX-M-55 among other resistant determinants (Yang *et al.*, 2016)).

In 2015, the incidence of infections with antibiotic-resistant bacteria in the European Union (EU) was estimated to be at 671,689 of which 33,110 were attributable deaths (Cassini *et al.*, 2019). In Spain, it has been estimated that in the next 35 years the number of deaths caused by MDR bacterial infections could reach 40,000 infections per year (AEMPS, 2019). It has also been warned that if measures are not implemented, MDR bacterial infections will become the first cause of death in the developed world, outcompeting cancer (O'Neill, 2016). MDR bacterial infections also represent a high burden in terms of economical cost to society, in the EU alone it is estimated that MDR bacterial infections cost EUR 1.5 billion annually (ECDC/EMEA, 2009). Therefore, antimicrobial resistance affects sectors beyond human health, such as animal health, agriculture, food production and economic development. In veterinary medicine, resistance to frequently used antibiotics affect animal health and welfare (OIE, 2003).

This problem is not isolated to Europe (Antibiotic resistance threats in the United States, 2019), as antibiotic resistance is now considered to be one of the most worrying global-health threats of our times (Mathers *et al.*, 2015). In fact, antibiotic treatments are evolving into a time-dependent resource due to the emergence of resistances and it has become urgent to develop new therapeutic strategies, whislt also attempting to slow down the adaptive process of bacteria.

Under the 'One Health' perspective and to ensure development and implementation of global strategies, the World Health Organisation (WHO), the World Organisation for Animal Health (OIE), the Food and Agriculture Organisation (FAO) and the Codex Alimentarius Commission presented overall goals that should be internationally followed (FAO-OIE-WHO, 2010, 2017). The proposed measures are summarised in the following points:

- To promote rational use of antibiotics by implementing good practices in all sectors.
- To reduce the spread of infections focusing on prevention.
- To strengthen health systems and monitoring capabilities globally.
- To strengthen regulation implementation and legislation globally.
- To boost research and encourage the development of new treatments against bacterial infections (ECDC/EMEA, 2009).

In consequence, the EU started the "European action plan" that aims to involve numerous sectors and actors, including human and veterinary medicine, agriculture, finance, environment, pharmaceutical industries and well-informed consumers. In the veterinary sector, the action plan against antimicrobial resistance emphasises the need to consider all other preventive and therapeutic options before using antibiotics. In food production, intervention strategies target each stage of the food chain from the perspective of "farm to table food safety"(FAO, 2011; WHO, 2011). In the research field, surveillance programs focus on the identification of antimicrobial emergence, reservoirs of pathogenic bacteria and their dissemination routes.

1.14.1. EUROPEAN POLICIES AGAINST THE EMERGENCE OF ANTIBIOTIC RESISTANCES

In 2001, in order to ensure the availability and prolong the effectiveness of antimicrobial agents some guidelines were implemented on the prudent use of antimicrobial agents in human medicine (2002/77/EC) (European Commission, 2001) and the directive on the community code relating to veterinary medicinal products, (2001/82/EC) (European Parliament and Council, 2001) with special mention to oral antimicrobial treatments (medicated feed or water) in herds. In this regard, a veterinary prescription is mandatory and individual treatments of the affected animal(s) should be preferable over group or mass treatments.

In 2011, a statement was published between the Commission to the European Parliament and the Council including a five-year action plan against the rising threats of Antimicrobial Resistance. The European plan supported and actively collaborated with the WHO (WHO, 2015), OIE (OIE, 2016), FAO (FAO, 2016) and the Codex Alimentarius Commission (Codex Alimentarius, 2011), as well as, the ECDC (European Centre for Diseases Control), EMA (European Medicines Agency) and EFSA (European food safety Authority) international organizations.

In 2015, the Official Journal of the European Union published the European Parliament resolution on "safer healthcare in Europe: improving patient safety and fighting antimicrobial resistance (2016/C 353/02)" (European Parliament, 2015) and the commission notice on "guidelines for the prudent use of antibiotics in veterinary medicine (2015/C 299/04)" (European Commission, 2015).

In 2017, an addional statement was published between the Commission to the European Parliament and the Council detailing a second proposal on "a European one health action plan against antimicrobial resistance (AMR)(2017/2254-INI)" (European Commission, 2017). This report includes specific disposals to preserve some antimicrobials for human medicine and prohibits the preventive use of antimicrobials through medicated feed in animals (the use of antimicrobial agents as growth promoters is banned in Europe since 2006).

1.14.2. NATIONAL POLICIES AGAINST THE EMERGENCE OF ANTIBIOTIC RESISTANCES

In 2014, a national plan against antibiotic resistance was launched in Spain, named "Plan Nacional frente a la Resistencia a los Antibióticos (PRAN)"(AEMPS, 2014).

In Spain, the implementation of PRAN has achieved a 7.2% reduction of antibiotic use in human medicine between 2015 and 2018 and 32.4% reduction of antibiotic sales for veterinary medicine between 2014 and 2017 (AEMPS, 2019).

2. OBJECTIVES

1. Evaluate the prevalence of sequence type 131 (ST131) and its subclones among *E. coli* strains consecutively isolated from patients with extraintestinal infections in 2016 and among extended spectrum β-lactamases ESBL-producing *E. coli* (ESBLEC) isolated in 2015, in two hospitals (Lucus Augusti University hospital and AP-HP Beaujon hospital), paying special attention to the emergence of new highly virulent and/or multidrug-resistant (MDR) clones.
2. Gain knowledge of phenotypic and genotypic characteristics of the ST131 strains and other high-risk clones isolated from humans and animals, determining phylogroups, clonotypes, sequence types (STs), O:H serotypes, Virulence factor (VF)-encoding genes, antibiotic resistance including the whole genome sequencing of a subset of strains to perform core genome (CG) and plasmidome analysis.
3. Gain knowledge about early biofilm formation (EBF) in *E. coli* strains implicated in extraintestinal infections and contrasted the biofilm-forming ability with genotypic characteristics of the strains.
4. Study production animals (pigs) and companion animals (dogs) as reservoirs for ST131 strains and other high-risk clones for humans and investigate potential clone transfer.

3. METHODS

3.1. EPIDEMIOLOGICAL BACKGROUND OF THE *E. COLI* COLLECTION

3.1.1. HUMAN ORIGIN BACTERIAL COLLECTIONS

Two collections of *E. coli* clinical isolates were studied. The first collection consisted of 196 non-duplicate (one isolate per patient) *E. coli* consecutively isolated (CEC) in 2016 from two hospitals [Lucus Augusti hospital in Lugo (Spain), *n* = 100, and Beaujon hospital in Paris (France), *n* = 96]. This collection that comprised 13 ESBL-producing isolates came from different sources: 146 from urine, 22 from blood, 5 from bile, 3 from ascitic fluid, 6 from abscesses, and 14 from various other sources. As ESBL production is currently one of the most important worldwide threatening mechanisms of antibiotic resistance in *E. coli* clinical isolates, we studied a second collection consisting of 188 non-duplicate ESBL-producing *E. coli* (ESBLEC) isolates obtained from the same two hospitals in 2015 (Lugo hospital, *n* = 92, and Beaujon hospital, *n* = 96). This collection comprised 139 isolates from urine, 25 from blood, seven from bile and 17 from various other sources.

3.1.2. ANIMAL ORIGIN BACTERIAL COLLECTIONS

3.1.2.1. Porcine bacterial collection

We have performed the WGS of 11 resistant *E. coli* O25b:H4-B2-ST131 clade B isolates, seven from piglets with diarrhoea and four from pork meat. The isolates belong to extensive epidemiological studies accomplished in Spain. The seven isolates from diarrhoeic piglets were isolated during the period of 2006-2016 (García-Meniño *et al.*, 2018) and the four from pork meat during the years 2011 and 2012 (Herrera Estévez, 2015).

3.1.2.2. Canine bacterial collection

We characterised a total of 197 non-duplicate *E. coli* isolated from faecal samples of 104 healthy dogs collected in Spain between 2013 and 2017 in collaboration with the Society for the Protection of Animals and Plants in Lugo (Spain) (*n*=26) and the Rof Codina Veterinary Hospital in Lugo (Spain) (*n*= 78).

3.1.3. REFERENCE BACTERIAL COLLECTION

In all procedures, strains from LREC-USC were included as positive controls.

3.2. BACTERIAL CONSERVATION

All bacteria were conserved in the same conditions following the LREC-USC protocol: the isolated strains were cultured in nutrient agar with 0.75% (w / v) and were stored at room temperature in Vacutainer ™ tubes. Under these conditions, *E. coli* strains maintain their viability for at least 10 years. For the preparation of the conservation medium, a mixture of nutrient agar (11.5 g/l) (Applichem-Panreac) and nutrient broth (4 g/l) (Applichem-Panreac) was used.

3.3. PHENOTYPIC CHARACTERIZATION

3.3.1. SEROTYPING

The determination of O and H -antigens was carried out using the method previously described by Guinée *et al.* (1981) and modified by Blanco *et al.* (1996). This method determines *E. coli* O and H -antigens through positive agglutination with available O (O1 to O181) and H (H1 to H56) antisera. The diagnostic antisera used in this thesis were produced and rendered monospecific in LREC-USC. The specificity was controlled by agglutination assays with known reference strains. The O-antigen types O31, O47, O67, O72, O93, O94 and O122 were excluded from this classification as it was proved they are not somatic antigens. For H-antigens, designations H13, H22, and H50 are no longer in use (Murray *et al.*, 1999; Ratiner *et al.*, 2010). Isolates that did not react with any antisera were classified as O non-typeable (ONT) or H non-typeable (HNT) and those that were non-motile were denoted as HNM.

The LREC-USC serotyping procedure includes 176 specific O antisera and 51 specific H antisera. To simplify the method, diagnostic antisera are grouped into polyvalent antisera constituted by a variable number of known monovalent antisera which are used in the first stage of antigen identification.

3.3.1.1. Determination of O-antigens (serogrouping)

Serogrouping includes two phases: (I) a presumptive determination of the O-antigen and (II) a serogroup confirmation by O-antigen titration.

The first step is to prepare two O-antigen suspensions per tested strain. For this purpose, strains are cultured [on tryptone soya agar (TSA) and incubate 18-24h at 37ºC] and further suspended in 2 ml of 0.85% NaCl saline solution (SS). The bacterial concentration should be adjusted to number 6 of the McFarland Barium Sulfate Scale in order to reach approximatively 1,8 x10^9 bacteria per ml. The O-antigen is a thermostable polysaccharide (it remains viable after being heated at 100ºC for 2h) which is the major component of the outer membrane of Gram-negative bacteria. In many cases, *E. coli* strains are covered by another layer named capsule which is thermolatable and contain the K-antigen. Thus, one suspension is boiled for 1 hour (1h at 100ºC) and the second is autoclaved at 121°C for 2.5 h to guaranty K-antigen inactivation (uncovering O-antigen). After cooling, 2 ml of formalinized (0.5%, v/v) SS containing gentian violet (0.005%, w/v) are added in each suspension. Suspensions remain viable for months if they are kept tightly closed at 4°C.

Presumptive determination of O-antigen: (I) dispense 50 µl of each polyvalent antiserum into the wells of a polystyrene microtiter plates and add 50 µl of the O-antigen suspension boiled at 100°C to each well, (II) cover and incubate the microtitre plates at 37°C overnight, (III) examine for agglutination. When a positive reaction is obtained with one or more polyvalent antisera, O-antigen suspension must be assayed with all the monovalent antisera included in the corresponding polyvalent. If the reaction is negative with all polyvalent antisera, O-antigen suspensions autoclaved at 121°C for 2.5 h are tested with O antisera O8, O9, O20 and O101, and if there are still no positive reaction, the problem strain is considered non-typeable (ONT).

The O-antigen determination is completed with the confirmation process of titration: (I) Dispense 50 µl of saline solution in wells 2 to 12 of a row of the microtitre plate, (II) Add 50 µl of a diluted O antiserum (1:40 or 1:80) to wells 1 and 2, (II) prepare a serial dilution (Transfer 50 µl from the second well to the next well, mix and continue until the well number 10, discard the last 50 µl of the mixture from well 10). Keep wells 11 and

12 as negative controls. Final dilution of the antiserum will be 1:80 (for antiserum 1:40) or 1:160 (for antiserum 1:80), (IV) add 50 ɀl of O-antigen suspension to each of the twelve wells of the corresponding row, from well 12 to well 1, (V) cover and incubate the microtitre plate at 37°C overnight, (VI) examine for agglutination. The titre represents the highest antiserum dilution that shows agglutination of the strain assayed.

3.3.1.2. Determination of H-antigens

H-antigens are found on the bacteria's flagellar filaments which are motile organelles of bacteria. In order to optimize the H-antigen–antiserum agglutination reactions, tested isolates require a previous flagellum growth induction. For this purpose, the isolates are inoculated and incubated overnight in 8 ml semi-solid medium contained in U-shaped tubes that allow bacterial mobility from one butt to the other (MIL medium) (Difco). This procedure is performed for each isolate at least three times until motility is fully expressed [strains that do not express motility after six repetitions are designated as non-motile strains (HNM) and cannot be phenotypically assign]. Once this stage is concluded, motile strains are inoculated into 7 ml of brain heart infusion (BHI) agar (Difco) and incubated at 37°C while shaking at 80 rpm for 6-8 h. Afterwards, 7 ml of formalinized (0.5%, v/v) SS are added. The final bacterial suspensions are left at room temperature overnight.

For H-antigen determination, 500 ɀl of the bacterial suspension plus 500 ɀl of polyvalent antiserum are mixed into kahn tubes and leaved for incubation into water at 45ºC for 2h. Subsequently, tubes are examined for agglutination. When a positive reaction is obtained with one or more polyvalent antisera, H-antigen suspension must be assayed with all the monovalent antisera included in the corresponding polyvalent.

3.3.2. Antimicrobial Susceptibility Testing and ESBL Typing

Antimicrobial susceptibility was determined by the minimal inhibitory concentrations (MICs) and/or the disc diffusion method. Resistance was interpreted based on the recommended breakpoints of the CLSI. Fifteen classes of antimicrobial agents were analyzed: penicillins (ampicillin, ticarcillin), penicillins and β-lactamase inhibitors (amoxicillin-clavulanic acid, ampicillin-sulbactam and piperacillin-tazobactam), 1st and 2nd generation of non-extended spectrum cephalosporins (cefazolin and cefuroxime), extended-spectrum cephalosporins (cefotaxime, ceftazidime and cefepime), cephamycins (cefoxitin), monobactams (aztreonam), carbapenems (imipenem, meropenem), aminoglycosides (gentamicin, tobramycin, amikacin), tetracyclines (doxycycline, minocycline and tigecycline), phenicols (chloramphenicol), nitrofurans (nitrofurantoin), quinolones (nalidixic acid, ciprofloxacin and levofloxacin), folate pathway inhibitors (trimethoprim-sulphamethoxazole), phosphonic acids (fosfomycin), and polymyxins (colistin).

E. coli MDR was defined as resistance to one or more agents in three or more classes of tested drugs (Magiorakos *et al.*, 2012).

ESBL production detected by the double disk synergy test (Jarlier *et al.*, 1988) was further confirmed by PCR. Genetic identification of ESBL and pAmpC types was carried out by PCR followed by amplicon sequencing as described in the molecular methods section.

3.3.3. DETERMINATION OF EARLY BIOFILM FORMATION (EBF)

The kinetics of early biofilm formation (EBF) were assessed using the BioFilm Ring Test® (BioFilm Control, Saint-Beauzire, France) according to the manufacturer's recommendations. Briefly, each isolate was sub-cultured twice on brain heart infusion (BHI) agar (Benton Dickinson, Le Pont-de-Claix, France) at 37°C for 24 h. Three colonies of the second subculture on BHI agar were suspended in BHI broth (Biofilm Control). Suspension was standardised to an optical density at 600 nm of 1.00 ± 0.05 (Ultrospec10: Biochrom, Cambridge, United Kingdom) and then diluted at 1:250 in BHI broth to obtain a final concentration of approximately 10^6 UFC/ml. The bacterial suspension was supplemented (1% vol/vol) with magnetic microbeads (TONER 4, Biofilm Control) and 200 µl of the mix was deposited in two wells of three polystyrene 96-well microtiter plates that were incubated for 2, 3, and 5 h, respectively. At the end of each incubation time, 100 µl of liquid contrast solution (LIC001, Biofilm Control) was added on the top of each well and the microplate was put on a magnetic block for 1 min. After magnet contact, free beads were attracted toward the centre of each well, forming a brown spot, while beads embedded in biofilms were blocked and remained undetectable. Each microplate was scanned using a BioFilm Control plate reader. The intensity of the spot was analysed using the BioFilm Ring Test®-software version 3.0.3 and expressed as a biofilm formation index (BFI) with values ranging from 20 (non-formation of biofilm) to 0 (high formation of biofilm) that is inversely proportional to attached bacteria that block the beads. BFI values ≤ 5 mean that the isolates are strong biofilm producers; BFI values between 6 and 14, moderate producers; BFI values between 15 and 19, weak producers; non-producers display a BFI value of 20. Each isolate was tested in three independent experiments and the average of the six measures was used as the final BFI value. For each experiment, isolates S250 and 39 previously described with the method carried out in the present study as strong and negative producers of biofilm, respectively, were used as controls (Nicolas-Chanoine *et al.*, 2017). We also included BHI broth without bacteria as negative control.

3.4. MOLECULAR CHARACTERISATION

The molecular characterisation of the bacterial strains included: (I) the description of virulence profiles through the identification of Virulence factor (VF)-encoding genes by Polymerase Chain Reaction (PCR) followed by the definition of the virulence status, (II) the description of the antimicrobial resistances profile through PCR followed by amplicon sequencing, (III) the determination of the phylogenetic groups of *E. coli* through PCR, (IV) the determination of clonotypes (CH) and sequence types (ST) through PCR followed by amplicon sequencing, (VI) the identification of the clades and subclades of ST131 *E. coli* by PCR and (VII) the identification of the *fimB* gene by PCR followed by amplicon sequencing in a subset of strains from the human origin collection of bacteria.

Polymerase Chain Reaction (PCR): LREC-USC protocol

Approximately 10 µM (using 10 µM sterile inoculating loops) of the cultured bacteria [on Tryptone Soy Agar (TSA) media] are suspended in 600 µl of sterile MiliQ water in an eppendorf tube and boiled at 100ºC for 5 min for DNA extraction. After centrifugation at 11.000 rpm for 2 min, 5 µl of supernatant are taken for PCR. All reactions are carried out in a 25 µl volume containing the extracted DNA plus 12.5 µl of buffer supplied with *Taq* polymerase, 2,5 mM MgCl2 and 200 µM of each nucleotide: dATP, dCTP, dGTP y dTTP (NZYtechTaq 2x Green Master Mix) and 0.5 to 1 µl (variable according to the

allele-specific primers used) of the appropriate primers at 20 ɀM concentration. We adjusted the final volume of the reaction to 25 ɀl by adding MiliQ sterile water. PCR reactions are performed under the following conditions: denaturation 3 min at 94ºC, 35 cycles of 1 min at 94ºC and 1 min at the specific primer-annealing temperature and a final extension step of 90 s at 72ºC. Plus, a final extension of 3 min at 72ºC. The lecture of the amplified DNA fragments is performed by conventional agarose gel electrophoresis under UV light [2% Seakem® LE agarose in 1X TAE buffer (1.6 M Tris, 0.8 M sodium acetate, 0.025 M EDTA) with 5 ɀl of Green Safe Premium marker (Nzytech) per 100 ml]. In all procedures negative and positive controls are systematically included.

Sanger sequencing

Amplicons were obtained by conventional PCR, carried out in a 50 ɀl volume containing 1 ɀl extracted DNA plus 10 ɀl of the 5x My Taq™ Reaction Buffer (Bioline) that include 5mM dNTPs, 15 mM MgCl2 plus 0.2 ɀl of MyTaqTM DNA polymerase (Bioline), and 1 ɀl of the appropriate primers at 20 ɀM concentration. With the following conditions: denaturation 3 min at 94ºC, 35 cycles of 1 min at 94ºC and 1 min at the specific primer-annealing temperature and 1 min at 72ºC. The final extension lasted for 3 min at 72ºC. In all procedures negative and positive controls were systematically included.

Afterwards, forward and reverse complementary strands were sequenced at the Sequencing Facility of University of Santiago de Compostela (Lugo). They implemented the "BigDye™ Terminator v3.1 Cycle Sequencing" protocol.

We further analysed the sequences using the BLAST sequence software and Genbank [registered trademark of the National Library of Medicine (NIH) and the National Center for Biotechnology Information (NCBI)].

3.4.1. SPECIES IDENTIFICATION AND PHYLOGENETIC GROUP ASSIGNMENT

E. coli species was confirmed by detection of the beta-D-glucuronidase encoding gene *uidA* as performed in other studies (Martins *et al.*, 1993). Assignment to the main phylogroups (A, B1, B2, C, D, E, F) was based on the protocol of Clermont *et al.*, (2013). We completed species identification with the *Escherichia* genus discrimination PCR protocol describe by Lindsey *et al.* (2017) if the *uidA* gene was absent or the phylogroups assignment failed. Tables 7 and 8 include the PCR conditions and primers. Figure 3 schematises the interpretation of the PCR results.

Table 7. Primers used for the determination of the phylogenetic groups A, B1, B2, C, D, E y F (Clermont *et al.*, 2013).

Target	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
chuA	ChuA.1b	ATGGTACCGGACGAACCAAC	288	58 ºC	Clermont *et al.* (2013)
	ChuA.2b	TGCCGCCAGTACCAAAGACA			Clermont *et al.* (2000)
yjaA	YjaA.1b	CAAACGTGAAGTGTCAGGAG	211		
	YjaA.2b	AATGCGTTCCTCAACCTGTG			
TSPE4.C2	TspE4C2.1b	CACTATTCGTAAGGTCATCC	152		Clermont *et al.* (2013)
	TspE4C2.2b	AGTTTATCGCTGCGGGTCGC			
arpA	AceK.f	AACGCTATTCGCCAGCTTGC	400		
	ArpA1.r	TCTCCCCATACCGTACGCTA			
TrpAgpC	trpAgpC.1	AGTTTTATGCCCAGTGCGAG	219	56 ºC	Lescat *et al.* (2013)
	trpAgpC.2	TCTGCGCCGGTCACGCCC			
*arp*A	ArpAgpE.f	GATTCCATCTTGTCAAAATATGCC	301	57 ºC	Lescat *et al.* (2013)
	ArpAgpE.r	GAAAAGAAAAAGAATTCCCAAGAG			
trpA	trpBA.f	CGGCGATAAAGACATCTTCAC	489	56 ºC	Clermont *et al.* (2008)
	trpBA.r	GCAACGCGGCCTGGCGGAAG			

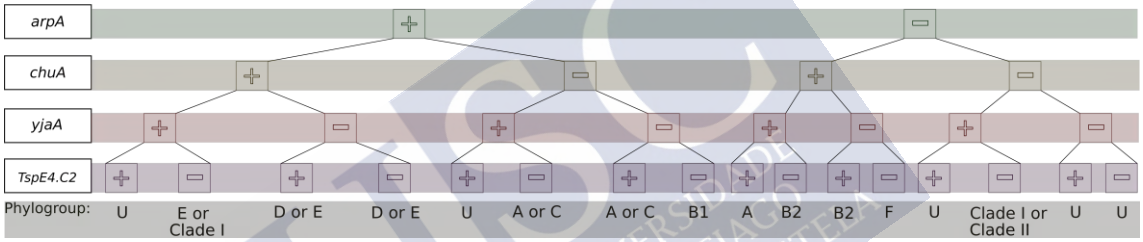


Figure 3 Interpretation of PCR results for phylogroups designation Clermont *et al.* (2013). Distintion between A and C phylogroups depends on the results obtained with trpAgpC primers (The C phylogroups is assigned whit trpAgpC positivity). Distinction between D and E phylogroups depends on the results obtained with ArpAgpE primers (The E phylogroups is assigned whit ArpAgpE positivity). Distinction between phylogroup E and Clade I and between Clade I and Clade II requires complemented characterizsation with specific primers described in Clermont *et al.*, (2011).

Table 8. Primers used for the determination *uidA* gene and distinction of *E. coli*, *E. albertii* and *E. fergusonii* (Lindsey *et al.*, 2017).

Target	Primer	annealing temperature	Nucleotide sequence (5´- 3´)	PCR product size (bp)	Reference
uidA	uidA-F	60 ºC	GCGTCTGTTGACTGGCAGGTGGTGG	503	Gómez-Duarte *et al.* (2010)
	uidA-R		GTTGCCCGCTTCGAAACCAATGCCT		
E. coli	EC_F	57 ºC	CCAGGCAAAGAGTTTATGTTGA	212	Lindsey *et al.* (2017)
	EC_R		GCTATTTCCTGCCGATAAGAGA		
E. albertii	EA_F		GTAAATAATGCTGGTCAGACGTTA	393	
	EA_R		AGTGTAGAGTATATTGGCAACTTC		
E. fergusonii	EF_F		AGATTCACGTAAGCTGTTACCTT	575	
	EF_R		CGTCTGATGAAAGATTTGGGAAG		

3.4.4. MLST, CH TYPING AND CLONE DETERMINATION

The sequence types (STs) were established following the multilocus sequence typing (MLST) scheme of Achtman by gene amplification and sequencing of the seven housekeeping genes (*adk*, *fum*C, *gyr*B, *icd*, *mdh*, *pur*A, and *rec*A), according to the protocol and primers specified at the *E. coli* MLST website (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli) (Wirth *et al.*, 2006). Additionally, porcine isolates were also typed according to the MLST Pasteur Institute scheme. Clonotype identification was determined by *fumC* and *fimH* (CH) sequencing (Tchesnokova *et al.*, 2013; Roer *et al.*, 2018). Clones were characterised by using the association of phylogroup, clonotype, and ST.

Table 9. Primers used for clonotypes and Achtman sequence type designation.

Gene	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
fimH (locus size 469)	fimH-wf	CACTCAGGGAACCATTCAGGCA	975	57 ºC	Weissman *et al.* (2012)
	fimH-wr	CTTATTGATAAACAAAAGTCAC			
fumC (locus size 469)	fumCF	TCACAGGTCGCCAGCGCTTC	806	56 ºC	
	fumCR	GTACGCAGCGAAAAAGATTC			
adk (locus size 536)	adkF	ATTCTGCTTGGCGCTCCGGG	975	56 ºC	Wirth *et al.* (2006)
	adkR	CCGTCAACTTTCGCGTATTT			
fumC (locus size 469)	fumCF	TCACAGGTCGCCAGCGCTTC	806	56 ºC	
	fumCR	GTACGCAGCGAAAAAGATTC			
gyrB (locus size 460)	gyrBF	TCGGCGACACGGATGACGGC	911	56 ºC	
	gyrBR	ATCAGGCCTTCACGCGCATC			
icd (locus size 518)	icdF	ATGGAAAGTAAAGTAGTTGTTCCGGCACA	878	56 ºC	
	icdR	GGACGCAGCAGGATCTGTT			
mdh (locus size 452)	mdhF	ATGAAAGTCGCAGTCCTCGGCGCTGCTGGCG G	932	66 ºC	
	mdhR	TTAACGAACTCCTGCCCCAGAGCGATATCTTT CTT			
purA (locus size 478)	purAF	CGCGCTGATGAAAGAGATGA	816	56 ºC	
	purAR	CATACGGTAAGCCACGCAGA			
recA (locus size 510)	recAR1	AGCGTGAAGGTAAAACCTGTG	780	56 ºC	
	recAF1	ACCTTTGTAGCTGTACCACG			

3.4.5. IDENTIFICATION OF ST131 CLADES AND SUBCLADES

The ST131 clades (A, B, C), subclades C2 (also known as subclone *H*30Rx) and C1, and cluster C1-M27 were established by PCR (Matsumura *et al.*, 2017b).

Table 10. Primers used for identification of ST131 clades and subclades (Matsumura *et al.,* 2017b).

Target	Primer	Nucleotide sequence (5-3)	PCR product size (bp)
ST131	ST131_R19-YF1	AGCAACGATATTTGCCCATT	580
	ST131_R19-YR1	GGCGATAACAGTACGCCATT	
Clade A	CladeAspe4-YF5	TGACGGGACGTGAGCAAATTA	707
	CladeAspe4-YR5	AGTCAGACCTAGCCACCCTT	
Clade B	prfC-1615spe0-YF1	CAACGTTGAAGCAGTGTATGAG	442
	prfC-d2034-YR1	TGACAATCGACGGCTTTAGA	
Subclade C1	C1-578spe-YF1	GGCCCCACAAATTGCTT	337
	C1-898-YR1	CGCACCTCCGATACCAAA	
Cluster C1-M27	M27PP1C-YF1	TGAATCAAAGGTCCGAGCTG	232
	M27PP1C-YR1	TATGGCTGGCAGATGCTTTA	
Subclade C2	nrdI-534spe2-YF1	ACGGATTCAGGTAGACGATT	164
	nrdI-678R	CCTCACCAAAGTTGCGATTAC	
Clade C	C-SNP1-700spe-YF1	CGCTGGCCAGTTATCTGAAAT	103
	C-SNP1-762spe-YR2	CCTTTCACCAACTGGGTTACT	

3.4.6. MOLECULAR CHARACTERISATION OF *FIMB* GENE

We randomly selected 127 isolates from the human bacterial collection (approximately one-third of the isolates) by using the RAND function (Microsoft Excel) system for the *fimB* gene analysis. The *fimB* gene was amplified with specific primers (FimB F: 5′-AGCATGGCGTTTGTATGG-3′; FimB R: 5′-CCCTGGTATCTCAACTATCTCT-3′) and sequenced as previously described (Nicolas-Chanoine *et al.*, 2017). When the *fimB* gene was disrupted, the detection of the previously described *IS*3-like in *H*30 R isolates (Totsika *et al.*, 2011; Nicolas-Chanoine *et al.*, 2017) was performed by PCR using specific primers designed in this study (FimBw F: 5′-AGCATGGCGTTTGTATGG-3′; IS R: 5′-CTGAATGTGATGTGCCGATG-3′).

Phylogenetic tree of FimB variants was constructed by the UPGMA method of MEGA 6. The 34 *fimB* gene sequences have been registered in the GenBank database under the following accession numbers: MK301552 to MK301585.

3.4.2 VIRULENCE GENOTYPING

Virulence factor (VF)-encoding genes of *E. coli* causing extraintestinal infections were screened by PCR (Clermont et al., 2013; Dahbi et al., 2014; Mamani et al., 2019). The used conditions and primers for PCR amplification of VF-encoding genes are indicated in Table 11.

The isolates were presumptively designed as extraintestinal pathogenic *E. coli* (ExPEC) if positive for ≥2 of 5 markers, including *papAH* and/or *papC*, *sfa/focDE*, *afa/draBC*, *kpsM II*, and *iutA* (Johnson *et al.*, 2015), and as uropathogenic *E. coli* (UPEC) if positive for ≥3 of 4 markers, including *chuA*, *fyuA*, *vat*, and *yfcV* (Spurbeck *et al.*, 2012) and as avian pathogenic *E. coli* (APEC) if positive for ≥ 4 of 5 markers (*hlyF*, *iutA*, *iroN*, *iss* and *ompT*) (Johnson *et al.*, 2008d). We established a virulence gene score with the number of extraintestinal virulence-associated genes, while the virotypes (A to F) of the ST131 isolates were established according to the scheme described in Table 12 according to Dahbi *et al.* (2014).

Table 12. Virotype designation scheme for ST131 *E. coli* (Adapted from Dahbi *et al.* 2014).

Genes:	afa/draBC	afa operon FM955459	iroN	sat	ibeA	papG II	papG III	cnf1	hlyA	cdtB	neuC-K1	kpsM II-K2	kpsM II-K5
Virotype A	+	+	−	+/−	−	−	−	−	−	−	−	+	−
Virotype B	−	−	+	+/−	−	+/−	−	−	−	−	−	−	+/−
Virotype C1	−	−	−	+	−	−	−	−	−	−	−	+	−
Virotype C2	−	−	−	+	−	−	−	−	−	−	−	−	+
Virotype C3	−	−	−	+	−	−	−	−	−	−	−	−	−
Virotype D1	−	−	+/−	−	+	−	−	−	−	+	−	−	+
Virotype D2	−	−	+/−	−	+	−	+	−	−	+	−	−	+
Virotype D3	+/−	+/−	+/−	+/−	+	−	−	−	−	−	−	−	+
Virotype D4	−	−	+/−	−	+	−	−	−	−	−	+	−	−
Virotype D5	−	−	+/−	−	+	−	+	+	+	−	−	−	+
Virotype E	−	−	−	+	−	+	−	+	+	−	−	−	+
Virotype F	−	−	−	+	−	+	−	−	−	−	−	−	+

Table 11. Primers used for detection of Virulence factor (VF)-encoding genes of *E. coli* causing extraintestinal infections.

Target	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
Adhesins					
fimH	FimH F	TGCAGAACGGATAAGCCGTGG	508	58-66ºC	Johnson and Stell, (2000)
	FimH R	GCAGTCACCTGCCCTCCGGTA			
fimAvMT78	fimA215	ACTTTAGGATGAGTACTG	266	52 ºC	Marc and Dho-Moulin, (1996)
	fimA201	TCTGGCTGATACTACACC			
papEF	PapEF F	GCAACAGCAACGCTGGTTGCATCAT	336	58-66	Yamamoto *et al.* (1995)
	PapEF R	AGAGAGAGCCACTCTTATACGGACA			
papGII	Pap-II f	GGGCATTGCTACGGTAACCTG	545	54-60 ºC	Mora *et al.* (2013)
	Pap-II r	CGCTATTAATAGACAGATCACC			
papGIII	Pap-III f	CGGCAACTTTAAGCTATGTG	720	60-68 ºC	Mora *et al.* (2013)
	Pap-III r	TGTACCATCTCATCGTTGTCTC			
papC	Forward	GTGGCAGTATGAGTAATGACCGTTA	205	60 ºC	Johnson *et al.* (2015)
	Reverse	ATATCCTTTCTGCAGGGATGCAATA			
papAH	papA-F	ATGGCAGTGGTGTCTTTTGGTG	720	70 ºC	Johnson and Stell, (2000)
	papA-R	CGTCCCACCATACGTGCTCTTC			
sfa/focDE	sfa 1	CTCCGGAGAACTGGGTGCATCTTAC	410	56-64 ºC	Le Bouguenec *et al.* (1992)
	sfa 2	CGGAGGAGTAATTACAAACCTGGCA			
yfcV	yfcV-F	ACATGGAGACCACGTTCACC	292	63 ºC	Spurbeck *et al.* (2012)
	YfcV-R	GTAATCTGGAATGTGGTCAGG			
afaFM955459	Afa-025F	GAGTCACGGCAGTCGCGGCGG	207	55 ºC	Blanco *et al.* (2009)
	Afa-025R	TTCACCGGCGACCAGCCATCTCC			
Toxins					
afa/draBC	afa1	GCTGGGCAGCAAACTGATAACTCTC	750	64 ºC	Le Bouguenec *et al.* (1992)
	afa2	CATCAAGCTGTTTGTTCGTCCGCCG			
cnf 1	CNF1-F2	CAGGAGGTACTTAGCAGCGT	468	48-58 ºC	Mora *et al.* (2013)
	CNF1-RC	TAATTTTGGGTTTGTATC			
cdtB	cdt-s1	GAAAGTAAATGGAATATAAATGTCCG	466	48-52 ºC	Tóth *et al.* (2003)
	cdt-as1	AAATCTCCTGCAATCATCCAGTTA			
	cdt-s2	GAAAATAAATGGAACACACATGTCCG			
	cdt-as2	AAATCACCAAGAATCATCCAGTTA			
hlyA	hly f	AACAAGGATAAGCACTGTTCTGGCT	1177	64 ºC	Yamamoto *et al.* (1995)
	hly r	ACCATATAAGCGGTCATTCCCGTCA			
hlyF	Forward	TCGTTTAGGGTGCTTACCTTCAAC	444	60 ºC	Morales *et al.* (2004)
	Reverse	TTTGGCGGTTTAGGCATTCC			
sat	SatF	GCAGCTACCGCAATAGGAGGT	937	60 ºC	Johnson *et al.* (2013)
	SatR	CATTCAGAGTACCGGGGCCTA			
vat	vat-F	TCAGGACACGTTCAGGCATTCAGT	1100	70 ºC	Spurbeck *et al.* (2012)
	vat-R	GGCCAGAACATTTGCTCCCTTGTT			
Siderophores					
*iro*N	Ironec-f	AAGTCAAAGCAGGGGTTGCCCG	665	62 ºC	Johnson *et al.* (2000)
	Ironec-r	GACGCCGACATTAAGACGCAG			
fyuA	fyuA-F	GTAAACAATCTTCCCGCTCGGCAT	850	63 ºC	Spurbeck *et al.* (2012)
	fyuA-R	TGACGATTAACGAACCGGAAGGGA			
*iuc*D	Aer f	TACCGGATTGTCATATGCAGACCGT	602	54 ºC	Yamamoto *et al.* (1995)
	Aer r	AATATCTTCCTCCAGTCCGGAGAAG			
*iut*A	aer-851f	GGCTGGACATCATGGGAACTGG	301	66 ºC	Johnson *et al.* (1997)
	aer-1152r	CGTCGGGAACGGGTAGAATCG			
Capsule					
*kps*M II	KpsII f	GCGCATTTGCTGATACTGTTG	272	60 ºC	Johnson and Stell, (2000)
	KpsII r	CATCCAGACGATAAGCATGAGCA			

Table 11. *Cont.*

Target	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
Capsule					
*kps*M II-K2	kpsII f KpsII-K2r	GCGCATTTGCTGATACTGTTG AGGTAGTTCAGACTCACACCT	570	60 ºC	Johnson and O'Bryan, (2004)
*kps*M II-K5	K5 f kpsII r	CAGTATCAGCAATCGTTCTGTA CATCCAGACGATAAGCATGAGCA	159	54 ºC	Johnson and Stell, (2000)
*neu*C-K1	neu1 neu2	AGGTGAAAAGCCTGGTAGTGTG GGTGGTACATCCCGGGATGTC	676	54 ºC	Moulin-Schouleur *et al.* (2006)
*kps*M III	kps III f kps III r	TCCTCTTGCTACTATTCCCCCT AGGCGTATCCATCCCTCCTAAC	392	54 ºC	Johnson and Stell, (2000)
Micellaneous					
*cva*C	CoIV-Cf CoIV-Cr	CACACACAAACGGGAGCTGTT CTTCCCGCAGCATAGTTCCAT	680	68 ºC	Johnson and Stell, (2000)
*ibe*A	Ibe10 f Ibe10 r	AGGCAGGTGTGCGCCGCGTAC TGGTGCTCCGGCAAACCATGC	170	58 ºC	Johnson and Stell, (2000)
iss	is-f is-r	CAGCAACCCGAACCACTTGATG AGCATTGCCAGAGCGGCAGAA	323	62 ºC	Johnson *et al.* (2008e)
*mal*X	MALX-F MALX-R	GCATGAGCAGTGCGATACATCGC AGGGCTGGGAAGTGGTTTAGCC	828	68 ºC	Mora *et al.* (2013)
ompT	Forward Reverse	ATCTAGCCGAAGAAGGAGGC CCCGGGTCATAGTGTTCATC	559	60 ºC	Johnson *et al.* (2015)
*tra*T	TraTf TraTr	GGTGTGGTGCGATGAGCACAG CACGGTTCAGCCATCCCTGAG	290	60 ºC	Johnson and Stell, (2000)
tsh	tsh03 tsh15	GGTGGTGCACTGGAGTGG AGTCCAGCGTGATAGTGG	640	54-56 ºC	Dozois *et al.* (2000)
usp	usp-f	ACATTCACGGCAAGCCTCAG	440	68 ºC	Bauer *et al.* (2002)

E. coli capsules are designated after interpretation of the PCR results following the scheme of Table 13.

Table 13. Capsule designation based on PCR results.

Capsule	*kpsM-II gene*	*neuC-K1 gene*	*kpsM-II-K2 gene*	*kpsM-II-K5 gene*
type II-K1	+	+	+	-
type II-K2	+	-	+	-
type II-K5	+	-	+	+

In order to detect hybrid pathotypes, the 196 human clinical *E. coli* consecutively isolated (CEC) in 2016 were also examined for ten VF-encoding genes, specific for pathotypes of intestinal pathogenic *E. coli* (InPEC): typical and atypical enteropathogenic *E. coli* (tEPEC and aEPEC) (*eae* and *bfpA* genes), enteroinvasive *E. coli* (EIEC) (*ipaH* gene), enterotoxigenic *E. coli* (ETEC) (*eltA* and *est* genes), Shiga toxin-producing *E. coli* (STEC) (*stx*1 and *stx*2 genes) and enteroaggregative *E. coli* (EAEC) (*aatA*, *aaiC* and *aggR* genes) (Blanco *et al.*, 2006). Table 14 includes the used conditions and primers.

Table 14. Primers used for detection of virulence factors encoding genes of intestinal pathogenic *E. coli* (InPEC)

Gene	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
STEC					
stx1	VT1-F VT1-R	TCGCTGAATGTCATTCGCTCTGC TCAGCAGTCATTACATAAGAAC	539	55 ºC	Mora et al. (2011)
stx2	VT2-F1 VT2-F2 VT2-R	TTTCTTCGGTATCCTATTCCC TGTCTTCAGCATCTTATGCAG CTGCTGTCCGTTGTCATGGAA	358	55 ºC	
EPEC					
eae	EAE-V3F EAE-MBR	CATTGATCAGGATTTTTCTGGT TCCAGAATAATATTGTTATTACG	510	54 ºC	Mora et al. (2011)
bfp	BFP-NF1 BFP-NR1	ATGGTTTCTAAAATCATGAATAAG ATTATTCCGGAATTGCAGATGTGT	262	55 ºC	Bennett, (2003); García-Meniño *et al.* (2018)
ETEC					
estA	STa-A STa-B	ATTTTTATTTCTGTATTGTCTTT GGATTACAACACAGTTCACAGCAGT	176	50-52 ºC	Penteado *et al.* (2002)
estB	Stb-F Stb-R	ATCGCATTTCTTCTTGCATC GGGCGCCAAAGCATGCTCC	175	54 ºC	Blanco *et al.*, (1997)
eltA	LT-A-1 LT-A-2	GGCGACAGATTATACCGTGC CCGAATTCTGTTATATATGTC	696	54-56 ºC	Schultsz *et al.* (1994)
EIEC					
ipaH	EI1 EI2	GCTGGAAAAACTCAGTGCCT CCAGTCCGTAAATTCATTCT	424	55 ºC	Tornieporth *et al.* (1995)
EAEC					
aatA	pCVD432/start pCVD432/stop	CTGGCGAAAGACTGTATCAT CAATGTATAGAAATCCGCTGTT	630	60 ºC	Schmidt *et al.* (1995)

3.4.3. ANTIMICROBIAL RESISTANCES GENOTYPING

Genetic identification of *bla*$_{ESBL}$ types (TEM, SHV and CTX-M enzymes), pAmpC types (*LAT-1* to *LAT-4, CMY-2* to *CMY-7* and *BIL-1)* and the *mcr* genes (*mcr-1, mcr-2, mcr-3, mcr-4 y mcr-5)* was carried out by PCR followed by amplicon sequencing (Perez-Perez and Hanson, 2002; Leflon-Guibout *et al.*, 2004a, 2004b; Blanco *et al.*, 2009). The used primers are indicated in Table 15.

The genetic environment of the *mcr* detected genes were manually revisited and compared to previously reported ones with CLC Sequence Viewer (Version 8.9, QIAGEN) and EasyFig tools (Sullivan *et al.*, 2011b).

Table 15. Primers used for resistance gene detection.

Gene	Primer	Nucleotide sequence (5´- 3´)	PCR product size (bp)	annealing temperature	Reference
bla$_{CTX-M}$	CTX-C3	ATGTGCAGCACCAGTAAAGTGATG	542	55 ºC	Mora *et al.* (2013)
	CTX-C4	ACCGCGATATCGTTGGTGGTGCC			
bla$_{CTX-M-15\ extremo\ 3'}$	CTX-MF1	ATAAAACCGGCAGCGGTG	483	60 ºC	Leflon-Guibout *et al.* (2004a)
	CTX-MF2	GAATTTTGACGATCGGGG			
bla$_{CTX-M-grupo\ 1}$	M13U	GGTTAAAAAATCACTGCGTC	863	60 ºC	Saladin *et al.* (2002)
	M13L	TTGGTGACGATTTTAGCCGC			
bla$_{CTX-M-grupo\ 1}$[a]	CTX-15-F1	GAAGCTAATAAAAAACACACGTGG	1044-1123	52 ºC	Mora *et al.* (2013)
	CTX-15-R	GTATGCGCAAGCGCAGGTGG			
bla$_{CTX-M-grupo9}$	CTX-M9-F	GTGACAAAGAGAGTGCAACGG	856	64 ºC	Simarro *et al.* (2000)
	CTX-M9-R	ATGATTCTCGCCGCTGAAGCC			
bla$_{CTX-M-grupo9}$ [a]	CTX-M9-14-14B-24F	GAATACTGATGTAACACGGA	998	44 ºC	García-Meniño *et al.* (2018)
	CTX-M9-RS	AGCTGAAGATGTATATCAAG			
bla$_{CTX-M-grupo9}$[a]	CTX-M9-14-14B-24F	GAATACTGATGTAACACGGA	989	52 ºC	García-Meniño *et al.* (2018)
	CTX-M14-24-R	CTGCGTTGTCGGGAAGATACG			
bla$_{CTX-M-grupo9}$ [a]	CTX-M9-14B-F	CCTATACCCGAGGCGCGACAG	1059	44 ºC	García-Meniño *et al.* (2018)
	CTX-M9-RS	AGCTGAAGATGTATATCAAG			
bla$_{CTX-M-grupo9}$[a]	CTX-M14-24-F	CTAAATTCTTCGTGAAATAGTG	1049	44 ºC	García-Meniño *et al.* (2018)
	CTX-M14-24-R	CTGCGTTGTCGGGAAGATACG			
bla$_{SHV}$	SHV-F2	TTGTCGCTTCTTTACTCGCC	879	64 ºC	Mora *et al.* (2013)
	SHV-R2	CCCGGCGATTTGCTGATTTCGC			
bla$_{SHV}$[a]	SHV-1	GGGTTATTCTTATTTGTCGC	930	48 ºC	Rasheed *et al.* (1997)
	SHV-2	TTAGCGTTGCCAGTGCTC			
bla$_{TEM}$	TEM-1F	ATGAGTATTCAACATTTCCG	868	58 ºC	Rasheed *et al.* (1997)
	TEM-1R	CTGACAGTTACCAATGCT TA			
LAT-1 a LAT-4, CMY-2 a CMY-7, BIL-1	CITMF	TGGCCAGAACTGACAGGCAAA	462	66 ºC	Perez-Perez and Hanson, (2002)
	CITMR	TTTCTCCTGAACGTGGCTGGC			
mcr-1	CLR5-F	CGGTCAGTCCGTTTGTTC	309	57 ºC	Liu *et al.*, (2016d)
	CLR5-R	CTTGGTCGGTCTGTAGGG			
mcr-1	mcr1_320bp_fw	AGTCCGTTTGTTCTTGTGGC	320	58 ºC	Rebelo *et al.* (2018)
	mcr1_320bp_rev	AGATCCTTGGTCTCGGCTTG			
mcr-1[a]	mcrS1-F	GGGATTGCGCAATGATTGC	548	58 ºC	García-Meniño *et al.* (2018)
	mcrS1-R	CACCCAAACCAATGATACG			
mcr-1[a]	mcrS1-F2	AAAGACGCGGTACAAGCAAC	1016	58 ºC	García-Meniño *et al.* (2018)
	mcrS1-R2	CCCACCGCCCATAATACGAA			

[a]Primers used exclusively for Sanger Sequencing.

3.5. WHOLE GENOME SEQUENCING (WGS), ASSEMBLY AND PRIMARY ANALYSIS

We performed the WGS of a total of 109 *E. coli* strains belonging to three different studies: (I) 75 human ESBLEC (51 genomes from Lucus Augusti hospital and 24 genomes from Beaujon hospital). These genomes come from different clinical samples: 57 isolates from urine, 11 from blood and 7 from various other sources including one from bile, one from ascitic fluid, one from a bone, two from wounds and two from digestive infections; (II) 23 ST372 strains, 21 ST372 *E. coli* isolates from canine origin and two ST372 *E. coli* isolates human origin and (III) 11 resistant *E. coli* O25b:H4-B2-ST131 clade B strains, seven from piglets with diarrhoea and four from pork meat. The whole genome sequencing and further analysis of the genomes were performed as follows.

Total DNA was extracted with the QIAmp DNA Mini Kit (Qiagen). Libraries were prepared by using the TruSeq DNA PCR-Free protocol (Illumina) at the Genomics and Bioinformatics Core Facility (Centre for Biomedical Research of La Rioja, CIBIR). Paired-end 100 bp reads on fragments of 550 bp insert size were sequenced in an Illumina HiSeq 1500.

Genomes were reconstructed by using PLACNET webserver (Vielva *et al.*, 2017). Identification of Open Reading Frames (ORFs) and genome annotation of the assembled genetic elements was performed by using Prokka (Seemann, 2014).

Genomes were *in silico* typed by the following databases: SerotypeFinder (Joensen *et al.*, 2015), ClermontTyping Web (Beghain *et al.*, 2018), MLSTtyper (Larsen *et al.*, 2012), CHtyper (Camacho *et al.*, 2009). Three different databases were used for the identification of antibiotic resistance genes: ResFinder (Zankari *et al.*, 2012a), CARD Resistance Gene Identifier (McArthur *et al.*, 2013) and ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) (Gupta *et al.*, 2014). Besides, the PointFinder database (Zankari *et al.*, 2017) was used to determine point mutations. The VirulenceFinder (Ren *et al.*, 2017) and the VFDB (Chen *et al.*, 2005) databases were used to explore virulence factors.

The assembly information of draft genomes, database sources and input parameters can be found in annex 1: Tables S1 and S2. The datasets of the 11 LREC ST131 porcine origin genomes can be found in the NCBI sequence databases as part of BioProject SUB5714329, with accession codes SAMN11936814 to SAMN11937940. The 23 LREC ST372 canine and human origin genomes can be found in the NCBI sequence databases as part of BioProject PRJNA627579, with accession codes SAMN14677009 to SAMN14677031.

3.5.1. CORE GENOME (CG) AND PHYLOGENETIC ANALYSIS

In this thesis, four core genome (CG) analysis were performed: (I) a first CG (CG-1) including the 75 ESBLEC genomes from human strains sequenced in this thesis, (II) a second CG (CG-2) including 30 ST131 ESBLEC genomes from human strains sequenced in this thesis plus other 51 human ST131 *E. coli* full-genomes references, (III) a third CG (CG-3) including the 11 genomes from porcine ST131 strains sequenced in this thesis plus 73 ST131 *E. coli* full-genomes references and (IV) a fourth CG (CG-4) including the 23 ST372 *E. coli* genomes sequenced in this study plus 174 ST372 *E. coli* full-genome references. Regarding CG-1, CG-2 and CG-3, the core genome was defined as the collection of genes present at least once in all the genomes analysed, with more than 90% similarity and 90% coverage, as defined by Lanza *et al.* (2014). In case of CG-4, the

single nucleotide position (SNP) tree analysis was done using the CSI Phylogeny 1.4 server from the CGE with J22 strain as reference (ID: GCA_009497315). After analysing the SNP matrix, we took all the ST372 genomes from human strains plus some representatives from canine strains to made a tree visualization using EnteroBase (Zhou *et al.*, 2020). References were retrieved from the Enterobase (http://enterobase.warwick.ac.uk/species/index/ecoli) and the NCBI Bioproject Database (https://www.ncbi.nlm.nih.gov/bioproject/). The accession number of the mentioned references can be found as Supplementary Tables 2 in Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms8111712 and Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387.

3.5.2. MOBILE GENETIC ELEMENTS (MGES) ANALYSIS

Plasmid reconstruction from WGS data was performed by PLACNETw method (Vielva *et al.*, 2017). In PLACNETw representation, most plasmids and integrative conjugative elements (ICEs) can be recognised by their replication initiation proteins (RIP) and/or RELaxase proteins (REL), both of them considered as MGEs markers.

Incompatibility groups (Inc) and pMLST subtypes were *in silico* determined with PlasmidFinder and pMLST (Carattoli *et al.*, 2014), respectively, via the CGE online services (http://genomicepidemiology.org/). Additionally, reconstructed plasmids were sub-typed according to the exact relaxase family subtype by phylogenetic comparison with previously defined relaxases subfamilies (Alvarado *et al.*, 2012). Reconstructed plasmids and references belonging to the same Inc groups were compared by using BRIG (Alikhan *et al.*, 2011) and EasyFig tools, a comparative genomic tool that allows for visualizing homologies and similarities between contigs using BLAST (Sullivan *et al.*, 2011b).

All reference plasmids used in this study were recovered from NCBI database (https://www.ncbi.nlm.nih.gov/pubmed) (annex 1, Table S3). Finally, ICEs typing from the ST372 genomes was complemented by *in silico* analysing the ICE-harbouring contigs with ICEberg (ICEfinder and VRprofile) (annex 1, Table S2). and comparing them with Easyfig (Sullivan *et al.*, 2011a).

3.6. STATISTICAL ANALYSIS

Analysis of variance (ANOVA) test was performed to compare the mean number of virulence genes and the association of variables with biofilm formation phenotypes. Normality test was performed (Shapiro-Wilk test). Dichotomous variables were described using enumeration and percentages, and compared using two-tailed Fisher's exact test. *P*-values < 0.05 were considered statistically significant. Correlation between quantitative variables was assessed by Pearson correlation coefficient, and the assessment of repeatability was via the Bland and Altman, (1986) statistical method. Additionally, non-parametric Mann-Whitney-Wilcoxon test on paired data was conducted. All analyses were carried out by XLSTAT statistical software (https://www.xlstat.com).

4. RESULTS

4.1. Study 1. 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

This study was presented in part at the 28th European Congress of Clinical Microbiology and Infectious Diseases, Madrid, 21-24 April 2018 and published in Antibiotics the 4th of April of 2020 https://doi.org/10.3390/antibiotics9040161.

4.1.1. Phylogenetic Groups

The most frequent phylogenetic group in both hospitals was B2 (48%-Spain vs. 58.3%-France) followed by the other six phylogenetic groups: A (14% vs. 15.6%), B1 (10% vs. 8.3%), C (11% vs. 4.2%), D (9% vs. 5.2%), E (5% vs. 5.2%) and F (3% vs. 3.1%). Although we observed a higher prevalence of B2 isolates in the French hospital and C isolates in the Spanish hospital, the differences were not statistically significant (annex 2, Table S4 and Figure 4).

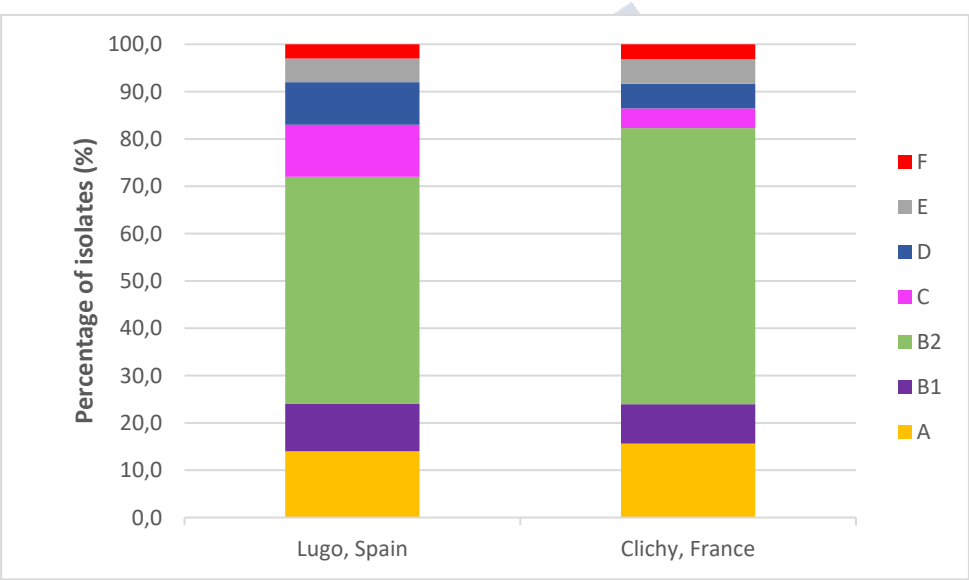


Figure 4. Comparison of the distribution of phylogenetic groups in the two hospitals.

4.1.2. Serotypes and Sequence Types

Forty O serogroups were found, but 118 (60.2%) of the 196 isolates belonged to one of the following eight serogroups: O1 (3.6%), O2 (11.2%), O4 (6.6%), O6 (10.2%), O8 (7.1%), O9 (5.1%), O18 (5.1%) and O25 (11.2%) (Figure 5). The isolates expressed 21 different H antigens, but 119 (60.7%) isolates showed only seven types of flagellar antigens: H1 (11.2%), H4 (23.5%), H5 (4.1%), H6 (7.7%), H7 (6.1%), H18 (5.1%) and H31 (3.1%) (Table 16).

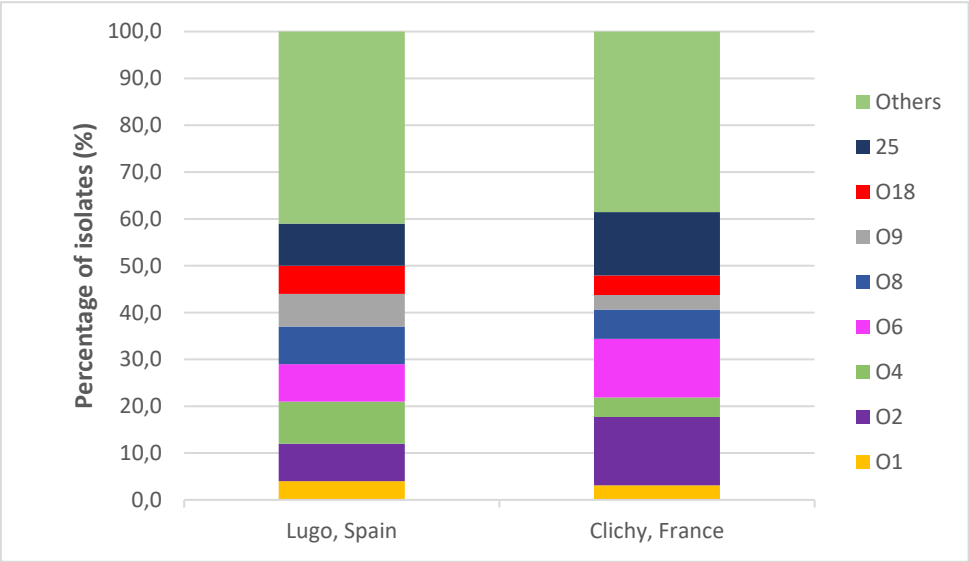


Figure 5. Comparison of the distribution of O serogroups in the two hospitals.

Seventy-one STs were found among the 196 studied isolates, but 61% and 63.5% of Spanish and French isolates, respectively, belonged to one of the ten following STs: ST10 (3% vs. 6.3%), ST12 (4% vs. 4.2%), ST69 (8% vs. 5.2%), ST58 (4% vs. 1%), ST73 (9% vs. 12.5%), ST88 (9% vs. 3.1%), ST95 (6% vs. 4.2%), ST127 (3% vs. 4.2%), ST131 (12% vs. 11.5%) and ST141 (3% vs. 11.5%). Statistically significant differences between the two hospitals were observed only with respect to ST141 (P = 0.026) (annex 2, Table S4 and Figure 6).

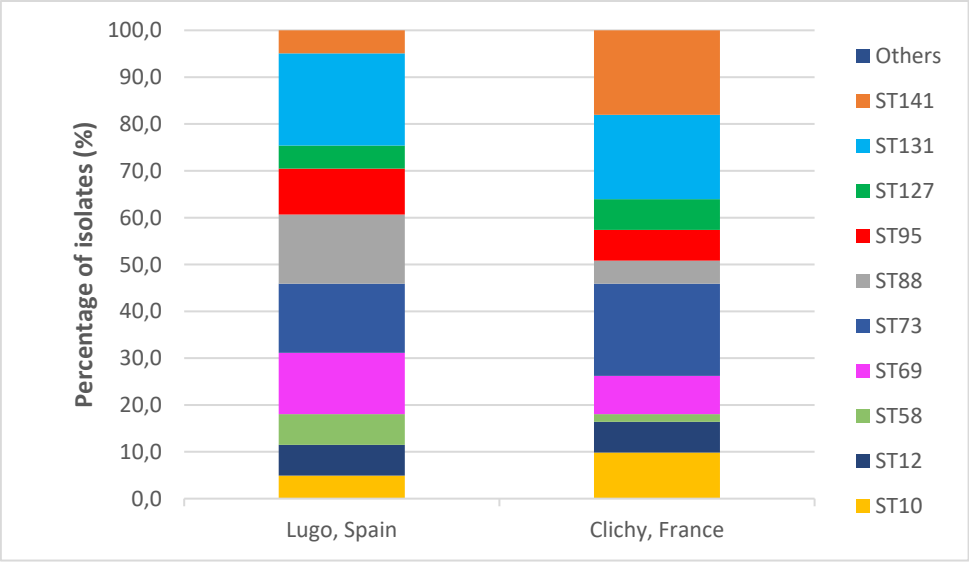


Figure 6. Comparison of the distribution of sequence types in the two hospitals.

A strong association between certain serotypes and STs was observed. Thus, the 11 isolates of serotype O2:H6 were ST141, the 12 O6:H1 isolates were ST73, 5 of the 6 O6:HNM isolates were ST127, 5 of the 6 O8:H4 isolates were ST88, 3 of the 4 O9:H4 isolates were ST88, and 17 of the 21 O25:H4 isolates were ST131 (Table 16).

Table 16. Clones, serotypes, and extraintestinal pathogenic *E. coli* (ExPEC), uropathogenic *E. coli* (UPEC), and multidrug resistant (MDR) status of the 196 isolates.

Clone (number of isolates from Spain and France)[a,b]	Serotype (number of isolates)	ExPEC status	UPEC status	MDR status
A-CH11-54-ST10 (1/2)	O48:HNM (1), O133:H40 (1), O153:HNT (1)	1	0	0
B1-CH4-27-ST58 (2/1)	O8:HNM (1), O9:H4 (1), O9:H25 (1)	0	0	1
B2-CH13-27-ST12 (1/1)	O4:H5 (1), O4:HNM (1)	2	2	0
B2-CH13-106-ST12 (2/2)	O4:H1 (3), O4:HNM (1)	4	4	0
B2-CH24-10-ST73 (2/4)	O2:H1 (1), O2:HNM (1), O6:H1 (3), O6:HNM (1)	6	6	1
B2-CH24-12-ST73 (2/1)	O6:H1 (1), O25:H1 (1), O25:H4 (1)	3	3	0
B2-CH24-30-ST73 (0/3)	O6:H1 (3)	3	3	1
B2-CH24-103-ST73 (4/2)	O6:H1 (4), O22:H1 (1), O22:HNM (1)	6	6	1
B2-CH38-15-ST95 (2/2)	O2:H5 (1), O2:H7 (1), O18:H7 (2)	4	4	0
B2-CH38-27-ST95 (1/1)	O2:H4 (1), O25:H4 (1)	2	2	1
B2-CH38-41-ST95 (1/1)	O1:H7 (2)	2	2	0
B2-CH14-2-ST127 (1/2)	O6:H31 (1), O6:HNM (2)	3	3	1
B2-CH14-136-ST127 (1/1)	O6:H31 (1), O6:HNM (1)	2	2	0
B2-CH40-22-ST131 (1/2)	O25:H4 (3)	3	3	2
B2-CH40-30-ST131 (5/8)	O25:H4 (13)	8	13	12
B2-CH40-41-ST131 (5/1)	O4:H5 (2), O4:HNM (1) O12:HNM (1), ONT:H5 (2)	3	3	4
B2-CH52-5-ST141 (3/8)	O2:H6 (8), O14:HNM (1), ONT:H4 (1), ONT:H6 (1)	11	11	0
B2-CH52-14-ST141 (0/2)	O2:H6 (2)	2	2	0
B2-CH103-9-ST372 (1/1)	O18:H31 (2)	0	2	0
B2-CH14-64-ST1193 (1/1)	O75:HNM (1), ONT:HNM (1)	1	2	2
C-CH4-39-ST88 (9/1)	O8:H4 (5), O9:H4 (3), O9:HNM (1), ONT:H4 (1)	5	0	9
D-CH35-27-ST69 (8/5)	O15:H18 (2), O15:HNM (1), O18:HNM (1), O44:H18 (2), O77:HNM (1), O106:H4 (2), O175:HNM (1), ONT:H18 (3)	9	0	6
E-CH100-96-ST362 (2/1)	O7:H6 (2), O21:HNT (1)	1	0	3
E-CH37-27-ST405 (1/1)	O18:H6 (1), O102:H4 (1)	1	0	1
F-CH32-41-ST59 (3/0)	O1:H7 (3)	3	3	0
F-CH4-58-ST648 (0/2)	O25:H4 (1), ONT:H42 (1)	2	1	1

[a] Clones represented by a single isolate: A-CH11-23-ST10 (ONT:HNM) (F), A-CH11-27-ST10 (O14:HNM) (F), A-CH11-30-ST10 (O101:H4) (S), A-CH11-43-ST10 (O8:H4) (S), A-CH11-137-ST10 (O65:H10) (F), A-CH11-168-ST10 (O10:HNM) (F), A-CH11-54-ST34 (O9:H1) (S), A-CH11-54-ST44 (ONT:H49) (S), A-CH11-41-ST93 (O14:HNM) (F), A-CH11-NT-ST93 (O25:H4) (S), A-CH11-NEG-ST93 (O51:H52) (S), A-CH11-54-ST167 (O101:H21) (S), A-CH99-54-ST361 (O9:HNM) (S), A-CH4-34-ST399 (O126:H27) (F), A-CH107-233-ST401 (O21:H25) (S), A-CH7-53-ST540 (O14:HNM) (F), A-CH11-54-ST744 (ONT:H9) (S) , A-CH7-86-ST746 (O171:HNT) (S), A-CH7-0-ST1139 (O9:HNM) (F), A-CH11-25-ST1141 (O73:H4) (S), A-CH4-NEG-ST1284 (O101:HNM) (F), A-CH4-0-ST2795 (O21:HNT) (F) , A-CH11-43-ST3596 (O8:HNM) (F), A-CH11-27-ST3877 (O38:HNM) (F), A-CH7-54-ST new 1-540 Like (O9:HNM) (F), A-CH27-23-ST new 2-437 like (O98:HNM) (S), B1-CH4-25-ST17 (O4:H2) (S), B1-CH4-24-ST58 (ONT:H21) (S), B1-CH4-25-ST58 (O75:H20) (S), B1-CH41-86-ST101 (O103:H21) (S), B1-CH4-NEG-ST155 (ONT:HNM) (F), B1-CH8-31-ST210 (O155:H19) (S), B1-CH6-31-ST448 (O148:H8) (F), B1-CH6-34-ST448 (ONT:H8) (S), B1-CH6-35-ST448 (O11:H10) (S), B1-CH6-31-ST453 (O18:HNM) (F) , B1-CH4-32-ST767 (ONT:H9) (S), B1-CH6-32-ST847 (ONT:H2) (F), B1-CH4-30-ST2025 (O8:HNM) (F), B1-CH4-NT-ST2077 (ONT:H2) (F), B1-CH4-NT-ST new 3-1071 like (O8:H8) (F), B2-CH13-6-ST12 (O4:H5) (S), B2-CH13-223-ST12 (O4:H5) (F), B2-CH14-27-ST14 (O18:HNM) (S), B2-CH14-NT-ST14 (O18:HNM) (F), B2-CH24-13-ST73 (O2:H1,12) (F), B2-CH24-27-ST73 (O6:H1) (F), B2-CH24-32-ST73 (O2:H1) (S), B2-CH24-NT-ST80 (O75:HNM) (F), B2-CH38-30-ST95 (ONT:H4) (S), B2-CH38-54-ST95 (O45:H7) (S), B2-CH24-2-ST104 (O22:H1) (S), B2-CH14-180-ST127 (O6:HNM) (F), B2-CH14-224-ST127 (O6:HNM) (S), B2-CH40-298-ST131 (O25:H4) (S), B2-CH52-76-ST141 (O2:H6) (F), B2-CH43-27-ST144 (O135:H4) (F), B2-CH14-27-ST404 (O75:HNM) (S), B2-CH38-92-ST421 (O1:H7) (F), B2-CH40-22-ST428 (O106:H4) (F), B2-CH40-21-ST555 (ONT:H4) (F), B2-CH43-13-ST567 (O83:HNM) (S), B2-CH38-5-ST569 (O46:H31) (S), B2-CH108-75-

ST636 (O83:H7) (S), B2-CH38-18-ST1231 (O18:H7) (S), B2-CH319-197-ST2015 (O2:H14) (F), B2-CH40-22-ST2556 (ONT:H4) (F), B2-CH43-197-ST2558 (O2:H14) (S), B2-CH13-5-ST3352 (O4:HNM) (S), C-CH4-32-ST23 (O5:H31) (F), C-CH4-27-ST88 (O8:H19) (F), C-CH4-NT-ST88 (O8:HNM) (F), C-CH4-27-ST new 4-88-like (O8:H40) (S), C-CH4-24-ST410 (O8:H9) (S), D-CH35-47-ST106 (O77:H18) (S), E-CH26-65-ST38 (O1:H15) (F), E-CH31-54-ST57 (O114:H10) (S), E-CH26-270-ST115 (O2:H9) (S), E-CH26-27-ST963 (O2:H18) (F), E-CH26-5-ST new 5 (ONT:H18) (F), F-CH45-97-ST117 (ONT:H4) (F). [b] Bold highlights those clones that presented at least one MDR isolate. S = Spain. F = France.

4.1.3. CLONES

A total of 107 clones (defined by the association of phylogroup, clonotype and ST) were identified among the 196 isolates, with 26 of them including at least two isolates and only nine at least four isolates: B2-CH13-106-ST12 (4 isolates), B2-CH24-10-ST73 (6), B2-CH24-103-ST73 (6), B2-CH38-15-ST95 (4), B2-CH40-30-ST131 (13), B2-CH40-41-ST131 (6), B2-CH52-5-ST141 (11), C-CH4-39-ST88 (10) and D-CH35-27-ST69 (13) (Table 16). Only statistically significant differences were observed with respect to clone B2-CH52-5-ST141, which was more prevalent in the French hospital ($P = 0.018$).

4.1.4. VIRULENCE FACTOR (VF)-ENCODING GENES

Of the 31 VF-encoding genes analyzed, four (*fimH*, *fyuA*, *chuA*, *ompT*) were detected in more than 60% of the 196 isolates and 11 (*papAH*, *papC*, *papEF*, *yfcV*, *vat*, *iucD*, *iutA*, *iroN*, *traT*, *malX*, *usp*) in at least 40%. In contrast, five VF-encoding genes (*afa/draBC*, *cdtB*, *tsh*, *kpsM II-K2*, *kpsM III*) were found in less than 10% of the isolates (Table 17).

There was a strong correlation between VF-encoding gene profiles and STs. A higher mean of VF-encoding gene score was observed in the isolates belonging to the following dominant B2- phylogenetic group STs (ST12, 17.1; ST73, 17.4; ST95, 18.2; ST127, 15.4; ST131, 12.3 and ST141, 16.2), compared with the dominant STs belonging to phylogroups A (ST10, 6.3), B1 (ST58, 8.4), C (ST88, 9.8) and D (ST69, 9.4) (Table 17 and Figure 7).

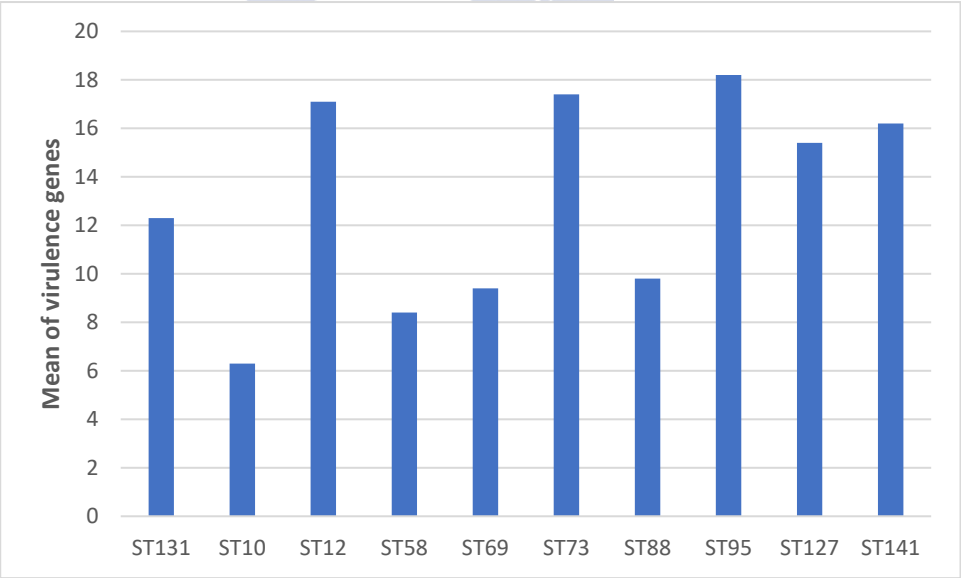


Figure 7. Mean of VF-encoding genes within dominant sequence types.

Of the 196 isolates, 61.7% were presumptively classified as ExPEC and 54.1% as UPEC. All ST12, ST73, ST95, ST127 and ST141 isolates and the majority of ST131 isolates were classified as UPEC. In contrast, none of the ST10, ST58, ST69 and ST88 isolates presented the virulence markers necessary to be classified as UPEC (Table 17).

4.1.5. Clonotypes, Clades, Subclades, Clusters and virotypes of ST131 Isolates

The 23 isolates of the dominant ST in both countries, i.e. ST131, were distributed in four clonotypes: CH40-30 (5 Spanish isolates vs. 8 French isolates), CH40-41 (5 vs. 1), CH40-22 (1 vs. 2) and CH40-298 (1 vs. 0) (annex 2, Table S4).

Isolates of clade A and non-C1-M27 subclade C1 were the most commonly detected (6 isolates for each), followed by those of subclade C2 (also known as subclone *H*30Rx) (5 isolates), clade B (4 isolates) and cluster C1-M27 (1 isolate) (annex 2, Table S4 and Figure 8).

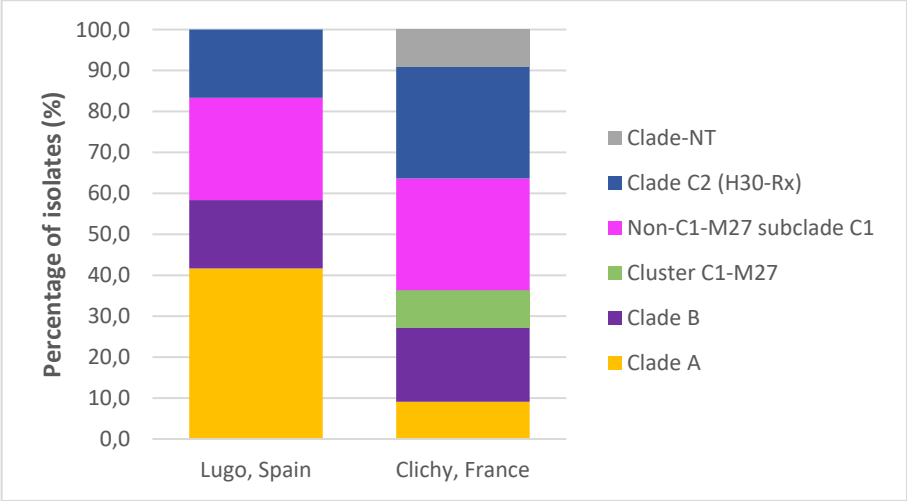


Figure 8. Comparison of the distribution of sequence type (ST131) clades and clusters in the two hospitals.

The previously determined ST131 virotypes were identified in 17 of the 23 ST131 isolates (annex 2, Table S4), among which virotypes A (3 isolates), C2 (3 isolates) and C3 (3 isolates) were most prevalent. The virotypes of six strains could not be determined, since they showed new combinations of virulence genes not included in the classification scheme used (annex 2, Table S4).

Table 17. Virulence factor (VF)-encoding genes observed from the 196 isolates and the isolates included in the 10 most frequent sequence types.

Number (%) of isolates											
VF Gene	Total (n=196)	B2 ST131 (n=23)	A ST10 (n= 9)	B2 ST12 (n=8)	B1 ST58 (n=5)	D ST69 (n=13)	B2 ST73 (n=21)	C ST88 (n=12)	B2 ST95 (n=10)	B2 ST127 (n=7)	B2 ST141 (n=14)
Adhesins											
fimH	193 (98.5)	23	9	8	5	13	21	12	10	7	14
fimAv$_{MT78}$	27 (13.8)	0	4	0	0	0	0	0	4	0	0
papAH	84 (42.9)	6	1	8	0	6	18	10	7	6	7
papC	87 (44.4)	6	1	8	0	7	18	10	7	6	7
papEF	92 (46.9)	6	1	8	0	11	18	10	7	6	7
sfa/focDE	58 (29.6)	1	0	8	0	0	20	0	2	5	13
afa/draBC	11(5.6)	7	2	0	0	0	0	0	0	0	0
yfcV	106 (54.1)	20	0	8	0	0	21	0	10	7	14
Toxins											
sat	47 (24.0)	13	3	0	0	9	12	0	0	0	0
cnf1	49 (25.0)	4	0	8	0	0	17	0	0	7	7
hlyA	55 (28.1)	3	1	8	0	0	20	0	0	7	7
hlyF	44 (22.4)	2	2	0	4	0	0	8	7	0	3
cdtB	12 (6.1)	3	0	0	0	0	6	0	2	0	0
tsh	7 (3.6)	0	0	0	0	0	0	0	2	0	1
vat	80 (40.8)	0	0	7	0	0	20	0	10	7	14
Iron uptake											
iucD	99 (50.5)	16	6	2	4	11	12	8	7	0	3
iutA	99 (50.5)	16	6	2	4	11	12	8	7	0	3
iroN	93 (47.4)	3	2	8	4	0	20	8	6	4	14
fyuA	162 (82.7)	23	7	8	4	12	21	10	10	7	14
chuA	135 (68.9)	23	0	8	0	12	21	0	10	7	14
Capsule											
kpsM II	119 (60.7)	17	2	4	0	10	20	0	10	7	14
neuC-K1	35 (17.9)	0	0	0	0	0	0	0	10	1	12
kpsM II-K2	17 (8.7)	7	0	0	0	3	0	0	0	0	0
kpsM II-K5	67 (34.2)	10	2	4	0	7	20	0	0	6	2
kpsM III	6 (3.1)	0	0	3	0	1	0	0	0	0	0
Miscellaneous											
cvaC	35 (17.9)	2	2	0	4	0	0	7	7	0	3
iss	46 (23.5)	2	2	1	4	0	2	8	5	0	5

Table 17. *Cont.*

VF Gene	Number (%) of isolates Total (n=196)	B2 ST131 (n=23)	A ST10 (n= 9)	B2 ST12 (n=8)	B1 ST58 (n=5)	D ST69 (n=13)	B2 ST73 (n=21)	C ST88 (n=12)	B2 ST95 (n=10)	B2 ST127 (n=7)	B2 ST141 (n=14)
Micellaneous											
traT	111 (56.6)	17	4	6	5	7	4	9	8	3	6
ibeA	23 (11.7)	4	0	0	0	0	0	0	4	0	3
malX	109 (55.6)	21	0	8	0	0	21	0	10	7	14
usp	107 (54.6)	23	0	8	0	0	21	0	10	7	14
ompT	151 (77.0)	23	2	8	4	12	21	10	10	7	14
ExPEC status	121 (61.7)	15	2	8	0	9	21	6	10	7	14
UPEC status	106 (54.1)	20	0	8	0	0	21	0	10	7	14
Range of VFs	1 to 23	8-20	1-14	16-19	2-10	5-12	13-20	6-13	13-22	13-17	12-23
Mean of VFs	11.6	12.3	6.3	17.1	8.4	9.4	17.4	9.8	18.2	15.4	16.2

Bold indicates the VF-encoding gene present in ≥ 60% of the isolates.

4.1.6. ANTIMICROBIAL RESISTANCE

The prevalence of resistance to ampicillin, doxycycline, nalidixic acid, ciprofloxacin and trimethoprim-sulfamethoxazole was >20%. In contrast, no isolates resistant to amikacin, colistin, fosfomycin or imipenem were detected (Table 18).

Only 13 (6.6%) isolates produced an ESBL enzyme: CTX-M-1 (4 isolates), CTX-M-14 (3), CTX-M-15 (3), CTX-M-27 (1), CTX-M-55 (1) and CTX-M-32 (1) (annex 2, Table S5). However, 73 (37.2%) of the 196 analyzed isolates were classified as MDR. MDR was especially associated with ST88 (83.3%) and ST131 (78.3%) isolates, but an important percentage of ST58 (40%) and ST69 (42.6%) isolates were also classified as MDR (Figure 9). The MDR differences between ST88 and ST131 with respect to ST10 (33.3%), ST12 (12.5%), ST73 (14.3%), ST95 (10%), ST127 (28.6%) and ST141 (0%) were statistically significant. The new global emergent MDR clonal group, i.e. ST1193 (Johnson *et al.*, 2019; Tchesnokova *et al.*, 2019), was displayed by two isolates, one from Spain and the other one from France (Table 16). Among the 73 MDR isolates, 31 (42.5%) belonged to only four clones: B2-CH40-30-ST131 (12 isolates), B2-CH40-41-ST131 (4), C-CH4-39-ST88 (9) and D-CH35-27-ST69 (6) (Table 16). Clone B2-CH40-30-ST131 was also the most prevalent among the ESBL-producing isolates (4 of 13 isolates) (annex 2, Table S5).

Fourteen (*papAH*, *papC*, *sfa/focDE*, *yfcV*, *cnf1*, *hlyA*, *vat*, *iroN*, *chuA*, *neuC-K1*, *ibeA*, *malX*, *usp* and *ompT*) of the 31 analyzed VF-encoding genes were found to be associated with isolates that did not show multidrug resistance, while only the *traT* gene was found to be associated with MDR isolates (annex 2, Table S6).

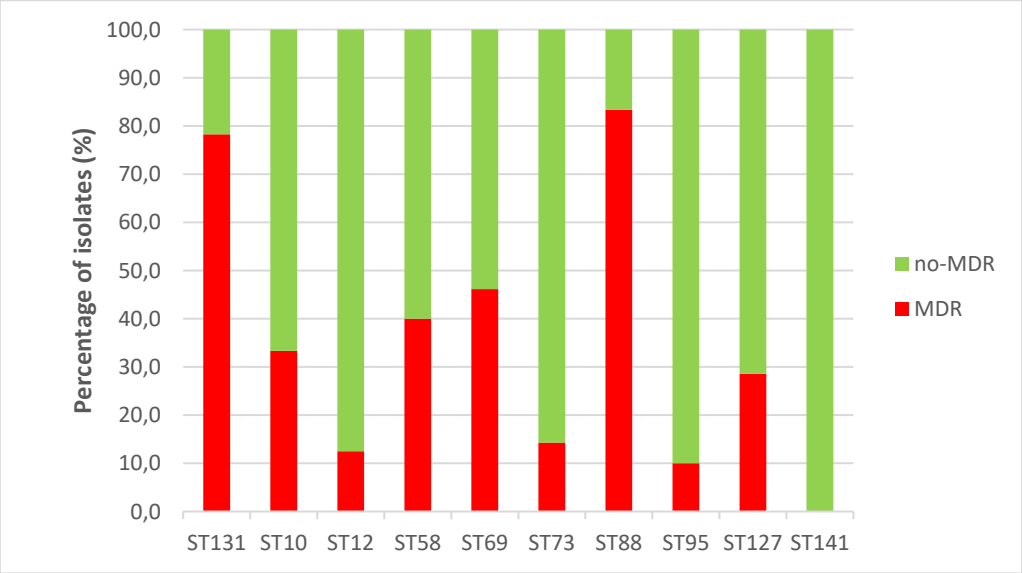


Figure 9. Multidrug resistance (MDR) within dominant sequence types.

4.1.7. HYBRID PATHOTYPES

In one of the 196 isolates that displayed serotype O153:HNT, the three (*aatA*, *aaiC* and *aggR*) genes specific for EAEC were detected. This hybrid ExPEC/EAEC isolate, which belonged to the A-CH11-54-ST10 clone, harbored 14 additional VF-encoding genes specific for ExPEC (*fimH*, fimA$_{VMT78}$, *papAH*, *papC*, *papEF*, *afa/draBC*, *sat, hlyA*, *iucD*, *iutA*, *fyuA*, *kpsM II-K5*, *traT*, *ompT*), and was resistant to ampicillin and doxycycline. None of the genes specific for the different diarrheagenic *E. coli* pathovars was detected in the remaining 195 isolates.

Table 18. Antimicrobial resistance observed from the 196 isolates and the isolates included in the 10 most frequent sequence types.

	Number (%) of resistant isolates										
Drug[ab]	Total (n=196)	B2 ST131 (n=23)	A ST10 (n=9)	B2 ST12 (n=8)	B1 ST58 (n=5)	D ST69 (n=13)	B2 ST73 (n=21)	C ST88 (n=12)	B2 ST95 (n=10)	B2 ST127 (n=7)	B2 ST141 (n=14)
Ampicillin AM10	**111 (56.6)** [b]	**19 (82.6)**	**4 (44.4)**	**7 (87.5)**	**5 (100)**	**8 (61.5)**	**6 (28.6)**	**12 (100)**	**3 (30)**	**4 (57.1)**	**3 (21.4)**
Amoxicillin-Clavulanate AMC30	27 (13.8)	**7 (30.4)**	0	**2 (25%)**	0	1 (7.7)	1 (4.8)	**9 (75)**	0	0	0
Cefazolin CZ30	27 (13.8)	**13 (56.5)**	0	0	1 (20)	0	0	**3 (25)**	0	0	0
Cefuroxime CXM30	21 (10.7)	**7 (30.4)**	1 (11.1)	0	1 (20)	0	0	**3 (25)**	0	0	0
Cefotaxime CTX30	15 (7.7)	**6 (26.1)**	0	0	1 (20)	0	0	0	0	0	0
Ceftazidime CAZ30	3 (1.5)	1 (4.3)	0	0	0	0	0	0	0	0	0
Cefepime FED30	9 (4.6)	2 (8.7)	0	0	1 (20)	0	0	1 (8.3)	0	0	0
Cefoxitin FOX30	7 (3.6)	2 (8.7)	0	0	0	0	0	1 (8.3)	0	0	0
Aztreonam ATM30	9 (4.6)	2 (8.7)	0	0	0	0	0	1 (8.3)	0	0	0
Gentamicin GM10	17 (3.6)	4 (17.4)	0	0	0	1 (7.7)	0	5 (41.7)	0	0	0
Tobramycin NN10	13 (6.6)	2 (8.7)	0	0	0	1 (7.7)	0	5 (41.7)	0	0	0
Doxycycline D30	**60 (30.6)**	**6 (26.1)**	**3 (33.3)**	1 (12.5)	**3 (60)**	**5 (38.5)**	4 (19)	**10 (83.3)**	1 (10)	**3 (42.9)**	0
Chloramphenicol C30	23 (11.7)	1 (4.3)	1 (11.1)	1 (12.5)	0	1 (7.7)	2 (9.5)	**9 (75)**	0	1 (14.3)	0
Nitrofurantoin FD300	2 (1.0)	0	0	0	0	0	0	0	0	0	0
Nalidixic Acid NAL30	**65 (33.2)**	**17 (73.9)**	**2 (22.2)**	**3 (37.5)**	**3 (60)**	**7 (53.8)**	0	**9 (75)**	2 (20)	0	0
Ciprofloxacin CIP5	**42 (21.4)**	**13 (56.5)**	1 (11.1)	0	1 (20)	0	0	**9 (75)**	1 (10)	0	0
Trimethoprim-Sulfamethoxazole SXT25	**57 (29.1)**	**9 (39.1)**	**2 (22.2)**	**2 (25%)**	**3 (60)**	**7 (53.8)**	3 (14.3)	**9 (75)**	1 (10)	1 (14.3)	1 (7.1)
Multidrug resistance	**73 (37.2)**	**18 (78.3)**	**3 (33.3)**	1 (12.5)	**2 (40)**	**6 (42.6)**	3 (14.3)	**10 (83.3)**	1 (10)	**2 (28.6)**	0
ESBL	13 (6.6)	**5 (21.7)**	0	0	1 (20.0)	0	0	0	0	0	0

[a]None of the 196 isolates were resistant to the following antimicrobials: Imipenem (IMP10), amikacin (AN30), fosfomycin (FOS200) and colistin (CL10). [b]When the percentage is greater than 20%, it is highlighted in bold.

4.2 Study 2. High Prevalence of ST131 Subclades C2-H30Rx and C1-M27 among Extended-Spectrum β-Lactamase-Producing *Escherichia coli* Causing Human Extraintestinal Infections in Patients from Two Hospitals of Spain and France during 2015

This study was presented in part at the 28th European Congress of Clinical Microbiology and Infectious Diseases, Madrid, 21-24 April 2018, at the 29th European Congress of Clinical Microbiology and Infectious Diseases, Amsterdam, 13-16 April 2019 and published in Frontiers in Cellular and Infection Microbiology the 24th of March of 2020 https://doi.org/10.3389/fcimb.2020.00125.

4.2.1. Types of ESBL Enzymes

A total of 89 (47.3%) isolates were positive for CTX-M-15, 39 (20.7%) for CTX-M-14, 27 (14.4%) for CTX-M-1, 21 (11.2%) for CTX-M-27, 11 (5.9%) for SHV-12, two (1.1%) for CTX-M-32 and one (0.5%) for CTX-M-55. CTX-M-15 was the commonest ESBL detected in the two hospitals (44.6% in Lucus Augusti and 50.0% in Beaujon). CTX-M-14 was significantly ($P = 0.0003$) more frequent in Lucus Augusti (31.5%) than in Beaujon (10.4%), whereas CTX-M-1 (20.8% versus 7.6%; $P = 0.008$) and CTX-M-27 (15.6% versus 6.5%; $P = 0.0389$) were more frequent in Beaujon than in Lucus Augusti (Table 19).

4.2.2. Phylogenetic Groups

The most frequent phylogenetic group was B2 (50.5%), followed by A (14.4%), C (11.7%), B1 (9.6%), F (6.9%), E (5.3%) and D (1.6%). Phylogenetic groups A, C and F were found more frequently among Lucus Augusti isolates, while phylogenetic group B2 was more frequent among Beaujon isolates ($P < 0.05$) (Table 19).

Among CTX-M-15 and CTX-M-27-producing isolates, the most frequent phylogenetic group was B2, while among those producing CTX-M-1, CTX-M-14 and SHV-12, the most frequent phylogenetic groups were B1, C and F respectively (Table 20).

4.2.3. Sequence Types

ST131 was detected in 33 (39.1%) and 46 (47.9%) of isolates in Lucus Augusti and Beaujon, respectively. The 109 remaining isolates displayed 57 different STs and the following STs displayed at least three isolates: ST10 (8 isolates), ST23 (3), ST38 (4), ST58 (3), ST88 (5), ST95 (4), ST167 (3), ST354 (5), ST361 (3), ST410 (6), ST648 (4), ST744 (3), and ST1615 (6). ST354, ST410 and ST1615 were significantly (P< 0.05) more frequent in Lucus Augusti (5.4%, 6.5% and 6.5%) than in Beaujon (0% for the three STs) (Table 19 and Table 20). The new emerging global clone ST1193 was identified in one isolate from France and one from Spain. The majority (121 of 188; 64.4%) of ESBLEC isolates belonged to only three clonal complexes: CC10 (19 isolates), CC23 (22 isolates) and CC131 (80 isolates) (Table S1 in Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125).

4.2.4. Clonotypes, Clades and Subclades of ST131 Isolates

The 79 ST131 isolates were distributed in three clonotypes: CH40-30 (73 isolates), CH40-35 (2), CH40-41 (4) (Table 19). Subclade C2 (also known as subclone *H*30Rx) was the commonest subclade detected among the 79 ST131 isolates (52 isolates; 65.8%), followed by cluster C1-M27 (16 isolates; 20.3%), cluster non-C1-M27 (6 isolates; 7.6%) and clade A (4 isolates; 5.1%). The 52 C2 subclade isolates were positive for CTX-M-15, whereas the 16 C1-M27 isolates were positive for CTX-M-27. Five non-C1-M27

isolates of C1 subclade were positive for CTX-M-15 and one for CTX-M-14. The four isolates belonging to clade A were positive for CTX-M-1 (2), CTX-M-15 (1) and CTX-M-27 (1).

Table 19. Characteristics of 188 ESBLEC isolated from extraintestinal infections.

Characteristic	No. (%) of isolates		
	Lucus Augusti Hospital, Spain (n= 92)	**Beaujon Hospital, France (n= 96)**	***P-Value***
ESBL enzyme			
CTX-M-1	7 (7.6)	**20 (20.8)**	0.008
CTX-M-14	**29 (31.5)**	10 (10.4)	0.0003
CTX-M-15	41 (44.6)	48 (50.0)	
CTX-M-27	6 (6.5)	**15 (15.6)**	0.0389
CTX-M-32	2 (2.2)	0	
CTX-M-55	0	1 (1.0)	
SHV-12	8 (8.7)	3 (3.1)	
Phylogenetic group			
A	**18 (19.6)**	9 (9.4)	0.0367
B1	7 (7.6)	11 (11.5)	
B2	38 (41.3)	**57 (59.4)**	0.0097
C	**16 (17.4)**	6 (6.3)	0.0151
D	1 (1.1)	2 (2.1)	
E	2 (2.2)	8 (8.3)	
F	**10 (10.9)**	3 (3.1)	0.0341
ST131			
Total	33 (39.1)	46 (47.9)	
Clade A	1 (1.1)	3 (3.1)	
Subclade C1 Cluster C1-M27	5 (5.4)	11 (11.5)	
Subclade C1 Cluster non-C1-M27	0	**6 (6.3)**	0.0164
Subclade C2	27 (29.3)	25 (26.0)	
Clade-Not typeable	0	1 (1.0)	
Clonotype CH40-30	30 (32.6)	43 (44.8)	
Clonotype CH40-35	2 (2.2)	0	
Clonotype CH40-41	1 (1.1)	3 (3.1)	
Virotype A	**7 (7.6)**	1 (1.0)	0.0280
Virotype C1	1 (1.1)	1 (1.0)	
Virotype C2	10 (10.9)	15 (15.6)	
Virotype C3	0	2 (2.1)	
Virotype E	10 (10.9)	8 (8.3)	
Virotype F	3 (3.3)	**11 (11.5)**	0.0293
Virotype A-like	1 (1.1)	6 (6.3)	
Virotype E-like	1 (1.1)	0	
Virotype-Not typeable	1 (1.1)	2 (2.1)	
Other ST[a]			
ST10	6 (6.5)	2 (2.1)	
ST23	1 (1.1)	2 (2.1)	
ST38	0	4 (4.2)	
ST58	2 (2.2)	1 (1.0)	
ST88	1 (1.1)	4 (4.2)	
ST95	0	4 (4.2)	
ST167	2 (2.2)	1 (1.0)	
ST354	**5 (5.4)**	0	0.0265
ST361	2 (2.2)	1 (1.0)	
ST410	**6 (6.5)**	0	0.0126
ST648	3 (3.3)	1 (1.0)	
ST744	1 (1.1)	2 (2.1)	
ST1615	**6 (6.5)**	0	0.0126

[a]Represented by at least 3 isolated. Significant differences are indicated in bold.

Table 20. Phylogenetic groups and sequence types according to ESBL enzymes.

ESBL enzyme (no. of isolates)	Phylogenetic groups	Sequence types
CTX-M-1 (27)	A (3) B1 (11) B2 (3) C (6) D (1) E (2) F (1)	ST10 (1), ST23 (2), ST34 (1), ST58 (1), ST69 (1), **ST88** (3), ST90 (1), ST131 (2), ST155 (1), ST162 (1), ST205 (1), ST224 (1), ST453 (2), ST1266 (1), ST1431 (1), ST2067 (1), ST2558 (1), ST2766 (1), ST3778 (1), ST8152 (1), ST new 2-6057like (1), ST new 3-1268like (1)
CTX-M-14 (39)	A (8) B1 (4) B2 (10) C (11) E (2) F (4)	**ST10** (4), ST12 (1), ST59 (1), ST73 (1), ST88 (2), ST93 (1), ST95 (2), ST131 (1), **ST167** (3), **ST354** (3), ST357 (1), ST362 (1), ST404 (2), ST405 (1), **ST410** (3), ST448 (1), ST602 (1), ST1154 (1), ST1193 (1), **ST1615** (6), ST5528 (1), ST10328 (1)
CTX-M-15 (89)	A (11) B2 (64) C (4) D (2) E (4) F (4)	ST4 (1), ST34 (1), ST38 (2), ST44 (1), ST69 (1), ST90 (1), ST95 (2), ST127 (1), **ST131** (59), ST141 (1), ST358 (1), **ST361** (3), ST405 (1), **ST410** (3), ST540 (1), **ST648** (4), ST744 (2), ST1284 (1), ST2279 (1), ST3075 (1), ST5214 (1)
CTX-M-27 (21)	B2 (18) C (1) E (2)	**ST131** (17), ST38 (2), ST90 (1), ST1193 (1)
CTX-M-32 (2)	B1 (2)	ST58 (2)
CTX-M-55 (1)	A (1)	ST744 (1)
SHV-12 (11)	A (4) B1 (1) C (1) F (5)	ST10 (2), ST23 (1), ST117 (1), ST156 (1), ST354 (2), ST1485 (1), ST3778 (1), ST new 1-10 like (2)

The most frequent STs are indicated in bold.

4.2.5. CLONES

A total of 71 clones (defined by the association of phylogroup, clonotype and ST) were identified among the 188 ESBLEC with 23 of them including at least two isolates and only five at least five isolates: A-CH11-54-ST10 (6 isolates), B2-CH40-30-ST131 (73), C-CH4-24-ST410 (6), C-CH263-32-ST1615 (6) and F-CH88-58-ST354 (5) (Table 21).

4.2.6. SEROTYPES

The 188 ESBLEC isolates belonged to 30 O serogroups and expressed 17 different H antigens, but 71 of the 79 ST131 isolates belonged to serotype O25:H4. The other prevalent serotypes were: O9:HNM (three ST361 isolates), O11:H9 (three ST1615 isolates), O16:H5 (three ST131 isolates), O18:H7 (four ST95 isolates), O20:H9 (three ST410 isolates), O75:HNM (two ST404 and two ST1193 isolates of clonal complex 14), and O86:H18 (three ST38 isolates). The H4 and H5 flagellar antigens were associated with ST131, the H7 with ST95, the H9 with ST10, ST744 and four STs of the clonal complex 23 (ST23, ST90, ST410, ST1615), the H18 with ST38 and ST69, and the H6 and the H45 antigens with ST648 (Table 21 and Table S1 in Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125).

Table 21. ESBL enzymes, serotypes and ExPEC and UPEC status according to the clones including at least two isolates.

Clones (no. of isolates from Spain and France)	ESBL enzymes	Serotypes	ExPEC	UPEC	APEC
A-CH11-54-ST10 (5/1)	CTX-M-14 (4), SHV-12 (2)	O6:HNM (1), O101:HNT (2), ONT:HNM (3)	0	0	2
A-CH11-negative-ST167 (2/1)	CTX-M-14 (3)	O101:HNM (2), O101:H25 (1)	1	0	0
A-CH11-54-ST744 (1/2)	CTX-M-15 (2), CTX-M-55 (1)	ONT:H9 (1), ONT:HNM (2)	0	0	0
A-CH99-54-ST361 (2/1)	CTX-M-15 (3)	O9:HNM (3)	0	0	0
B1-CH4-32-ST58 (1/1)	CTX-M-1 (1), CTX-M-32 (1)	O9:H25 (1), ONT:HNM (1)	1	0	2
B1-CH6-31-ST453 (0/2)	CTX-M-1 (2)	O23:HNM (1), ONT:HNM (1)	0	0	1
B2-CH38-15-ST95 (0/2)	CTX-M-14 (2)	O18:H7 (2)	2	2	2
B2-CH38-294-ST95 (0/2)	CTX-M-15 (2)	O18:H7 (2)	2	2	2
B2-CH40-30-ST131 (30/43)	CTX-M-15 (56), CTX-M-27 (16), CTX-M-14 (1)	O25:H4 (69), O14:H4 (1), ONT:H4 (3)	68	73	0
B2-CH40-35-ST131 (2/0)	CTX-M-15 (2)	O25:H4 (2)	2	2	0
B2-CH40-41-ST131 (1/3)	CTX-M-1 (2), CTX-M-15 (1), CTX-M-27 (1)	O16:H5 (3), O153:H5 (1)	3	2	0
B2-CH14-27-ST404 (1/1)	CTX-M-14 (2)	O75:HNM (2)	2	2	0
B2-CH14-64-ST1193 (1/1)	CTX-M-14 (1), CTX-M-27 (1)	O75:HNM (2)	2	2	0
C-CH4-35-ST23 (1/2)	CTX-M-1 (2), SHV-12 (1)	O55:H9 (1), O78:HNM (2)	1	0	2
C-CH4-41-ST88 (0/2)	CTX-M-1 (1), CTX-M-14 (1)	O8:HNM (1), O86:HNT (1)	2	0	0
C-CH4-303-ST88 (0/2)	CTX-M-1 (2)	O9:H17 (2)	2	0	2
C-CH4-142-ST90 (2/0)	CTX-M-1 (1) CTX-M-15 and CTX-M-27 (1)	O8:H9 (2)	0	0	0
C-CH4-24-ST410 (6/0)	CTX-M-14 (3), CTX-M-15 (3)	O9:HNM (2), O20:H9 (3), ONT:HNM (1)	0	0	0
C-CH263-32-ST1615 (6/0)	CTX-M-14 (6)	O11:H9 (3), O153:H9 (1), ONT:H9 (2)	0	0	0
E-CH26-negative-ST38 (0/3)	CTX-M-15 (1), CTX-M-27 (2)	O86:H18 (3)	3	0	0
E-CH37-27-ST405 (1/1)	CTX-M-14 (1), CTX-M-15 (1)	O102:H4 (1), O102:HNM (1)	0	0	0
F-CH88-58-ST354 (5/0)	CTX-M-14 (3), SHV-12 (2)	O1:H34 (1), O1:HNM (1), O11:H4 (2), O153:HNT (1)	1	2	0
F-CH4-171-ST648 (2/0)	CTX-M-15 (2)	O45:H45 (2)	0	0	0

4.2.7. VIRULENCE FACTOR (VF)-ENCODING GENES

Of the 188 ESBLEC isolates, 57.4% were classified as ExPEC, 52.7% as UPEC and 12.8% as APEC. The prevalence of ExPEC (92.4% vs. 32.1%) (P<0.001) and UPEC (97.5% vs. 20.2%) (P<0.001) status were higher within ST131 isolates than within non-ST131 isolates. In contrast, the prevalence of APEC (0% vs. 22%) status was higher among non-ST131 isolates (P<0.001) (Table 22).

The ST131 isolates showed a higher VF-encoding gene score (mean 13.367) compared with the non-ST131 isolates (mean 7.661) (P<0.001). However, four isolates belonging to clones B2-CH38-15-ST95 and B2-CH38-294-ST95 were those with the highest number of virulence genes (mean 21.000).

Nineteen VF-encoding genes (*papAH, papC, papEF, afa/draBC*, *yfcV*, *sat*, *cnf1*, *hlyA, iucD*, *iutA*, *fyuA*, *chuA*, *kpsM II*, *kpsM II-K2*, *kpsM II-K5*, *traT, malX*, *usp* and *ompT*) were significantly associated with ST131 isolates, whereas that ten (*fimAv*$_{MT78}$, *sfa/focDE*, *hlyF*, *tsh, vat*, *iroN*, *kpsM II-K1*, *cvaC*, *iss and ibeA*) were significantly associated with non-ST131 isolates.

The C2 subclade isolates showed a higher virulence score (mean 14.250) compared with C1-M27 isolates (mean 10.875) (P<0.001). The genes *papAH, papC, papEF, cnf1 and hlyA* were associated with the C2 subclade isolates.

The most prevalent virotypes identified in ST131 isolates were A (8 isolates), C2 (25), E (18) and F (14) and a new virotype similar to A (virotype A-like) displayed by seven isolates. Further, a second new virotype similar to E (virotype E-like) was found in one isolate (Table S1 in Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125). The virotype A was found more frequently among Lucus Augusti isolates (P = 0.0280), while virotype F was more frequent among Beaujon isolates (P = 0.0293).

Table 22. Virulence factor-encoding genes in the studied 188 isolates and according to ST131 lineage.

Genes	No. (%) of isolates					*P-Value*[a]	*P-Value*[a]
	Total (n = 188)	ST131 Cluster C1-M27 (n = 16)	ST131 Subclade C2 (n = 52)	ST131 (n = 79)	Non-ST131 (n = 109)	C1-M27 versus C2	ST131 versus non-ST131
Adhesin							
fimH	180 (95.7)	16 (100)	52 (100)	79 (100)	101 (92.7)		
fimAv$_{MT78}$	27 (14.4)	0	0	0	**27 (24.8)**		< 0.00001
papAH	50 (26.6)	0	**30 (57.7)**	**31 (39.2)**	19 (17.4)	< 0.00001	0.0008
papC	51 (27.1)	0	**30 (57.7)**	**31 (39.2)**	20 (18.3)	< 0.00001	0.0013
papEF	52 (27.7)	0	**31 (59.6)**	**32 (40.5)**	20 (18.3)	< 0.00001	0.0007
sfa/focDE	10 (5.3)	0	0	0	**10 (9.2)**		0.0036
afa/draBC	20 (10.6)	0	10 (19.2)	**15 (19.0)**	5 (4.6)		0.0018
yfcV	104 (55.3)	16 (100)	52 (100)	**77 (97.5)**	27 (24.8)		< 0.00001
Toxin							
sat	82 (43.6)	15 (93.8)	50 (96.2)	**74 (93.7)**	8 (7.3)		< 0.00001
cnf1	25 (13.3)	0	**17 (32.7)**	**18 (22.8)**	7 (6.4)	0.0049	< 0.00001
hlyA	25 (13.3)	0	**17 (32.7)**	**18 (22.8)**	7 (6.4)	0.0049	< 0.00001
hlyF	31 (16.5)	0	0	0	**31 (28.4)**		< 0.00001
cdtB	5 (2.7)	0	1 (1.9)	1 (1.3)	4 (3.7)		
tsh	12 (6.4)	0	0	0	**12 (11)**		0.0011
vat	17 (9)	0	0	0	**17 (15.6)**		0.00005
Iron uptake							
iucD	132 (70.2	15 (93.8)	50 (96.2)	**74 (93.7)**	58 (53.2)		< 0.00001
iutA	134 (71.3)	15 (93.8)	50 (96.2)	**74 (93.7)**	60 (55)		< 0.00001
iroN	32 (17)	0	0	0	**32 (29.4)**		< 0.00001
fyuA	128 (68.1)	16 (100)	52 (100)	**79 (100)**	49 (45)		< 0.00001
chuA	121 (64.4)	16 (100)	52 (100)	**79 (100)**	42 (38.5)		< 0.00001
Capsule							
kpsM II	106 (56.4)	15 (93.8)	49 (94.2)	**73 (92.4)**	33 (30.3)		< 0.00001
kpsM II-K2	13 (6.9)	0	7 (13.5)	**11 (13.9)**	2 (1.8)		0.0015
kpsM II-K5	83 (44.1)	15 (93.8)	42 (80.8)	**62 (78.5)**	21 (19.3)		< 0.00001
neuC-K1	10 (5.3)	0	0	0	**10 (9.2)**		0.0036
kpsM III	2 (1.1)	0	0	0	2 (1.8)		
Miscellaneous							
cvaC	21 (11.2)	0	0	0	**21 (19.3)**		< 0.00001
iss	23 (12.2)	0	0	0	**23 (21.1)**		< 0.00001
traT	139 (73.9)	13 (81.3)	42 (80.8)	**64 (81)**	75 (68.8)		0.0422
ibeA	12 (6.4)	0	0	0	**12 (11)**		0.0011
malX	113 (60.1)	16 (100)	52 (100)	**79 (100)**	34 (31.2)		< 0.00001
usp	101 (53.7)	16 (100)	52 (100)	**79 (100)**	22 (20.2)		< 0.00001
ompT	126 (67)	16 (100)	52 (100)	**79 (100)**	47 (43.1)		< 0.00001
ExPEC status	108 (57.4)	15 (93.8)	50 (96.2)	**73 (92.4)**	35 (32.1)		< 0.00001
UPEC status	99 (52.7)	16 (100)	52 (100)	**77 (97.5)**	22 (20.2)		< 0.00001
APEC status	24 (12.8)	0	0	**0**	**24 (22)**		< 0.00001

[a] *P-Value*s (by Fisher´s exact test) are shown where $P < 0.05$. Significant differences are indicated in bold.

4.3. STUDY 3. ASSOCIATION BETWEEN KINETICS OF EARLY BIOFILM FORMATION AND CLONAL LINEAGE IN *ESCHERICHIA COLI*

This study was presented in part at the 8ème Colloque du Réseau National Biofilms, Clermont-Ferrand, 5-6 December 2017 and published in Frontiers in Microbiology the 31th of May of 2019 https://doi.org/10.3389/fmicb.2019.01183.

4.3.1. KINETICS OF EARLY BIOFILM FORMATION

Here we study the early biofilm formation (EBF) of the two human origin bacterial collection of *E. coli* clinical isolates.

To assess the repeatability of the duplicate determination of biofilm formation performed on each of the 394 isolates at each time, we calculated Pearson's correlation coefficient and Bland and Altman, (1986) analysis between the Biofilm formation index (BFI) obtained for each determination. Pearson's correlation coefficient was high and constant over time: 0.96 at 2 h, 0.96 at 3 h, and 0.95 at 5 h (annex 2, Figure S1A), and the Bland and Altman, (1986) analysis showed that only 3.4% of values at 2 h, 5.5% at 3 h, and 6.8% at 5 h exceed the established standard deviation (annex 2, Figure S1B).

Biofilm formation index values obtained at each time point (2, 3, and 5 h) showed a distribution of both CEC and ESBLEC isolates in the three levels of biofilm production (strong, moderate, and weak production) defined according to the BFI value ranges. After 2 h of incubation, strong producers (BFI: 0–5) and moderate producers (BFI: 6–14) were more frequent among CEC isolates than ESBLEC isolates (P = 0.004 and P = 0.0189, respectively), whereas weak and non-producers (BFI: 15–20) were more frequent among ESBLEC isolates than among CEC isolates (P < 0.0001) (Table 23). After 3 h of incubation, such significant differences were still observed between CEC and ESBLEC isolates for the strong and weak producers but not for the moderate producers. After 5 h of incubation, no significant difference was observed between CEC and ESBLEC for the three types of producers.

Table 23. Distribution of CEC and ESLEC isolates according to the biofilm index (BFI) values after 2, 3 and 5h of incubation.

	Number (%) of isolates								
BFI range	**At 2h CEC n = 196**	**ESBLEC n = 198**	***P-Value*[a]**	**At 3h CEC n = 196**	**ESBLEC n = 198**	***P-Value*[a]**	**At 5h CEC n = 196**	**ESBLEC n = 198**	***P-Value*[a]**
0-5	28 (14.3)	8 (4.0)	0.0004	36 (18.4)	11 (5.6)	<0.0001	71 (36.2)	70 (35.4)	0.9164
6-14	20 (10.2)	8 (4.0)	0.0189	39 (19.9)	27 (13.6)	0.1064	41 (20.9)	45 (22.7)	0.7150
15-20	148 (75.5)	182 (91.9)	<0.0001	121 (61.7)	160 (80.8)	<0.0001	84 (42.9)	83 (41.9)	0.9188

[a] Two-tailed *P-Value*s by Fisher´s exact test.

The study of individual BFI trajectories of the isolates of the three 2-h groups over time showed that both CEC and ESBLEC isolates displayed three distinct EBF kinetics. As indicated in Figures 10A, A', most of the 28 CEC (27 non-ESBL producers) and the 8 ESBLEC strong biofilm producers at 2 h (average BFI: 1 for the two collections) remained strong producers after 3 h (average BFI: 1 and 2, respectively) and 5 h (average BFI: 4 and 3, respectively) of incubation. Accordingly, these isolates were classified into group 1 (G1) corresponding to quickly and persistently strong biofilm producers. As indicated in Figures 10B, B', the 20 CEC and 8 ESBLEC moderate producers at 2 h, with average BFI values of 11 and 13, respectively, displayed a wider variability of BFI values

at 3 and 5 h than the strong producers. However, their average BFI at 3 h (7 and 11, respectively) and at 5 h (10 and 6, respectively) led us to classify them as mostly moderate producers over the study period [group 2 (G2)]. As indicated in Figures 10C, C', the 148 CEC and 182 ESBLEC weak producers at 2 h (average BFI: 18 for CEC and ESBLEC) also displayed a wide variability of BFI values at 3 and 5 h with a notable part of isolates becoming strong producers at 5 h. However, the average BFI values were 17 for CEC and 18 for ESBLSE isolates at 3 h, and 12 for CEC and ESBLEC isolates at 5 h. Accordingly, we classified these isolates into group G3 corresponding to weak biofilm producers. We compared the individual BFI trajectories of all CEC isolates and CEC isolates over time without the 13 ESBL-producing isolates identified in this collection and found no significant difference (data not shown).

Figure 10 shows that time point 2 h was the most representative time point of the division of our *E. coli* population into three groups. Accordingly, the comparisons between G1, G2, and G3 isolates will be made at time point 2 h in the rest of the study. Furthermore, the similarity of the individual biofilm formation kinetics of the CEC and ESBLEC isolates within each group led us to characterize altogether the CEC and ESBLEC isolates of each group in the rest of the study.

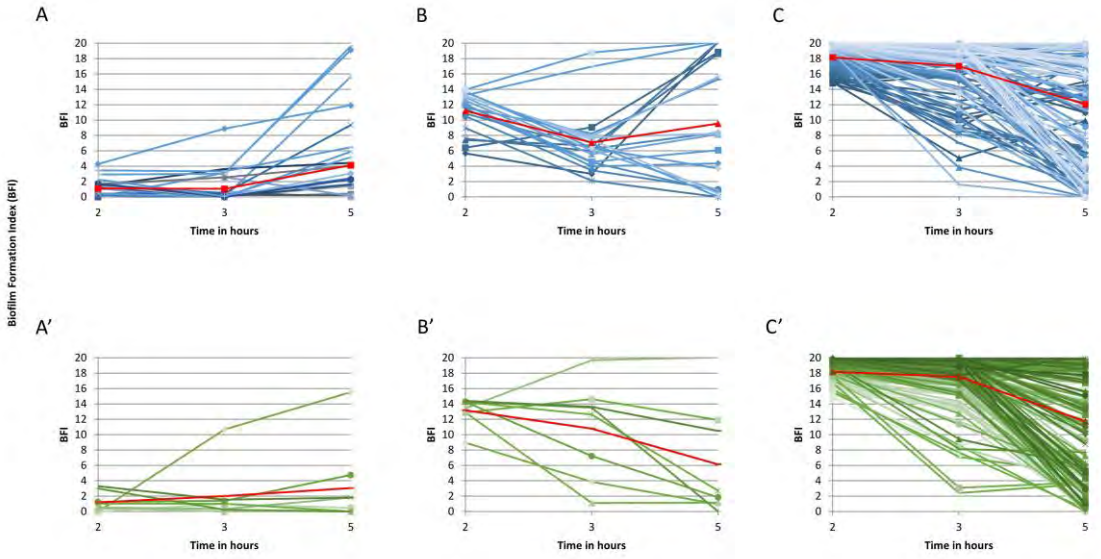


Figure 10. Individual biofilm formation trajectories over time for the 394 isolates. Individual biofilm formation index (BFI) and average BFI values at each time point (2, 3, and 5 h) are represented as blue lines and a red connected line, respectively, for CEC isolates (A–C) and as green lines and a red connected line, respectively, for ESBLEC isolates (A'–C'). (A,A') present the biofilm kinetics of the strong biofilm producers at 2 h (BFI≤5), (B,B'), the biofilm kinetics of the moderate producers at 2 h (BFI>5–≤14), and (C,C'), the biofilm kinetics of the weak (BFI>14 and <20) and the non-producers (BFI = 20) at 2 h.

4.3.2. CHARACTERIZATION OF PHYLOGROUPS, SEQUENCE TYPES, AND CLONES

4.3.2.1 Phylogroups

Most of the 394 isolates (51.3%) were assigned to phylogroup B2. The remaining isolates were distributed into phylogroups A (15.0%), B1 (9.6%), C (9.6%), E (5.1%), F (5.1%), and D (4.3%). Isolates belonging to phylogroup B2 were predominant among G1, G2, and G3 isolates (Figure 11). The seven phylogroups (A, B1, B2, C, D, E, and F) were detected among G3 isolates, four (A, B1, B2, and D) among G2 isolates, and four (B1, B2, E, and F) among G1 isolates (Figure 11). Isolates belonging to phylogroup B2 were more frequent among G1 isolates (77.8%) than among G3 isolates (47.6%) (P = 0.0006) (annex 2, Table S7). By contrast, isolates belonging to phylogroups A and C were significantly more frequent among G3 isolates than among G1 isolates (A: 16.1% vs. 0%, P = 0.0098; C: 11.5% vs. 0%, P = 0.0419) (annex 2: Table S7).

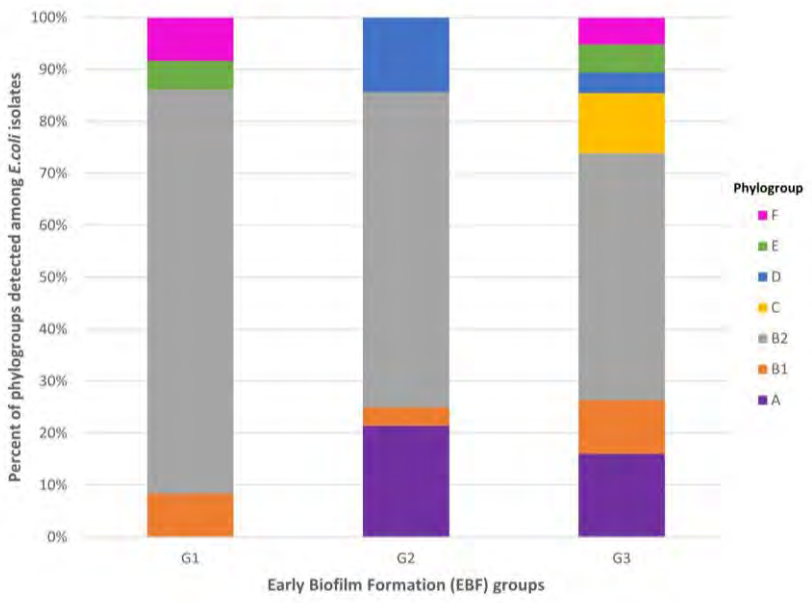


Figure 11. Distribution of phylogroups A, B1, B2, C, D, E, and F among G1, G2, and G3 isolates. G1: strong producers, G2: moderate producers, and G3: weak producers. Isolates belonging to phylogroup B2 were more frequent among G1 isolates (77.8%) than among G3 isolates (47.6%) (P = 0.0006). By contrast, isolates belonging to phylogroups A and C were significantly more frequent among G3 isolates than among G1 isolates (A: 16.1 vs. 0%, P = 0.0098; C: 11.5 vs. 0%, P = 0.0419).

4.3.2.2. Sequence Types

The 394 *E. coli* isolates displayed 99 STs. The most frequent STs accounting for more than half (52.0%) of the isolates were ST131 (26.4%), ST73 (5.6%), ST10 (4.3%), ST88 (4.3%), ST141 (4.1%), ST69 (3.8%), and ST95 (3.6%). The distribution of the seven most frequent STs among G1, G2, and G3 isolates is presented in Figure 12. These seven STs were displayed by G3 isolates whereas five were displayed by G2 isolates (absence of ST88 and ST95) and only three (ST131, ST141, and ST73) were displayed by G1 isolates. Among the three STs shared by the three groups, ST131 was significantly more frequent among G3 isolates than among G1 isolates (29.4% vs. 11.1%, $P = 0.0289$). By contrast, ST141 was more frequent among G1 isolates than among G3 (25.0% vs. 1.5%, $P < 0.0001$) (annex 2, Table S8).

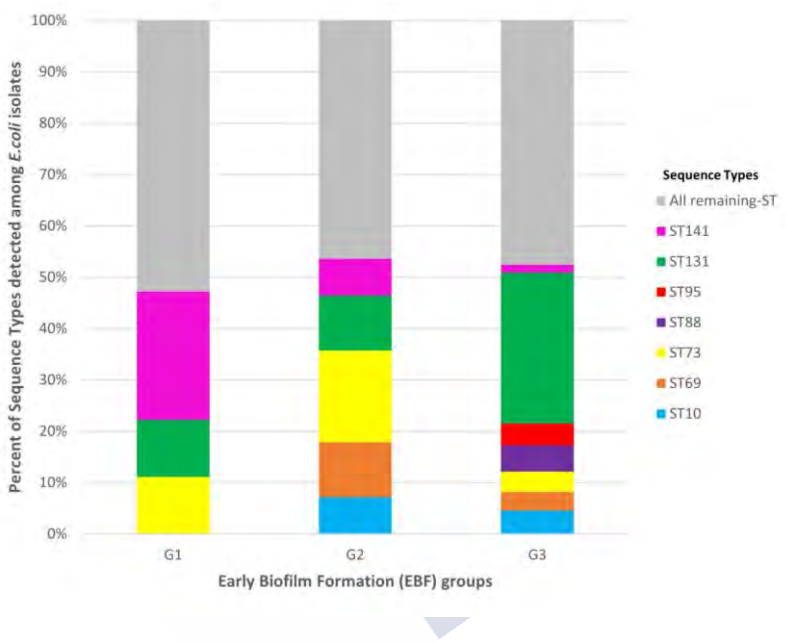


Figure 12. Distribution of the most frequent sequence types (STs) among G1, G2, and G3 isolates. G1: strong producers, G2: moderate producers, and G3: weak producers. ST141 was more frequent among G1 isolates than among G3 (25.0 vs. 1.5%, $P < 0.0001$). By contrast, ST131 was significantly more frequent among G3 isolates than among G1 isolates (29.4 vs. 11.1%, $P = 0.0289$).

4.3.2.3. Clones

A total of 153 clones were identified among the 394 isolates, with 31 of them including at least 3 isolates (Figure 13 and annex 2, Table S9) and 6 including at least 10 isolates: B2-CH40-30-ST131 (87 isolates), D-CH35-27-ST69 (14 isolates), B2-CH52-5-ST141 (13 isolates), C-CH4-39-ST88 (11 isolates), A-CH11-54-ST10 (10 isolates), and B2-CH40-41 ST131 (10 isolates). The 36 G1 isolates belonged to 26 clones (annex 2, Table S10). These 26 clones were also found in 25% of the G2 isolates ($P < 0.0001$) and in only 5.2% of the G3 isolates ($P < 0.0001$) (annex 2, Table S10).

Five clones (B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5-ST14, B2-CH52-14-ST141, and E-CH100-96-ST362) were associated with G1 and accounted for 41.7% of G1 isolates, 10.7% of G2 isolates ($P = 0.0105$), and 2.4% of G3 isolates ($P < 0.0001$) (annex 2, Table S9). Only the B2-CH40-30-ST131 clone was associated with G3, accounting for 25.5% of G3 isolates, 10.7% of G2 isolates ($P = 0.1068$), and 0% of G1 isolates ($P < 0.0001$) (annex 2, Table S9).

Figure 13. Distribution of the 31 clones including at least three isolates among G1, G2, and G3. G1: strong producers, G2: moderate producers, and G3: weak producers. Five clones (B2-CH14-2-ST127, B2-CH40-22-ST131, B2-CH52-5-ST14, B2-CH52-14-ST141, and E-CH100-96-ST362) were associated with G1 and accounted for 41.7% of G1 isolates, 10.7% of G2 isolates ($P = 0.0105$), and 2.4% of G3 isolates ($P < 0.0001$). Only the B2-CH40-30-ST131 clone was associated with G3, accounting for 25.5% of G3 isolates, 10.7% of G2 isolates ($P = 0.1068$), and 0% of G1 isolates ($P < 0.0001$).

4.3.3. Comparison of the Virulence Factor (VF)-Encoding Genes Among G1, G2, and G3 Isolates

As indicated in Table 24, some VF-encoding genes were mostly observed among G1. These genes consisted of genes encoding adhesins (*papAH, papC, papEF, sfa/focDE*, and *yfcV*), genes encoding toxins (*cnf1, hlyA*, and *vat*), genes encoding two variants of group II capsule (*kpsM II-K5* and *neuC-K1*), and miscellaneous genes (*ibeA, malX*, and *ups*). Inversely, the *sat* and *traT* genes were mostly observed among G3 isolates. Concerning the genes encoding proteins involved in iron uptake, those encoding ferric aerobactin receptors (*iucD* and *iutA*) were significantly more frequent among G3 isolates than among G1 isolates, whereas it was the contrary for the *iroN* and *chuA* genes. Mean

of VF-encoding genes was significantly higher among G1 isolates than among G3 isolates (13.8 vs. 10.4, $P < 0.001$)

Table 24. Virulence factor-encoding genes displayed by G1, G2 and G3 isolates.

Gene/Status	Number (%) of isolates				*P-Value*	
	Total n=394	G1 n=36	G2 n=28	G3 n=330	G1 vs. G2	G1 vs. G3
Adhesine						
fimH	383 (97.2)	36 (100)	27 (96.4)	320 (97)		
fimAVMT78	56 (14.2)	4 (11.1)	4 (14.3)	48 (14.5)		
papAH	135 (34.3)	19 (52.8)	12 (42.9)	104 (31.5)		0.0149[b]
papC	140 (35.5)	20 (55.6)	12 (42.9)	108 (32.7)		0.0093[b]
papEF	146 (37.1)	20 (55.6)	13 (46.4)	113 (34.2)		0.0168[b]
sfa/focDE	69 (17.5)	22 (61.1)	9 (32.1)	38 (11.5)	0.0259[a]	<0.0001[b]
afa/dra BC	31 (7.9)	0 (0.0)	1 (3.6)	30 (9.1)		
yfcV	213 (54.1)	31 (86.1)	17 (60.7)	165 (50)	0.0396[a]	<0.0001[b]
Toxin						
sat	130 (33)	3 (8.3)	7 (25)	120 (36.4)		0.0006[b]
cnf1	74 (18.8)	17 (47.2)	12 (42.9)	45 (13.6)		<0.0001[b]
hlyA	80 (20.3)	17 (47.2)	13 (46.4)	50 (15.2)		<0.0001[b]
hlyF	77 (19.5)	6 (16.7)	4 (14.3)	67 (20.3)		
cdtB	17 (4.3)	4 (11.1)	0 (0.0)	13 (3.9)		
tsh	21 (5.3)	2 (5.6)	2 (7.1)	17 (5.2)		
vat	99 (25.1)	23 (63.9)	15 (53.6)	61 (18.5)		<0.0001[b]
Iron uptake						
iucD	237 (60.2)	8 (22.2)	13 (46.4)	216 (65.5)		<0.0001[b]
IutA	239 (60.7)	8 (22.2)	13 (46.4)	218 (66.1)		<0.0001[b]
iroN	129 (32.7)	25 (69.4)	14 (50)	90 (27.3)		<0.0001[b]
fyuA	295 (74.9)	29 (80.6)	21 (75)	245 (74.2)		
chuA	260 (66)	33 (91.7)	21 (75)	206 (62.4)		0.0008[b]
Capsule						
KpsM II	229 (58.1)	30 (83.3)	18 (64.3)	181 (54.8)		0.0011[b]
KpsM II-K2	31 (7.9)	0 (0.0)	1 (3.6)	30 (9.1)		
KpsM II-K5	151 (38.3)	21 (58.3)	13 (46.4)	117 (35.5)		0.0105[b]
neuC-K1	47 (11.9)	9 (25)	4 (14.3)	34 (10.3)		0.0153[b]
KpsM III	8 (2.0)	0 (0.0)	1 (3.6)	7 (2.1)		
Miscellaneous						
cvaC	58 (14.7)	6 (16.7)	3 (10.7)	49 (14.8)		
iss	72 (18.3)	8 (22.2)	3 (10.7)	61 (18.5)		
traT	256 (65)	15 (41.7)	13 (46.4)	228 (69.1)		0.0014[b]
ibeA	36 (9.1)	12 (33.3)	1 (3.6)	23 (7.0)	0.004[b]	<0.0001[b]
malX	226 (57.4)	31 (86.1)	17 (60.7)	178 (53.9)	0.0396[a]	0.0003[b]
usp	211 (53.6)	28 (77.8)	17 (60.7)	166 (50.3)		0.0024[b]
ompT	282 (71.6)	30 (83.3)	22 (78.6)	230 (69.7)		
Range	1-23	1-23	1-20	1-22		
Mean	10.8	13.8	11.8	10.4	0.001[c]	0.001[c]

[a] Two-tailed P values by Fisher´s exact test are shown where P< 0.05. [b] Significant differences after applying the Bonferroni correction because of the $P < 0.025$. [c] Analysis of variance (ANOVA test) was performed to compare the mean of virulence genes between the three groups.

4.3.4. *FIMB* ALLELES

Given that the ST131 clone with a *fimH*30 allele was the dominant clone among our *E. coli* population and that previous results had shown (I) an *IS*3-like-linked defective *fimB* gene encoding a co-factor of the regulation of the type 1 fimbriae synthesis in this clone (Totsika *et al.*, 2011) and (II) the key role played by type 1 fimbriae in promoting biofilm formation in this clone (Sarkar *et al.*, 2016), we analyzed the *fimB* gene in 127 isolates randomly selected among the 394 isolates. PCR experiments showed no *fimB* amplicon in 10 (7.9%) isolates (annex 2, Table S11) belonging to the following clones:

A-CH11-0-ST93 (*n* = 2), A-CH11-0-ST167 (*n* = 1), A-CH99-54-ST361 (*n*= 1), A-CH7-53-ST540 (*n*= 1), A-CH7-54-ST540 (*n*= 1), A-CH4-0-ST1284 (*n*= 1), B1-CH4-0-ST155 (*n*= 1), B2-CH38-92-ST421 (*n*= 1), and E-CH26-0-ST38 (*n*= 1). An expected amplicon was detected in 89 (70.1%) isolates and an amplicon with a higher size was detected in 28 (22.0%) isolates (annex 2, Table S11). Sequencing experiments showed the insertion of *IS*3-like in the *fimB* gene in these 28 (22%) isolates comprising 25 of the 26 selected CH40-30 ST131 isolates and the three selected CH24-30 ST73 isolates. Overall, an intact *fimB* gene was significantly more frequent among G1 isolates than among G3 isolates (*P*=0.03) (annex 2, Table S11).

Through *fimB* gene sequencing of the 89 isolates with an expected *fimB* amplicon, 34 *fimB* alleles were detected (annex 2, Table S12), of which 11 new variants compared with those currently registered in the GenBank database. These 34 alleles encoded 10 FimB proteins (P1 to P10) (annex 2, Table S12). Phylogenetic tree built with these 10 proteins (annex 2, Figure S2) revealed two large clusters. One of these clusters included the 14 alleles encoding FimB P1 and the other one included the eight alleles encoding FimB P2. FimB P2 and the closely related FimB P3 (annex 2, Figure S2) were associated with G1 and accounted for 69.2% of G1 isolates vs. 10.6% of G3 isolates (*P*< 0.0001). Inversely, FimB P1 was associated with G3, accounting for 69.7% of G3 isolates vs. 23.1% of G1 isolates (*P* = 0.0023) (Figure 14 and in annex 2, Table S13).

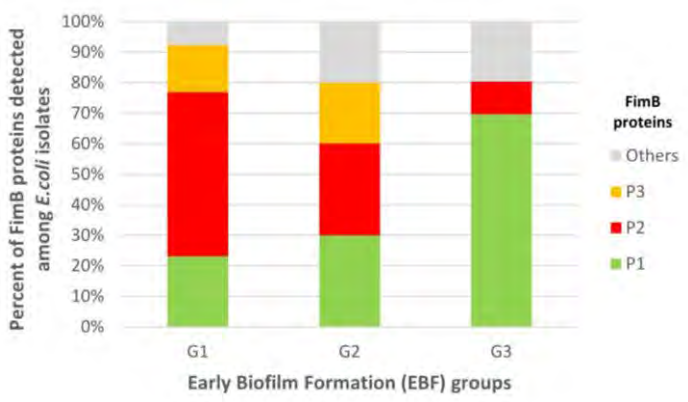


Figure 14. Distribution of the FimB protein (P) variants among G1, G2, and G3 isolates. G1: strong producers, G2: moderate producers, and G3: weak producers. FimB P2 and the closely related FimB P3 were associated with G1 and accounted for 69.2% of G1 isolates vs. 10.6% of G3 isolates (*P* < 0.0001). Inversely, FimB P1 was associated with G3 accounting for 69.7% of G3 isolates vs. 23.1% of G1 isolates (*P* = 0.0023).

4.3.5. Correlation Between Biofilm Production and Bacterial Characteristics

The one-way ANOVA test using the individual BFI of 394 isolates showed that clonal lineage explained 72% of the EBF variability and FimB, 28%. Among the VF-encoding genes significantly associated with biofilm production, the *sfa/focDE, cnf1, hlyA, vat*, and *iroN* genes were those explaining from 8 to 16% of the variability of early biofilm production. Isolate source was not associated with the EBF.

4.4. STUDY 4. WHOLE GENOME ANALYSIS OF 75 ESBLEC FROM HUMAN ORIGIN

The WGS-ESBLEC collection includes 30 genomes from sequence type complex 131 (STC131) and 45 that are not from STC131, grouped in the categories STC131 and non-STC131 respectively. Clones were defined by the combination of phylogroup, clonotype (*fumC* and *fimH* alleles) and sequence type (ST).

The STC131 group includes 21 genomes from subclade C2 distributed into 3 clones (n=19, B2-CH40-30-ST131; n=1, B2-CH40-35-ST131 and n=1, B2-CH40-5-ST2279), five genomes from clade A (B2-CH40-41-ST131) and four genomes from cluster C1M27 subclade C1 (B2-CH40-30-ST131). Besides, the non-STC131 group includes three ST10 (A-CH11-54-ST10), two ST23 (C-CH4-35-ST23), four ST38 (n=3, E-CH26-*fimH* negative-ST38 and n=1, E-CH26-65-ST38), four ST58 (n=2, B1-CH4-27-ST58 and n=2, B1-CH4-32-ST58), five ST88 (n=2, C-CH4-41-ST88; n=2, C-CH4-303-ST88 and n=1, C-CH4-39-ST88), four ST95 (n=2, B2-CH38-15-ST95 and n=2, B2-CH38-294-ST95), two ST167 (A-CH11-*fimH* negative-ST167), three ST354 (F-CH88-58-ST354), three ST361 (A-CH99-54-ST361), three ST410 (C-CH4-24-ST410), three ST648 (n=2, F-CH4-171-ST648 and n=1, F-CH4- *fimH* negative-ST648), three ST744 (A-CH11-54-ST744), three ST1193 (B2-CH14-64-ST1193) and three ST1615 (C-CH263-32-ST1615) genomes.

The serotype composition of the STC131 group includes 25 O25:H4 genomes, four O16:H5 and one ONT:H5. The serotype composition of the non-STC131 group is diverse.

4.4.2. COREGENOME (CG) AND PHYLOGENETIC ANALYSIS

The 75 genomes were taken to calculate the CG and infer the phylogeny of the STC131 and non-STC131 ESBLEC genomes (Figure 15). The CG length was 1551735+/-88 pb and comprised 1659 total of CDS (CG-1). The SNP (single nucleotide polymorphism) analysis revealed six clusters related to phylogroups. There are over 29000 SNPs distance between the different phylogroup-clusters with a maximum of 38854 SNPs. The ST38 genomes belong to phylogroup E (cluster 1) and were clustered in the first branch of the phylogenetic tree (starting from the top). Cluster 1 genomes have 0 to 4331 SNPs distance between them with a mean of 2796 SNPs. The second branch was constituted by 11 genomes from phylogroup A (cluster 2) including ST10, ST167, ST361 and ST744. Cluster 2 genomes have 2 to 12452 SNPs distance between them with a mean of 6357 SNPs.The third branch was constituted by 17 genomes, which includes four from phylogroups B1 (cluster 3) and 13 from phylogroup C (cluster 4). Cluster 3 includes the ST58 genomes and has 91 to 1066 SNPs distance between them with a mean of 625 SNPs. Cluster 4 includes ST23, ST88, ST410 and ST1615. They have 1 to 6154 SNPs distance between them with a mean of 2806 SNPs. The fourth branch clustered genomes from phylogroup F (cluster 5) including ST354 and ST648. Cluster 5 has 0 to 20601 SNPs distance between genomes with a mean of 12547 SNPs. Finally, the last branch was constituted by 37 genomes from phylogroup B2 (cluster 6) including ST95, ST1193, ST131 and ST2279 (clonal complex ST131). Cluster 6 has 0 to 16965 SNPs distance between genomes with a mean of 6037 SNPs. There are more than 12000 SNPs distance between ST95 and ST1193 with a mean of 7039 SNPs. The STC131 group of genomes was clustered together attending to ST131 clades and has 1 to 3535 SNPs distance between them with a mean of 1065 SNPs. Clade A has 1 to 37 SNPs distance between genomes. Clade C1-M27 has 8 to 20 SNPs distance between genomes and were clustered close to genomes from subclade C2 (H30Rx). Subclade C2 has 3 to 85 SNPs

distance between genomes. The distance between cluster C1-M27 and subclade C2 is 21 to 83 SNPs while between clades C1-M27 and A, the distance is 3439 to 3463 SNPs.

Subsequently, we performed a second CG (CG-2) that includes 81 STC131 genomes (Figure 16). The 30 STC131 ESBLEC group of genomes from this study plus other 51 ST131 genomes from human origin (the accession number of the mentioned 51 references are in Supplementary Table 2 in Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387). The CG length was 1994444 +/- 283 pb and comprised 2147 total of CDS. The SNP analysis revealed six clusters related to the ST131 clades with a maximum distance of 5794 SNPs between genomes. The first cluster includes 49 genomes from clade C (*fimH*30) and they have 1 to 124 SNPs distance between them with a mean of 44 SNPs. The second cluster constitutes a new clade that includes one genome from ST2279 (LREC293) with *fimH*5 allele that belongs to clonal complex 131 and has 320 to 390 SNPs distance with a mean of 336 SNPs to clade C. The third cluster includes 17 genomes from clade B (*fimH*22) and has 7 to 888 SNPs distance between them with a mean of 440 SNPs. Clade B is wide, including here seven subclusters: (I) five genomes from subcluster B0 that have 7 to 47 SNPs distance between them with a mean of 31 SNPs (the distance to the remaining genomes from cluster B is 404 to 888 with a mean of 605 SNPs), (II) two genomes from subcluster B1 that have 48 SNPs distance (the distance to the remaining genomes from cluster B is 84 to 688 with a mean of 356 SNPs), (III) two genomes from subcluster B3 that have 60 SNPs distance (the distance to the remaining genomes from cluster B is 396 to 822 with a mean of 574 SNPs), (IV) two genomes frome subcluster B4 that have 45 SNPs distance (the distance to the remaining genomes from cluster B is 670 to 888 with a mean of 771 SNPs), (V) one genome from subcluster B5 (the distance to the remaining genomes from cluster B is 264 to 828 with a mean of 472 SNPs), (VI) three genomes frome subcluster B6 that have 21 to 35 SNPs distance between them with a mean of 29 SNPs (the distance to the remaining genomes from cluster B is 164 to 687 with a mean of 421 SNPs) and (VII) two genomes frome subcluster B7 that have 23 SNPs distance (the distance to the remaining genomes from cluster B is 164 to 688 with a mean of 391 SNPs). The fourth cluster includes one genome from clade I1 (f*imH54*) and the distance to genomes from cluster B is 1153 to 1991 with a mean of 1520 SNPs. The fifth cluster constitutes a second new clade that includes one genome (LREC287) with *fimH*35 allele and has 1038 to 2464 SNPs distance with a mean of 1329 SNPs to genomes from clade C and clade B. The last cluster is constituted by 12 genomes from clade A (*fimH*41) that includes three subclusters: (I) subcluster A1 represented by one genome (SE15) that has 363 to 652 SNPs distance with a mean of 397 SNPs to the remaining genomes form cluster A, (II) a subcluster represented by one genome (LREC288) that has 317 to 652 SNPs distance with a mean of 355 SNPs to the remaining genomes form cluster A and (III) 10 genomes from subcluster A3 that have 2 to 59 SNPs distance with a mean of 39 SNPs between them.

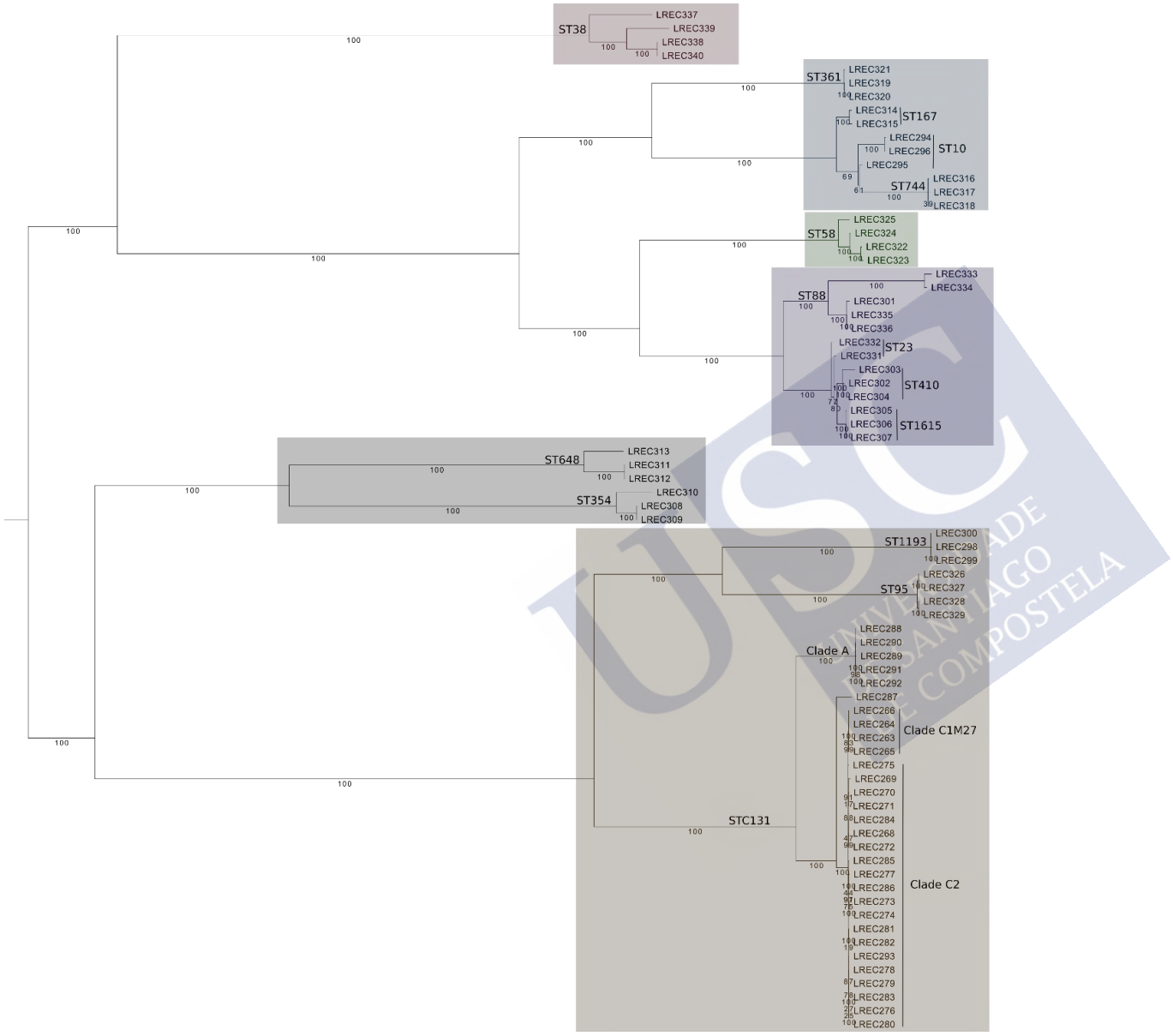


Figure 15. Coregenome of 75 ESBLEC.

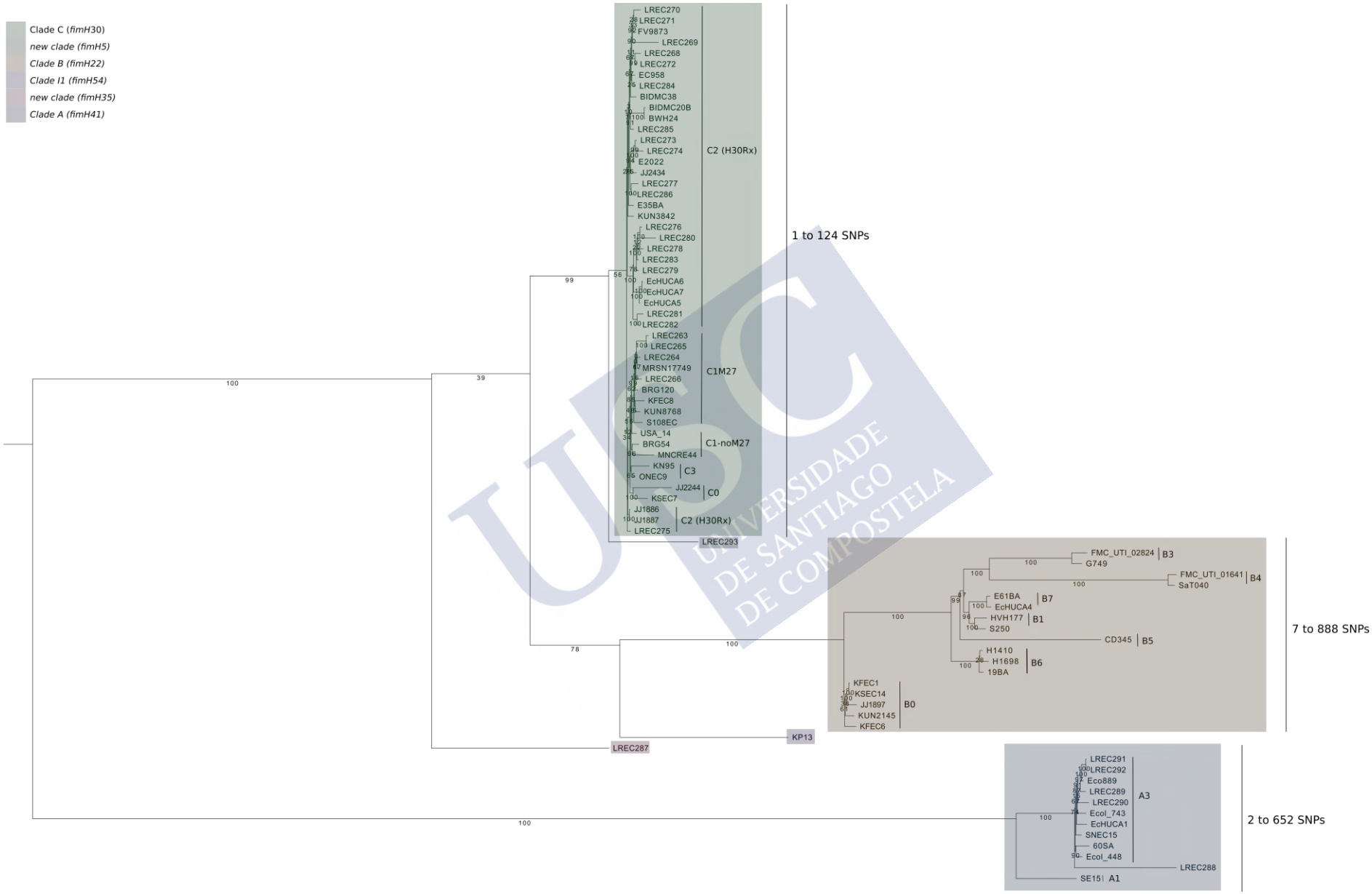


Figure 16. Coregenome of 81 Clonal Complex ST131 *Escherichia coli* genome.

4.4.3. VIRULENCE PROFILE

We *in silico* investigated the presence of 164 Virulence factor (VF)-encoding genes in the chromosome of 75 ESBLEC genomes.

Comparatively, STC131 group of genomes has a higher average of virulence genes (VGs) than the non-STC131 group of genomes (62 vs. 47) (Table 25). However, the non-STC131 group is constituted by 14 different STs. We observed that the mean of VGs is 79 in ST95; 68 in ST38 and ST1193; 59 in ST648; 58 in ST354; 50 in ST88; 44 in ST1615; 40 in ST23; 38 in ST58; 34 in ST167; 30 in ST410; 26 in ST10 and ST361 and 18 in ST744. There was a strong correlation between VF-encoding gene profiles and STs (Table 26).

Table 25. Distribution of Virulence factor (VF)-encoding genes within the total of ESBLEC and within the STC131 and non-STC131 groups.

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[a]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 *vs* NON-STC131**
Pilus, fimbrias and regulators of adhesion					
afaA	fimbrial major subunit AfaA [Afimbrial adhesin AFA-I mannose-resistant adhesin (AI010)]	16 (21%)	7 (23%)	9 (20%)	
afaB-I	chaperone AfaB [Afimbrial adhesin AFA-I mannose-resistant adhesin (AI010)]	10 (13%)	**9 (30%)**	1 (2%)	0.001
afaC-I	usher AfaC [Afimbrial adhesin AFA-I mannose-resistant adhesin (AI010)]	9 (12%)	**9 (30%)**	0 (0%)	0.001
afaD	AfaD [Afimbrial adhesin AFA-I mannose-resistant adhesin (AI010)]	18 (24%)	7 (23%)	11 (24%)	
afaE-I	AFA-I adhesin [Afa/Dr family (VF0212)]	10 (13%)	1 (3%)	9 (20%)	
afaE-II	fimbrial adhesin AfaE-II [AFA-II chloramphenicol-resistant adhesin (AI011)]	1 (1%)	0 (0%)	1 (2%)	
afaA-VIII	AfaA-VIII [Afimbrial adhesin AfaE-VIII (AI015)]	8 (11%)	0 (0%)	**8 (18%)**	0.019
afaB-VIII	chaperone protein AfaB-VIII [Afimbrial adhesin AfaE-VIII (AI015)]	8 (11%)	0 (0%)	**8 (18%)**	0.019
afaC-VIII	AfaC-VIII usher protein [Afimbrial adhesin AFA-I mannose-resistant adhesin (AI010)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
afaD-VIII	AfaD-VIII [Afimbrial adhesin AfaE-VIII (AI015)]	8 (11%)	0 (0%)	**8 (18%)**	0.019
afaE-VIII	AfaE-VIII adhesin [Afimbrial adhesin AfaE-VIII (AI015)]	8 (11%)	0 (0%)	**8 (18%)**	0.019
daaD	DaaD precursor [F1845 fimbrial adhesin (AI016)]	2 (3%)	2 (7%)	0 (0%)	
daaE	F1845 fimbrial adhesin [Afa/Dr family (VF0212)]	2 (3%)	2 (7%)	0 (0%)	
draA	cytoplamic protein DraA [Dr adhesins (VF0223)]	4 (5%)	2 (7%)	2 (4%)	
draB	chaperone protein DraB [Dr adhesins (VF0223)]	2 (3%)	0 (0%)	2 (4%)	
draC	usher protein DraC [Dr adhesins (VF0223)]	3 (4%)	0 (0%)	3 (7%)	
draP	DraP [Dr adhesins (VF0223)]	12 (16%)	**9 (30%)**	3 (7%)	0.010
afaCD/draP	Dr antigen-specific adhesin operons (ID: JN688153)	12 (16%)	**9 (30%)**	3 (7%)	0.010
air	*inverse autotransporter adhesin EaeX/Air*	3 (4%)	0 (0%)	3 (7%)	
f17d-C	F17 fimbrial uscher [F17 pili (AI037)]	3 (4%)	0 (0%)	3 (7%)	
f17d-D	F17 fimbrial chaperone [F17 pili (AI037)]	3 (4%)	0 (0%)	3 (7%)	
f17d-G	F17 fimbrial adhesin [F17 pili (AI037)]	3 (4%)	0 (0%)	3 (7%)	

Table 25. *Cont.*

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[a]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 vs. NON-STC131**
Pilus, fimbrias and regulators of adhesion					
faeE	fimbrial chaperone FaeE [K88 pili/F4 fimbriae (AI040)]	1 (1%)	0 (0%)	1 (2%)	
faeJ	fimbrial protein FaeJ [K88 pili/F4 fimbriae (AI040)]	1 (1%)	0 (0%)	1 (2%)	
fimA	Type-1 fimbrial protein A chain precursor [Type 1 fimbriae (VF0221)]	11 (15%)	1 (3%)	**10 (22%)**	0.042
fimB	Type 1 fimbriae Regulatory protein fimB [Type 1 fimbriae (VF0221)]	44 (59%)	7 (23%)	**37 (82%)**	<0.0001
fimC	Chaperone protein fimC precursor [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimD	Outer membrane usher protein fimD precursor [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimE	Type 1 fimbriae Regulatory protein fimE [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimF	FimF protein precursor [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimG	FimG protein precursor [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimH	FimH protein precursor [Type 1 fimbriae (VF0221)]	69 (92%)	30 (100%)	39 (87%)	
fimI	Fimbrin-like protein fimI precursor [Type 1 fimbriae (VF0221)]	65 (87%)	**30 (100%)**	35 (78%)	0.012
iha	bifunctional enterobactin receptor/adhesin protein	31 (41%)	**22 (73%)**	9 (20%)	<0.0001
nfaE	Diffuse adherence fibrillar adhesin gene	12 (16%)	**9 (30%)**	3 (7%)	0.010
papB	regulatory protein PapB [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papC	usher protein PapC [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papD	chaperone protein PapD [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papE	P pilus minor subunit PapE [P fimbriae (VF0220)]	2 (3%)	2 (7%)	0 (0%)	
papF	P pilus minor subunit PapF [P fimbriae (VF0220)]	11 (15%)	**10 (33%)**	1 (2%)	0.0002
papEF[c]	P pilus minor subunit (ID:X61239)	17 (23%)	10 (33%)	7 (16%)	
papG	P pilus tip adhesin PapG [P fimbriae (VF0220)]	11 (15%)	**10 (33%)**	1 (2%)	0.0002
papH	P pilus termination subunit PapH [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papI	regulatory protein PapI [P fimbriae (VF0220)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
papJ	P pilus assembly protein PapJ [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papK	P pilus minor subunit PapK [P fimbriae (VF0220)]	17 (23%)	10 (33%)	7 (16%)	
papX	PapX protein regulates flagellum synthesis to repress motility [P fimbriae (CVF425)]	33 (44%)	**28 (93%)**	5 (11%)	<0.0001
sfaA	S-fimbrial major subunit SfaA [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaB	S fimbrial switch regulatory protein SfaB [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaE	S fimbriae periplasmic chaperone SfaE [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaF	S fimbriae outer membrane usher SfaF [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	

Table 25. *Cont.*

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[a]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 vs. NON-STC131**
Pilus, fimbrias and regulators of adhesion					
sfaG	S fimbriae minor subunit SfaG [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaH	S fimbriae adhesin SfaH [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaS	S fimbriae minor subunit SfaS [S fimbriae (VF0222)]	4 (5%)	0 (0%)	4 (9%)	
sfaX	regulatory protein SfaX [S fimbriae (VF0222)]	10 (13%)	0 (0%)	**10 (22%)**	0.012
sfaY	regulatory protein SfaY [S fimbriae (VF0222)]	3 (4%)	0 (0%)	3 (7%)	
sfaDE	*Sfa* (S fimbriae) (ID: MN022635)	4 (5%)	0 (0%)	4 (9%)	
yfcV	Putative chaperone-usher fimbria (ID: CP054236.1)	40 (53%)	**27 (90%)**	13 (29%)	<0.0001
yagV/ecpE	*E. coli* common pilus chaperone EcpE [ECP (VF0404)]	73 (97%)	29 (97%)	44 (98%)	
yagW/ecpD	polymerized tip adhesin of ECP fibers [ECP (VF0404)]	71 (95%)	29 (97%)	42 (93%)	
yagX/ecpC	*E. coli* common pilus usher EcpC [ECP (VF0404)]	74 (99%)	29 (97%)	45 (100%)	
yagY/ecpB	*E. coli* common pilus chaperone EcpB [ECP (VF0404)]	74 (99%)	29 (97%)	45 (100%)	
yagZ/ecpA	*E. coli* common pilus structural subunit EcpA [ECP (VF0404)]	73 (97%)	28 (93%)	45 (100%)	
ykgK/ecpR	regulator protein EcpR [ECP (VF0404)]	72 (96%)	29 (97%)	43 (96%)	
Transport and secretion systems					
espL1	Type III secretion system effector espL1 [LEE encoded T3SS (SS020)]	23 (31%)	0 (0%)	**23 (51%)**	<0.0001
espL4	Type III secretion system effector EspL4 [LEE encoded T3SS (SS020)]	11 (15%)	0 (0%)	**11 (24%)**	0.005
espP	autotransporter serine protease [EspP (VF0208)]	1 (1%)	0 (0%)	1 (2%)	
espR1	Type III secretion system effector espR1 [LEE encoded T3SS (SS020)]	10 (13%)	0 (0%)	**10 (22%)**	0.012
espR3	Type III secretion system effector espR3 [LEE encoded T3SS (SS020)]	1 (1%)	0 (0%)	1 (2%)	
espX1	Type III secretion system effector EspX1 [LEE encoded T3SS (SS020)]	16 (21%)	0 (0%)	**16 (36%)**	<0.0001
espX4	Type III secretion system effector EspX4 [LEE encoded T3SS (SS020)]	9 (12%)	0 (0%)	**9 (20%)**	0.009
espX5	Type III secretion system effector EspX5 [LEE encoded T3SS (SS020)]	27 (36%)	0 (0%)	**27 (60%)**	<0.0001
espY1	Type III secretion system effector EspY1 [LEE encoded T3SS (SS020)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
espY2	Type III secretion system effector EspY2 [LEE encoded T3SS (SS020)]	9 (12%)	0 (0%)	**9 (20%)**	0.009
espY3	Type III secretion system effector EspY3 [LEE encoded T3SS (SS020)]	4 (5%)	0 (0%)	4 (9%)	
espY4	Type III secretion system effector EspY4 [LEE encoded T3SS (SS020)]	4 (5%)	0 (0%)	4 (9%)	
gspC	general secretion pathway protein C [T2SS (VF0333)]	55 (73%)	18 (60%)	37 (82%)	
gspD	general secretion pathway protein D [T2SS (VF0333)]	54 (72%)	17 (57%)	**37 (82%)**	0.020
gspE	general secretion pathway protein E [T2SS (VF0333)]	53 (71%)	17 (57%)	**36 (80%)**	0.040

Table 25. *Cont.*

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[a]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 vs. NON-STC131**
Transport and secretion systems					
gspF	general secretion pathway protein F [T2SS (VF0333)]	55 (73%)	18 (60%)	37 (82%)	
gspG	general secretion pathway protein G [T2SS (VF0333)]	57 (76%)	18 (60%)	**39 (87%)**	0.012
gspH	general secretion pathway protein H [T2SS (VF0333)]	55 (73%)	18 (60%)	37 (82%)	
gspI	general secretion pathway protein I [T2SS (VF0333)]	57 (76%)	18 (60%)	**39 (87%)**	0.012
gspJ	general secretion pathway protein J [T2SS (VF0333)]	57 (76%)	18 (60%)	**39 (87%)**	0.012
gspK	general secretion pathway protein K [T2SS (VF0333)]	60 (80%)	18 (60%)	**42 (93%)**	0.0007
gspL	general secretion pathway protein L [T2SS (VF0333)]	60 (80%)	18 (60%)	**42 (93%)**	0.0007
gspM	general secretion pathway protein M [T2SS (VF0333)]	66 (88%)	25 (83%)	41 (91%)	
mchF	ABC transporter protein MchF	11 (15%)	0 (0%)	**11 (24%)**	0.005
mchC	MchC protein	5 (7%)	0 (0%)	5 (11%)	
Toxins					
astA	heat-stable enterotoxin 1 [EAST1 (VF0216)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
cba	colicin B activity protein	1 (1%)	0 (0%)	1 (2%)	
cma	colicin M activity protein Cma	1 (1%)	0 (0%)	1 (2%)	
cnf1	cytotoxic necrotizing factor 1 [CNF-1 (VF0240)]	7 (9%)	**7 (23%)**	0 (0%)	0.001
hlyA	Hemolysin A [Hemolysin (VF0225)]	7 (9%)	**7 (23%)**	0 (0%)	0.001
hlyB	Hemolysin B [Hemolysin (VF0225)]	6 (8%)	**6 (20%)**	0 (0%)	0.003
hlyD	Hemolysin D [Hemolysin (VF0225)]	7 (9%)	**7 (23%)**	0 (0%)	0.001
sat	secreted autotranspoter toxin (ID: AF289092)	28 (37%)	**23 (77%)**	5 (11%)	<0.0001
senB	enterotoxin [ShET2 (VF0258)]	1 (1%)	0 (0%)	1 (2%)	
tsh	Temperature-sensitive hemagglutinin Tsh autotransporter-Serine protease	1 (1%)	0 (0%)	1 (2%)	
vat	Vacuolating autotransporter toxin. Serine protease (ID: KR094957.1)	5 (7%)	0 (0%)	5 (11%)	
Capsule[b,c]					
kpsM	KpsM subunit [group II capsule (VF0239)]	40 (53%)	**25 (83%)**	15 (33%)	<0.0001
kpsM-K1	K1 type of groupe II capsules	7 (9%)	0 (0%)	**7 (16%)**	0.037
kpsM-K2	K2 type of groupe II capsules	17 (23%)	**16 (53%)**	1 (2%)	<0.0001
kpsM-K5	K5 type of groupe II capsules	16 (21%)	9 (30%)	7 (16%)	

Table 25. *Cont.*

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[a]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 vs. NON-STC131**
Iron uptake					
ireA	TonB-dependent siderophore receptor IreA	7 (9%)	0 (0%)	**7 (16%)**	0.037
iroB	glucosyltransferase IroB [Salmochelin (IA013)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
iroC	ATP binding cassette transporter [Salmochelin (IA013)]	4 (5%)	0 (0%)	4 (9%)	
iroD	esterase [Salmochelin (IA013)]	6 (8%)	0 (0%)	6 (13%)	
iroE	esterase [Salmochelin (IA013)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
iroN	salmochelin receptor IroN [IroN (VF0230)]	3 (4%)	0 (0%)	3 (7%)	
iucA	aerobactin synthesis protein IucA [Aerobactin (VF0123)]	32 (43%)	**23 (77%)**	9 (20%)	<0.0001
iucB	aerobactin synthesis protein IucB [Aerobactin (VF0123)]	34 (45%)	**24 (80%)**	10 (22%)	<0.0001
iucC	aerobactin siderophore biosynthesis protein IucC [Aerobactin (VF0229)]	33 (44%)	**23 (77%)**	10 (22%)	<0.0001
iucD	L-lysine 6-monooxygenase IucD [Aerobactin (VF0229)]	28 (37%)	**18 (60%)**	10 (22%)	0.001
iutA[c]*	ferric aerobactin receptor precusor IutA (ID: CP033091)	30 (40%)	**21 (70%)**	9 (20%)	<0.0001
chuA[c]	Outer membrane heme/hemoglobin receptor ChuA (ID: CP054236.1)	47 (63%)	**30 (100%)**	17 (38%)	<0.0001
chuS	heme oxygenase ChuS [Chu (VF0227)]	47 (63%)	**30 (100%)**	17 (38%)	<0.0001
chuT	periplasmic heme-binding protein ChuT [Chu (VF0227)]	43 (57%)	**30 (100%)**	13 (29%)	<0.0001
chuU	heme permease protein ChuU [Chu (VF0227)]	47 (63%)	**30 (100%)**	17 (38%)	<0.0001
chuV	ATP-binding hydrophilic protein ChuV [Chu (VF0227)]	50 (67%)	**30 (100%)**	20 (44%)	<0.0001
chuW	Putative oxygen independent coproporphyrinogen III oxidase [Chu (VF0227)]	47 (63%)	**30 (100%)**	17 (38%)	<0.0001
chuX	putative heme-binding protein ChuX [Chu (VF0227)]	37 (49%)	**30 (100%)**	7 (16%)	<0.0001
chuY	ChuY [Chu (VF0227)]	47 (63%)	**30 (100%)**	17 (38%)	<0.0001
shuA	outer membrane heme/hemoglobin receptor ShuA [Shu (VF0256)]	10 (13%)	0 (0%)	**10 (22%)**	0.012
shuT	periplasmic binding protein ShuT [Shu (VF0256)]	4 (5%)	0 (0%)	4 (9%)	
shuV	ATP-binding component of hemin transport system [Shu (VF0256)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
shuX	shu locus protein ShuX [Shu (VF0256)]	10 (13%)	0 (0%)	10 (22%)	0.012
fyuA	pesticin/yersiniabactin receptor protein [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
irp1	yersiniabactin biosynthetic protein Irp1 [Yersiniabactin (VF0136)]	53 (71%)	**30 (100%)**	23 (51%)	<0.0001
irp2	yersiniabactin biosynthetic protein Irp2 [Yersiniabactin (VF0136)]	48 (64%)	**28 (93%)**	20 (44%)	<0.0001
ybtA	transcriptional regulator YbtA [Yersiniabactin (VF0136)]	50 (67%)	**27 (90%)**	23 (51%)	0.001
ybtE	yersiniabactin siderophore biosynthetic protein [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
ybtP	lipoprotein inner membrane ABC-transporter [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001

Table 25. *Cont.*

Virulence factor (VF)-encoding genes[a]		Number (%) of isolates			*P-value*[d]
		Total (n = 75)	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**STC131 vs. NON-STC131**
Iron uptake					
ybtQ	inner membrane ABC-transporter YbtQ [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
ybtS	salicylate synthase Irp9 [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
ybtT	yersiniabactin biosynthetic protein YbtT [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
ybtU	yersiniabactin biosynthetic protein YbtU [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
ybtX	putative signal transducer [Yersiniabactin (VF0136)]	52 (69%)	**29 (97%)**	23 (51%)	<0.0001
Miscellaneous					
aslA	putative arylsulfatase [AslA (VF0238)]	55 (73%)	**30 (100%)**	25 (56%)	<0.0001
capU	putative hexosyltransferase CapU	7 (9%)	1 (3%)	6 (13%)	
gad	Glutamate decarboxylase	32 (43%)	16 (53%)	16 (36%)	
eilA	HilA/EilA family virulence transcriptional regulator	10 (13%)	0 (0%)	**10 (22%)**	0.012
ibeA	invasion protein IbeA [Ibes (VF0237)]	7 (9%)	0 (0%)	**7 (16%)**	0.037
malX	Pathogenicity-associated island marker (PAI) (CP054236.1)	43 (57%)	**30 (100%)**	13 (29%)	<0.0001
usp	Uropathogenic-specific protein (bacteriocin) (NC_017651)	40 (53%)	**30 (100%)**	10 (22%)	<0.0001
ompT	Outer membrane protein (protease) T (CP054227.1)	51 (68%)	**29 (97%)**	22 (49%)	<0.0001
Status clasification					
ExPEC	If positive for ≥2 of 5 markers, including *papAH* and/or *papC*, *sfa/focDE*, *afa/draBC*, *kpsM II* and *iutA*.	44 (59%)	**26 (87%)**	18 (40%)	<0.0001
UPEC	If positive for ≥3 of 4 markers, including *chuA*, *fyuA*, *vat* and *yfcV*.	36 (48%)	**26 (87%)**	10 (22%)	<0.0001

[a]We *in silico* investigated the presence of 164 Virulence factor (VF)-encoding genes in the chromosome of 75 ESBLEC genomes with a minimum of 90% identity and 80% of coverage. [b]Group II capsules of *E. coli* have been determined by the detection of 21 genes, including the *ksl* genes that have been associated with KpsM-II-K1 (7 genes from the UTI89 sequence, ID: CP000243), with KpsM-II-K2 (6 genes from *Escherichia coli* CFT073, ID: AE014075) and with KpsM-II-K5 (1 gene from the Nissle 1917 isolate, ID: AJ58688) plus seven common genes from group II capsules (*Kps*C, *Kps*D, *Kps*E, *Kps*F, *Kps*S, *Kps*T and *Kps*U). [c]85% identity and 60% of minimum coverage. *Detection of *iutA* by PCR cover more variants (n=58; 77% positives). [d]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$. Significant differences are indicated in bold.

Table 26. Distribution of some Virulence Factor (VF)-encoding genes within the total of ESBLEC, the STC131 group and the non-STC131 group which includes ST10, ST23, ST38, ST58, ST88, ST95, ST167, ST354, ST361, ST410, ST648, ST744, ST1193 and ST1615.

Genes	**Total (n = 75)**	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**ST10 (n=3)**	**ST23 (n = 2)**	**ST38 (n = 4)**	**ST58 (n = 4)**	**ST88 (n = 5)**	**ST95 (n = 4)**	**ST167 (n = 2)**	**ST354 (n=3)**	**ST361 (n = 3)**	**ST410 (n = 3)**	**ST648 (n = 3)**	**ST744 (n = 3)**	**ST1193 (n = 3)**	**ST1615 (n = 3)**	***P-value*[a] STC131 vs. NON-STC131**
afaA-VIII	8 (11%)	0	**8 (18%)**	0	1 (50%)	0	0	3 (60%)	0	1 (50%)	0	0	0	0	0	0	3 (100%)	0.019
afaB-VIII	8 (11%)	0	**8 (18%)**	0	1 (50%)	0	0	3 (60%)	0	1 (50%)	0	0	0	0	0	0	3 (100%)	0.019
afaC-VIII	7 (9%)	0	**7 (16%)**	0	1 (50%)	0	0	3 (60%)	0	1 (50%)	0	0	0	0	0	0	2 (67%)	0.037
afaD-VIII	8 (11%)	0	**8 (18%)**	0	1 (50%)	0	0	3 (60%)	0	1 (50%)	0	0	0	0	0	0	3 (100%)	0.019
afaE-VIII	8 (11%)	0	**8 (18%)**	0	1 (50%)	0	0	3 (60%)	0	1 (50%)	0	0	0	0	0	0	3 (100%)	0.019
astA	7 (9%)	0	**7 (16%)**	0	0	1 (25%)	0	2 (40%)	0	0	2 (67%)	0	0	0	0	0	2 (67%)	0.037
eilA	10 (13%)	0	**10 (22%)**	0	0	4 (100%)	0	0	0	0	3 (100%)	0	0	3 (100%)	0	0	0	0.012
ibeA	7 (9%)	0	**7 (16%)**	0	0	0	0	0	4 (100%)	0	3 (100%)	0	0	0	0	0	0	0.037
sat	28 (37%)	**23 (77%)**	5 (11%)	0	0	1 (25%)	0	0	0	0	0	0	0	1 (33%)	0	3 (100%)	0	<0.0001
malX	43 (57%)	**30 (100%)**	13 (29%)	0	0	0	0	0	4 (100%)	0	3 (100%)	0	0	3 (100%)	0	3 (100%)	0	<0.0001
afa/draP	12 (16%)	**9 (30%)**	3 (7%)	0	0	3 (75%)	0	0	0	0	0	0	0	0	0	0	0	0.010
iutA	30 (40%)	**21 (70%)**	9 (20%)	0	0	2 (50%)	0	0	1 (25%)	0	0	0	0	1 (33%)	0	3 (100%)	2 (67%)	<0.0001
papC	17 (23%)	10 (33%)	7 (16%)	0	0	0	1 (25%)	4 (80%)	0	1 (50%)	0	0	0	1 (33%)	0	0	0	
papH	17 (23%)	10 (33%)	7 (16%)	0	0	0	1 (25%)	4 (80%)	0	1 (50%)	0	0	0	1 (33%)	0	0	0	
sfaDE	4 (5%)	0	4 (9%)	0	0	0	0	0	4 (100%)	0	0	0	0	0	0	0	0	
Group II capsules																		
K1	7 (9%)	0	**7 (16%)**	0	0	0	0	0	4 (100%)	0	0	0	0	0	0	3 (100%)	0	0.037
K2	17 (23%)	**16 (53%)**	1 (2%)	0	0	0	0	0	0	0	1 (33%)	0	0	0	0	0	0	<0.0001
K5	16 (21%)	9 (30%)	7 (16%)	0	0	4 (100%)	0	0	0	0	0	0	0	3 (100%)	0	0	0	

Table 26. *Cont.*

Genes[a]	**Total (n = 75)**	**STC131 (n = 30)**	**non-STC131 (n = 45)**	**ST10 (n=3)**	**ST23 (n = 2)**	**ST38 (n = 4)**	**ST58 (n = 4)**	**ST88 (n = 5)**	**ST95 (n = 4)**	**ST167 (n = 2)**	**ST354 (n=3)**	**ST361 (n = 3)**	**ST410 (n = 3)**	**ST648 (n = 3)**	**ST744 (n = 3)**	**ST1193 (n = 3)**	**ST1615 (n = 3)**	***P-value*[d] STC131 *vs* NON-STC131**
chuA	47 (63%)	**30 (100%)**	17 (38%)	0	0	4 (100%)	0	0	4 (100%)	0	3 (100%)	0	0	3 (100%)	0	3 (100%)	0	<0.0001
fyuA	52 (69%)	**29 (97%)**	23 (51%)	0	1 (50%)	3 (75%)	2 (50%)	4 (80%)	4 (100%)	0	2 (67%)	0	3 (100%)	1 (33%)	0	3 (100%)	0	<0.0001
vat	5 (7%)	0	5 (11%)	0	0	0	0	0	2 (50%)	0	0	0	0	0	0	3 (100%)	0	
yfcV	40 (53%)	**27 (90%)**	13 (29%)	0	0	0	0	0	4 (100%)	0	3 (100%)	0	0	3 (100%)	0	3 (100%)	0	<0.0001
ExPEC	44 (59%)	**26 (87%)**	18 (40%)	0	0	4 (100%)	1 (25%)	4 (80%)	4 (100%)	0	1 (33%)	0	0	1 (33%)	0	3 (100%)	0	<0.0001
UPEC	36 (48%)	**26 (87%)**	10 (22%)	0	0	0	0	0	4 (100%)	0	2 (67%)	0	0	1 (33%)	0	3 (100%)	0	<0.0001
Mean of VGs	53	62	47	26	40	68	38	50	79	34	58	26	30	59	18	68	44	

[a]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$. Significant differences are indicated in bold.

4.4.4. RESISTANCE ASSOCIATED POINT MUTATIONS (PMS)

Thirteen point mutations (PMs) related to the expression of resistance to antibiotics were investigated. The STC131 group showed a higher mean of PMs than the non-STC131 group (5 vs. 3). In regards to the non-STC131 group, the higher mean of PM was observed in ST354 (mean=5). Followed by ST10, ST167, ST361, ST410, ST648, ST744 and ST1193 (mean=4), ST1615 (mean=3), ST58 and ST88 (mean=1). ST38 and ST95 did not have any of the investigated PMs (Table 27).

Table 27. Chromosomal point mutations screened in 75 ESBLEC genomes, STC131 and non-ST131 groups.

Point Mutations[a]	Amino acid change	Number (%) of isolates			*P*-value[b] STC131 vs. NON-STC131
		Total (n = 75)	STC131 (n = 30)	non-STC131 (n = 45)	
gyrA p.S83L	S -> L	59 (79%)	**29 (97%)**	30 (67%)	0.003
gyrA p.D87N	D -> N	53 (71%)	25 (83%)	28 (62%)	
parC p.S80I	S -> I	56 (75%)	25 (83%)	31 (69%)	
parC p.E84V	E -> V	30 (40%)	**25 (83%)**	5 (11%)	<.0001
parC p.A56T	A -> T	3 (4%)	0 (0%)	3 (7%)	
parC p.G78C	G -> C	1 (1%)	0 (0%)	1 (2%)	
parE p.S458A	S -> A	12 (16%)	0 (0%)	**12 (27%)**	0.002
parE p.I355T	I -> T	3 (4%)	0 (0%)	3 (7%)	
parE p.L416F	L -> F	6 (8%)	0 (0%)	6 (13%)	
parE p.I529L	I -> L	29 (39%)	**29 (97%)**	0 (0%)	<.0001
16S (rrsH) r.1192C>T		1 (1%)	1 (3%)	0 (0%)	
16S (rrsB) r.1192C>T		2 (3%)	2 (7%)	0 (0%)	
16S (rrsC) r.1519A>T		1 (1%)	0 (0%)	1 (2%)	
Mean		3	5	3	

[a]Screened with a minimum of 90% identity and 80% of coverage. [b]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$. Significant differences are indicated in bold.

We observe that S83L and D87N *gyr*A PMs and S80I and E84V *par*C PMs are the most common. These PMs can be found in some genomes of the STC131 group and also in some genomes from the non-STC131 group such as ST10, ST167, ST354, ST361, ST410, ST648, ST744, ST1193 and ST1615. In contrast, the I529L *par*E PM is exclusively related to STC131 (P=<0.0001) while A56T and G78C *par*C PMs and S458A, I355T, L416F *par*E PMs are only present in the non-STC131 group. The remaining PMs were very rare among the genomes analysed (Table 27).

4.4.5. MOBILE GENETIC ELEMENTS (MGES)

The chromosome of the 75 ESBLEC genomes harbours 21 presumptive integrative conjugative elements (ICEs). We found four genomes which presented a MOB_P codifying protein (one in clone A-CH11-54- ST10, one in clone B1-CH4-32-ST58, one in clone D-CH26-65-ST38 and one in clone D-CH26-negative-ST38), four genomes that presented a MOB_H codifying protein (in clone B2-CH40-41-ST131), eight genomes that presented a MOB_Q codifying protein (three in clone B2-CH14-64-ST1193, two in clone C-CH4-24-ST410, two in clone F-CH4-171-ST648 and one in clone B2-CH40-41-ST131), two that presented a MOB_H and a MOB_Q codifying protein (in clone D-CH26-*fimH* negative-ST38) and three that presented a RptA2 type of replication protein (in clone A-CH11-54-ST744).

4.4.5.1. Plasmidome Analysis of 75 ESBLEC genomes

Reconstruction by PLACNET allowed us to describe a presumptive plasmidome for each genome. The 75 ESBLEC genomes harbour a total of 228 plasmids. Allowing to describe a presumptive plasmidome which includes 97 large (>19 kb; mean of 164 kb) conjugative plasmids (Mpf proteins detected), 90 mobilizable plasmids (MOB proteins detected) and 41 presumptively no mobilizable infective plasmids (in which any relaxase and/or Mpf protein was detected). We identified 182 plasmids that harbour a relaxase protein and 46 in which no relaxase was detected (annex 2). Briefly, 86 of 228 plasmids were found within the 30 genomes that constituted the STC131 group. There were 34 conjugatives, 33 mobilizables and 19 presumptively no mobilizable plasmids. The remaining 142 plasmids were found within the 45 genomes that constituted the non-STC131 group. There were 63 conjugatives, 57 mobilizables and 22 presumptively no mobilizable plasmids (unpublished results).

We explore the whole plasmidome of the 75 ESBLEC genomes and found the following relaxase (MOB) families: 74 MOB_P type (41%) including 20 MOB_{P51} (11%), 19 MOB_{P12} (10%); 10 MOB_{P3} (5%), one MOB_{P11} (1%) and 24 not subtype MOB_P (13%); 68 MOB_F type (37%) including 40 MOB_{F12} (22%), 4 MOB_{F11} (2%) and 24 not subtype MOB_F (13%); 26 MOB_Q type (14%) including 12 MOB_{Q12} (7%), 9 MOB_{Qu} (5%) and 5 not subtype MOB_Q (3%); 6 MOB_{V2} (3%); 2 MOB_{H11} (1%) and 6 other minority MOB types (4%) (Table 28).

In addition, by incompatibility (Inc.) group typing the plasmidome of the 75 ESBLEC genomes include: 98 Col-like, 78 IncF (75 multireplicon plasmids with many different FAB combinations, where [F2:A1:B-] (n=7), [F36:A4:B1] (n=5), [F18:A-:B1] (n=4) and [F29:A-:B10] (n=4) were the most frequent), 17 IncQ (14 IncQ1 and 3 IncQ not subtyped), 15 IncI (2 IncI1-ST3-CC3, 1 IncI1-ST26-CC2, 1 IncI1-ST275-CC2, 1 IncI1-ST16, 1 IncI1-ST95-CC9, 1 IncI1-ST80-CC31, 1 IncI1-ST37-CC3, 5 IncI1-STnd and 2 IncI2), 13 IncX (7 IncX1, 5 IncX4 and 1 IncX not subtyped), 12 IncB/O/K/Z (3 IncK, 2 IncBO, and 7 IncB/O/K/Z not subtyped), 7 Inc13, 6 IncY, 5 IncP, 5 IncN (1 IncN-ST1, 1 IncN-ST9, 3 IncN not subtyped), 2 IncHI2-ST1 and 1 IncR. Of which, 48 Col-like, 30 IncF, 6 IncX, 4 Inc13, 4 IncN and 2 IncB/O/K/Z were found among the STC131 group of genomes and 50 Col-like, 48 IncF, 17 IncQ, 15 IncI, 10 IncB/O/K/Z, 7 IncX, 6 IncY, 5 IncP, 3 Inc13, 2 IncHI2-ST1, 1 IncN, and 1 IncR were found among the non-STC131 group of genomes. It should be considered that some plasmids harbour more than one Inc. group (this could be due to the existence of cointegrated plasmids or due to an impossible separation by manual pruning during the reconstruction of the sequenced genome).

Table 28. Distribution of relaxase proteins (MOB) within 75 ESBLEC genomes, the STC131 and non-STC131 groups.

MOB typing	Total (182 MOB/75 genomes)	STC131 (63 MOB/30 genomes)	non-STC131 (119 MOB/45 genomes)	*P-value*[a]
MOB_{H11}	2 (1%)	0 (0%)	2 (2%)	
MOB_{F}	68 (37%)	28 (44%)	40 (34%)	
MOB_{F11}	4 (2%)	2 (3%)	2 (2%)	
MOB_{F12}	40 (22%)	12 (19%)	28 (24%)	
Not subtyped MOB_{F}	24 (13%)	14 (22%)	10 (8%)	
MOB_{P}	74 (41%)	15 (24%)	**59 (50%)**	0.001
MOB_{P11}	1 (1%)	0 (0%)	1 (1%)	
MOB_{P12}	19 (10%)	1 (2%)	18 (15%)	
MOB_{P3}	10 (5%)	6 (10%)	4 (3%)	
MOB_{P51}	20 (11%)	5 (8%)	15 (13%)	
Not subtyped MOB_{P}	24 (13%)	3 (5%)	11 (9%)	
MOB_{Q}	26 (14%)	**17 (27%)**	9 (8%)	0.001
MOB_{Q12}	12 (7%)	7 (11%)	5 (4%)	
MOB_{Qu}	9 (5%)	5 (8%)	4 (3%)	
Not subtyped MOB_{Q}	5 (3%)	5 (8%)	0 (0%)	
MOB_{V2}	6 (3%)	3 (5%)	3 (3%)	
Other MOB	6 (3%)	0 (0%)	6 (3%)	

[a]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$. Significant differences are indicated in bold.

We also investigated the presence of virulence and resistance associated genes carried by plasmids. Therefore, we found the toxins codifying genes *ast*A (n=1), *cba* (n=4), *cma* (n=6), *senB* (n=9), sat (n=1) and *tsh* (n=2); the salmochelin receptor *iro*N(n=8); the increased serum survival protein *iss* (n=11); the bifunctional adhesin protein *iha* (n=4) and the transporter proteins codifying genes *mch*F (n=7) and *esp*P(n=3). Additionally, we found the beta-lactamases codifying genes $bla_{CTX\text{-}M\text{-}27}$ (n=4), $bla_{CTX\text{-}M\text{-}32}$ (n=1), $bla_{CTX\text{-}M\text{-}55}$ (n=2), $bla_{CTX\text{-}M\text{-}15}$ (n=24), $bla_{CTX\text{-}M\text{-}14}$ (n=11), $bla_{SHV\text{-}12}$ (n=2), $bla_{CTX\text{-}M\text{-}1}$ (n=4), $bla_{OXA\text{-}1}$ (n=16), $bla_{TEM\text{-}1A}$ (n=1), $bla_{TEM\text{-}1B}$ (n=29), $bla_{TEM\text{-}190}$ (n=1), $bla_{TEM\text{-}78}$ (n=1), $bla_{TEM\text{-}141}$ (n=1), $bla_{TEM\text{-}206}$ (n=1), $bla_{TEM\text{-}214}$ (n=1) and $bla_{TEM\text{-}30}$ (n=1); the aminoglycoside-resistant codifying genes *aad*A1 (n=15), *aad*A2 (n=5), *aad*A5 (n=24), *aad*A9 (n=1), *aad*A13 (n=1), *aad*A17 (n=1), *ant*(2")-Ia (n=2), *aac*(3)-IIa (n=1), *aac*(3)-IId (n=11), *aac*(3)-IV (n=3), *aph*(3')-Ia (n=14), *aph*(3")-Ib (n=29), *aph*(6)-Id (n=32), *aph*(4)-Ia (n=3) and *aac*(6')-Ib-cr (n=16); the trimethoprim-resistant codifying genes *dfr*A1 (n=6), *dfr*A5 (n=5), *dfr*A7(n=1), *dfr*A12(n=2), *dfr*A8(n=1), *dfr*A14 (n=7) and *dfr*A17(n=27); the sulfonamide-resistant codifying genes *sul*1 (n=32), *sul*2 (n=34) and *sul*3 (n=6); the macrolide-resistant codifying genes *erm*(B) (n=1), *lnu*(F) (n=1), *mdf*(A) (n=1), *mph*(A) (n=25) and *mph*(B) (n=3); the phenicol-resistant codifiying genes *cml*A1 (n=4), *flo*R (n=3), *cat*A1(n=7) and *cat*B3 (n=16); the tetracycline-resistant codifying genes *tet*(A) (n=32), *tet*(B) (n=11) and *tet*(M) (n=2) and the fosfomycin-resistant codifying gene *fos*A3 (n=1).

The IncF plasmids vary greatly in their composition and carried almost all the virulence and resistance genes found in the plasmidome of the 75 ESBLEC genomes analysed (annex 2). The ESBL resistant phenotype was a common feature in the 75 genomes that were selected for WGS and the location of the genes that confer this attribute was investigated. We retrieve the ESBL codifying enzymes in 67 of the 75

genomes sequenced. We found five *bla*CTX-M-1, nine *bla*CTX-M-14, 23 *bla*CTX-M-15, four *bla*CTX-M-27, one *bla*CTX-M-32, two *bla*CTX-M-55 and two *bla*SHV-12 carried by plasmids and three *bla*CTX-M-1, six *bla*CTX-M-14, ten *bla*CTX-M-15, one *bla*CTX-M-27 and one *bla*CTX-M-32 located in the chromosome. The contigs that harboured the ESBL enzyme codifying genes were manually revised to confirm the plasmid location.

4.4.5.2. *E. coli* ST131 clonal complex plasmidome

The plasmidome of the 30 genomes from STC131 group showed some common features related with the ST131 clades that are summarised in Table 29. The major part of the genomes (n=24, 80%) harboured Col-like and IncF plasmids either alone or with other plasmids families (annex 2).

The plasmidome harboured 16 ESBL enzymes encoding *bla* genes. Clade A harboured one *bla*CTX-M-1 carried by an IncF [F29:A-:B10] plasmid, cluster C1M27 harboured two *bla*CTX-M-27 carried by IncF [F1:A2:B20] plasmids and subclade C2 harboured 13 *bla*CTX-M-15 of witch 5 carried by IncF [F2:A1:B-] and 8 by other FAB type IncF plasmids.

Table 29. Correlation between clades, ICEs, plasmid types and CTX-M genes in the STC131 group of genomes.

Features	ST131 Clade			
	A (n=5)	C1M27 (n=4)	C2 (n=21)	STC131
MOB types				
MOB_F	6	4	18	28
MOB_P	0	5	9	14
MOB_Q	1	5	11	17
MOB_{V2}	0	0	3	3
Incompatiblity groups (Inc.) [FAB formula]				
Col-like	5	16	27	48
Inc13	0	1	3	4
IncX	0	2	4	6
IncK (IncB/O/K/Z)	0	0	2	2
IncN	0	0	4	4
IncF [F-:A-:B10]	1	0	0	1
IncF [F1:A6:B66]	0	1	0	1
IncF [F1:A2:B20]	0	2[b]	0	2
IncF [F2:A-:B-]	0	0	0	0
IncF [F2:A1:B-]	0	0	7[c]	7
IncF [F4:A2:B20]	0	1	0	1
IncF [F29:A-:B10]	4[a]	0	0	4
IncF [F36:A4:B1]	0	0	2	2
Other IncF[d]	0	0	11	11
IncFIB (H89-PhagePlasmid)	1	0	0	1
ICEs	5	0	0	5

[a]One *bla*CTX-M-1 carrier. [b]Two *bla*CTX-M-27 carriers. [c]Five *bla*CTX-M-15 carriers. [d]singel FAB representants.

4.5. STUDY 5. MOLECULAR CHARACTERISTICS OF EXTRAINTESTINAL PATHOGENIC *E. COLI* (EXPEC), UROPATHOGENIC *E. COLI* (UPEC), AND MULTIDRUG RESISTANT *E. COLI* ISOLATED FROM HEALTHY DOGS IN SPAIN. WHOLE GENOME SEQUENCING OF CANINE ST372 ISOLATES AND COMPARISON WITH HUMAN ISOLATES CAUSING EXTRAINTESTINAL INFECTIONS

This study was presented in part at the 29th European Congress of Clinical Microbiology and Infectious Diseases, Amsterdam, 13-16 April 2019 and published in Microorganisms the 31th of October of 2020 https://doi.org/10.3390/microorganisms8111712.

4.5.1. PHYLOGENETIC GROUPS OF THE 197 CANINE ISOLATES

The most common phylogenetic group displayed by the 197 canine faecal *E. coli* isolates was B2 (42.6%), followed by A (16.2%), B1 (13.2%), F (9.1%), E (7.1%), C (5.1%), and D (3.0%) (annex 2, Table S14).

4.5.2. VIRULENCE FACTOR (VF)-ENCODING GENES IN THE 197 CANINE ISOLATES

Of the 28 VF-encoding genes analysed, eight (*fimH*, *yfcV*, *vat*, *iroN*, *fyuA*, c*huA*, *malX*, and *usp*) were detected in more than 40% of the 197 canine isolates and nine (*papAH*, *papC*, *sfa/focDE*, *cnf1*, *hlyA*, *kpsM II*, *kpsM II-K5*, *traT*, *ibeA*) in at least 20%. In contrast, six VF-encoding genes (*afa/draBC*, *sat*, *cdtB*, *neuC-K1*, *kpsM II-K2*, *kpsM III*) were found in less than 10% of these isolates (Table 30).

A higher mean of VF-encoding gene score was observed in the 84 canine isolates belonging to the dominant B2-phylogenetic group (mean of 12.79) ($P < 0.05$) compared with the isolates belonging to phylogroups A (2.31), B1 (2.96), C (6.60), D (5.67), E (4.14), and F (7.94) (Table 30).

Of the 197 canine isolates, 74 (37.6%) were presumptively classified as ExPEC and 82 (41.6%) as UPEC (Table 30) resulting in 91 ExPEC and/or UPEC isolates. The majority (85.7%; 78 of 91) of ExPEC and/or UPEC isolates belonged to phylogenetic group B2. In contrast, only 5.7% (6 of 106) of non-ExPEC and non-UPEC isolates were assigned to this phylogenetic group ($P < 0.00001$). The A, B1, C, and E phylogenetic groups were significantly associated with non-ExPEC and non-UPEC isolates (annex 2, Table S15).

4.5.3. ANTIMICROBIAL RESISTANCE IN THE 197 CANINE ISOLATES

In total, 28 (14.2%) of the 197 analysed canine faecal *E. coli* isolates were classified as MDR. Multidrug resistance was significantly associated with isolates belonging to B1, D and E phylogenetic groups (annex 2, Table S16). Furthermore, only eight (28.6%) of MDR isolates showed the ExPEC and/or the UPEC status (annex 2, Table S17).

In total, 10 of the 28 MDR isolates produced an ESBL enzyme: CTX-M-1 (four isolates), CTX-M-14 (four isolates), CTX-M-55 (one isolate) and SHV12 (one isolate). Besides, 10 other isolates produced a plasmid-mediated AmpC β-lactamase of CMY-2 type.

Table 30. Virulence factor (VF)-encoding genes detected in the 197 canine *E. coli* isolates. Relationship with phylogenetic groups.

VF-Encoding Gene	Number of Isolates (%)									
	Total (n = 197)	**A (n = 32)**	**B1 (n = 26)**	**B2 (n = 84)**	**C (n = 10)**	**D (n = 6)**	**E (n = 14)**	**F (n = 18)**	**Clade V (n = 1)**	**NT [1] (n = 6)**
Adhesins										
fimH	192 (97.5)	29	24	84	10	6	14	18	1	6
fimAvMT78	20 (10.2)	12	1	1	0	0	2	3	0	1
papAH	62 (31.5)	2	3	54	0	1	1	1	0	0
papC	62 (31.5)	2	3	54	0	1	1	1	0	0
sfa/focDE	66 (33.5)	0	2	63	0	0	0	1	0	0
afa/draBC	1 (0.5)	0	0	0	0	1	0	0	0	0
yfcV	96 (48.7)	4	0	80	0	0	0	11	0	1
Toxins										
sat	3 (1.5)	0	0	2	0	0	0	1	0	0
cnf1	54 (27.4)	0	0	53	0	0	0	1	0	0
hlyA	56 (28.4)	0	2	54	0	0	0	0	0	0
cdtB	8 (4.1)	0	0	6	0	0	0	2	0	0
tsh	21 (10.7)	2	2	0	8	0	2	6	0	1
vat	79 (40.1)	1	0	76	0	0	0	2	0	0
Iron uptake										
iutA	33 (16.8)	3	3	7	8	1	2	9	0	0
iroN	93 (47.2)	3	6	66	8	0	2	8	0	0
fyuA	106 (53.8)	7	2	77	8	5	1	5	0	1
chuA	122 (61.9)	0	0	84	0	6	14	18	0	0
Capsule										
kpsM II	64 (32.5)	1	0	41	0	4	4	10	1	3
neuC-K1	12 (6.1)	0	0	10	0	0	0	2	0	0
kpsM II-K2	7 (3.6)	0	0	2	0	3	0	1	0	1
kpsM II-K5	45 (22.8)	1	0	29	0	1	4	7	1	2
kpsM III	6 (3.0)	1	0	2	0	2	1	0	0	0
Miscellaneous										
cvaC	21 (10.7)	1	4	2	8	0	0	6	0	0
iss	31 (15.7)	2	5	6	8	0	2	8	0	0
traT	59 (29.9)	2	12	17	8	2	7	10	0	1
ibeA	54 (27.4)	0	0	49	0	0	0	3	0	2
malX	93 (47.2)	0	5	79	0	0	0	7	0	2
usp	86 (43.7)	1	3	77	0	1	1	2	0	1
ExPEC status	74 (37.6)	2	2	61	0	1	1	7	0	0
UPEC status	82 (41.6)	1	0	77	0	0	0	4	0	0
Range of VFs	0 to 18	0 to 10	1 to 10	2 to 17	1 to 8	5 to 7	2 to 7	3 to 18	3	1 to 9
Mean of VFs	7.87	2.31	2.96	12.79	6.60	5.67	4.14	7.94	3.00	3.67

[1]Using the revised protocol developed by Clermont *et al.* (2013) six isolates were not typeable (NT). These six isolates belonged to phylogroup A using the first protocol developed by Clermont *et al.* (2013) that classifies isolates into only four phylogenetic groups (A, B1, B2, D).

4.5.4. Sequence Types, Clones and Serotypes Displayed by the 91 Canine ExPEC and/or UPEC Isolates and 28 MDR isolates

Sequences types (STs), clones (defined by the association of phylogroup, clonotype and ST) and O:H serotypes were established only for the 91 canine isolates classified as ExPEC and/or UPEC and the 28 canine MDR isolates

A total of 34 STs were identified in the canine ExPEC and/or UPEC isolates and 22 in the MDR isolates. Among these STs, 18 were previously undescribed (annex 2, Table S18). Each of these 18 new STs were displayed by one isolate. Seven dominant STs (ST12, ST38, ST73, ST127, ST141, ST372, and ST648) were observed among the 91

canine ExPEC and/or UPEC and the 28 canine MDR isolates. There was a strong correlation between VF-encoding gene profiles and the dominant STs (Table 31).

Table 31. Virulence factor (VF)-encoding genes detected in the 65 canine *E. coli* isolates included in the 7 most frequent sequence types identified in ExPEC, UPEC and MDR isolates.

VF-Encoding Gene	Number of Isolates						
	B2-ST12 (n = 9)	**D-ST38 (n = 4)**	**B2-ST73 (n = 4)**	**B2-ST127 (n = 8)**	**B2-ST141 (n = 5)**	**B2-ST372 (n = 29)**	**F-ST648 (n = 6)**
Adhesins							
fimH	9	4	4	8	5	29	6
fimAv$_{MT78}$	0	0	0	0	0	0	0
papAH	9	0	4	7	1	21	0
papC	9	0	4	7	1	21	0
sfa/focDE	9	0	3	8	4	26	1
afa/draBC	0	0	0	0	0	0	0
yfcV	9	0	4	8	5	29	6
Toxins							
sat	0	0	0	0	0	0	0
cnf1	8	0	4	7	0	23	0
hlyA	9	0	4	7	1	23	0
cdtB	0	0	2	0	0	0	1
tsh	0	0	0	0	0	0	5
vat	8	0	4	8	5	29	2
***Iron* uptake**							
iutA	1	0	0	1	0	0	6
iroN	9	0	2	7	4	26	5
fyuA	8	4	4	8	5	29	3
chuA	9	4	4	8	5	29	6
Caps*ul*e							
kpsM II	6	4	4	7	5	2	6
neuC-K1	0	0	0	0	5	0	1
kpsM II-K2	0	3	0	0	0	0	1
kpsM II-K5	6	1	4	7	0	2	4
kpsM III	2	0	0	0	0	0	0
Miscellaneous							
cvaC	0	0	0	0	0	0	5
iss	0	0	0	1	0	0	5
traT	1	1	2	1	0	4	6
ibeA	0	0	0	0	4	29	1
malX	9	0	4	8	5	29	2
usp	9	1	4	7	5	29	1
ExPEC status	9	0	3	8	4	20	6
UPEC status	8	0	3	8	5	29	3
Range of VFs	12 to 16	5 to 7	14 to 16	12 to 17	10 to 14	8 to 16	9 to 18
Mean of VFs	14.40	5.50	15.30	14.40	12.00	13.10	12.17

A total of 50 clones were identified among the 91 canine isolates classified as ExPEC and/or UPEC, with 11 of them including at least two isolates and only four, at least four isolates *i.e*. B2-CH14-180-ST127 (four isolates), B2-CH52-14-ST141 (four isolates), B2-CH103-9-ST372 (25 isolates) and F-CH4-58-ST648 (five isolates) (Table 32). In recent studies conducted by our research group (Flament-Simon *et al.*, 2019, 2020b, 2020a), we had identified, as indicated in Table 32, 15 of the 50 canine ExPEC/UPEC clones comprising 49 isolates among the 261 human ExPEC and/or UPEC isolates included in a collection of 394 *E. coli* isolates causing extraintestinal infections. However, only 31 of the 49 human ExPEC and/or UPEC isolates presented the same clone-related O:H serotypes as the canine isolates (Table 32). Among these 31 human isolates, 28 belonged to B2 phylogroup clones and three to F phylogroup clones identified among canine isolates. These B2 clones were distributed into five ST lineages including four lineages currently dominant in humans (ST73, ST127, ST141 and ST1193,) and the lineage

currently established as the dominant lineage in dogs, namely lineage ST372. In dogs, we found three clones in the lineage ST73 with the same serotype (O6:H1). The eight human isolates sharing this lineage with dogs were distributed into the same three clones and displayed serotype O6:H1. In dogs, we found two clones in the lineage ST127 displaying two serotypes with one (O6:HNM) of them present in the two clones. The four human isolates sharing this lineage with dogs were distributed into the same two clones but displayed the common serotype (O6:HNM). In dogs, we found two clones in the lineage ST141 with the same serotype (O2:H6). The 11 human isolates that shared this lineage with dogs were distributed into the same two clones and displayed the same serotype as human isolates. In dogs, we found one clone in lineage ST1193 with one serotype (O75:HNM). Three human isolates shared this clone and serotype with dogs. Concerning the lineage ST372, we found five clones in dogs with one of them including isolates displaying six serotypes. The two human isolates sharing the lineage ST372 with dogs belonged to this multiple-serotype clone and both displayed one of the six serotypes (O83:H31). Concerning the three F group human isolates, they belonged to one of the three F group clones (F-CH32-41-ST59) identified in dogs and showed the same serotype (O1:H7).

Among the 28 canine MDR isolates, we observed 24 different clones, of which nine had also been identified among the above cited 394 isolates causing infections in humans (Table 33) (Flament-Simon *et al.*, 2019, 2020b, 2020a)

4.5.5. Whole Genome Sequencing (WGS) and Molecular Characterisation of ST372 Isolates

For WGS, we selected 23 of the above studied ST372 isolates. They comprised 21 of the 29 Spanish canine faecal ST372 strains that were isolated in 2013 (n=9) and 2017 (n=12) and two previously published human ST372 strains isolated in 2016 (Flament-Simon *et al.*, 2019): strains LREC_341 isolated in Spain from an abscess and LREC_342 isolated in France from a bone infection. Both human strains showed serotype O18:H31 and clonotype CH103-9 whereas the 21 canine strains showed six different serotypes [O4:H31 (seven isolates), O83:H31 (four isolates), O25:H31 (four isolates), O15:H31 (three isolates), O21:H31 (two isolates) and O117:H28 (one isolate)] and four clonotypes [CH103-9 (18 isolates), CH103-10, CH103-17, and CH103-240].

The main objectives were to get more insights into the *E. coli* ST372 lineage that appears as one of the most prevalent *E coli* lineages among the canine faeces *E. coli* populations and to elucidate if there is any relation between canine and human ST372 strains.

To infer the phylogeny, we performed an SNP-tree with 197 genomes of ST372 strains [23 from this study (labelled LREC strains) and 174 obtained from public databases] corresponding to 151 genomes from canine strains and 46 genomes from human strains (CG-4). A total of 70% of these genomes corresponded to strains collected between 2017 and 2019 while the remaining 30% corresponded to strains isolated between 1995 and 2016. Regarding geographical distribution, 46 genomes (23.4%) were from strains collected in Europe and 143 (72.6%) from strains collected in North America.

The SNP analysis of the *E. coli* ST372 lineage revealed a wide and heterogeneous population allowing us to describe six clusters. Figure 17 only includes 97 representative genomes (including the 23 LREC genomes sequenced in this study and the 46 genomes from human strains) of the 197 analysed so that it is possible to visualise all the information.

Table 32. Clones and clonal-related serotypes of 91 canine ExPEC and/or UPEC isolates. Prevalence of the canine clones and clonal-related serotypes among ExPEC and/or UPEC isolates causing extraintestinal infections in humans (Flament-Simon *et al.*, 2019, 2020b, 2020a).

Clone of Canine ExPEC and/or UPEC Isolates	Clone-Related Serotype of Canine ExPEC and/or UPEC Isolates (Number of Isolates)	Number of Human ExPEC and/or UPEC Isolates with Same Clone of Canine Isolates (49 of 261)	Number of Human ExPEC and/or UPEC Isolates with Same Clone and Serotype of Canine Isolates (31 of 261)
A-CH11-NEG-ST93	O5:H4 (1)	3	0
A-CH11-27-ST new 1	O4:H27 (1)	0	0
B1-CH4-27-ST58	O8:H25 (1), O9:H25 (1)	0	0
B2-CH13-5-ST12	O4:HNM (1), O18:H5 (1)	0	0
B2-CH13-7-ST12	O4:H1 (1), O4:HNM (1)	0	0
B2-CH13-130-ST12	O18:H5 (1)	0	0
B2-CH13-223-ST12	O18:H5 (2)	1	0
B2-CH13-430-ST12	O4:H5 (1)	0	0
B2-CH13-431-ST12	O4:H5 (1)	0	0
B2-CH24-9-ST73	O120:H31 (1)	0	0
B2-CH24-27-ST73	**O6:H1** (1)	1	1
B2-CH24-30-ST73	**O6:H1** (1)	4	3
B2-CH24-103-ST73	**O6:H1** (1)	6	4
B2-CH24-1-ST80	O75:H7 (1)	0	0
B2-CH38-30-ST95	O1:H7 (1)	1	0
B2-CH14-2-ST127	**O6:HNM** (2)	4	3
B2-CH14-180-ST127	**O6:HNM** (3), O6:H11 (1)	1	1
B2-CH14-fimH_{TR}new 1-ST127	O6:H31 (1), O6:HNM (1)	0	0
B2-CH40-NEG-ST131	O25:H4 (1)	0	0
B2-CH52-5-ST141	**O2:H6** (1)	13	9
B2-CH52-14-ST141	**O2:H6** (4)	2	2
B2-CH103-9-ST372	O4:H31 (9), O15:H31 (1), O21:H31 (3), O25:H31 (4), **O83:H31** (7), O117:H28 (1)	2	2
B2-CH103-10-ST372	O15:H31 (1)	0	0
B2-CH103-17-ST372	O15:H31 (1)	0	0
B2-CH103-240-ST372	O83:H31 (1)	0	0
B2-CH103-706-ST372	O83:H31 (1)	0	0
B2-CH96-433-ST646	O22:HNM (1)	0	0
B2-CH43-fimH_{TR}new 2-ST929	O138:H14 (1)	0	0
B2-CH13-175-ST961	O4:HNM (1)	0	0
B2-CH52-428-ST998	O2:H6 (1)	0	0
B2-CH14-64-ST1193	**O75:HNM** (2)	4	3
B2-CH363-75-ST2622	O83:H6 (1)	0	0
B2-CH195-2-ST5644	O175:H5 (1)	0	0
B2-CH13-fimH_{TR}new 3-ST new 2	O18:HNM (1)	0	0
B2-CH13-429-ST new 3	O4:H5 (1)	0	0
B2-CH103-9-ST new 4	O4:H31 (1)	0	0
B2-CH103-12-ST new 5	O6:HNM (1)	0	0
B2-CH11-34-ST new 6	O5:H11 (1)	0	0
B2-CH40-20-ST new 7	O1:H4 (1)	0	0
B2-CH363-75-ST new 8	O83:H4 (1)	0	0
B2-CH24-2-ST new 9	ONT:H1 (1)	0	0
B2-CH24-1473-ST new 10	O120:H5 (1)	0	0
B2-CH23-31-ST new 11	O103:H4 (1)	0	0
B2-CH40-20-ST new 12	O1:H4 (1)	0	0
B2-CH40-21-ST new 13	O13:H4 (1)	0	0
D-CH35-27-ST new 14	O77:H18 (1)	0	0
E-CH132-65-ST501	ONT:H1 (1)	0	0
F-CH32-41-ST59	**O1:H7** (1)	4	3
F-CH4-27-ST648	O4:H6 (1)	1	0
F-CH4-58-ST648	O153:H42 (5)	2	0

Bold highlights those canine clones and clone-related serotypes also detected among ExPEC and/or UPEC isolates causing extraintestinal infections in humans.

Table 33. Clones and clonal-related serotypes of 28 canine multidrug resistant (MDR) *E. coli* isolates. Prevalence of the canine clones and clone-related serotypes among *E. coli* isolates causing extraintestinal infections in humans (Flament-Simon *et al.*, 2019, 2020b, 2020a).

Clone of Canine MDR Isolates	**Clone-Related Serotype of Canine MDR Isolates (Number of Isolates)**	**Type of ESBL and pAmpC Enzymes Produced by Canine MDR Isolates**	**Number of Human *E. coli* Isolates with Same Clone of Canine MDR Isolates (Number and Type ESBL Produced by Human Isolates) (35 of 394)**	**Number of Human *E. coli* Isolates with Same Clone and Serotype of Canine MDR Isolates (1 of 394)**
A-CH11-54-ST10	O128:HNM (1)	SHV12	10 (3 SHV12)	0
A-CH11-NEG-ST93	O5:H4 (1)	none	4 (1 CTX-M-14)	0
A-CH11-54-ST8953	O101:HNM (1)	CMY-2	0	0
A-CH11-27-ST new 1	O4:H27 (1)	none	0	0
B1-CH4-27-ST58	O8:H25 (1), **O9:H25** (1)	none	4 (1 CTX-M-14 and 1 CTX-M-32)	1
B1-CH4-121-ST155	O5:H11 (1)	none	0	0
B1-CH4-366-ST155	O9:H10 (1)	CMY-2	1 (1 CTX-M-1)	0
B1-CH4-425-ST new 15	O123:H11 (1)	CTX-M-1	0	0
B1-CH4-31-ST new 16	O8:H7 (1)	CTX-M-1	0	0
B1-CH29-38-ST new 17	O8:H49 (1)	CTX-M-1	0	0
B1-CH30-38-ST new 18	O12:H8 (1)	CMY-2	0	0
B2-CH13-223-ST12	O18:H5 (1)	CMY-2	1	0
B2-CH13-429-ST new 3	O4:H5 (1)	CMY-2	0	0
C-CH4-39-ST88	O45:HNM (1)	CTX-M-1	11 (1 CTX-M-14)	0
D-CH26-5-ST38	O86:H18 (3)	CTX-M-14	0	0
D-CH26-65-ST38	O1:H34 (1)	CMY-2	2 (1 CTX-M-15)	0
D-CH35-27-ST new 14	O77:H18 (1)	none	0	0
E-CH31-54-ST57	O27:H40 (1)	CMY-2	1	0
E-CH11-167-ST695	O99:H38 (1)	none	0	0
E-CH4-31-ST1011	O166:H45 (1)	CTX-M-55	0	0
E-CH23-221-ST1140	O44:H39 (1)	none	0	0
E-CH485-426-ST3774	O9:H31 (1)	CMY-2	0	0
F-CH88-145-ST457	O11:H25 (2)	CMY-2	0	0
F-CH4-27-ST648	O4:H6 (1)	CTX-M-14	1 (1 CTX-M-15)	0

Bold highlights those canine clones and serotypes also detected among *E. coli* isolates causing extraintestinal infections in humans.

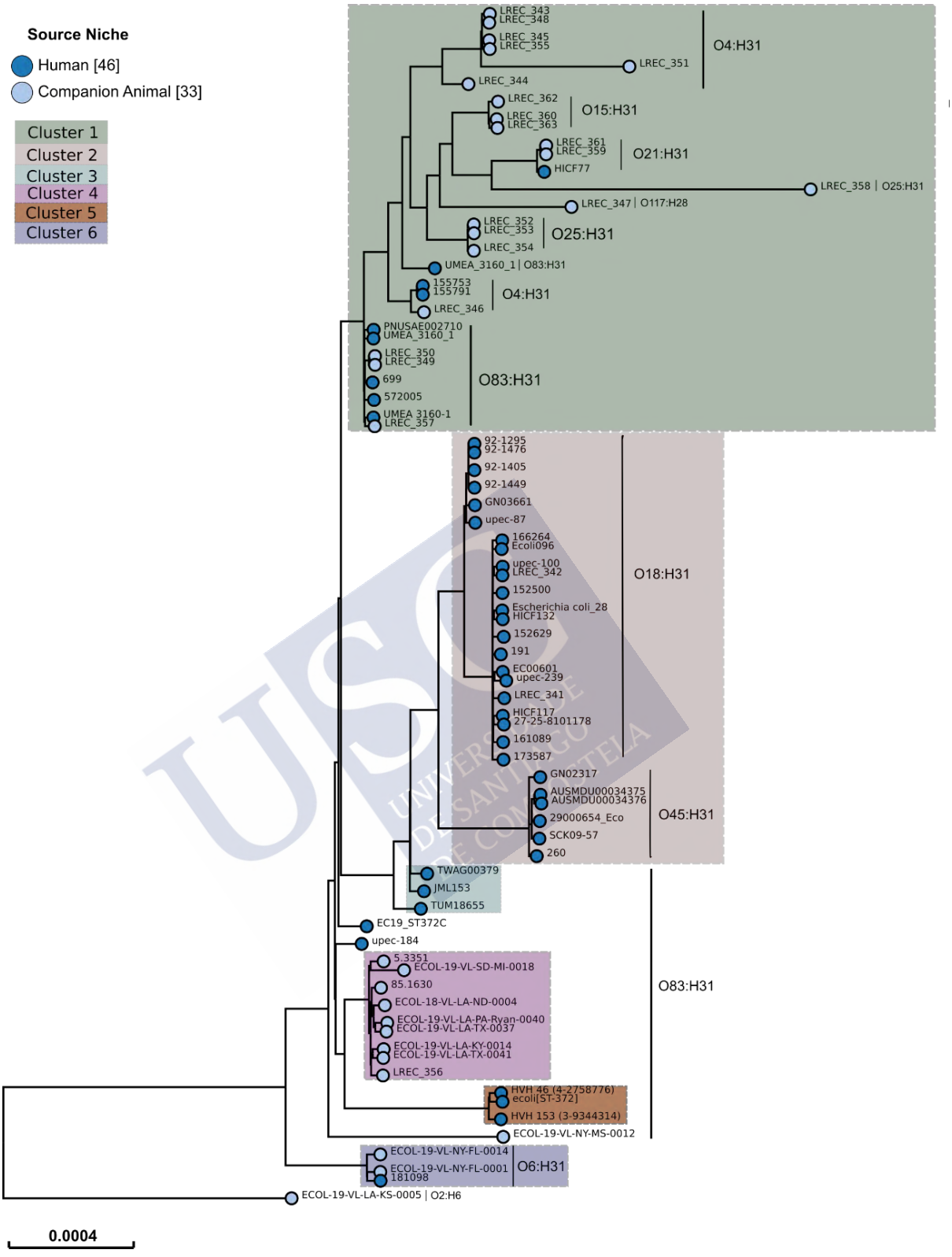


Figure 17. SNP-tree of 79 representative ST372 *E. coli* genomes from 46 human strains and 33 canine strains.Tree visualization by EnteroBase (Zhou *et al.*, 2020). The 33 genomes from canine strains are representatives of the different clusters identified in the SNP matrix of a previous SNP-tree performed with the 197 genomes analysed in this study. The identified serotypes are listed beside the vertical line.

The criterion established to define a cluster was that it should include genomes with less than 200 SNPs distance between them. An exception to this rule was the inclusion of the genome ECOL-19-VL-SD-MI-0018, with a maximum of 391 SNP distance, in cluster 4. Five genomes did not reach this criterion having more than 400 SNP distance between them and could form five other clusters. However, we have included those genomes in only one category (undefined) to simplify the following analysis.

According to the phylogenetic tree built from the genome of the 197 strains, cluster 1 comprised 147 (74.6%) of the 197 analysed genomes. This cluster was mostly composed of genomes of canine strains (138 genomes; 93.9%). Genomes of canine strains were also included in clusters 4 (9 genomes) and 6 (2 genomes) while only human strain genomes were included in clusters 2 (28 genomes), 3 (3 genomes) and 5 (3 genomes) (Table 34). Thus, cluster 1 comprised significantly more canine strain genomes ($P < 0.00001$) while clusters 2 ($P < 0.00001$), 3 ($P = 0.01209$) and 5 ($P = 0.01209$) comprised significantly more human strain genomes. Twenty of the 21 genomes of the Spanish canine strains belonged to cluster 1 whereas, the genome of the remaining Spanish canine strain (LREC_356) belonged to cluster 4. The genomes of the Spanish and French human strains (LREC_341 and LREC_342) belonged to cluster 2.

Table 34. Distribution into the phylogenetic clusters of the 197 canine and human ST372 strains.

Cluster	Number of Strains (%)		*p*-Value [1]
	Canine (n = 151)	Human (n = 46)	
1	138 (91.4)	9 (19.6)	<0.00001
2	0	28 (60.9)	<0.00001
3	0	3 (6.5)	0.01209
4	9 (6.0)	0	
5	0	3 (6.5)	0.01209
6	2 (1.3)	1 (2.2)	
Undefined	2 (1.3)	2 (4.3)	

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

Both clusters 1 and 2 were the most frequent clusters observed among the studied *E. coli* ST372 strains (canine and human) isolated in Europe and North America. However, cluster 1 was significantly associated with North America strains ($P = 0.02476$), while cluster 2 was especially associated with Europe strains ($P = 0.01233$) (Table 35).

Table 35. Cluster distribution of the 197 studied ST372 strains according to countries.

Cluster (Number of Strains)	Number of Strains (%)		*p*-Value [1] Europe vs. North America	Countries (Number of Strains)
	Europe (n = 46)	North America (n = 143)		
1 (n = 147)	30 (65.2)	117 (81.8)	0.02476	USA (109), Spain (20), Canada (6), UK (4), Sweden (3), France (2), Germany (1), North America (2)
2 (n = 28)	11 (23.9)	13 (9.1)	0.01233	USA (13), UK (8), Spain (1), France (1), The Netherlands (1), Australia (2), unknown (2)
3 (n = 3)	0	0		Japan (1), Kenya (1), unknown (1)
4 (n = 8)	1 (2.2)	7 (4.9)		USA (7), Spain (1), unknown (1)
5 (n = 3)	3 (6.5)	0	0.01371	Denmark (2), France (1)
6 (n = 3)	1 (2.2)	2 (1.4)		USA (2), UK (1)
Undefined (n = 5)	0	4 (2.8)		USA (4), Australia (1)

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

To compare the virulence profile of the 197 canine and human ST372 strains, we *in silico* investigated the presence of 32 VF-encoding genes in the 197 strains and defined their ExPEC and UPEC status. We also investigated the distribution of those VF-encoding genes according to the classification of the strains into the six defined clusters. Table 36 summarized the results obtained from the mentioned analysis. Microbiological, geographical and genomic data of each of the 197 studied strains are available in Supplementary Table S2 in Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms8111712.

The canine ST372 strains showed a higher VF-encoding gene score (mean 16.79) compared with the human ST372 strains (mean 13.76). However, three human stains belonging to cluster 5 were those with the highest number of VF-encoding genes (mean 21.67). Eight VF-encoding genes (*papAH*, *papC*, *papEF*, *focCD*, *focG*, *cnf1*, *hlyA*, and *iroN*) were significantly associated with canine ST372 isolates, whereas that five (*hlyF*, *iutA*, *kpsM II*, *kpsM II-K5*, and *iss1*) were significantly associated with human ST372 isolates. Interestingly, the ExPEC status was found more frequently among canine ST372 strains (74.8%) than human strains (21.7%) ($P < 0.00001$) (Table 36).

The more prevalent serotype was O83:H31 which represents 36.0 % of the 197 ST372 strains followed by O4:H31 (17.8%), O15:H31 (15.2%), O18:H31 (10.2%), O45:H31 (3.0%), O117:H28 (3.0%), O21:H14 (2.5%), O21:H31 (2.5%), O75:H31 (2.5%), O-unknown:H31 (2.5%), O25:H31 (2.0%), O2:H6 (0.5%) and O-unknown:H28 (0.5%). The serotypes O4:H31 ($P = 0.02631$) and O15:H31 ($P = 0.00062$) were significantly associated with canine ST372 strains, whereas the serotypes O18:H31 ($P < 0.00001$) and O45:H31 ($P = 0.00012$) were significantly more frequent among human ST372 strains. The 65 canine strains of serotypes O4:H31 and O15:H31 belonged to cluster 1 and the 26 human strains of serotypes O18:H31 and O45:H31 belonged to cluster 2 (Table 37). In contrast, the dominant serotype O83:H31 was frequently identified among canine (38.4%) and human (28.3%) strains, and, although the majority of the strains with this serotype belonged to cluster 1, O83:H31 strains were also found in clusters 3, 4 and 5.

Table 37. Distribution of serotypes among the 197 ST372 strains according to origins (canine and human) and cluster types.

Serotype *In Silico*	**Number of Strains (%)**		***p*-Value [1] Canine vs. Human**	**Number of Isolates Belonging to Cluster**						
	Canine (n = 151)	**Human (n = 46)**		**1**	**2**	**3**	**4**	**5**	**6**	**Undefined**
O2:H6	1 (0.7)	0								1
O4:H31	32 (21.2)	3 (6.5)	0.02631	35						
O6:H31	2 (1.3)	1 (2.2)							3	
O15:H31	30 (19.9)	0	0.00062	30						
O18:H31	0	20 (43.5)	<0.00001		20					
O21:H14	5 (3.3)	0		5						
O21:H31	4 (2.6)	1 (2.2)		5						
O25:H31	4 (2.6)	0		4						
O45:H31	0	6 (13.0)	0.00012		6					
O75:H31	5 (3.3)	0		5						
O83:H31	58 (38.4)	13 (28.3)		53		3	9	3		3
O117:H28	6 (4.0)	0		6						
O-unknown:H31	3 (2.0)	2 (4.3)		3	2					
O-unknown:H28	1 (0.7)	0		1						

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

Table 36. Distribution of the VF-encoding genes detected among the 197 ST372 *E. coli* genomes according to strain origins (canine/human) and cluster types.

VF-Encoding Gene	Number of Strains (%)									*p*-Value [1] Canine vs. Human
	Canine (n =151)	Human (n = 46)	Cluster 1 (n = 147)	Cluster 2 (n = 28)	Cluster 3 (n = 3)	Cluster 4 (n = 9)	Cluster 5 (n = 3)	Cluster 6 (n = 3)	Undefined (n = 4)	
Adhesins										
fimH	151 (100)	46 (100)	147	28	3	9	3	3	4	
papAH	88 (58)	5 (11)	89	0	0	0	0	3	1	<0.00001
papC	113 (75)	9 (20)	113	0	0	0	3	3	3	<0.00001
papEF	112 (74)	9 (20)	113	0	0	0	3	2	3	<0.00001
sfaDE	3 (2)	0 (0)	1	0	0	0	0	2	0	
sfaS	6 (4)	0 (0)	6	0	0	0	0	0	0	
focCD	136 (90)	12 (26)	144	0	0	0	1	1	2	<0.00001
focG	138 (91)	12 (26)	144	0	0	0	1	3	2	<0.00001
afaBCD/draP	0	0 (0)	0	0	0	0	0	0	0	
yfcV	150 (99)	46 (100)	146	28	3	9	3	3	4	
Toxins										
sat	0	0 (0)	0	0	0	0	0	0	0	
cnf1	117 (77)	11 (24)	120	0	0	0	3	3	2	<0.00001
hlyA	117 (77)	11 (24)	119	0	0	0	3	3	3	<0.00001
hlyF	1 (1)	3 (7)	1	0	0	0	3	0	0	0.04034
cdtB	0	0 (0)	0	0	0	0	0	0	0	
vat	148 (98)	45 (98)	144	27	3	9	3	3	4	
Iron uptake										
iutA	1 (1)	3 (7)	1	0	0	0	3	0	0	0.04034
iroN	139 (92)	14 (30)	144	0	0	0	3	3	3	<0.00001
fyuA	151 (100)	46 (100)	147	28	3	9	3	3	4	
chuA	151 (100)	46 (100)	147	28	3	9	3	3	4	
ireA	5 (3)	1 (2)	5	0	0	0	0	0	1	
Capsule										
kpsM II	16 (11)	36 (78)	5	28	3	9	3	0	4	<0.00001
kpsM II-K1	0 (0)	0 (0)	0	0	0	0	0	0	0	
kpsM II-K2	0 (0)	0 (0)	0	0	0	0	0	0	0	
kpsM II-K5	16 (11)	36 (78)	5	28	3	9	3	0	4	<0.00001
Miscellaneou										
iss1	1 (1)	3 (7)	1	0	0	0	3	0	0	0.04034
iss2	140 (93)	45 (98)	136	27	3	9	3	3	4	
traT	34 (23)	10 (22)	25	5	0	9	3	1	1	
ibeA	149 (99)	46 (100)	146	28	3	9	3	3	3	
malX-PAI	150 (99)	46 (100)	146	28	3	9	3	3	4	
usp	151 (100)	46 (100)	147	28	3	9	3	3	4	
ompT	151 (100)	46 (100)	147	28	3	9	3	3	4	
ExPEC status	113 (75)	10 (22)	113	0	1	0	3	3	3	<0.00001
UPEC status	151 (100)	46 (100)	147	28	3	9	3	3	4	
Mean of VFs	16.79	13.76	16.93	12.11	12.00	13.00	21.67	18.00	17.00	

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

The three most prevalent serotypes in Europe were O4:H31 (23.9%), O83:H31 (21.7%) and O18:H31 (19.6%), while in North America, they were O83:H31 (39.9%), O15:H31 (18.2%) and O4:H31 (16.8%). The serotypes O18:H31 (P = 0.02203) and O25:H31 (P = 0.00317) were more frequently observed in Europe, whereas the serotype O83:H31 (P = 0.03291) was more prevalent in North America (Table 38).

Table 38. Serotypes of the 197 ST372 strains according to countries.

Serotype *In Silico*	**Number of Strains (%)**		***p*-Value [1] Europe vs. North America**	**Countries (Number of Strains)**
	Europe (n = 46)	**North America (n = 143)**		
O2:H6	0	1 (0.7)		USA (1)
O4:H31	11 (23.9)	24 (16.8)		USA (23), Spain (7), UK (2), France (1), Sweden (1), Canada (1)
O6:H31	1 (2.2)	2 (1.4)		USA (2), UK (1)
O15:H31	4 (8.7)	26 (18.2)		USA (25), Spain (3), Canada (1), France (1)
O18:H31	9 (19.6)	10 (7.0)	0.02203	USA (10), UK (7), Spain (1), France (1)
O21:H14	1 (2.2)	4 (2.8)		USA (4), Sweden (1)
O21:H31	3 (6.5)	2 (1.4)		Spain (2), UK (1), USA (1), Canada (1)
O25:H31	4 (8.7)	0	0.00317	Spain (4)
O45:H31	1 (2.2)	2 (1.4)		USA (2), Australia (2), Netherlands (1)
O75:H31	0	5 (3.5)		USA (5)
O83:H31	10 (21.7)	57 (39.9)	0.03291	USA (52), Spain (4), Canada (2), Denmark (2), UK (1), France (1), Sweden (1), Germany (1)
O117:H28	1 (2.2)	5 (3.5)		USA (5), Spain (1)
O-unknown:H31	1 (2.2)	4 (2.8)		USA (3), Canada (1), UK (1)
O-unknown:H28	0	1 (0.7)		USA (1)

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

The 23 LRCE genomes sequenced in this study were investigated in greater depth. These genomes were reconstructed to analyse the chromosome and plasmidome separately. The size of the chromosomes had an average of 5,043,308 pb and were encompassed in 55 to 178 contigs. We found an integrative conjugative element (ICE) with relaxase type MOB_Q in all the genomes except for LREC_347 genome. These ICEs belong to the ICEKp1 family, a yersiniabactin synthesis-associated ICE type (similar to ICEEcoUMN026-1). The contigs that harboured the ICE region were revised allowing us to detect the presence of a pathogenic island (PAI) and some VF-encoding genes. The ICE contig retrieved from LREC_356 genome was the longest (2,540,863 pb) and showed a high percentage of homology with the contigs of the other genomes harbouring the ICE region (annex 2, Figure S3). Interestingly, the contig from LREC_356 harboured the *ompT*, *iss*, *vat*, *fyvA,* and *yfcV* virulence- encoding genes, the last three mentioned genes being those used (in addition to c*huA*) to define UPEC status. We also identified the secretion system effector homolog type T6SS and the PAI_AET37190. In second place, in terms of ICE contig length, was LREC_357 (1,677,509 pb) genome that harboured the *iroN, iss, fyvA* and *yfcV* virulence-encoding genes. The ICE contigs had not the same length [varying from 2,540,863 pb to 307,789 pb (short read sequencing limitations)] and all the genes mentioned as found in the ICE contig from LREC_356 were not retrieved in the other 20 genomes harbouring ICE contigs: three harboured the *fyvA* and *yfcV* genes (LREC_359, LREC_361 and LREC_344) and the remaining 17 genomes harboured the *fyvA* gene. We concluded that the presence of this type of ICE was a common feature in the ST372 genomes from the 22 of 23 studied Spanish strains and may be involved in the acquisition of their UPEC status.

We also described 11 plasmids (four conjugative plasmids, six mobilizable plasmids and one plasmid with no relaxase suggesting that it is not mobilizable) which belonged

to the following relaxase families (MOB) and incompatibility groups (Inc.): MOB_{P3}/IncX1 (n=3); MOB_{P1}/nd (n=2); MOB_{F12}/IncFII-pCD1 (n=2); MOB_{F12}/IncFII-IncFIB (n=1); MOB_{H11}/IncHI2 (n=1); MOB_{Qu}/ColRNAI (n=1); nd/p0111 (n=1). To predict plasmid transferability, we investigated the presence of mating pair formation (Mpf) system proteins. These proteins were present in all the previously described MOB_{F12} and MOB_{H11} conjugative plasmids. Furthermore, *in silico* analysis showed that these plasmids did not carry resistance or virulence encoding genes except for the *cba* and *cma* genes that were found in plasmid pLREC354_1 and a bla_{TEM} gene found in pLREC346_1. Table 39 summarizes the MGE content of the 23 ST372 genomes.

We *in silico* investigated the presence of 189 VF-encoding genes, 87 antibiotic-resistance encoding genes (ARGs) and 18 types of point mutations (Supplementary Table S8 in Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms8111712.). Through this analysis, the 23 ST372 strains were shown with an UPEC status and harbouring a wide variety of VF-encoding genes, reaching an average number of 80. In contrast, these 23 ST372 strains were shown as carrying very few ARGs. However, genes encoding drug efflux were detected but only in the two human strain genomes (LREC_341 and LREC_342) that also harboured antibiotic-resistance encoding genes: bla_{TEM-1A}, *sul1*, *aadA1*, *dfrA1* and *mdf(A)*. These results were in agreement with those previously obtained by conventional methods.

Table 39. Description of mobile genetic elements (MGE)s found in the 23 ST372 strain genomes sequenced in this study.

Genome of Strain	Serotype of Strain	Origin	Cluster	MGEs		MOB/Inc Typing; Size (kb)	
				ICEs (kb of Contig)	Number of Plasmids	Plasmid_1	Plasmid_2
LREC_341	O18:H31	Human	2	MOB_{Q} (308)	2	MOB_{P1}/nd[1]; (164)	nd/p0111; (92)
LREC_342	O18:H31	Human	2	MOB_{Q} (90)	2	MOB_{P1}/nd; (720)	MOB_{Qu}/ColRNAI; (4)
LREC_343	O4:H31	Canine	1	MOB_{Q} (474)	1	MOB_{F12}/IncFII [F-:A-:B-], pCD1; (66)	
LREC_344	O4:H31	Canine	1	MOB_{Q} (1208)	0		
LREC_345	O4:H31	Canine	1	MOB_{Q} (158)	1	MOB_{F12}/IncFII [F2:A-:B-], pCD1; (75)	
LREC_346	O4:H31	Canine	1	MOB_{Q} (653)	1	MOB_{P3}/IncX1; (47)	
LREC_347	O117:H28	Canine	1		1	MOB_{P3}/IncX1; (38)	
LREC_348	O4:H31	Canine	1	MOB_{Q} (472)	0		
LREC_349	O83:H31	Canine	1	MOB_{Q} (707)	0		
LREC_350	O83:H31	Canine	1	MOB_{Q} (707)	0		
LREC_351	O4:H31	Canine	1	MOB_{Q} (653)	0		
LREC_352	O25:H31	Canine	1	MOB_{Q} (658)	0		
LREC_353	O25:H31	Canine	1	MOB_{Q} (658)	0		
LREC_354	O25:H31	Canine	1	MOB_{Q} (658)	0		
LREC_355	O4:H31	Canine	1	MOB_{Q} (157)	0		
LREC_356	O83:H31	Canine	4	MOB_{Q} (2541)	2	MOB_{F12}/IncFIB, IncFII [F-:A-:B52]; (162)	MOB_{P3}/IncX1; (36)
LREC_357	O83:H31	Canine	1	MOB_{Q} (1678)	0		
LREC_358	O25:H31	Canine	1	MOB_{Q} (660)	0		
LREC_359	O21:H31	Canine	1	MOB_{Q} (1317)	0		
LREC_360	O15:H31	Canine	1	MOB_{Q} (86)	0		
LREC_361	O21:H31	Canine	1	MOB_{Q} (1317)	0		
LREC_362	O15:H31	Canine	1	MOB_{Q} (641)	0		
LREC_363	O15:H31	Canine	1	MOB_{Q} (172)	1	MOB_{H11}/IncHI2; (204)	

[1] nd, not detected.

4.6 STUDY 6. PRODUCTION ANIMAL: WHOLE GENOME SEQUENCING AND CHARACTERISTICS OF MCR-1-HARBORING PLASMIDS OF PORCINE *ESCHERICHIA COLI* ISOLATES BELONGING TO THE HIGH-RISK CLONE O25B:H4-ST131 CLADE B

This study was presented in part at the 29th European Congress of Clinical Microbiology and Infectious Diseases, Amsterdam, 13-16 April 2019 and published in Frontiers in Microbiology the 24th of March of 2020 https://doi.org/10.3389/fmicb.2020.00387.

4.6.1. CHARACTERIZATION OF THE ISOLATES BY CONVENTIONAL METHODS

4.6.1.1. Pasteur Sequence Types (PST) and *fimH* alleles

All of the 11 ST131 isolates belonged to serotype O25b:H4 and the phylogenetic group B2. The seven isolates from piglets with diarrhoea showed the PST9 whereas the four isolates from pork meat showed the PST43 (Table 40).

Also all of the 11 ST131 isolates belonged to clade B and showed the *fimH*22 allele or mutational derivatives of the *fimH22* allele (*fimH*161, *fimH*298, *fimH*326, *fimH*332, *fimH*336, *fimH*338, *fimH*374). Resulting in clonotypes CH40-22 (3 isolates), CH40-161 (1), CH40-298 (2), CH40-326 (1), CH40-332 (1), CH40-336 (1), CH40-338 (1) and CH40-374 (1).

4.6.1.2 Virotypes

The 11 ST131 isolates harboured between 17 and 20 virulence gene markers screened by PCR. All isolates harboured the following genes: *iro*N, *iss, hly*F, *omp*T, *pap*EF, *papG* III, *kpsM*-II-K5, *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA,* and *usp*. However, some genes were variables: *pap*C and *pap*AH (10 isolates), *cnf*1 (six isolates), *hly*A (five isolates) and *cdt*B (four isolates). All isolates belonged to virotype D (Table 40). Six of seven isolates recovered from piglets with diarrhoea with PST9 showed the virotype D5 (*ibeA*, *papG* III, *cnf1, hlyA, kpsM*-II-K5), while the four PST43 isolates from meat showed the virotype D2 (*ibeA*, *papG* III, *cdtB, kpsM*-II-K5). All the 11 porcine ST131 isolates exhibited a virulence profile that satisfied the APEC, UPEC and ExPEC status and due to their origin were classified as APEC, UPEC and ExPEC like isolates (Table 40 and Supplementary Table S1 in Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387).

4.6.1.3 Resistance Phenotype

The resistance profile of the 11 ST131 porcine isolates was determined, showing that all but one exhibited a MDR phenotype. Amoxicillin/clavulanic acid resistance was present in 10 isolates, quinolone resistance in seven, colistin resistance in five isolates (with *mcr*-1 gene), cotrimoxazole resistance in four isolates, chloramphenicol resistance in two isolates, aminoglycoside resistance in two isolates and tetracycline resistance in one isolate (Table 40).

4.6.2. Whole Genome Sequencing (WGS)

The draft genomes of the 11 ST131 porcine isolates yielded 32 to 78 contigs larger than 1 Kb, with assembly sizes ranging from 4.940 Mb to 5.451Mb (annex 1, Table S1). The WGS data were in complete agreement with our previous characterization performed with conventional methods.

4.6.2.1. Core-genome (CG) and Phylogenetic Analysis

A total of 84 *E. coli* genomes were taken from different clades and subclades of ST131 in order to calculate the ST131 core-genome (CG). The main objectives were in regards to insights of ST131 lineage and the relationship between isolates of different

origins, with special attention to clade B of ST131 lineage. Due of that the CG analysis was dominated by clade B (n=49, 58%), and the remainder isolates belonged to clade C (n=26, 31%), clade A (n=8, 10%) and clade I (n=1, 1%). The final collection represented sequences from human (n=51, 61%), porcine (n=19, 23%), avian (n=8, 10%) and wild animals (Antarctic pinnipeds) (n=6, 7%) sources.

The CG encompasses 2.98 Mb (length=2,983,333 +/- 7,338 pb), comprised by 3,100 total CDS (CG-3). The SNPs analysis of the CG performed within ST131 lineage revealed a well-defined clade structure, similar to those previously described, including the three defined clades: A, B and C. In total, 14,838 SNPs were present in the alignment. Furthermore, there was a distance of 7,742 to 8,565 SNPs between isolates from clade A and clade C; 6,508 to 8,675 between isolates from clade A and clade B; and 2,214 to 4,280 between isolates from clade B and clade C. The distance within clades varies greatly: being 36 to 1,782 SNPs in clade A, 8 to 2,051 SNPs in clade B and 1 to 723 SNPs in clade C (Figure 18 and supplementary Table S2 in Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387).

All isolates of porcine origin analysed in this study were grouped in clade B. As shown in Figure 18, clade B was characterized by the presence of a mixture of human, livestock and food origin genomes and showed the greatest intra-group diversity. We identified 11 clusters or subclades (B0, B1, B2, B3, B4, B5, B6, B6-like, B7, B8, B9), including three subclades (B3, B6, B7) with isolates of both human and animal origin. In this study, we described five new subclades, named B6, B6-like, B7, B8, and B9. There was a distance of 167 to 232 SNPs between isolates from B6 and B6-like and 172 to 304 SNPs between isolates from B6 and B7. Further, B7 showed 119 to 145 SNPs distance with B1. Those subclades appeared as the closer within each other in clade B (Supplementary Table S2 in Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387). The 11 ST131 porcine isolates from this study were grouped in the new subclades B6 (LREC_160/PST43, LREC_161/PST43, LREC_168/PST43, LREC_176/PST43 from pork meat and virotype D2) and B7 (LREC_153/PST9, LREC_154/PST9, LREC_155/PST9, LREC_157/PST9, LREC_158/PST9, LREC_159/PST9, LREC_162/PST9 from piglets with diarrhoea and virotype D5). The new subclade B6-like includes two isolates from porcine origin. Otherwise, new subclades B8 and B9 only include ST131 isolates from Antarctic pinnipeds.

Notably, there were three (H1410, 19BA, H1698) isolates from human origin in subclade B6 and two (EcHUCA4, E61BA) in subclade B7. The distance between human isolates and porcine isolates from subclades B6 and B7 had an average of 60 SNPs and of 46 SNPs (i.e 20 SNP/Mb and 15 SNP/Mb), respectively.

4.6.2.2. Plasmidome Analysis

The 11 ST131 porcine isolates showed high heterogeneity in their plasmid content, harbouring between two to seven different plasmids (Table 41 and in annex 2, Figure S4). We find members of eight of the 17 main MOB plasmid groups found in γ-proteobacteria. Ordered by prevalence: MOB_{F12} (n=13), MOB_{P51} (n=8), MOB_{P3} (n=3), MOB_{V2} (n=2), MOB_{H11} (n=2), MOB_{P12} (n=2), MOB_{P131} (n=1) and MOB_{Q12} (n=1). Finally, we described six plasmids (16%) in which any MOB protein was found (non-mobilizable-infective plasmids). Twenty-two plasmids were considered large (> 33 kb) and presumptively conjugative, and sixteen were small multicopy plasmids (< 7 kb).

In summary, 38 plasmids were described in the 11 genomes analysed. The plasmids belonged to the following relaxase families (MOB) and incompatibility groups (Inc): MOB_{P51}/ColE1-like (n=8), MOB_{F12}/IncF [F2:A-:B1] (n=8), MOB_{F12}/IncF [F2:A-:B1] plus and presumptively cointegrated IncHI2-ST4/ST-nt (not typable) (n=3),

MOB_{F12}/IncF [F10:A-:B1] (n=2), MOB_{P12}/IncI1-ST27/ST171 (n=2), MOB_{P3}/IncX4 (n=2), MOB_{V2}/nd (not detectable) (n=2), MOB_{H11}/IncHI2-ST4 (n=1), MOB_{H11}/IncHI2-ST4 plus and presumptively cointegrated IncQ1 (n=1), MOB_{P131}/IncL/M plus and presumptively cointegrated IncQ1 (n=1), MOB_{P3}/IncX1 (n=1), MOB_{Q12}/nd (n=1). We also localized six plasmids that could not be affiliated with any categories.

All 13 MOB_{F12}/IncF plasmids were carriers of virulence genes. The genes *iro*BCDEN, *iss, mch*F, *iut*A, *iuc*ABCD were constant (Table 41). All the MOB_{F12}/IncF plasmids found in our study were structurally compared by BRIG software, using the pJIE186_2 (NC_020271) plasmid as reference (annex 2, Figure S5). The plasmids analysed showed extensive sequence similarity over at least 100 kb, which includes not only backbone genes. The comparison emphasizes that all of them share an 80 kb conserved region comprising the virulence genes described previously as well as *cvaABC, sitABC, ompT and hlyF*.

The eight MOB_{P51}/ColE1-like plasmids identified in this study exhibited a wide size range (1,7-6,8 kb) and were analysed using the ColE1 plasmid (J01566.1) as reference (annex 2, Figure S6). The plasmids pLREC153_3, pLREC157_3, pLREC158_2 and pLREC161_3 (all of them greater than 6.8 Kb) showed the highest homology to the reference plasmid. All of them carried the colicin E1 *cea* and immunity (*imm*) genes, and the entry exclusion (*exc1-exc2*) system genes. On the other hand, plasmids from isolate LREC_160 were different among them and with the others, exhibiting a wide size range (1.8 to 5.6 Kb), and only pLREC160_4 carried the colicin E1 gene.

Three MOB_{P3}/IncX plasmids were present among our isolates. We found pLREC157_2 that belong to the IncX1 subcluster and had a similar backbone than p2ESCUM (NC_011739.1) but codes for a completely different replication initiation protein (annex 2, Figure S7). While pLREC154_2 and pLREC160_2 belong to the IncX4 subcluster and were analysed using pSH696_34 (JX258654.1) as a reference (annex 2, Figure S8). Since pLREC160_2 carried the *mcr.1-1* gene, some previous reported IncX4 reference plasmids harboring the *mcr* gene were added. Compared to pSH696_34 plasmid pLREC160_2 and the other *mcr* carrier plasmids (pESTMCR and pICBEC7Pmcr) lacked of two important backbone genes: the conjugative coupling protein gene (*traG*) and the replication protein encoding region (*rep*).

Structural comparison of the two MOB_{P12}/IncI1-ST171 plasmids and the single MOB_{P131}/IncL/M plasmid found within our isolates was also performed using pEK204 (EU935740) and pEC743-OXA48 (CP015071.1) as references, respectively. Both showed a high coverage of homology with the reference (annex 2, Figure S9 and **S10**).

We also described five small (1.5 kb) multicopy plasmid in which any relaxase protein was found, categorized as cryptic. Those cryptic no-MOB plasmids were analysed by using pEC10D (NC_017650.1) as reference (annex 2, Figure S11). BRIG comparison showed that four plasmids were highly similar to pEC10D and codes for the same replication protein (cryptic 1, from the HTH36 superfamily replication proteins). While pLREC159_6 codes for a different replication protein and was called cryptic 2 (Table 41).

Finally, we analysed the MOB_{V2} plasmids found in the isolates LREC_153 and LREC_157. A BLAST research was performed to identify a reference plasmid, being pEC0674 (MF684783.1) as the closest match. Two MOB_{V2}-like related plasmids were also included for structural comparison (annex 2, Figure S12). Although no significant similarity was found between these latter two references and the pEC0674, the pLREC153_4 showed similarity with the replication protein, an hypothetical one and the Rec protein from the reference plasmid; whereas the pLREC157_4 only showed similarity to the latter two proteins.

4.6.2.3. Antimicrobial Resistance Genes (ARGs)

Many ARGs were found in the 11 ST131 porcine isolates (Table 41), including acquired resistance genes, point mutations and efflux/transporter genes.

Besides constitutive genes (PBP and AmpH, that were present in all isolates), ARGs conferring resistance to penicillin (variants of *bla*$_{TEM}$ gene, accounting: *bla*$_{TEM-1C}$, n=10; *bla*$_{TEM-40}$, n=10; *bla*$_{TEM-135}$, n=10; *bla*$_{TEM-150}$, n=2; *bla*$_{TEM-1A}$, n=2; *bla*$_{TEM-54}$, n=1), tetracycline (variants of *tet* gene, accounting: *tet*A, n=9; *tet*C, n=9; *tet*R, n=8; *tet*M, n=1), quinolones (*gyr*A_S83L, n=7) and colistin (*mcr*-1.1, n=5) were the most common. In addition, we described genes conferring resistance to sulphonamides (variants of *sul* gene, accounting: *sul*2, n=4; *sul*3, n=4; *sul*1, n=3), trimetoprim (variants of *drf* gene, accounting: *dfr*A1, n=3; *dfr*A15, n=3; *dfr*A12, n=1), aminoglycosides (different ACCs, APHs and ANTs genes, accounting: APH-3''-Ib, n=3; APH-6'-Id, n=3; *aad*A1, n=3; ANT-3''-Ia, n=3; *str*A, n=3; *str*B, n=3; APH-3'-Ia, n=2; *aad*A2, n=2; AAC-3'-IIa, n=2; AAC-3'-IIc, n=2; *aad*A17, n=1), phenicols (variants of *cml*A gene, *cat*A1 and *mdf*A gene, accounting: *mdfA,* n=11; *cml*A1, n=1; *cml*A6, n=1; *cat*A1, n=1) and to amonium quaternary biocide (*qac*Edelta1, n=3; *qac*HR, n=1) (Table 41).

The phenotypic resistance profile in the studied genomes was due to genes carried by plasmids, except for quinolone resistance, which in our study is mediated by a Ser83Leu point mutation found in the chromosomal *gyr*A gene. In the case of LREC_158 isolate it exhibited the *bla*$_{TEM}$ genes (-1c, -40 and -135) variants set into a chromosomal assembled contig. However, these genes were located into IncL/M, IncX4, IncI1, IncF and IncHI2 in the other analysed isolates (Table 41).

As shown in Table 40, all but one isolate exhibited a MDR profile. Large conjugative MDR plasmids were identified in four of the ten isolates. These were typed as MOB$_{H11}$/IncHI2-ST4 (pLREC159_1 and pLREC176_1), MOB$_{F12}$/IncF (pLREC159_2, pLREC161_1) and MOB$_{P131}$/IncL/M (pLREC153_2), and a non-typeable one (pLREC161_2). Most common antibiotic resistance profiles encode by these plasmids are ampicillin/penicillin, aminoglycosides, tetracyclines, sulphonamides, cotrimoxazole and polymyxins. Finally, we described some drug efflux genes in pLREC159_4 and pLREC161_2 (Table 41).

Class 1 integrons were present in pLREC153_2 (*IntI-attI* [*-sul2+aph-(3'')-Ib+aph-(6)-Id+dfrA1+aadA1*]*-qacE1-sul1*), pLREC159_2 (*IntI-attI* [*dfrA1+aadA1*]*-qacE1-sul1*) and pLREC176_1 (*IntI-attI* [*dfrA1+aadA1*]*-qacE1-sul1*). No class 2 or 3 integrons were detected in our genomes.

4.6.2.4. Colistin Resistance Vehicles and Genetic Environment

Colistin resistance by *mcr* gene was found in five of the analysed isolates. Therefore, the *mcr*-carrying plasmids were investigated in greater depth. The *mcr*-1.1 gene was located in two MOB$_{H11}$/IncHI2-ST4, a MOB$_{P3}$/IncX4 and a MOB$_{F12}$/IncF [F2:A-:B1], that was theoretically cointegrated with an IncHI2-ST-nt plasmid. Besides, the LREC_154 isolate showed the *mcr*-1.1 in a chromosomal location (Table 41). In this study, any point mutation in the chromosomally encoded genes *pmr*A or *pmr*B was found.

The complete sequence of pLREC160_2 (MOB$_{P3}$/IncX4) was 40,221 bp in size and contains 69 predicted ORFs. A structural comparison against other reference IncX4 plasmids is shown in Figure 19, where pICBEC7P*mcr* (34,992 pb, NZ_CP017246.1) was used as the internal reference. The structure of the *mcr-1* carriers IncX4 plasmid remains stable if we exclude one ORF encoding for an IS91 family transposase present in pLREC160_2 and pICBEC7P*mcr*. The BRIG comparison showed that the *mcr-1* negative plasmid pJIE143 differs from pICBEC7P*mcr* in a region of approximately 6 kb (encoding for *par*A, *mcr*, and *pap*2 genes). All the analysed plasmids showed a typical IncX4

backbone, including the replication initiation protein gene (*pir)*, conjugal transfer protein genes (*tra*) plus *taxABC* and *pilX* operons **(**annex 2, Figure S13 to S16).

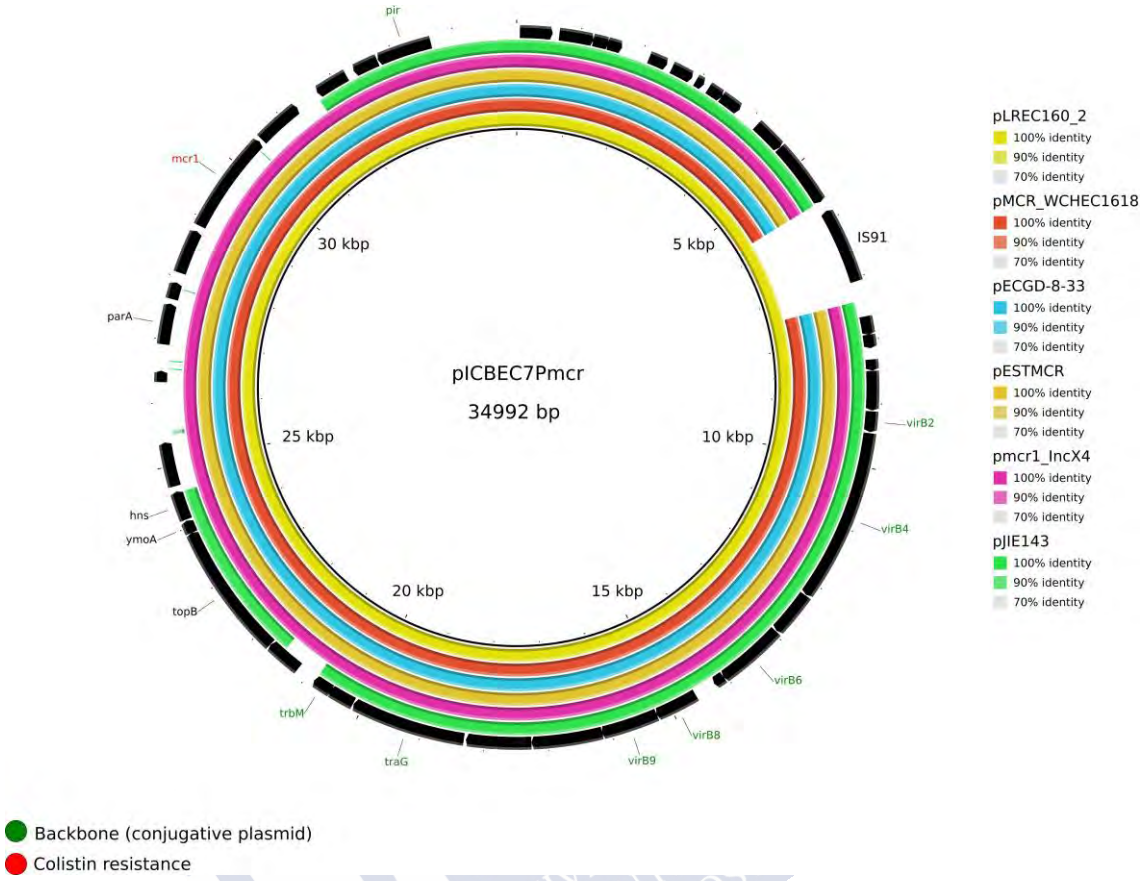


Figure 19. Structural comparison between *mcr*-1 plasmids and IncX4 references. The alignment includes five references and the one IncX4 *mcr*-1-bearing plasmid found in our study. The plasmid pICBEC7*mcr* (IncX4) was used as a reference to match with the other plasmids with [pMCR_WCHEC1618, pECGD-8-33, pESTMCR and pmcr1_IncX4] and without the *mcr*-1 gene [pIJE143]. The outer circle with black arrows denotes the annotation of reference sequence pICBEC7mcr. The image was generated using BRIG (default parameters with 90/70 as upper/lower threshold).

Plasmid pLREC161_1 was 330,357 bp in size and contain 616 predicted ORFs. This MOB_{F12} plasmid belongs to the IncF [F2:A-:B1] incompatibility group and was theoretically cointegrated with an IncHI2 plasmid. A structural comparison with plasmid pMR0516*mcr* [F18:A-:B1] (KX276657.1) as an internal reference (225,069 bp) showed 100% identity in more than 70 % of query coverage from BRIG (Figure 20). Genes involved in the conjugative transfer (*tra, trb* and pili operons), in the toxin-antitoxin system (*vagC*) and the inhibition of SOS response (*psi*AB) were present. Most of the differences were observed in the resistance module. It is worthy of note that pLREC161_1 and pMR0516*mcr* plasmids contained APEC virulence genes and that the surrounding region of the *mcr* gene includes heavy metal resistance genes to tellurium (*ter* genes). Moreover, in annex 2, Figure S17, where pKP81_BE [F2:A-:B] (KU994859.1) was used as the reference, we observed 100% of identity in more than 78% of query coverage from BRIG with pLREC161_1. Interestingly plasmid R478 (IncHI2 reference plasmid, U62007.2) also shared with pKP81_BE some backbone genes and the tellurium resistance module like pMR0516*mcr* (annex 2, Figure S17). A wide part of the backbone in pLREC161_1 (approx. 44 kb) was common with the widely studied IncF plasmid from

E. coli-K12, used as non-*mcr* plasmid reference (from coord. 55 kb to 99 kb). The ancestral IncF plasmid of ST131 (pECSF1) was also included in BRIG comparisons (annex 2, Figure S18).

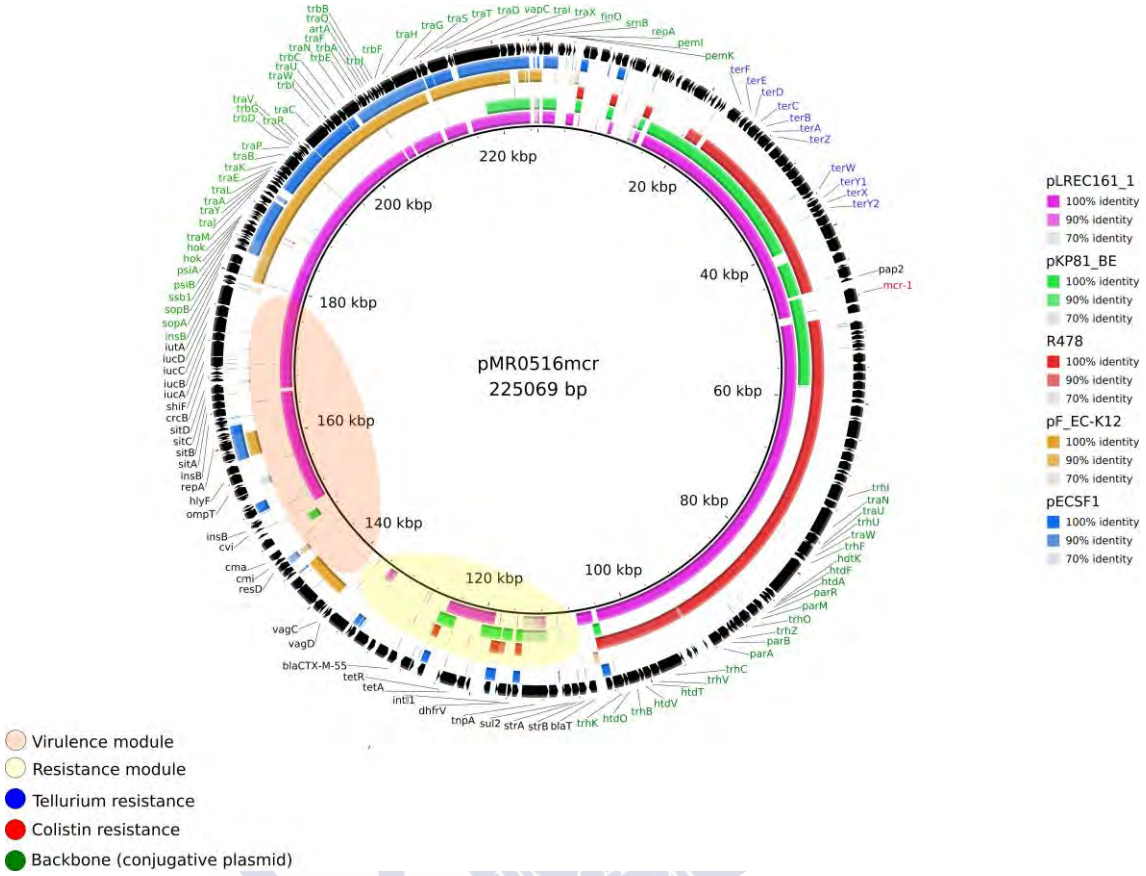


Figure 20. Structural comparison between *mcr-1* plasmids, IncF and IncHI2 references. The alignment includes four references and the one IncF *mcr-1*-bearing plasmid found in our study. pMR0516mcr (IncF) was used as a reference to match with the other plasmids with [pKP81_BE (IncF)] and without *mcr-1* gene [IncF plasmid from *E. coli* K12, pECSF1 (IncF) and R478 (IncHI2)]. The outer circle with black arrows denotes the annotation of pMR0516mcr. The image was generated using BRIG (default parameters with 90/70 as upper/lower threshold).

There were two MOB_{H11}/IncHI2-ST4 *mcr-1* bearing plasmids: (I) pLREC159_1 which was 248,461 bp in size and contain 420 predicted ORFs and, (II) pLREC176_1 which was 278,079 bp in size and contain 523 predicted ORFs. Both showed 100% of identity over more than 75% of query coverage in BRIG comparison with pHNSHP45-2 (KU341381.1) (Figure 21). The backbone was almost identical; nevertheless, the resistance module was distinct between them. The result of this comparative analysis showed that a wide range of resistance genes could be found in the MDR module of IncHI2 plasmids. Moreover, all the IncHI2 plasmids analysed presented a tellurium resistance module close to the *mcr*-1 cassette. The similarities between IncHI2 plasmids were high even though their origin and parental isolate characteristics differ.

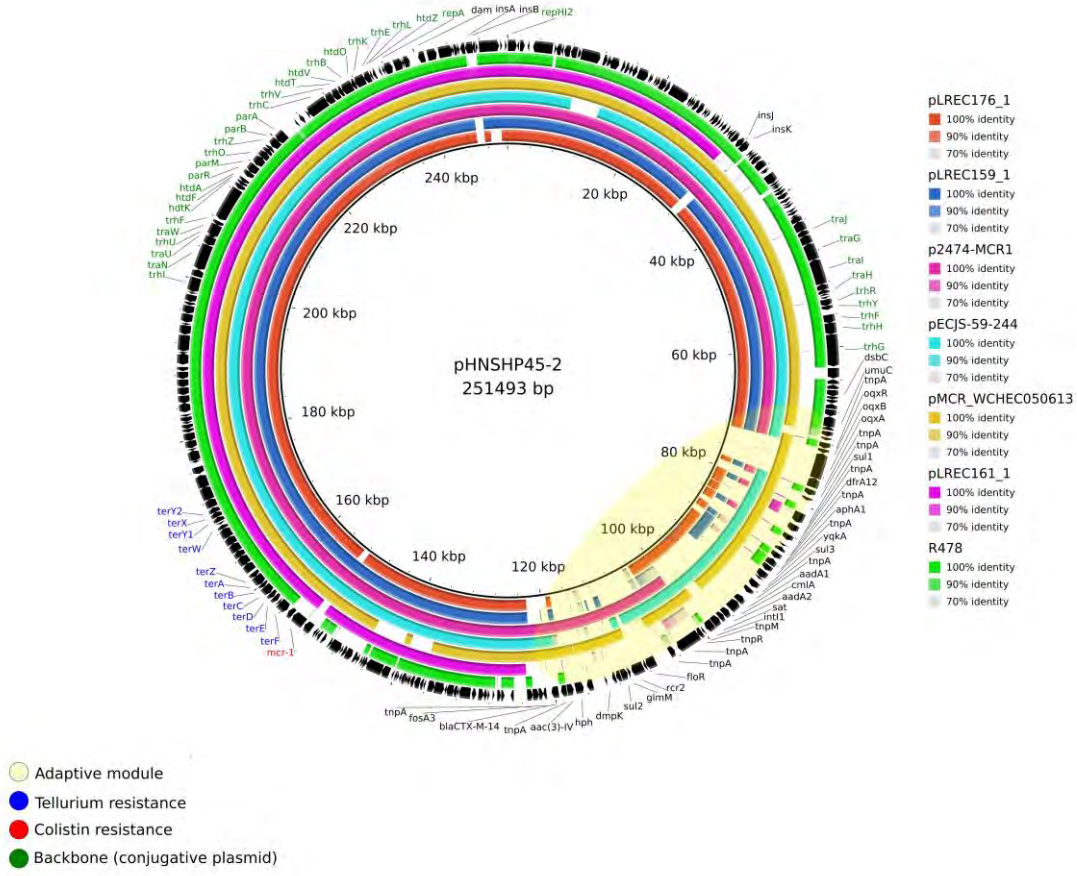


Figure 21. Structural comparison between *mcr-1* plasmids and IncHI2 references. The alignment includes four references and the *mcr-1*-bearing plasmids from our study pLREC176_1 and pLREC159_1 which are IncHI2 and pLREC161_1 which is an IncF plasmid with a presumptively IncHI2 cointegrated. The plasmid pHNSHP45-2 (IncHI2) was used as a reference to match with the other plasmids with [p2474-MCR1 (IncHI2), pECJS-59-244 (IncHI2) and pMCR_WCHEC050613 (IncHI2)] and without the *mcr-1* gene [R478 (IncHI2)]. The outer circle with black arrows denotes the annotation of reference sequence pHNSHP45-2. The image was generated using BRIG (default parameters with 90/70 as upper/lower threshold).

Afterward, a comparison between all our *mcr*-1 bearing plasmid was performed with pHNSHP45-2 (MOB_{H11}/IncHI2) as a reference, and as expected, almost none homologous region was found between plasmids from different families (annex 2, Figure S19). Nevertheless, pLREC161_1 ($MOBF_{12}$/IncF [F2:A-:B1] plus IncHI2-ST-nt) plasmid shared a part of his backbone with the reference one, including the module encoding for tellurium resistance and the *mcr* gene. Suggesting that the *mcr*-1 cassette was introduced into pLREC161_1 through the cointegrated IncHI2 plasmid.

We had also made a sequence alignment to analyze the surrounding region of our *mcr* bearing contigs and compare their structure against plasmids references (Figure 22). Regarding plasmid families, all the IncX4 plasmids were identical in terms of the genetic arrangement in the vicinity of the *mcr*-1 gene, while the insertion site of *mcr*-1 gene in

IncHI2 plasmid was variable. Interestingly, the *mcr* region in plasmids pMCR_WCHEC050613 (IncHI2) and plasmid pMR0516*mcr* (IncF) shared 100% of identity. It is also worth to note that the IncF *mcr* harboring plasmid pKP81_BE carried the tellurite resistance module surrounding the *mcr* gene like the IncHI2 plasmids.

Finally, the genetic arrangement of the *mcr*-1 cassette was investigated (annex 2, Figure S20). All sequences share a 2,542 bp DNA segment, containing *mcr*-1.1 and a putative ORF, encoding a Pap2 superfamily protein. However, some variability in the last 3' end of the *pap2* gene was observed.

The genome reconstruction of LREC_154 showed that *mcr* gene was inserted into the chromosome. However, the *mcr* harboring contig only included *mcr* and *pap2* genes. It was flanked at both ends by a partial ISA*pl1* sequence of 99 bp upstream and 405 pb downstream the *mcr*-1 cassette. In pLREC159_1 the *mcr* harboring contig was disrupted after a segment of 87 bp of DNA homologous to ISA*pl1* upstream the *mcr* gene. Probably the presence of the IS (repeated element) has interfered with the *de novo* reconstruction of the sequenced genome by the assembly of Illumina short reads. In addition, we identified the IS*26* and an IS*1294* surrounding the *mcr*-1 cassette in pLREC160_2.

Plasmids pLREC159_1 and pLREC176_1 identified as IncHI2-ST4, had identical genetic environments downstream the *mcr*-1 cassette.

4.6.3 STATISTICAL ANALYSIS

The variables (*cdt*B, *cnf*1, *hly*A, virotypes, PST and subclades) did not follow a Normal distribution (Shapiro-Wilk test: *P-Value* < 0.0001). We performed the Mann-Whitney-Wilcoxon test to assessed significative differences in the distribution of features (virotype D5|subclade B7|PST9|*cnf1*|*hly*A vs. virotype D2|subclade B6|PST43|*cdt*B variables; *P-Value* 0.007) between ST131 isolates from piglets with diarrhoea and pork meat.

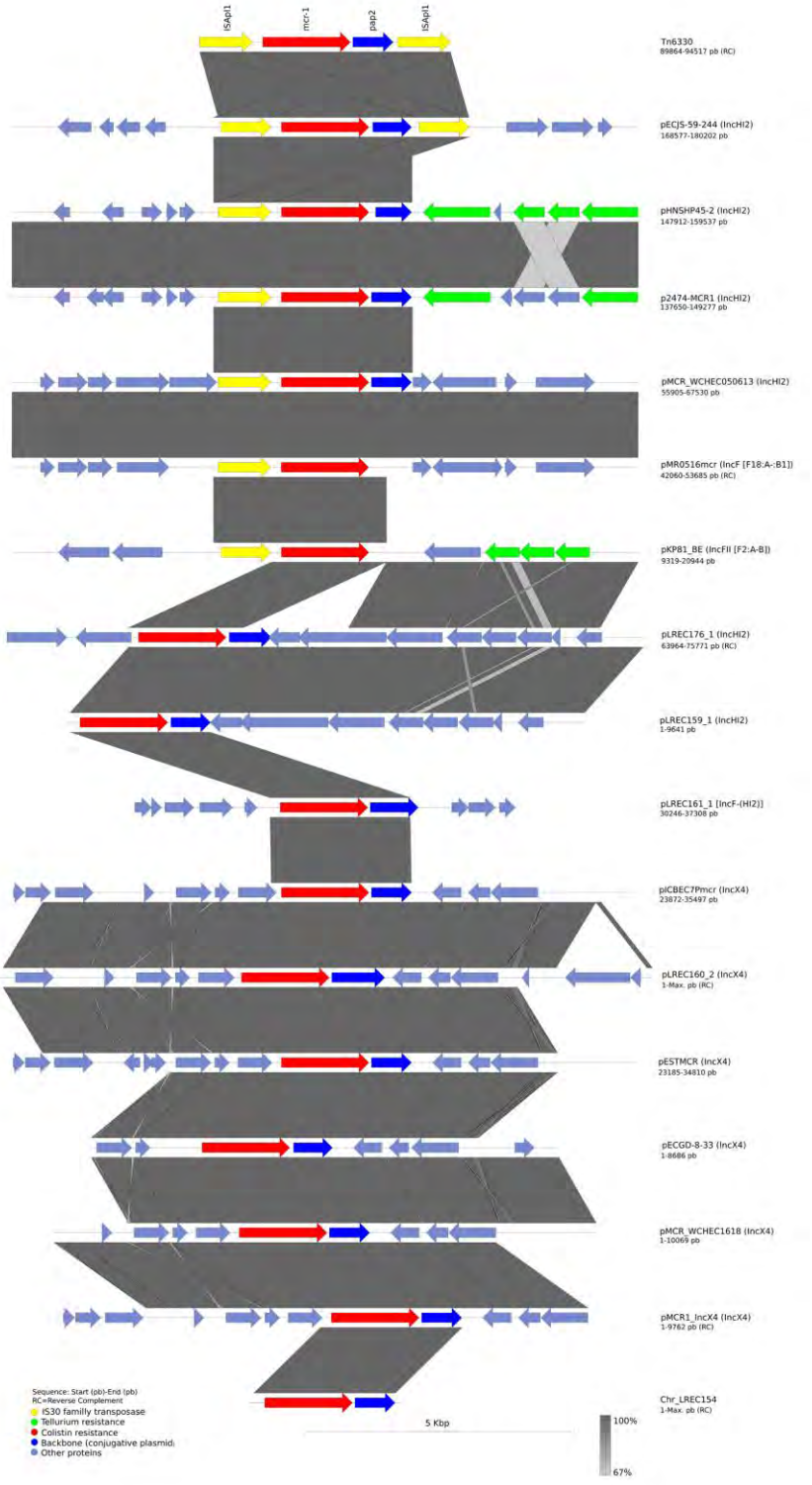


Figure 22. Comparison of homologous region of the composite transposon Tn6330 (CP029493.1) containing *mcr*-1 gene with our sequences and other frequently used reference sequences analysed in this manuscript. Open arrows represent coding sequences (green for tellurium resistance; yellow for IS*Apl*1; red for *mcr*-1 and blue for *pap*2) and indicate direction of transcription. The arrow size is proportional to the gene length. The shadow parallelograms denote genetic regions that exhibit sequence homology among different segments. Light shadow denotes regions with a lower level of sequence identity (99%) by BLAST. The image was generated using EasyFig (default parameters).

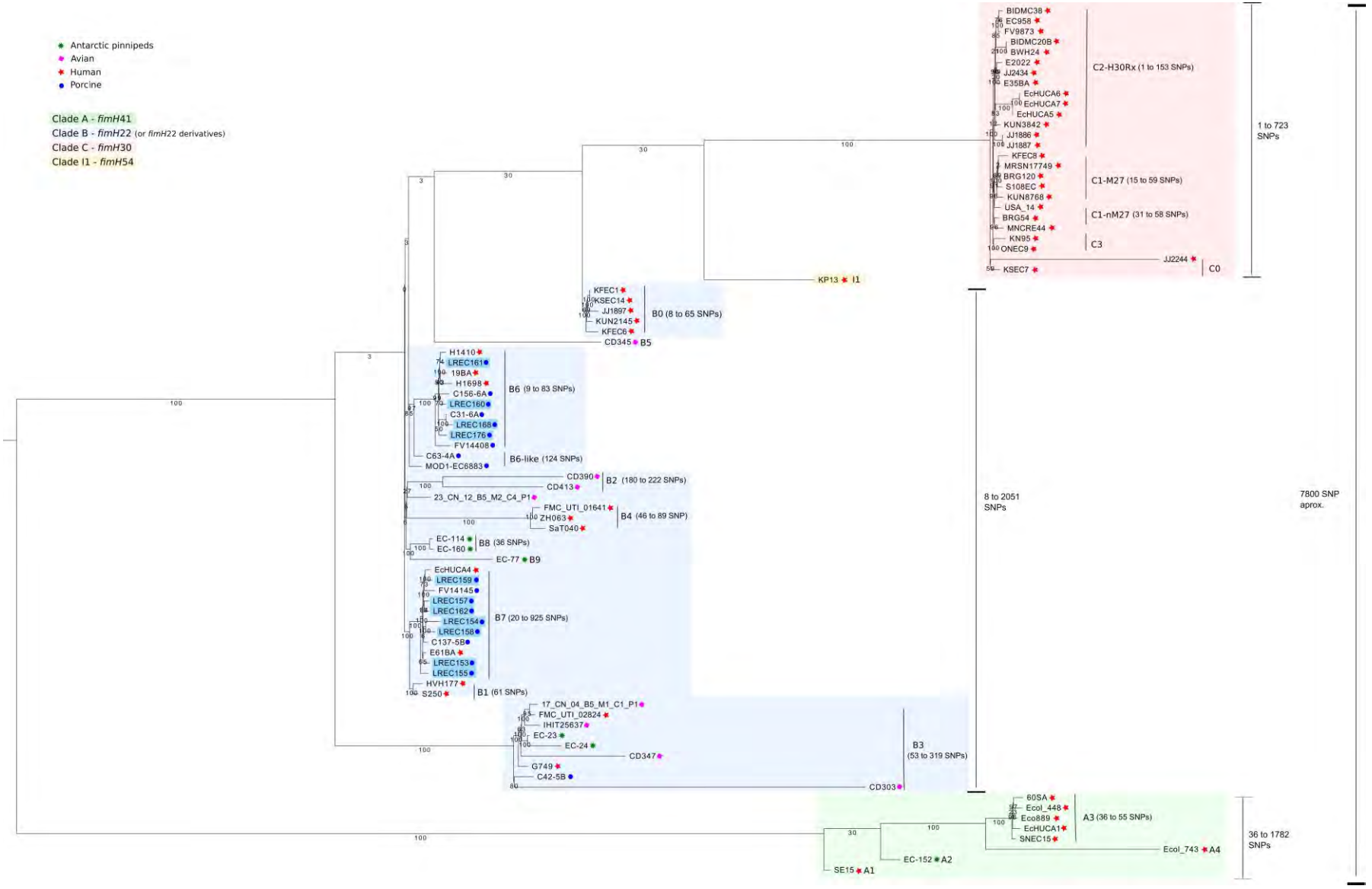


Figure 18. Phylogenetic analysis by core genome of the *E.coli* sequenced in this study (n=11) plus other references (n=73) from ST131 clone.

Table 40. Features, molecular typing, virulence and resistance profiles of ST131 *E. coli* isolates from porcine origin.

Isolate	Origin (year)	Serotype	Phylogroup [a]	Clonotype [b]	MLST [c]	Virulence profile [d, e]	Virotype [f]	Resistance phenotype [g]	MDR+I [h]
LREC_153 (FV11838)	Diahrroea (2008)	O25:H4	B2	CH40-332	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1, hly*A.	D5	AMC, AMP, AMP/SAM, CZ, GEN, NAL, TI, TOB, TMP/SMX	+
LREC_154 (FV9067)	Diahrroea (2006)	O25:H4	B2	CH40-22	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp,*	D-nt [i]	AMC, AMP, AMP/SAM, CHL, COL, NAL, TI	+
LREC_155 (FV14441)	Diahrroea (2010)	O25:H4	B2	CH40-374	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1, hly*A.	D5	AMC, AMP, AMP/SAM, TI	-
LREC_157 (FV14983)	Diahrroea (2010)	O25:H4	B2	CH40-161	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1, hly*A.	D5	AMC, AMP, AMP/SAM, NAL, TI	+
LREC_158 (FV15156)	Diahrroea (2010)	O25:H4	B2	CH40-326	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1, hly*A.	D5	AMC, AMP, AMP/SAM, NAL, TI	+
LREC_159 (FV12310)	Diahrroea (2009)	O25:H4	B2	CH40-338	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1 hly*A.	D5	AMC, AMP, AMP/SAM, COL, GEN, MI, NAL, TI, TOB, TMP/SMX	+
LREC_160 (C153-3A)	Meat (2012)	O25:H4	B2	CH40-298	ST131/PST43	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cdt*B.	D2	AMC, AMP, AMP/SAM, CZ, COL, PI/TZP, TI	+
LREC_161 (C187-6A)	Meat (2012)	O25:H4	B2	CH40-22	ST131/PST43	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cdt*B.	D2	AMC, AMP, AMP/SAM, COL, TI, TMP/SMX	+
LREC_162 (FV14984)	Diahrroea (2010)	O25:H4	B2	CH40-336	ST131/PST9	*iro*N, *iss*, *hly*F, *omp*T, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cnf1.*	D5	AMC, AMP, AMP/SAM, CHL, NAL, TI	+
LREC_168 (C41-4A)	Meat (2011)	O25:H4	B2	CH40-298	ST131/PST43	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cdt*B, *tsh.*	D2	AMC, AMP, AMP/SAM, NAL, TI	+
LREC_176 (C84-4A)	Meat (2011)	O25:H4	B2	CH40-22	ST131/PST43	*iro*N, *iss*, *hly*F, *omp*T, *pap*C, *pap*AH, *pap*EF (GIII), *kpsM*-II (K5), *iuc*D, *iut*A, *chu*A, *fyu*A, *yfc*V, *cva*C, *tra*T, *malX, ibeA, usp, cdt*B.	D2	AMP/SAM, COL, TI, TMP/SMX	+

(a) Phylogroups by Clermont *et al.* (2013); (b) CH, clonotype (*fumC-fimH* alleles); (c) ST, sequence type according to Achtman scheme; PST, sequence type according to the Pasteur scheme; (d) Extraintestinal virulence genes found by PCR; (e) All isolates exhibited the ExPEC, UPEC and APEC status; (f) The virotype was determined by PCR, based on the presence or absence of 13 virulence genes (Dahbi *et al.* 2014); (g) AMP, ampicillin; AMC , amoxicillin-clavulanic acid; AMP/SAM, ampicillin-sulbactam; CZ, ceftazidime; CHL, chloramphenicol; COL, colistin; GEN, gentamicin, MI, minocycline; NAL, nalidixic acid; TI, ticarcillin; TMP/SMX, trimethoprim-sulfamethoxazole; TOB, tobramycin; PI/TZP, Piperacillin-tazobactam; (h) Multidrug drug resistance (MDR); being resistant to at least one agent of ≥3 different antimicrobial categories including resistance to β-lactamase inhibitors (I); (i) nt, non-typeable.

Table 41. Plasmid content of the ST131 *E. coli* isolates from porcine origin and location of resistance and virulence genes.

Isolate	Genetic element[a]	No. of Contigs	Size (bp)	No. of copies[b]	Relaxase protein[c]	Replication protein[c]	Inc group (pMLST)[d]	Virulence genes [e]	Resistence genes[f], point mutations (blue) and efflux/transporter genes (green)
LREC_153	chr_LREC153	40	4,904,938	1	nd[h]	nd	na[g]	*papACDEFHJKX, hlyABD, cnf1, chuA*	PBP , *AmpH, gyrA_S83L, mdfA*
	pLREC153_1	8	330,910	1	MOB_{F12}, MOB_{H11}	RptA1	IncF [F2:A-:B1][IncHI2-ST4][i]	*iroBCDEN, iss, mchF, iutA, iucABCD*	
	pLREC153_2	16	102,056	1	MOB_{P131}	RptZ, RptC	IncL/M [IncQ1][i]		*bla_{TEM-1A}/bla_{TEM-54}/$bla_{TEM-150}$, AAC(3')-IIa, AAC(3')-IIc, APH(3')-Ia, APH(3'')-Ib, APH(6')-Id, aadA1, strA, strB, sul1, sul2, sul3, dfrA1, dfrA15, tet(A), tet(C), qacEdelta1*
	pLREC153_3	1	6,851	5	MOB_{P51}	nd	ColE1-like		
	pLREC153_4	1*	2,461	4	MOB_{V2}	RepL	nd		
LREC_154	chr_LREC154	111	4,946,997	1	nd	nd	na	*papBCDEFHIJKX, iss, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, gyrA_S83L, mdfA , mcr-1.1*
	pLREC154_1	2	124,513	1	MOB_{F12}	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	
	pLREC154_2	27	56,729	3	MOB_{P3}	RptF	IncX4		*bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, tet(A), tet(C), tet(M), tet(R)*
LREC_155	chr_LREC155	60	4,947,621	1	nd	nd	na	*papACDEFHJKX, hlyABD, cnf1, chuA, mchC, iha, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, mdfA*
	pLREC155_1	25	316,521	1	MOB_{F12}	RptA1	IncF [F2:A-:B1][IncHI2-ST-nt][i]	*iroBCDEN, iss,cma, cba, mchF, iutA, iucABCD*	
	pLREC155_2	23	208,964	1	MOB_{P12}	RptZ	IncI1-ST27		*bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, tet(A), tet(C), tet(R)*
LREC_157	chr_LREC157	79	4,896,484	1	nd	nd	na	*papACDEFHJKX, hlyABD, cnf1, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, gyrA_S83L, mdfA*
	pLREC157_1	12	137,775	1	MOB_{F12}	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	*bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$*
	pLREC157_2	1	46,172	1	MOB_{P3}	RptF	IncX1		
	pLREC157_3	1	6,851	3-4	MOB_{P51}	nd	ColE1-like		
	pLREC157_4	1*	2,047	5	MOB_{V2}	nd (RepA partial hit)	nd		
LREC_158	chr_LREC158	58	4,853,444	1	nd	nd	na	*papACDEFHJKX, hlyABD, cnf1, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, gyrA_S83L, mdfA*
	pLREC158_1	5	135,569	1	MOB_{F12}	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	*tet(A), tet(C), tet(R)*

Table 41. *Cont.*

Isolate	Genetic element[a]	No. of Contigs	Size (bp)	No. of copies[b]	Relaxase protein[c]	Replication protein[c]	Inc group (pMLST)[d]	Virulence genes [e]	Resistence genes[f], point mutations (blue) and efflux/transporter genes (green)
LREC_159	pLREC158_2	1*	6,648	4	MOB_{P51}	nd	ColE1-like		
	pLREC158_3	1*	1,554	5	nd	RptL2	nd (Cryptic_1)		
	chr_LREC159	97	4,869,944	1	nd	nd	na	*papACDEFHJKX, hlyABD, cnf1, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, gyrA_S83L, mdfA*
	pLREC159_1	10	248,461	1	MOB_{H11}	RptC	IncHI2-ST4 [IncQ1] [i]		*APH(3')-Ia, APH(3'')-Ib, APH(6)-Id, strA, strB, sul2, sul3, mcr-1.1*
	pLREC159_2	9	177,070	1	MOB_{F12}	RptA1	IncF[F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, *aadA1, AAC(3)-IIa, AAC(3)-IIc, sul1, sul2, dfrA1, drfA15,tet(A), tet(C), tet(R), qacEdelta1*
	pLREC159_3	1	68,523	1	MOB_{P12}	RptZ	IncI1-ST171		
	pLREC159_4	1	2,014	2	MOB_{Q12}	nd	nd		*aadA2, ANT(3'')-Ia, linG, lnu(F)*
	pLREC159_5	1*	1,553	4	nd	RptL2 (HTH36 family)	nd (Cryptic_1)		
	pLREC159_6	1*	1,507	3	nd	nd (RepA partial hit)	nd (Cryptic_2)		
LREC_160	chr_LREC160	107	5,049,024	1	nd	nd	na	*papCDEFJK, sfaX, mchCF, iha, iss, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, mdfA*
	pLREC160_1	18	145,187	1	MOB_{F12}	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, *tet(A), tet(C), tet(R)*
	pLREC160_2	5	40,221	2	MOB_{P3}	RptF	IncX4		bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, *mcr-1.1*
	pLREC160_3	1	5,630	3	MOB_{P51}	nd	ColE1-like		
	pLREC160_4	1	5,006	4-5	MOB_{P51}	nd	ColE1-like		
	pLREC160_5	1	4,515	3-4	MOB_{P51}	nd	ColE1-like		
	pLREC160_6	1	1,765	3-4	MOB_{P51}	nd	ColE1-like		
	pLREC160_7	1*	1,551	4	nd	RptL2	nd (Cryptic_1)		

Table 41. *Cont.*

Isolate	Genetic element[a]	No. of Contigs	Size (bp)	No. of copies[b]	Relaxase protein[c]	Replication protein[c]	Inc group (pMLST)[d]	Virulence genes [e]	Resistence genes[f], point mutations (blue) and efflux/transporter genes (green)
LREC_161	chr_LREC161	64	5,072,645	1	nd	nd	na	*papCDEFJK, sfaX, iss, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, mdfA*
	pLREC161_1	8	330,357	1	MOB_{F12}	RptA1	IncF [F2:A-:B1][IncHI2-ST-nt][i]	*iroBCDEN, iss, mchF, iutA, iucABCD*	bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, *mcr-1.1, tet(A), tet(C), tet(R)*
	pLREC161_2	15	33,473	2	nd	RptF	nd		bla_{TEM-1C}, *aadA2, aadA17, ANT(3'')-Ia, sul3, drfA12, linG, lnu(F), mef(B)*
	pLREC161_3	1	6,826	3	MOB_{P51}	nd	ColE1-like		
	pLREC161_4	1*	1,553	4	nd	RptL2 (HTH36 family)	nd (Cryptic_1)		
LREC_162	chr_LREC162	46	4,720,511	1	nd	nd	na	*chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, gyrA_S83L, mdfA*
	pLREC162_1+2[J]	29	216,025	1	MOB_{F12} (2)[J]	RptA1	IncF [F10:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$
LREC_168	chr_LREC168	113	5,117,208	1	nd	nd	na	*papCDEFJK, sfaX, iss, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, gyrA_S83L, mdfA*
	pLREC168_1+2[J]	13	144,861	1	MOB_{F12} (2)[J]	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	bla_{TEM-1C}/bla_{TEM-40}/$bla_{TEM-135}$, *tet(A), tet(C), tet(R)*
LREC_176	chr_LREC176	145	5,039,239	1	nd	nd	na	*papDFJKX, sfaX, iss, chuA, iroBCDEN, iucABCD, iutA*	PBP, *AmpH, mdfA*
	pLREC176_1	15	278,079	1	MOB_{H11}	RepFIB_RepA (x2)	IncHI2-ST4		bla_{TEM-1A}/$bla_{TEM-150}$, *aadA1, aadA2, APH(6)-Id, APH(3'')-Ib, strA, strB, sul1, sul2, sul3, drfA1, drfA15, mcr-1.1, tet(A), tet(C), tet(R), cmlA1, cmlA6, catA1, qacEdelta1, qacHR*
	pLREC176_2	9	132,494	1	MOB_{F12}	RptA1	IncF [F2:A-:B1]	*iroBCDEN, iss, mchF, iutA, iucABCD*	

[a]Chr, chromosome; p, plasmid. [b]Approximative number of plasmid copies based on sequencing coverage identified by PLACNET. [c]Relaxase and replication proteins were identified by the use of PLACNET databases. [d]Incompatibility groups were determined according to the PBRT scheme and pMLST subtypes according to the allele scheme in http://pubmlst.org/plasmid/. [e]Virulence genes were detected through VirulenceFinder and VFDB databases (minimum coverage of 90%, minimum identity of 75%); All genomes presented in their chromosome the following genes: *aslA, cheBWY, chuSTUVWXY, csgBDEFG, entABCEFS, fdeC, fepABCDG, fes, fimABCDEFGHI, flgG, flhAC, fliGIMP, fyuA, gad, gspCDEFGHIJKLM, ibeA, irp1, irp2, kpsDM, OmpA, yagVWXYZ/ecpEDCBA, ybtAEPQSTUX, ykgK/ecpR*. [f]Antimicrobial resistance genes were detected by using ARG-ANNOT, CARD and ResFinder databases (minimum coverage of 93%, minimum identity of 75%); All genomes presented in their chromosome the following genes (efflux pumps and transport modulators): *acrAEFS, bacRS, cpxA, CRP, emrABEKRY, eptA, evgAS, gadWX, H-NS, kdpE, marA, mdtABCEFGHKNOP, msbA, pmrF, tolC, ugd, yojI*. [g]na, not-applicable; [h]nd, not-detected. [i]Cointegrated plasmid. [j]Two plasmids that could not be separated (two MOB_{F12}). *Closed plasmids.

5. DISCUSSION

5.1. STRAINS OF HUMAN ORIGIN

5.1.1 CLONAL STRUCTURE OF *E. COLI* CAUSING EXTRAINTESTINAL INFECTIONS IN HUMANS

To our knowledge, this is the first study conducted in Spain in which the clonal structure of extraintestinal pathogenic *E. coli* was analyzed from clinical *E. coli* non-redundant and consecutively isolated. Indeed, the previous Spanish studies focused on selected clinical isolates. Thus, Blanco *et al.* (2011) evaluated the incidence of only three high-risk clones among 500 consecutively collected *E. coli* isolates, causing extraintestinal infections in five Spanish hospitals in 2009. They found that ST131 accounted for 12%, ST393 for 3% and ST69 for 4%, and these three clones accounted for 30% of the MDR isolates. In the present study, ST131 was one of the most prevalent ST in the enrolled Spanish hospital and accounted for 12%. This result strongly suggests that the ST131 rate has remained stable between 2009 and 2016 in Spain. The ST69 rate seems to have increased as this ST accounted for 8% in 2016, compared to 4% in 2009. Inversely, the non-detection of any ST393 isolate in the current study seems to indicate that its rate has declined dramatically in Spain between 2009 and 2016. Oteo *et al.* (2014) analyzed the ST distribution among OXA-1-producing *E. coli* isolates resistant to amoxicillin-clavulanate collected from clinical samples (>70% from urine), from seven Spanish hospitals in 2010 and found that ST88 (37.3%) and ST131 (32.8%) were the most prevalent STs in this *E. coli* subgroup. Interestingly, we found in the present study that ST88 (9%) was one of the most common STs after ST131 in Spain but not in France (3.1%).

In France, Van der Mee-Marquet *et al.* (2015) studied the genetic diversity of 412 *E. coli* bloodstream isolates recovered during 2014 and found 12 major ST complexes (STCs 10, 12, 14, 23, 31, 69, 73, 92, 95, 131, 141 and 155), containing 77.9% of the collected isolates. Clermont *et al.* (2017) analyzed 243 bacteremia isolates obtained in 2010 in the Paris area and found four most prevalent STCs: STC95 (14.0%), STC73 (13.2%), STC131 (8.6%) and STC69 (7.8%). La Combe *et al.* (2019) analyzed 260 respiratory tract isolates obtained from patients who had pneumonia and were hospitalized between 2012 and 2014 in intensive care units located throughout France. The most commonly identified lineages were STC73 (15.4%), STC131 (9.2%), STC69 (7.7%), STC141 (6.2%), STC127 (6.2%) and STC95 (5.4%). They compared this STC distribution to that observed in bacteremia and commensal isolates and noted that STC127 and STC141 were overrepresented and STC95 underrepresented in pneumonia isolates compared with bacteremia isolates. Brisse *et al.* (2012) compared the phylogenetic diversity of 152 CTX-M-producing and 152 non-ESBL-producing clinical *E. coli* isolates, obtained between 2008 and 2009 from ten hospitals located in the Paris area. The five most prevalent STs were distributed differently between CTX-M-producing and non-ESBL-producing *E. coli* isolates: ST131 (36.2% vs. 9.9%), ST10 (8.6% vs. 5.9%), ST73 (0% vs. 10.5%), ST95 (0% vs. 5.3%) and ST141 (0% vs. 4.6%). In the present study, the most prevalent STs in the enrolled French hospital were ST73 (12.5%), ST131 and ST141 (11.5% for each), ST10 (6.3%), ST69 (5.2%), ST12, ST95 and ST127 (4.2% for each). The comparative analysis of the present study and the four previously published studies shows that the distribution of the most prevalent STs appears to have been stable in France during the period-time covered by the five studies (2008-2016). However, our study highlighted that ST141, which was already present in 2008 (Gati *et al.*, 2019), has not only persisted in any type of source, except for blood, but has increased to become as prevalent as ST131, although it does not belong to the MDR-ST group.

The clonal structure of different collections of *E. coli* clinical isolates has also been investigated in other countries of the European continent. Thus, different studies were

conducted in UK, focusing either on UTI isolates (2007-2009) (Gibreel *et al.*, 2012) or bacteremia isolates (2001-2012) (Horner *et al.*, 2014; Day *et al.*, 2016; Kallonen *et al.*, 2017). Independently of the sample source and the location of the enrolled centers, the distribution of the predominant STs was the same in UK, with some variations in terms of frequency. Thus, ST73 varied from 16.6% to 18.6%, ST131 from 12% to 16.8%, ST69 from 5.4% to 10.5%, ST95 from 6.3% to 10.6% and ST12 from 4.4% to 4.6%. ST10 (4.3%) and ST127 (3.6%) were identified as prevalent ST among only the UTI isolates and ST12 only among bacteremia isolates. Kallonen *et al.* (2017) highlighted that after the emergence of ST69 (2002) and ST131 (2003) and their spread, a new equilibrium of *E. coli* populations was observed, resulting in a relative stability of the major STs. The notable difference between the UK studies and ours is the absence of ST88 and ST141 among the predominant STs in UK, whereas the rate of ST88 was high in Spain (9%) and that of ST141 in France (11.5%). In contrast, both ST88 and ST141 were identified in Germany among the prevalent STs and each accounted for 4.2% of 265 UTI isolates collected between 2004 and 2006 (Toval *et al.*, 2014). ST141 was also identified among 44 *E. coli* isolated from UTIs in Switzerland during 2016. It accounted for 11.4% of the isolates, whereas ST131 and ST69 accounted, each, for 13.6%, and ST73 for 6.8% (Nüesch-Inderbinen *et al.*, 2017). These features suggest that the clonal structure of *E. coli* in Germany and Switzerland seems to be closer to that of Spain and France than to that of UK.

Some studies were also conducted in North America. In Canada, Fibke *et al.* (2019) analyzed the genome of *E. coli* isolates responsible for UTI, in 385 women between 2012 and 2015. The major STs included ST95 (18.4%), ST73 (10.1%), ST127 (9.1%), ST131 (8.8%) and ST69 (7.5%). Among the studies performed in the USA on UTI isolates, different ST distributions were observed. Thus, Banerjee *et al.* (2013a) studied 299 clinical extraintestinal *E. coli* isolates (90% from UTIs) obtained during 2011 in Minnesota, and identified the five most prevalent STs, consisting of ST131 (27%), ST95 (11%), ST73 (8%), ST127 (6%) and ST69 (5%). On their side, Yamaji *et al.* (2018), who studied 233 *E. coli* isolated from UTIs in California during 2016 and 2017, and 225 isolates collected similarly between 1999 and 2000, showed that the ST131 isolates were less frequently identified than the ST95, ST127, ST73, and ST69 isolates in these two periods. Similar to UTI isolates, bacteremia isolates were differently distributed in ST according to the studies. Thus, Adams-Sapper *et al.* (2013), analyzing 220 bloodstream isolates obtained in San Francisco between 2007 and 2010, found only five prevalent STs and clonal complexes, accounting for 65% of the isolates. They comprised ST131 (23%), ST95 (18%), STC73 (8%), ST69 (9%) and STC12 (6%). On their side, Cole *et al.* (2019), studying 43 bacteremia *E. coli* isolates from newborns obtained between 2006 and 2016 in Oklahoma, found that ST95 was the most prevalent ST (11.3%) followed by ST131 (9.2%) and ST1193 (3.7%). Cole *et al.* (2019) noted that ST95 and ST131 isolates were present throughout the studied years, while ST1193 was only seen in the recent years. Overall, neither ST88 nor ST141 belong to the most prevalent clones in North America. In North America, ST95 is more prevalent than ST73, while, in Europe, the opposite is true.

Similar to Cole *et al.* (2019), we recently identified, i.e. in 2016, the global emergent MDR ST1193. To our knowledge, it is the first identification of this clone in Spain and the second in France. Indeed, Birgy *et al.* (2019) found that among 218 ESBL-producing *E. coli* causing febrile UTI in children between 2014-2017, ST1193 was one of the most prevalent clones during the most recent period of the study.

ST1193 seems to be more prevalent in Asia than in Europe and the USA. Indeed, Chen *et al.* (2019a), who characterized 100 bloodstream *E. coli* isolates from Zhejiang

(China) in 2015 showed that, among the most prevalent clones, ST131 (15.3%) was followed by ST1193 (7.1%), then, by ST95 (5.9%) and ST69 (5.9%). This ST distribution suggests that strong antibiotic pressure has displaced the frequently antimicrobial susceptible STs among blood *E. coli* isolates, such as ST73 and ST95. However, in China, ST95 isolates have not been fully displaced, since some isolates have acquired plasmids encoding ESBL enzymes.

To our knowledge, the analysis of the clonotypes (analysis of the *fumC* and *fimH* alleles: CH), in addition to the phylogroup and ST types, allowing clone determination, was only carried out by Tchesnokova *et al.* (2013) in the USA and Matsumura *et al.* (2017a) in Japan. Tchesnokova *et al.* (2013) found 222 CH clonotypes among 1518 *E. coli* isolates (93% from UTIs), recovered between 2010 and 2011. Matsumura *et al.* (2017a) found 103 clonotypes among 329 *E. coli* (65% from UTIs and 11.9% from bacteremia), collected from 10 Japanese hospitals in 2014. In our study, 107 clonotypes were identified among 196 isolates, which suggests a higher diversity of clones in our study than in the American and Japanese studies. However, the most remarkable difference between the three studies is the distribution of the most prevalent STs among the MDR isolates: B2-CH40-30-ST131 and D-CH35-27-ST69 clones in the USA, B2-CH40-30-ST131 clone in Japan, and B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 and D-CH35-27-ST69 clones in Spain and France. Consequently, it can be noted that ST88 is not a predominant ST in Japan, like in UK and the USA.

Fibke *et al.* (2019) provided data on the *fimH* alleles displayed by the 34 Canadian ST131, that they identified among UTI isolates: *fimH*22 (20.6%), *fimH*27 (5.9%), *fimH*30 (23.5%) and *fimH*41 (44.1%). In our study, the distribution of the *fimH* alleles among our 23 ST131 isolates was very different: *fimH*22 (13.0%), *fimH*30 (56.5%), *fimH*41 (26.1%) and *fimH*298 (4.3%). Kallonen *et al.* (2017), based on single nucleotide polymorphisms (SNPs), identified eight clades among English ST73 isolates associated with different O:H serotypes. Using the *fimH* allele type, we identified six *fimH* alleles among our 23 ST73 isolates (*fimH* 10, 12, 13, 27, 30, 32 and 103), that defined six clones, among which B2-CH24-10-ST73 and B2-CH24-103-ST73 clones predominated. Gordon *et al.* (2017b) studied the phylogeny of 200 STC95 isolates from humans living in France, Australia and the USA. The SNPs analysis revealed five main clusters associated to different *fimH* alleles and serotypes. In the present study, we have also identified five clones among our 10 ST95 isolates with different *fimH* alleles (*fimH* 15, 27, 30, 41 and 54) and serotypes (O1:H7, O2:H4, O2:H5, O2:H7, O18:H7, O25:H4 and O45:H7). Here, the most common clone was the B2-CH38-15-ST95, displaying serotypes O2:H5 (1 isolate), O2:H7 (1), O18:H7 (2).

Taking into consideration that Gati *et al.* (2019) identified hybrid shigatoxin-producing *E. coli* /UPEC isolates among ST141 isolates, we searched for these hybrid isolates among our 14 ST141 isolates, but no hybrid isolates were detected. However, we detected a French ExPEC isolate of serotype O153:HNT, that carried three VFs-encoding genes (*aatA*, *aaiC* and *aggR*), specific for EAEC. This hybrid belonged to A-CH11-54-ST10 clone and harbored 14 VFs typically found in ExPEC. Interestingly, Olesen *et al.* (2012) found that the EAEC isolates of serotype O78:H80 and ST10 were responsible for an outbreak of UTI in Denmark. On their side, Abe *et al.* (2008) and Lara *et al.* (2017) found some UPEC isolates with properties of EAEC in Brazil, including hybrid UPEC/EAEC ST69, ST73 and ST131 isolates.

In summary, approximately 10% of the extraintestinal *E. coli* infections that had occurred in 2016 in the two studied hospitals were caused by ST131 isolates, and approximately 60% of these infections were caused by isolates belonging to only 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131, ST141). ST88 and ST141

isolates were particularly frequent in the Spanish and French hospitals, respectively, while, so far, these two STs are absent among the dominant STs in UK, the USA and Asia but present in Germany and Switzerland. Our results confirm that in Europe, ST73 is much more prevalent than ST95, while in North America it is the opposite. The majority of ST12, ST73, ST95 and ST141 isolates were susceptible to most antibiotics, indicating that MDR was not the reason for their success. The results of the present study support the idea that their success is mainly due to the high number of VF-encoding genes that they possess. This study also shows that among the MDR isolates, four clones are predominant, especially: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 and D-CH35-27-ST69. Clone B2-CH40-30-ST131 was also the most prevalent clone among the ESBL-producing isolates. Lastly, this study confirmed the presence of the new MDR global emergent ST1193 in both France and Spain. All these results suggest that surveillance of the clonal structure and antibiotic susceptibility of ExPEC is required at both local and global levels, notably to evaluate the evolutive impact of the antibiotic overuse on one of the most important human bacterial pathogens.

5.1.2 EXTENDED-SPECTRUM B-LACTAMASE-PRODUCING *E. COLI* CAUSING EXTRAINTESTINAL INFECTIONS IN HUMANS

The management of urinary tract and bloodstream infections due to *E. coli* has been complicated by the emergence of multidrug-resistance, especially of that related to the expansion of high-risk clones such as ST131 (Nicolas-Chanoine *et al.*, 2007b; de Toro *et al.*, 2017). Since 2006, the prevalence of ESBLEC among *E. coli* causing bacteremia has raised in Lucus Augusti hospital. This increase has been due to the spread of the multidrug-resistant ST131 subclade C2 associated with the production of CTX-M-15. Thus, the number of ESBLEC isolates increased from 1.0% during 2000-2005 to 5.5% during 2006-2011. While during the first period 0% of the ESBLEC isolates belonged to subclade C2, during the second period this subclone represented 39.8% (Mamani *et al.*, 2019). A similar situation has been reported in France in different hospitals (Brisse *et al.*, 2012; Sauget *et al.*, 2016) and worldwide (Peirano *et al.*, 2012).

The main change with respect to previous studies conducted at the Lucus Augusti hospital, is the emergence of isolates producing CTX-M-27. Indeed, this enzyme was not produced by any of the 105 ESBLEC isolates recovered from extraintestinal infections between 2006 and 2007 (Blanco *et al.*, 2009) and by any of the 92 ESBLEC bloodstream isolates collected from 2001 to 2011 (Mamani *et al.*, 2019) and only by one of 47 ST131 ESBLEC isolated from urinary tract infections in 2012 (Dahbi *et al.*, 2013). Furthermore, the CTX-M-27 was only detected in one of the 44 Spanish hospitals analyzed in a study carried out in 2006 (Diaz *et al.*, 2010), in one of 94 clinical ESBLEC collected during 2008 in Vall d´Hebron hospital of Barcelona (Coelho *et al.*, 2011), and in none of the 92 ESBLEC obtained in eight Spanish hospitals during 2010 and 2011 (Merino *et al.*, 2016). In the present study, CTX-M-27 was significantly more frequent in Beaujon (15.6%) than in Lucus Augusti (6.5%). However, the Beaujon's CTX-M-27 percentage appeared as remarkably higher than those found before 2014 (4.5%-5.4%) in other 19 hospitals located as Beaujon in the Paris area (Robin *et al.*, 2017; Surgers *et al.*, 2018). Inversely, this percentage was closer to that found between 2014 and 2017 in 24 pediatric centers located in six French regions, *i.e.* 12.4% among 251 ESBLEC isolated from febrile urinary tract infections (FUTIs) (Birgy *et al.*, 2019). In addition, it has to be noted that CTX-M-27-producing isolates had already been found in 2012 in the feces of children in day care centers (DDCs) in France (Blanc *et al.,* 2014) and also in feces of patients hospitalized in Madrid during a European survey conducted between 2014 and 2015 (Merino *et al.*, 2018).

The increased prevalence of the CTX-M-27 in the two hospitals enrolled in the present study was mainly due to the expansion of cluster C1-M27 since 17 of the 21 positive CTX-M-27 isolates belonged to this cluster. The four remaining isolates belonged to three different STs, including ST38 (2 isolates), ST90 and ST1193. In the FUTI study, Birgy *et al.* (2019) also showed that their 31 CTX-M-27-producing isolates mostly belonged to cluster C1-M27 (10 isolates) and ST38 (8 isolates), and also that two belonged to ST1193.

The two ST1193 isolates identified in the present study belonged to clonotype CH14-64 and serotype O75:HNM and were positive for CTX-M-14 (the isolate from Spain) and CTX-M-27 (the isolate from France). As far as we know, this is the first report of the new emerging global clone ST1193 among clinical ESBLEC isolates from Spain. Inversely, the ST1193 has already been described in France and shown producing either CTX-M-15 or CTX-M-27 in the pediatric FUTIs study (Birgy *et al.*, 2019) and CTX-M-14 among the fecal isolates obtained from children in DCCs in France (Blanc *et al.*, 2014). ST1193 has also been detected in two of 243 third-generation-cephalosporin-resistant *E. coli* isolates obtained from patients with bloodstream infection in Denmark during 2015 (Roer *et al.*, 2018), in three of 51 clinical ESBLEC isolated in Germany during 2015 and 2016 (Valenza *et al.*, 2019) and in 11 of 225 cefotaxime-resistant *E. coli* isolated from UTIs in South-West England during 2017 and 2018 (Findlay *et al.*, 2020). ST 1193 is currently much more expanded in China (Xia *et al.*, 2014; Wu *et al.*, 2017) and the USA (Johnson *et al.*, 2019; Tchesnokova *et al.*, 2019).

Despite the emergence of ST1193 and ST131 cluster C1-M27, it is clear that ST131 subclade C2 associated to CTX-M-15 remains the most prevalent sublineage among ESBLEC in the two hospitals studied here, as it is the case in most of the European hospitals (Merino *et al.*, 2016; Sauget *et al.*, 2016; Roer *et al.*, 2018; Birgy *et al.*, 2019; Valenza *et al.*, 2019; Findlay *et al.*, 2020). The C2 subclade isolates showed a higher virulence score (mean 14.250) compared with the subclade C1-M27 ST131 isolates (mean 10.875) ($P<0.001$) and non-ST131 isolates (mean 7.661) ($P<0.001$). Interestingly, the *papAH, papC, papEF, cnf1* and *hlyA* genes were associated with the C2 subclade isolates, which mostly displayed the vitotypes A, A-like, C2, E and F. The virotype A-like is new and differs from the virotype A in the type of capsular *kpsM II* gene that is *K*5 instead of *K*2 (Dahbi *et al.*, 2014). In future studies, it would be very interesting to determine the whole genome sequence of the C2 subclade isolates belonging to the new virotype A-like.

Of note, four isolates belonging to clones B2-CH38-15-ST95 and B2-CH38-294-ST95 were those with the highest number of VF-encoding genes (mean 21.000). These four UTI isolates were classified as APEC and could be of avian origin and foodborne pathogens (Vincent *et al.*, 2010; Singer, 2015; Liu *et al.*, 2018a). ST95 is one of the most frequently ST identified among *E. coli* causing human extraintestinal infections, but it is rarely producer of ESBL enzymes (Kallonen *et al.*, 2017). Nevertheless, recently Birgy *et al.* (2019) detected nine ST95 isolates among 251 ESBLEC from pediatric FUTIs. The combination of so many virulence genes and resistance-encoding genes in this successful ST is very worrying.

Despite the enormous genetic diversity observed in our ESBLEC collection (71 clones amongst 188 ESBLEC), it can be concluded that the majority of the isolates belong to only three clonal complexes (CC10, CC23, and CC131) and that ST131 subclade C2 associated with the production of CTX-M-15 remains the most prevalent *E. coli* lineage among the ESBLEC isolates identified in the studied Spanish and French hospitals.

5.1.3 BIOFILM PRODUCTION IN *E. COLI* CAUSING EXTRAINTESTINAL INFECTIONS IN HUMANS

Here, we assessed the early biofilm production, i.e., adhesion stage of 394 *E. coli* clinical isolates obtained from different sources, at the same time periods, in two distant hospitals (Spanish and French hospitals). This study is in accordance with previous researches with regard to the wide spectrum of biofilm formation within ExPEC (Ponnusamy *et al.*, 2012; Agarwal *et al.*, 2013; Shokouhi Mostafavi *et al.*, 2018). Using the BioFilm Ring Test® like us, Olivares *et al.* (2016) found that *P. aeruginosa* causing infection in cystic fibrosis patients segregated into three groups (strong, moderate, and weak producers) at the early stage of biofilm formation. Accordingly, difference in ability to quickly adhere seems to exist within isolates of different bacterial species. The biofilm production displayed by our ESBLEC collection is also in accordance with that displayed by the ESBLEC collection recently published by Surgers *et al.*, (2019), i.e., predominance of weak biofilm producers among ESBLEC.

The novelty provided by our study consisted of the molecular characterization of the 394 isolates with regard to phylogroups, clonal lineages, VF-encoding genes, and FimB, a co-factor of the regulation of the synthesis of type 1 fimbriae that are critical adhesins for *E. coli* biofilm formation (Beloin *et al.*, 2008). Thus, we showed that among the characteristics significantly associated with EBF, clonal lineage was the most suitable characteristics to explain the variability of EBF. Association between clonal lineage and EBF phenotype has recently been shown among bone and joint infection *Staphylococcus aureus* isolates analyzed by using the BioFilm Ring Test® (Tasse *et al.*, 2018).

Among the limited number (<10%) of our *E. coli* that quickly and persistently produced early biofilm at high levels (G1 isolates), five clones (B2-CH14-2-ST127, CH52-5, and CH52-14 sub-lineages of B2-ST141, B2-CH40-22-ST131, and E-CH100-96-ST362) accounted for 41.7% of the 36 G1 isolates. Notable features have previously been provided by different studies for these clones. The B2-ST127 clone, which is one of the dominant clones among uropathogenic isolates (Yamaji *et al.*, 2018), was shown to be shared by humans, dogs, and cats (Johnson *et al.*, 2008b) and to be a dominant gut colonizer of humans (Ulleryd *et al.*, 2015) and fruit bats (Nowak *et al.*, 2017). The B2-ST141 clone was characterized by Clermont *et al.* (2017) as a commensal clone, i.e., digestive tract resident with low level of human invasiveness, through a study assessing the pathogenesis of bacteriaemic *E. coli* by matching large collections of bacteriaemic (Lefort *et al.*, 2011) and commensal isolates (Massot *et al.*, 2016). Other studies showed that this clone accounted for the gut dominant *E. coli* population of some healthy subjects (Leflon-Guibout *et al.*, 2008; Nicolas-Chanoine *et al.*, 2013) and was one of the dominant clones in the digestive tract of Antarctic pinnipeds (Mora *et al.*, 2018). The B2-CH40-22 ST131 clone was shown to have similar features to the B2-CH52-5 ST141 clone with regard to healthy subjects (Nicolas-Chanoine *et al.*, 2017) and Antarctic pinnipeds (Mora *et al.*, 2018). It was also found in poultry digestive tract and retail chicken meat and was shown to be a foodborne uropathogen (Liu *et al.*, 2018a). Concerning the E-CH100-96 ST362 clone, our study is, to our knowledge, the first one reporting the presence of this clone in humans. Indeed, it was identified so far in chicken (Cortés *et al.*, 2010) and as an agent causing bovine mastitis (Freitag *et al.*, 2017). Accordingly, most of these clones appear to be both ExPEC and intestinal dominant colonizers in some mammals and birds.

We found that some VF genes, including *papC, sfa/focDE, cnf1, hlyA*, and *ibeA*, were mostly identified in G1 isolates, i.e., strong biofilm producers. All these VFs had already been shown to be associated with *E. coli* strong biofilm producers (Naves *et al.*, 2008), and two of them, *sfa/focDE* and *ibeA*, had been shown to be characteristic traits of adherent/invasive *E. coli* (AIEC) that are strong biofilm producers (Martinez-Medina *et*

al., 2009b). The implication of AIEC in Crohn's disease (Darfeuille-Michaud *et al.*, 2004) was notably based on their high biofilm production. Indeed, in an inflamed intestinal environment, their biofilm-related mucosal attachment could allow them to penetrate the inner mucus layer and adhere to the epithelial surface, thus enabling more direct stimulation of epithelial and immune cells (Glasser *et al.*, 2001; Sartor, 2018). It is interesting to note here that Martinez-Medina *et al.* (2009a) found two ST131 isolates carrying the *ibeA* gene in the AIEC collection that they compared with ExPEC isolates. This finding suggests that it would be of interest to investigate all the AIEC phenotypic and genotypic features in our strong biofilm producers (Camprubí-Font *et al.*, 2019).

The present study in which a wide collection of *E. coli* isolates was analyzed confirms the results that we previously obtained with few isolates about the difference in early biofilm production between B2-CH40-22 ST131 and B2-CH40-30-ST131 clones, i.e., strong and weak production, respectively (Nicolas-Chanoine *et al.*, 2017). Knowing that the B2-CH40-22 ST131 clone is the ancestor of the globally disseminated B2-CH40-30-ST131 clone (Ben Zakour *et al.*, 2016), the loss of the ability to strongly produce biofilm at the early step of biofilm formation by the B2-CH40-30-ST131 clone could appear as an evolutionary adaptative trait in ST131 lineage.

In conclusion, this study highlights that the ability to quickly and persistently produce biofilm at high levels is a property displayed by a limited number of clones, some of which have been found in previous studies as dominant colonizers in some mammals and birds.

5.1.4 WHOLE GENOME ANALYSIS OF 75 ESBLEC

Knowledge on the bacterial population structure is necessary to better understand outbreaks and the dynamics of bacterial changes and spreading (Turrientes *et al.*, 2014). Here, we investigate the insights of ST131 clonal complex ESBLEC (STC131 group) and other emerging ESBL lineages (non-STC131 group) through the analysis of a total of 75 ESBLEC genomes.

The phylogenetic research includes two core genome analysis (CG-1 and CG-2). CG-1 includes the 75 ESBLEC genomes and represents a total of 16 different STs. In those genomes, the core includes 1,659 common CDS and reaches a distance of 38,854 SNPs between them. This analysis foregrounds the genomic diversity of ESBLEC population. *E. coli* is the better known example of host-adaptable bacteria species and the remarkable diversity of *E. coli* has been associated with its aptitude to colonize different environments (becoming for example an extraintestinal pathogen) and become a very important contributor to most mammal's microbiome outcompeting other bacteria. Therefore, *E. coli* is the dominant specie amongst the human gut *Enterobacteriae* bacterial population in healthy adults (Martinson and Walk, 2020).

Besides, CG-1 exhibit six clusters with a clear correlation with the phylogroups structure of *E. coli* genomes. The phylogroups structure of *E. coli* has been historically linked with the lifestyle of the strains. The most ancient lineages include strains belonging to B2, D, F and G phylogroups which are associated with extraintestinal infections, afterward phylogroup E had emerged and diverged on phylogroups A, B1 and C which are associated with severe intestinal pathologies such as haemolytic uremic syndrome and dysentery (Escobar-Páramo *et al.*, 2004; Clermont *et al.*, 2019). However, a recent study supports the hypothesis of long-lived resident *E. coli* of the human gut belonging to the phylogroups B2, A and F (Martinson and Walk, 2020). What is clear is that *E. coli* strains are not equal in regard to their ability to colonize different environments or become a gut resident (Martinson and Walk, 2020). We must take into consideration that resident *E.coli* are subject to the exposure of antibiotics during each medical treatment that the host may have required during its life. To date, the question of which factors made some resident

E. coli clones become pathogens it still open for debate. Thus, more research is needed to understand this important aspect of the *E. coli* ecology.

In CG-2 we included 81 genomes from the ST131 clonal complex (our 30 STC131 genomes plus other 51 ST131 genome references). The 81 genomes shared 2,147 CDS that constitute their core genome and reach a distance of 5,794 SNPs between them. ST131 phylogeny is characterised by the presence of a main O, H and *fim*H type related to each ST131 clade. CG-2 exhibit six clusters, cluster 1 includes 49 genomes with *fim*H30 from clade C and reach a distance of 124 SNPs between them, cluster 3 includes 17 genomes with *fim*H22 from clade B and reaches a distance of 888 SNPs between them, cluster 4 includes one genome with *fim*H54 from clade I1 and cluster 6 includes 12 genomes with *fim*H41 from clade A and reaches a distance of 652 SNPs between them. Cluster 2 and 5 constituted two new clades. Cluster 2 includes one genome with the *fim*H5 allele variant and a mean distance of 336 SNPs to the closest branch of genomes in CG-2 (clade C genomes) and cluster 5 includes one genome with the *fim*H35 allele variant and a mean distance of 1,329 SNPs to the closest branches of genomes in CG-2 (clade C and clade B genomes). It is interesting to note how cluster 1, which includes genomes from clade C, exhibits very short intern branches in comparison with its sister clades A and B. Clade C represents the major part of human clinical isolates from the ST131 clone (Nicolas-Chanoine *et al.*, 2014; Pitout and DeVinney, 2017) and it seems that those genomes had a more conserved core genome structure than those from other clades of the ST131 lineage. This population structure certainly reflects some aspects of its ecogenetics, which is highly dependent on recombination and mutational events involving adaptative mobile genes (Wirth *et al.*, 2006; Jaureguy *et al.*, 2008; Tenaillon *et al.*, 2010).

We also analysed the virulome, resistome and plasmidome of the sequenced genomes and confirmed our previous molecular characterisation. There was a clear association between ExPEC status, UPEC status and the STC131 group of genomes (p=<0.0001). Nonetheless, some VF-encoding genes, significantly associated with the STC131 group and related with the ExPEC-UPEC status attribution, were also found in genomes belonging to ST38, ST88, ST95, ST354 and ST1193. Therefore, all ST95 and ST1193 are designated as ExPEC and UPEC. ST38 and ST88 are only ExPEC.

Furthermore, some VF-encoding genes were found to be significantly associated with each defined group of genomes (STC131 or non-STC131). This refers to the toxin encoding gene *sat*, the adhesion encoding genes *afa*CD/*dra*P and *yfc*V, the type II capsule encoding gene *kps*M, the iron uptake encoding genes *iuc*D, *iut*A, *chu*A, *fyu*A and other miscellaneous genes such as *mal*X (pathogenicity-associated island marker), *usp* (uropathogenic-sepcific protein) and *omp*T (outer membrane protein) that were more frequent in the STC131 group (p=<0.010). Besides, a distinct distribution was also observed for some infrequent genes such as the toxin-encoding genes *cnf*1, *hly*A, *hly*B and *hly*D that were only present in the STC131 group (p= <0.003) or the fimbrial adhesion encoding genes *afa*A-VIII, *afa*B-VIII, *afa*C-VIII, *afa*D-VIII, *afa*E-VIII, the enteroaggregative *E. coli* HilA homolog encoding gene *eil*A and the invasion of the brain endothelium encoding gene *ibe*A that were only present in the non-STC131 group (p=<0.037). Going deeper, the *afa* genes codifying for fimbrial adhesin AfaE-VIII were found in the ST1615 genomes, the *eil*A gene was found in the ST38, ST354 and ST648 genomes and the *ibe*A gene in the ST95 and ST354 genomes. Moreover, group II capsule of *E. coli* was present in 53% of the screened genomes and was more frequent in the STC131 group (p<0.0001). K1 type was detected in seven genomes, three ST1193 and four ST95; K2 type in 17 genomes, 16 ST131 (12 from subclade C2 and four from clade

A) and one ST354; and K5 type in 16 genomes, nine ST131 (five from subclade C2 and four from cluster C1M27), four ST38 and three ST648.

In regards to the resistome, we first investigated the chromosome. We outline a high percentage of quinolone resistance-associated point mutation in the chromosome of both groups of genomes.

It is a well-supported hypothesis that the selective pressure exerted by the widespread introduction of fluoroquinolones and extended-spectrum cephalosporins in clinical medicine around the nineties have contributed to the persistence of the chromosomal fluoroquinolone resistance mutations and the acquisition of a specific IncF-*bla*$_{\text{CTX-M}}$ harbouring plasmid among clade C of ST131. Although the evolutionary dynamics of plasmid-host combination remain to be clearly elucidated, it is evident that this association has subsequently favoured the global dispersion of this high risk clone (Stoesser *et al.*, 2016b). Thus, a similar phenomenon could plausibly be replicated in other emergent resistant lineages of *E. coli*.

The plasmidome of the 75 ESBLEC genomes harboured a total of 228 plasmids, of which 86 (32%) were found within the 30 genomes that constituted the STC131 group and 142 (62%) were found within the 45 genomes that constituted the non-STC131 group. Regarding the relaxase families, MOB$_{\text{P}}$, MOB$_{\text{F}}$ and MOB$_{\text{Q}}$ types were the most frequent including 92% of the plasmids. We described a wide variety of incompatibility groups (Inc.) and this diversity was higher within the non-STC131 than the STC131 group (12 Inc. including Col-like, F, Q, I, X, Y, P, 13, HI2, N, R and B/O/K/Z versus 6 Inc. including Col-like, F, X, 13, N, and B/O/K/Z, respectively). Interestingly, the plasmid diversity was much higher within subclade C2 than within cluster C1M27 and A of ST131. Generally, the virulence and resistance encoding genes are carried by plasmids belonging to the IncF family.

Antibiotic resistance often disseminates in bacteria through MGEs. Therefore, in this study most of the resistance associated genes were found carried by plasmids. Nonetheless, plasmid-bacterium association within bacterial lineages have also been described (Price *et al.*, 2013; Johnson *et al.*, 2016c; Stoesser *et al.*, 2016c; Pitout and DeVinney, 2017; San Millan, 2018; Kondratyeva *et al.*, 2020). Kondratyeva *et al.* 2020 had recently described an existing plasmid-clade association within ST131 lineage: I) CTX-M-27-encoding IncF [F1:A2:B20] is strictly found in clade C1, II) CTX-M-15-encoding IncF [F2:A1:B-] exist only in subclade C2 and III) Col-like replicons Col156, Col(MG828) and Col8282 are associated with clade C1 (Kondratyeva *et al.*, 2020). In our study, we found two *bla*$_{\text{CTX-M-27}}$ genes carried by IncF [F1:A2:B20] plasmids and five *bla*$_{\text{CTX-M-15}}$ genes carried by IncF [F2:A1:B-] plasmids associated with the same clades than Kondratyeva *et al.* 2020. Thus, a plasmid-clade association can also be retrieved from our study. Regarding clade A, we describe the presence of IncF [F29:A-:B10] and ICEs as adistintive features from genomes from clade C.

In addition, we noticed that Col-like and IncF plasmids were equally dominant within both groups of genomes. While Kondratyeva *et al.* 2020 had only reported this dominance within the ST131 plasmidome. It is evident that IncF plasmids represent a clinically relevant vehicle of virulence and resistance genes within ESBLEC. The role of Col-like plasmids is still unknown (Kondratyeva *et al.*, 2020). However, it can be hypothesised that those families of plasmids are the most influent regarding the evolutionary success of ESBLEC lineages witch is mediated by a stable plasmid-bacterium association that favours clonal expansion (Johnson *et al.*, 2016b; Stoesser *et al.*, 2016b; San Millan, 2018). Moreover, the quantity and diversity of plasmids found within the ESBLEC genomes highlight the important role that MGEs plays in the evolution of *E. coli*, in particular with the acquisition of ESBL enzymes. These results are in line with other researches (Stoesser

et al., 2016b; Kondratyeva *et al.*, 2020). In this study, we highlight the diversity of plasmids existing within different ESBLEC emerging lineages.

5.2. STRAINS OF ANIMAL ORIGIN

5.2.1 COMPANION ANIMAL: DOG RESERVOIR

To get more insights into the population structure of canine *E coli*, we investigated those harboured in the intestinal tract of 104 healthy Spanish dogs by using different approaches, knowing that the gut is the reservoir of the great majority of *E. coli* causing extraintestinal infections. The phylogenetic group, VF-encoding gene and antibiotic susceptibility analyses, showed that among the 197 canine faecal isolates obtained from the 104 dogs, 84 (42,6%) belonged to B2 phylogroup, 91 (46,2%), mostly B2 group isolates, were classified as ExPEC and/or UPEC, and 28 (14.2%), mostly non-B2 group isolates, as MDR. This strongly suggests that the intestinal tract of healthy dogs might be an important reservoir of ExPEC and/or UPEC isolates, and in a lesser extent, of MDR *E. coli* isolates. However, some studies that focused on antibiotic-resistant canine isolates suggested that dogs might also be an important reservoir for antibiotic-resistant strains (Hordijk *et al.*, 2013; Belas *et al.*, 2014; Haenni *et al.*, 2014; Rocha-Gracia *et al.*, 2015; Schaufler *et al.*, 2015; Damborg *et al.*, 2015; Pires *et al.*, 2016; Seni *et al.*, 2016; Hansen *et al.*, 2016; Wedley *et al.*, 2017; Melo *et al.*, 2018; Liakopoulos *et al.*, 2018; Suay-García *et al.*, 2019; Umeda *et al.*, 2019; Dupouy *et al.*, 2019; Hong *et al.*, 2019; Karkaba *et al.*, 2019; Abreu-Salinas *et al.*, 2020; Sevilla *et al.*, 2020), notably for those producing ESBLs or CMY-2 (Hordijk *et al.*, 2013; Belas *et al.*, 2014; Haenni *et al.*, 2014; Liakopoulos *et al.*, 2018; Dupouy *et al.*, 2019). Although there was a low prevalence of MDR isolates among the 197 studied isolates, we found, as previously described that they produced ESBLs or CMY-2.

MLST assigned the 91 Spanish canine faecal isolates with an ExPEC and/or UPEC status to 34 STs among which six were displayed by 67% of the 91 isolates: ST372 (31.9%), ST12 (9.9%), ST127 (8.8%), ST648 (6.6%), ST141 (5.5%) and ST73 (4.4%). Few studies have been carried out so far to characterise the ST structure of canine ExPEC and/or UPEC isolates. In the USA, LeCuyer *et al.* (2018) analysed 295 *E. coli* isolates from canine UTI. They found that ST372, which is uncommon among the human *E. coli* pathogens (Adler *et al.*, 2012; Izdebski *et al.*, 2013; Flament-Simon *et al.*, 2019, 2020b; Manges *et al.*, 2019), was the predominant ST in canine UTI isolates (21.7%), and this was, well ahead of the five other most frequent STs: ST12 (6.4%), ST73 (6.4%), ST127 (4.1%), ST131 (4.1%) and ST297 (3.7%). A total of 170 (57.4%) of these isolates met the criterion to be classified as ExPEC, and, except for ST297, the most prevalent STs were associated with ExPEC status. In France, Valat *et al.* (2020) analysed 618 canine *E. coli* isolates collected from diagnostic laboratories, including 403 (65.2%) from UTIs. B2 phylogroup was over-represented (79.6%) and positively associated with the presence of numerous VFs, including those defining the ExPEC status. MLST of a randomly chosen subset of 89 isolates belonging to B2 phylogroup revealed five dominant STs: ST372 (17.9%), ST73 (17.9%), ST12 (10.1%), ST141 (7.9%) and ST961 (5.6%). In Australia, Kidsley *et al.* (2020b) focused their study on the canine fluroquinolone-susceptible *E. coli* clinical isolates (n=449) that were identified during a nation-wide survey of antibiotic resistance in Australian animals between January 2013 and January 2014. They found that these isolates mostly (n=317; 71%) belonged to B2 phylogroup. By using the RAPD typing system, they found a distribution of the 317 B2 group isolates into 35 main clusters. To pursue their molecular investigation, they sequenced and analysed the whole genome of 77 representatives of the B2 group fluoroquinolone-susceptible isolates. Thus, they found that the 77 sequenced isolates were assigned to 24 STs among which four were

dominant: ST372 (31%), ST73 (17%), ST12 (7%) and ST80 (7%). In sum, the present study and those previously published show that three STs (ST372, ST12 and ST73) are the dominant ST in healthy and infected dogs irrespective of the countries (the USA, France, Australia and Spain) and strain sources (clinical samples and faeces). Such a finding might argue for the "prevalence" theory with regard to UTI pathogenesis (most UTIs are opportunistic infections caused by bacteria that predominate in the faecal microbiota) in dogs. Nevertheless, the fact that the canine isolates belonging to the most dominant ST, either present in all studied countries (ST372, ST12 and ST73) or present in some studied countries (ST127 in the USA and Spain, and ST141 in France and Spain) were shown to harbour numerous VF-encoding genes might also argue for the "special pathogenicity" theory in dogs.

Concerning the ST structure of canine MDR isolates, there seems to exist a more important difference between the countries than for the non-MDR isolates. MLST assigned the 28 Spanish canine MDR isolates to a great diversity of STs comprising 15 established STs (ST10, ST12, ST38, ST57, ST58, ST88, ST93, ST155, ST457, ST648, ST695, ST1011, ST1140, ST3774, and ST8953) and seven new STs. The first 10 STs here listed have been identified in canine MDR isolates from different countries (Dierikx *et al.*, 2012; Wagner *et al.*, 2014; Ewers *et al.*, 2014; Rocha-Gracia *et al.*, 2015; Schaufler *et al.*, 2015; Liu *et al.*, 2016c, 2016b, 2017; Seni *et al.*, 2016; Falgenhauer *et al.*, 2016; Hansen *et al.*, 2016; Wedley *et al.*, 2017; Kawamura *et al.*, 2017; Boehmer *et al.*, 2018; LeCuyer *et al.*, 2018; Liakopoulos *et al.*, 2018; Maeyama *et al.*, 2018; Marques *et al.*, 2018; Melo *et al.*, 2018; Zhang *et al.*, 2018; Zogg *et al.*, 2018; Bortolami *et al.*, 2019; Chen *et al.*, 2019b; Umeda *et al.*, 2019; Dupouy *et al.*, 2019; Karkaba *et al.*, 2019; Loncaric *et al.*, 2020; Pepin-Puget *et al.*, 2020; Valat *et al.*, 2020). In contrast to some studies, we did not detect neither canine MDR isolates displaying the five important emerging MDR STs in humans: ST69, ST127, ST131, ST410 and ST1193 (Platell *et al.*, 2011; Harada *et al.*, 2012; Haenni *et al.*, 2014; Schaufler *et al.*, 2015; Bogaerts *et al.*, 2015; Liu *et al.*, 2016b, 2016c; Seni *et al.*, 2016; Falgenhauer *et al.*, 2016; Kawamura *et al.*, 2017; Belas *et al.*, 2018; Boehmer *et al.*, 2018; LeCuyer *et al.*, 2018; Liakopoulos *et al.*, 2018; Maeyama *et al.*, 2018; Marques *et al.*, 2018; Zhang *et al.*, 2018; Zogg *et al.*, 2018; Melo *et al.*, 2019; Umeda *et al.*, 2019; Bortolami *et al.*, 2019; Bourne *et al.*, 2019; Chen *et al.*, 2019b; Pepin-Puget *et al.*, 2020; Valat *et al.*, 2020), canine isolates harbouring the *mcr*-1 gene encoding resistance to colistin as described in China (Wang *et al.*, 2018a) nor canine isolates producing carbapeneamases, notably OXA-48, as described in Germany (Pulss *et al.*, 2018), France (Melo *et al.*, 2016), and United States (Liu *et al.*, 2016c). Finally, we found that none of the 29 here studied ST372 was MDR while previous studies have found ST372 isolates producing different types of ESBLs and CMY-2 (Wagner *et al.*, 2014; Hansen *et al.*, 2016; Liu *et al.*, 2016c; Boehmer *et al.*, 2018; Zhang *et al.*, 2018; LeCuyer *et al.*, 2018; Maeyama *et al.*, 2018; Melo *et al.*, 2018; Chen *et al.*, 2019b; Karkaba *et al.*, 2019; Pepin-Puget *et al.*, 2020; Valat *et al.*, 2020).

To get more insight into the potential link between the canine ExPEC and/or UPEC and the *E. coli* isolates causing extraintestinal infections in humans, we determined which clone (defined by the association of phylogroup, clonotype and ST) and which serotypes characterised the 91 canine ExPEC and/or UPEC in order to compare them with human *E. coli* clinical isolates collected in 2015 and 2016 in Spain and France and characterised for these two traits (Flament-Simon *et al.*, 2019). This approach allowed us to found that among the 50 clones identified in the 91 Spanish canine ExPEC and/or UPEC isolates, 15 were present in the human collection accounting for 49 (18,8%) of the human isolates. However, only 31 of the 49 human ExPEC and/or UPEC isolates presented the same O:H serotype as the canine ones. By coupling clonal type and serotype for each *E. coli* ST

lineage shared by dogs and humans, we observed various features about the distribution of the human isolates when the lineages included several clones and several clone-related serotype in dogs. This feature shows that it is difficult to make hypotheses about the relationship between canine and human isolates sharing a given clone-serotype couple in a given lineage without knowing the structure of the clone-serotype couples in the given lineage in humans. For example, we had found (Flament-Simon *et al.*, 2020b) that the human ST73 isolates were distributed into four clones of which two here were identified in dogs. In human, three of the four clones comprised, each, isolates with different serotypes but one serotype (O6:H1) was exhibited by isolates distributed into the four clones. In dogs, the ST73 isolates exhibited only serotype O6:H1. This suggests that serotype might be an ecological niche marker meaning, in this case, that isolates of the lineage ST73 exhibiting serotype O6:H1 are adapted to both dogs and humans. However, the Kidsley *et al.* (2020b) study in which a phylogenetic tree was built with the genome of ST73 strains from dogs, cats and humans, seems to contradict this hypothesis. Indeed, the 13 studied Australian canine isolates of the lineage ST73 exhibited four serotypes among which serotype O6:H1 was exhibited by only one isolate that formed an animal-specific cluster (containing cat O6:H1 ST73 isolates) distinct from the four main clusters of human O6:H1 ST73 isolates. Nevertheless, the hypothesis that we made for serotype O6:H1 with regard to Spanish canine and human ST73 isolates could be made for serotype O2:H1 with regard to Australian canine and human ST73 isolates as this serotype was shared by clustered canine and human isolates. By extending the comparison of the structure of clone-serotype couples to the other human-specific-human ST lineages (ST127, ST141 and ST1193) shared by the Spanish studied dogs and humans, we observed that the clone-serotype couples shared by dogs and humans comprised mostly the serotype the most frequent in the human clones. This feature seems to indicate that serotype frequency might be a variable involved in the *E. coli* exchanges between dogs and humans. Concerning the dog-specific lineage ST372, we had found only one clone comprising to isolates in Spanish humans while we found here five clones in the 29 Spanish dogs. Among the six serotypes exhibited by these 29 canine isolates, the serotype exhibited by the two human ST372 isolates corresponded to one (O18:H31) of the two dominant serotypes in dogs that was, on the other hand, exhibited by canine isolates belonging to three different clones. Thus, the suggestions that we made about the fact that serotype could be an ecological niche marker and that serotype frequency could shape the *E coli* exchange between dogs and humans seems to be able to be applied to the lineage ST 372.

Interestingly, concerning the 24 clones identified in the 28 canine MDR isolates, which were mostly non-B2-group isolates, we observed that if there were some clones (n=9) shared by the Spanish human (35 of 394) and canine (10 of 197) isolates there was only one isolate that shared the same clone and the same serotype as one canine isolate.

To better understand the potential relationship between canine and human *E. coli* isolates with regard to the lineage ST372, we turned to the whole genome sequencing and analysis of 197 ST372 strains (151 from dogs and 46 from humans). The SNP analysis of the core genome of these 197 strains revealed an extensive phylogenetic diversity of the ST372 isolates that was segregated into six clusters. Cluster 1 comprised 91.4% of canine strains while cluster 2 comprised 60.9% of human strains. Cluster 2 was specific of human strains associated with serotypes O18:H31 and O45:H31, the latter serotype being exclusively found in human ST372 strains. Three other serotypes were the most prevalent serotype among strains belonging to cluster 1, including O4:H31 and O15:H31 associated with canine strains, and O83:H31 identified in similar proportion among canine and human strains. Overall, the WGS analysis suggests that canine strains of clone B2-

CH103-9-ST372, belonging to cluster 1 and having serotype O83:H31 might cause extraintestinal infections in humans and dogs, as already suggested by the clone-serotype couple analysis, whereas strains of this clone belonging to cluster 2 and having serotypes O18:H31 and O45:H31 might cause only human extraintestinal infections. Molecular epidemiological studies on *E. coli* ST372 in human extraintestinal infections are required to confirm these suggestions.

Furthermore, we localised ICEs in the chromosome of 22 of the ST372 sequenced genomes and confirmed that all ICEs belong to a yersiniabactin synthesis-associated ICE type (ICEKp1 family) with relaxase type MOB_Q. In contrast, we found very few plasmids. Moreover, we found that the number of plasmids retrieved from the human ST372 strains was higher than that of plasmids found in the canine strains (four plasmids in the two human strains versus seven plasmids in the 21 canine strains). Interestingly, the genome of canine LREC_356 strain from cluster 4 carried two plasmids, and was the canine strain genome the most similar to human strain genomes. Those plasmids were not rich with genes encoding of antibiotic resistance and virulence-factors. Nonetheless, a high number of virulence factor encoding-genes were found in the chromosome of the ST372 genomes and we hypothesised that the origin of the UPEC status of ST372 strains is due to the acquisition of ICEs harbouring the genes associated with this status. Although there is still limited knowledge about the origin of genomic islands, like ICEs or pathogenicity islands (PAIs), it has been speculated that they derive from the integration of plasmids or phages into the chromosome. Further, genomic research has shown that genomic islands have played a major role in the transformation of avirulent into virulent bacteria. Besides, most VFs of ExPEC are encoded by ICEs and PAIs (Schubert *et al.*, 1998; Dobrindt *et al.*, 2004; Robinson *et al.*, 2018; Denamur *et al.*, 2020; Desvaux *et al.*, 2020).

In summary, the intestinal tract of healthy dogs appears as an important reservoir of ExPEC and/or UPEC and, to a lesser extent, of MDR *E. coli* isolates. However, the canine MDR isolates could be a good reservoir of ESBLs and CMY-2 because most of them produce these enzymes. Among the canine isolates displaying an ExPEC and/or UPEC status, clone B2-CH103-9-ST372 was dominant. This canine clone and 14 others, also displaying an ExPEC and/or UPEC status, had been identified in isolates previously published as causing extraintestinal infections in human suggesting a zoonotic potential of these clones. WGS analysis suggests that canine strains of clone B2-CH103-9-ST372, belonging to cluster 1 and having serotype O83:H31 might cause extraintestinal infections in both humans and dogs, whereas those strains of this clone belonging to cluster 2 and serotypes O18:H31 and O45:H31 might cause only human infections. Taking into consideration that Kidsley *et al.* (2020b) have recently characterised the phylogenetic relationship between canine, cat, and human isolates of the lineage ST73, such studies are still required for the other ST lineages and clones that we showed in this study to be shared by canine and human isolates in order to clarify their potential role in infection occurrence in both dogs and humans.

5.2.2. PRODUCTION ANIMAL: SWINE RESERVOIR

As far as we know, there are only six reports on ST131 from swine: (I) Trobos *et al.* (2009) studied 68 sulphonamide-resistant *E. coli* from different hosts and identified the first swine ST131 isolate obtained from pork meat in Denmark in 2003; (II) Schink *et al.* (2013) detected a CTX-M-1 ST131 isolate recovered from a porcine gastrointestinal tract infection among 495 ESBL-producing *E. coli* analysed in Germany during the years 2006–2007; (III) Herrera Estévez (2015) detected 18 ST131 isolates from 200 pork meat samples obtained in Spain during the years 2011 and 2012; (IV) Liu *et al.* (2018c) collected an isolate from a diarrheal pig among 44 ESBL-producing *E. coli* from Northwest China (2015 to 2017); (V) García-Meniño *et al.*, (2018) identified 18 ST131

isolates recovered from the 1823 piglets with diarrhoea sampled in Spain between the years 2009 and 2017; and (VI) Reid *et al.* (2019) reported two highly related ST131 *E. coli* isolates, the first from a healthy pig and the second from a human urinary tract infection, obtained in New South Wales (Australia) in 2007 and 2009, repectively.

In the present study, we performed WGS analysis of a collection of 11 representative ST131 isolates obtained from piglets with diarrhoea and pork meat in Spain. All of the 11 ST131 isolates belonged to serotype O25b:H4 and clade B, and presented a high number of virulence genes typical of *E.coli* that cause extraintestinal infections. In addition, the 11 porcine isolates harboured traits that genetically satisfied the criteria for designation as ExPEC, UPEC and APEC status being designated as APEC, UPEC and ExPEC like isolates due to their site of isolation. Although seven of the 11 ST131 isolates were isolated from piglets with diarrhoea, none of them were positive for virulence genes of intestinal pathogenic *E. coli* (Mora *et al.,* 2011).

In annex 2, Figure S21, based on the results published by Matsumura *et al.* (2017) and Ben Zakour *et al.* (2016) we had represented the already known ST131 lineages plus the new lineages found in this study. These sublineages were designated according to their *fimH* allele (type 1 fimbriae adhesin gene), phylogenetic clade (A, B, B0, I1, C0, C, C1, C2, and C3) and resistance profile. Briefly, B-*H*22 clade, a fluoroquinolone (FQ) and cephalosporin susceptible ancestor, evolved in C-*H*30R (where "R" indicates resistance to FQ) clade that was the predecessor of two sister subclades C1-*H*30R and C2-*H*30Rx (often carrier of the *bla*$_{\text{CTX–M–15}}$ enzyme). Later, two additional ST131 sublineages referred to C1-M27 (positive for the *bla*$_{\text{CTX–M–27}}$ enzyme) and C1-nM27 (non ESBL or positive for the *bla*$_{\text{CTX–M–14}}$ enzyme) have been described. The ST131 clonal group accounted for 490 (16%) of 2995 isolates obtained from extraintestinal clinical samples collected in five Spanish hospitals during the period 2005 to 2012 (Dahbi *et al.*, 2014). The most common lineages in this clinical collection were the C2-*H*30Rx (61.6%), B (17%) and C1-nM27 (16%). Recently, Mamani *et al.* (2019) also described the subclade C2-*H*30Rx as the most prevalent (85%) among 41 ST131 ESBL-producing *E. coli* causing bacteraemia in a Spanish hospital over a 12-year period (2000 to 2011). In contrast, almost all Spanish ST131 isolates obtained from animal origin belonged to clade B. In our laboratory we have tested a collection of 98 ST131 isolates from animals (57 from avian, 36 porcine and five from Antarctic pinnipeds) and all but one from pinnipeds belonged to clade B. Although clade B isolates share numerous virulence genes, virotypes vary depending on the origin of the isolates. Thus all avian isolates have virotype D4 (*ibeA*, neuC-K1) while swine isolates have virotypes D2 (*ibeA*, *papG* III, *cdtB, kpsM*-II-K5) and D5 (*ibeA*, *papG* III, *cnf1, hlyA, kpsM*-II-K5). Among the 45 clade B isolates causing bacteraemia in humans in Spain, the most frequent virotype was D2 (n=19), followed by virotypes D3 (n=12), D1 (n=8), D4 (n=6) and D5 (n=6) (Blanco *et al.* unpublished data).

The core genome SNP-based phylogenetic tree of 84 human and animal ST131 isolates analysed in this study showed (Figure 18) that they are mostly clustered into three well-defined clades (A, B and C), and similarly to Petty *et al.* (2014) study, we have observed that clade A is the most divergent from clades B and C. Our results also agree with those published by other authors (Ben Zakour *et al.*, 2016; Matsumura *et al.*, 2017b; Liu *et al.*, 2018a; Decano and Downing, 2019) regarding that clade B isolates are displayed into multiple subclades. We identified the six subclades (B0, B1, B2, B3, B4, B5, B6) described by Ben Zakour *et al.* (2016) and five new ones (B6-like, B7, B8, B9). Notably, the 11 ST131 porcine isolates from this study were grouped in the new subclades B6 (four isolates from pork meat and virotype D2) and B7 (seven isolates from piglets with diarrhoea and virotype D5). The new subclade B6-like also includes two isolates of

porcine origin, while the new subclades B8 and B9 included only ST131 isolates from Antarctic pinnipeds. In contrast, avian isolates of virotype D4 were distributed in the subclades B2, B3 and B5.

The newly described B6 and B7 subclades were primarily associated with porcine isolates, but also include some isolates causing bacteraemia and other extraintestinal infections in humans. Therefore, it is remarkable how those human isolates were closely related (46-60 SNPs average distance) with porcine isolates. Those results suggest the ecological overlap of human and animal isolates. Recently, Reid *et al.*, (2019) compared the genomes of two ST131 clade B *fimH22* isolates obtained in Australia, the first one isolated in 2007 from a healthy piglet and the second one obtained in 2009 from a human urinary tract infection. This study showed that the two isolates were highly related, separated by only 20 core SNPs (Reid *et al.*, 2019). Thus, the analysis of the phylogeny confirmed *E. coli* ST131 as pathogen capable of frequent inter-species movements.

The plasmidome of *E. coli* ST131 clade B from porcine origin had shown variability and some interesting differences concerning the actual published data. IncF plasmids are a common feature in all ST131 isolates (Lanza *et al.*, 2014), having at least one per isolate in this study. Specifically, F1:A2:B20 and F2:A1:B- plasmids are more commonly associated with the C1-*H*30R clade and C2-*H*30Rx, respectively (Johnson *et al.*, 2016b). In contrast, F2:A-:B1 plasmids were dominant in the porcine clade B isolates used in this study. This type of plasmid was absent in 35 non-ST131 *E. coli* porcine isolates from colibacillosis (García-Meniño *et al.*, 2019), suggesting that IncF plasmids vary greatly within the different lineages. Reid *et al.*, (2019) reported 23 F-types among 282 isolates from a mixture of origins belonging to the clade B of ST131 and the F2:A-:B1 IncF-subtype was one of the commonest (32%).

In accordance with previous studies, plasmids carrying virulence genes belong to the IncF family (Lanza *et al.*, 2014; Nicolas-Chanoine *et al.*, 2017). All the IncF plasmids found in this study conserved the ColV region, frequently detected among avian pathogenic *E.coli* (APEC) (Liu *et al.*, 2018a). This region comprises the *iss* (increased serum survival) gene, the *iroBCDEN* gene cluster (encoding the salmochelin siderophore system present in *Salmonella enterica* spp. and some ExPEC isolates), the *iucABCD* operon (encoding the enzymes of the biosynthetic pathway for aerobactin), *iutA* (outer membrane receptor of ferri-aerobactin complexes) gene, *cvaB and cvaC* (production of colicin) genes, *sitC* (Salmonella iron transporter C) and an *ompT-hlyF-mig14* cassette (outer membrane related genes). Some APEC cause human infection (Liu *et al.*, 2018a), nevertheless ColV plamids were absent in specific human clades, so they may not be a necessary feature to cause human infection (Reid *et al.*, 2019). It is currently unknown which advantages could the IncF plasmids confer to the bacterial cell apart from virulence and resistance to antibiotics. However, Johnson *et al.* (2016) has hypothesised that they must play multiple roles in the host success because ST131clone has adapted to harbour these plasmids at lower fitness cost, and has stabilized them through mutational events. We assume that the presence of both, resistance and virulence genes, on the same MGE is worrying for public health, as shown in our study in IncF plasmids like pLREC161_1 and pLREC 159_2. For instance, the evolutionary analysis performed by Ben Zakour *et al.* (2016) proved that ST131 emergence was adapted by isolates that had acquired first a subset of virulence genes followed by the gain of antibiotic resistances.

ColE1-like plasmids were the second more prevalent in our collection and the replication protein could not be determined, these results are in accordance with the findings of Lanza *et al.* (2014). Furthermore, ColE plasmids have been reported as *mcr-4* and *mcr-5* carriers (Borowiak *et al.*, 2017; Carattoli *et al.*, 2017; García-Meniño *et al.*, 2019). The ColE1-like plasmids from this study did not carry resistance genes. However,

they might be an important adaptive weapon. Moreover, we described one MOB_{P131} and two MOB_{V2} plasmids, which are not frequently documented and therefore absent among the description given by Lanza *et al.*, (2014) over ST131 plasmidome. Although, another MOB_{P131} plasmid has been recently found from a bla_{OXA-48} *E. coli* ST131 isolate collected from a patient with urinary infection (Stoesser *et al.*, 2016c), and MOB_{V2} plasmids have been described in other STs, such as ST69 from a human urine sample and ST93 from chicken meat samples (de Been *et al.*, 2014). Besides, MOB_{V2} plasmids pLREC153_4 and pLREC157_4 showed structural homology (around coordinates 3.7 kb to 7 kb in annex 2, Figure S12) with the pEC0674 plasmid (6 kb), an *mcr-5* carrier plasmid belonging from an ST29 *E. coli* isolate recovered from porcine faeces in Germany (Hammerl *et al.*, 2018). Finally, and also barely documented, were the cryptic plasmids found in our study. Similar small no-MOB plasmids than pLREC158_3, pLREC159_5, pLREC159_6, pLREC160_7 and pLREC161_4, have been reported among ST131 isolates. The role of these plasmids is still unknown but researchers have suggested that they could be providing useful mobilization tools for ST131 adaptation (Lanza *et al.*, 2014).

ARGs conferring resistance to penicillin/ampicillin, tetracycline, quinolones and colistin were the most common determinants. García-Meniño *et al.*, (2019) described chromosomal mutation (PmrB V161G or PmrA S39I) in addition to the plasmid mechanism (*mcr*) in 16 (46%) of non-ST131 colistin-resistant porcine *E.coli* isolates. Here, any complementary chromosomal mechanism related with system PmrAB was found. Nonetheless, some drug efflux genes were found in plasmids. Members of the *lnu* (previously *lin*) gene family (mediate lincosamide resistance) (Achard *et al.*, 2005; Roberts, 2008) and the *mef(B)* gene, which coexists with the *aadA* and *sul3* resistance genes (pLREC161_2). This combination has previously been described in plasmids from porcine *E. coli* isolates and mediates macrolide resistance (Liu *et al.*, 2009).

The recent analysis of 35 colistin-resistant non-ST131 *E. coli* isolates from porcine colibacillosis showed that 18 isolates (51%) were carriers of the *mcr*-1.1 gene variant. Those isolates were all MDR and the 50% belonged to Clonal Complex 10 (6 ST10, 1 ST5786, 2 ST7323). However, the remaining isolates belonged to different linages with single representants (ST1, ST29, ST42, ST93, ST100, ST118, ST156, ST398, ST4247). The *mcr*-1.1 gene was located in plasmids IncHI2 (33%) and IncX4 (22%), but it was also found chromosomally encoded in four isolates (García-Meniño *et al.*, 2019).

Currently, the *mcr*-1 cassette was found in a large diversity of plasmids and clonal lineages of *E. coli*. The broad adaptation of colistin resistance all over the world could be due with the ease with which the *mcr*-1 gene integrates into various regions (Hadjadj *et al.*, 2017). There is an association with successful plasmid families (Matamoros *et al.*, 2017; Zhong *et al.*, 2018). On the other hand, any association with clonal lineages of *E. coli* is described (Li *et al.*, 2018). However, the description of *mcr*-1 positive *E. coli* ST131 isolates is exceptionally worrying. Because ST131 has once been responsible for the worldwide appearance ESBL resistances due to the acquisition of specific IncF epidemic plasmids harbouring $bla_{CTX-M-15}$ gene (Nicolas-Chanoine *et al.*, 2007b, 2014; Mathers *et al.*, 2015) and therefore, major therapeutic failure could come due to the association of *mcr*-1 gene with other broad-spectrum resistance mechanisms, such as ESBLs and/or carbapenemases (Du *et al.*, 2016; Haenni *et al.*, 2016; Li *et al.*, 2016; Olaitan *et al.*, 2016; Skov and Monnet, 2016). Further, clade B of ST131 has been signalled out as a potential foodborne uropathogen and as a stronger early biofilm producer than clade C of ST131 (Nicolas-Chanoine *et al.*, 2017; Liu *et al.*, 2018a; Flament-Simon *et al.*, 2019).

In this study the genetic context of *mcr*-1 gene was diverse. Recent studies propose that the transposon Tn6330, flanked at both ends by ISA*pl1*, has lead mobilization of *mcr*-1 gene. Therefore, four structures of mobile elements carrying *mcr*-1 have been described: (I) The composite transposon Tn6330 (ISA*pl1*–*mcr*-1–orf–ISA*pl1* structure); (II) a single ISA*pl1* upstream (ISA*pl1*–*mcr*-1–orf structure); (III) a single ISA*pl1* downstream (*mcr*-1–orf–ISA*pl1* structure); and (IV) and sequences lacking ISA*pl1* altogether (*mcr*-1–orf structure). Although a few other truncated or interrupted versions have been reported (Pham Thanh *et al.*, 2016; Snesrud *et al.*, 2016). In the study carried by Snesrud *et al.* (2018) a prevalence of 11 % for structure I, 22% for structure II, 1% for structure III and 66% for structure IV (with no associated IS) have been described. Besides, the upstream ISA*pl1* element was observed in 78% of IncHI2 plasmids (Matamoros *et al.*, 2017) while the *mcr*-1–orf structure has been described as dominant in small *mcr*-1-IncX4 plasmids(Li *et al.*, 2017b; Sun *et al.*, 2017). These results are in line with our findings. The five ST131 colistin-resistant isolates from this study presented the *mcr*-1.1 allele (KP347127.1) and showed the *mcr*-1–orf structure. However, pLREC159_1(IncHI2) presumptively harbored the ISA*pl1* upstream while any copy of ISApl1 was present around the *mcr*-1 gene in pLREC160_2 (IncX4). Nevertheless, *pap*2 gene and insertion sequences IS26 and IS1924 were present in pLREC160_2. Zhong *et al.* (2018) suggested that for IncX4 plasmids IS26/26 structures could be involved in *mcr*-1 gene mobilization. The BRIG comparison showed that the *mcr*-1 harbouring IncX4 plasmids were genetically homologous between each other as previously described (Sun *et al.*, 2017), suggesting that mobilization of the *mcr*-1 cassette was quite stable within this type of plasmids. Furthermore, pLREC159_1 and pLREC176_1 had almost identical in structure around the *mcr*-1 gene and were related to the presence of tellurium resistance genes, as well as, for pLREC161_1. This heavy metal resistance lend HI2 plasmids to success in wastewater environments (Reid *et al.*, 2019).

Moreover, a highly relevant finding of the present study was the appearance of chromosomally encoded colistin resistance. We found a presumptive ISA*pl1*–*mcr*-1–orf–ISA*pl1*-like structure. This genetic arrangement was described as the most prevalent by Li *et al.* (2018) in bacterial chromosomes. It has been proposed that the lack of one or both of the ISA*pl1*-elements, implies stabilization of the *mcr*-1 gene (non-transposable) (Snesrud *et al.*, 2016, 2018). Thus, chromosomally encoded colistin resistance might mean a more stable inheritance in ST131 porcine lineage. Nevertheless, it also implies less mobilization among isolates.

Some studies have proved that stopping using colistin in positive *mcr*-1 carriers porcine farms has resulted in the disappearance of the resistance (Randall *et al.*, 2018). However, due to the MDR profile of our ST131 isolates, this might not be that easy, due to co-selection by other antimicrobials and host fitness adaptation. The fitness effects of carrying the *mcr*-1 gene were evaluated by Wu *et al.* (2018). They proved that IncI2 and IncX4 *mcr*-1 positive plasmids conferred fitness advantage for its host. In contrast, IncHI2 plasmids imposed a slight fitness cost and competitive disadvantage.

To sum up, it seems that epidemic plasmids type IncX4 and IncHI2, which have been reported as the more prevalent carriers of *mcr*-1 gene, are responsible for the acquisition of the colistin resistance in the high-risk clone ST131 of *E.coli* in our isolates from porcine origin. Worryingly, the *mcr*-1 gene has also integrated an IncF plasmid form ST131. The clinical significance of this category of plasmids is therefore highlighted because they are responsible of the epidemic dissemination of ESBL resistance worldwide (Johnson *et al.*, 2016b). We hypothesise that IncF plasmids like pLREC161_1 had acquired the *mcr* gene via cointegrating an IncHI2 *mcr*-1 harbouring plasmid. The *mcr*-carring IncF plasmid pMR0516*mcr* that has been described by McGann *et al.* (2016)

also shared 89 kb with pHNSHP45-2 (*mcr*-1-carrying IncHI2 plasmid) and possibly the same phenomenon had occurred.

The impact that humans have on the environment and other species is evident. A multidisciplinary point of view, taking into account the environment, animals, and human for the surveillance of health, is given by the concept of "One Health" (Destoumieux-Garzón *et al.*, 2018). This philosophy fits perfectly the actual global situation. Measures to reinforce prevention and proper management for overall antimicrobials might be the right strategy to fight against antibiotic resistance.

Further studies to understand why some plasmids evolve to become genetically stable are necessary to be able to anticipate the evolutionary dynamic of high risk clones.

As a limitation, mentioning that short reads obtained by WGS do not always allow a perfect genome reconstruction due to the high amount of repeated sequences in the bacterial genomes. Long read sequencing technologies will probably bring new perspectives in future researches.

In summary, the clade B of ST131 showed a huge genetic diversity, and five new subclades were defined (B6, B6-like, B7, B8, and B9). The majority of ST131 porcine isolates belong to new subclades B6 and B7. An association between the phylogeny, the virotype, and the origin of porcine isolates was established. Some porcine and human clinical isolates were highly related. Most porcine ST131 isolates are MDR (91%) and carry many ARGs. Colistin resistance was introduced through MGEs and had been able to stabilize chromosomally. IncX4 and IncHI2 epidemic mcr-1–harboring plasmids are responsible for acquired colistin resistance encoded by mcr-1.1 gene. The surrounding environment of the mcr-1 cassette is variable but within the same family of plasmids, insertion, and stabilization had common structures. The plasmidome of ST131 clade B is distinct from other clades within ST131 clone and contains an impressive variety of different plasmids. Characteristically ST131 porcine isolates have APEC [F2:A-:B1]–IncF plasmids. ColE1-like plasmids and IncX plasmids are also frequently observed. This behaviour could be related to the clonal success due to its ability to easily rearrange DNA by horizontal gene transfer.

6. CONCLUSIONS

1. Approximately 10% of the extraintestinal *E. coli* infections that had occurred in 2016 in the two studied hospitals were caused by ST131 strains, and approximately 60% of these infections were caused by strains belonging to only 10 STs (ST10, ST12, ST58, ST69, ST73, ST88, ST95, ST127, ST131 and ST141).
2. The majority of ST12, ST73, ST95 and ST141 human extraintestinal strains were susceptible to most antibiotics, indicating that MDR was not the reason for their success. The results of the present study support the idea that their success is mainly due to the high number of VF-encoding genes that they possess.
3. Among the MDR human extraintestinal strains, four clones are predominant, especially: B2-CH40-30-ST131, B2-CH40-41-ST131, C-CH4-39-ST88 and D-CH35-27-ST69.
4. Despite the enormous genetic diversity observed in our ESBLEC collection (71 clones amongst 188 ESBLEC), it can be concluded that the majority of the strains belong to only three clonal complexes (CC10, CC23, and CC131) and that ST131 subclade C2 associated with the production of CTX-M-15 remains the most prevalent *E. coli lineage* among the ESBLEC strains identified in the studied Spanish and French hospitals.
5. Among our 394 extraintestinal *E. coli* strains, <10% are able to quickly and persistently produce high biofilm levels over 5 h. These strains belong to a few clones previously described in various studies as dominant gut colonizers in mammalians and birds, and comprised the B2-CH40-22-ST131 clone, i.e., the ancestor of the globally disseminated B2-CH40-30-ST131 clone that is the dominant clone among the weak biofilm producers.
6. The core genome analysis of ESBLEC highlights genomic diversity among ESBL *E. coli* population. That includes representatives from phylogroups A, B1, B2, C, E and F. We described two new clades in ST131 clonal complex represented by genomes harbouring the *fim*H5 and *fim*H35 alleles. Clade C is characterised by showing less SNPs distance in the core genome than its sister clades A and B.
7. There is a correlation between VF-encoding gene profiles and STs.
8. There is an impressive diversity of plasmids among ESBLEC. However, MOB_F, MOB_P and MOB_Q families of relaxases represent 92% of the plasmidome and Col-like and IncF are the most frequent incompatibility groups. Besides, most of the virulence and resistance associated encoding genes are carried by plasmids belonging to the IncF family.
9. In ST131 lineage, the plasmid incompatibility group diversity is higher among subclade C2 than among clade A and cluster C1-M27 subclade C1. Clade A is related with the presence of ICEs and [F29:A-:B10]-IncF plasmids.
10. The intestinal tract of healthy dogs appears as an important reservoir of ExPEC and/or UPEC, and, in a lesser extent, of MDR *E. coli* strains.
11. Among the canine strains displaying an ExPEC and/or UPEC status, clone B2-CH103-9-ST372 was dominant.
12. WGS analysis suggests that canine strains of clone B2-CH103-9-ST372, belonging to cluster 1 (serotypes O4:H31, O21:H31 and O83:H31) and cluster 6 (serotype O6:H31) might cause extraintestinal infections in both humans and dogs, whereas those strains of this clone belonging to cluster 2 and serotypes O18:H31 and O45:H31 might cause only human infections.
13. Clade B of ST131 showed a huge genetic diversity, and five new subclades were defined (B6, B6-like, B7, B8, and B9). The majority of ST131 porcine strains belong to new subclades B6 and B7.

14. Most porcine ST131 strains are MDR (91%) and carry many antibiotic resistance genes. Colistin resistance was introduced through MGEs and had been able to stabilize chromosomally. IncX4 and IncHI2 epidemic *mcr-1*–harboring plasmids are responsible for acquired colistin resistance encoded by *mcr-1.1* gene. The surrounding environment of the *mcr-1* cassette is variable but within the same family of plasmids, insertion, and stabilization had common structures.

7. REFERENCES

Abe, C. M., Salvador, F. A., Falsetti, I. N., Vieira, M. A. M., Blanco, J., Blanco, J. E., et al. (2008). Uropathogenic Escherichia coli (UPEC) strains may carry virulence properties of diarrhoeagenic E. coli. *FEMS Immunol. Med. Microbiol.* 52, 397–406. doi:10.1111/j.1574-695X.2008.00388.x.

Abreu-Salinas, F., Díaz-Jiménez, D., García-Meniño, I., Lumbreras, P., López-Beceiro, A. M., Fidalgo, L. E., et al. (2020). High Prevalence and Diversity of Cephalosporin-Resistant Enterobacteriaceae Including Extraintestinal Pathogenic E. coli CC648 Lineage in Rural and Urban Dogs in Northwest Spain. *Antibiotics* 9, 468. doi:10.3390/antibiotics9080468.

AbuOun, M., Stubberfield, E. J., Duggett, N. A., Kirchner, M., Dormer, L., Nunez-Garcia, J., et al. (2017). mcr-1 and mcr-2 variant genes identified in Moraxella species isolated from pigs in Great Britain from 2014 to 2015. *J. Antimicrob. Chemother.* 72, 2745–2749. doi:10.1093/jac/dkx286.

Achard, A., Villers, C., Pichereau, V., and Leclercq, R. (2005). New lnu(C) Gene Conferring Resistance to Lincomycin by Nucleotidylation in Streptococcus agalactiae UCN36. *Antimicrob. Agents Chemother.* 49, 2716–2719. doi:10.1128/AAC.49.7.2716-2719.2005.

Adams-Sapper, S., Diep, B. A., Perdreau-Remington, F., and Riley, L. W. (2013). Clonal Composition and Community Clustering of Drug-Susceptible and -Resistant Escherichia coli Isolates from Bloodstream Infections. *Antimicrob. Agents Chemother.* 57, 490–497. doi:10.1128/AAC.01025-12.

Adler, A., Gniadkowski, M., Baraniak, A., Izdebski, R., Fiett, J., Hryniewicz, W., et al. (2012). Transmission dynamics of ESBL-producing Escherichia coli clones in rehabilitation wards at a tertiary care centre. *Clin. Microbiol. Infect.* 18, E497–E505. doi:10.1111/j.1469-0691.2012.03999.x.

AEMPS (2014). Plan Nacional Resistencia Antibioticos (PRAN). Plan estratégico y de acción para reducir el riesgo de selección y diseminación de la resistencia a los antibióticos.

AEMPS (2019). Plan nacional frente a la resistencia a los antibióticos 2019-2021.

Agarwal, J., Mishra, B., Srivastava, S., and Srivastava, R. (2013). Genotypic characteristics and biofilm formation among Escherichia coli isolates from Indian women with acute cystitis. *Trans. R. Soc. Trop. Med. Hyg.* 107, 183–187. doi:10.1093/trstmh/trs090.

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12, 402. doi:10.1186/1471-2164-12-402.

Allsopp, L. P., Beloin, C., Ulett, G. C., Valle, J., Totsika, M., Sherlock, O., et al. (2012). Molecular Characterization of UpaB and UpaC, Two New Autotransporter Proteins of Uropathogenic Escherichia coli CFT073. *Infect. Immun.* 80, 321–332. doi:10.1128/IAI.05322-11.

Allsopp, L. P., Totsika, M., Tree, J. J., Ulett, G. C., Mabbett, A. N., Wells, T. J., et al. (2010). UpaH Is a Newly Identified Autotransporter Protein That Contributes to Biofilm Formation and Bladder Colonization by Uropathogenic Escherichia coli CFT073. *Infect. Immun.* 78, 1659–1669. doi:10.1128/IAI.01010-09.

Alvarado, A., Garcillán-Barcia, M. P., and de la Cruz, F. (2012). A degenerate primer MOB typing (DPMT) method to classify gamma-proteobacterial plasmids in clinical and environmental settings. *PLoS One* 7, e40438. doi:10.1371/journal.pone.0040438.

Ambler, R. P. (1980). The structure of β-lactamases. *Philos. Trans. R. Soc. London. B, Biol. Sci.* 289, 321–331. doi:10.1098/rstb.1980.0049.

Anastasi, E. M., Matthews, B., Stratton, H. M., and Katouli, M. (2012). Pathogenic Escherichia coli Found in Sewage Treatment Plants and Environmental Waters. *Appl. Environ. Microbiol.* 78, 5536–5541. doi:10.1128/AEM.00657-12.

Andersen, T. E., Khandige, S., Madelung, M., Brewer, J., Kolmos, H. J., and Møller-Jensen, J. (2012). Escherichia coli Uropathogenesis In Vitro : Invasion, Cellular Escape, and Secondary Infection Analyzed in a Human Bladder Cell Infection Model. *Infect. Immun.* 80, 1858–1867. doi:10.1128/IAI.06075-11.

Anderson, G. G. (2003). Intracellular Bacterial Biofilm-Like Pods in Urinary Tract

Infections. *Science (80-.).* 301, 105–107. doi:10.1126/science.1084550.

Antýo, E.-M., Glodde, S., Li, G., Sharifi, R., Homeier, T., Laturnus, C., et al. (2008). The chicken as a natural model for extraintestinal infections caused by avian pathogenic Escherichia coli (APEC). *Microb. Pathog.* 45, 361–369. doi:10.1016/j.micpath.2008.08.005.

Antibiotic resistance threats in the United States (2019). Atlanta, Georgia doi:10.15620/cdc:82532.

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., and Pevzner, P. A. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics* 32, btw493. doi:10.1093/bioinformatics/btw493.

Arber, W. (2000). Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol. Rev.* 24, 1–7. doi:10.1111/j.1574-6976.2000.tb00529.x.

Arnaud, I., Maugat, S., Jarlier, V., and Astagneau, P. (2015). Ongoing increasing temporal and geographical trends of the incidence of extended-spectrum beta-lactamase-producing Enterobacteriaceae infections in France, 2009 to 2013. *Eurosurveillance* 20, 30014. doi:10.2807/1560-7917.ES.2015.20.36.30014.

Arredondo-Alonso, S., Willems, R. J., van Schaik, W., and Schũrch, A. C. (2017). On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb. Genomics* 3. doi:10.1099/mgen.0.000128.

Bahadori, M., Motamedifar, M., Derakhshandeh, A., Firouzi, R., Motamedi Boroojeni, A., Alinejad, M., et al. (2019). Genetic relatedness of the Escherichia coli fecal population and strains causing urinary tract infection in the same host. *Microbiologyopen* 8, e00759. doi:10.1002/mbo3.759.

Banerjee, R., Johnston, B., Lohse, C., Chattopadhyay, S., Tchesnokova, V., Sokurenko, E. V., et al. (2013a). The Clonal Distribution and Diversity of Extraintestinal Escherichia coli Isolates Vary According to Patient Characteristics. *Antimicrob. Agents Chemother.* 57, 5912–5917. doi:10.1128/AAC.01065-13.

Banerjee, R., Robicsek, A., Kuskowski, M. A., Porter, S., Johnston, B. D., Sokurenko, E., et al. (2013b). Molecular Epidemiology of Escherichia coli Sequence Type 131 and Its H30 and H30-Rx Subclones among Extended-Spectrum-β-Lactamase-Positive and -Negative E. coli Clinical Isolates from the Chicago Region, 2007 to 2010. *Antimicrob. Agents Chemother.* 57, 6385–6388. doi:10.1128/AAC.01604-13.

Baquero, F., Tedim, A. P., and Coque, T. M. (2013). Antibiotic resistance shaping multi-level population biology of bacteria. *Front. Microbiol.* 4. doi:10.3389/fmicb.2013.00015.

Baroud, M., Dandache, I., Araj, G. F., Wakim, R., Kanj, S., Kanafani, Z., et al. (2013). Underlying mechanisms of carbapenem resistance in extended-spectrum β-lactamase-producing Klebsiella pneumoniae and Escherichia coli isolates at a tertiary care centre in Lebanon: role of OXA-48 and NDM-1 carbapenemases. *Int. J. Antimicrob. Agents* 41, 75–79. doi:10.1016/j.ijantimicag.2012.08.010.

Basmaci, R., Bonacorsi, S., Bidet, P., Biran, V., Aujard, Y., Bingen, E., et al. (2015). Escherichia Coli Meningitis Features in 325 Children From 2001 to 2013 in France. *Clin. Infect. Dis.* 61, 779–786. doi:10.1093/cid/civ367.

Bauer, R. J., Zhang, L., Foxman, B., Siitonen, A., Jantunen, M. E., Saxen, H., et al. (2002). Molecular Epidemiology of 3 Putative Virulence Genes for Escherichia coli Urinary Tract Infection— usp, iha, and iroN E. coli. *J. Infect. Dis.* 185, 1521–1524. doi:10.1086/340206.

Beachey, E. H. (1981). Bacterial Adherence: Adhesin-Receptor Interactions Mediating the Attachment of Bacteria to Mucosal Surfaces. *J. Infect. Dis.* 143, 325–345. doi:10.1093/infdis/143.3.325.

Beghain, J., Bridier-Nahmias, A., Le Nagard, H., Denamur, E., and Clermont, O. (2018). ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. *Microb. Genomics* 4. doi:10.1099/mgen.0.000192.

Belas, A., Marques, C., Aboim, C., and Pomba, C. (2018). Emergence of Escherichia coli ST131 H30/H30-Rx subclones in companion animals. *J. Antimicrob. Chemother.* doi:10.1093/jac/dky381.

Belas, A., Salazar, A. S., Gama, L. T. d., Couto, N., and Pomba, C. (2014). Risk factors for

faecal colonisation with Escherichia coli producing extended-spectrum and plasmid-mediated AmpC -lactamases in dogs. *Vet. Rec.* 175, 202–202. doi:10.1136/vr.101978.

Beloin, C., Roux, A., and Ghigo, J.-M. (2008). "Escherichia coli Biofilms," in, 249–289. doi:10.1007/978-3-540-75418-3_12.

Ben Zakour, N. L., Alsheikh-Hussain, A. S., Ashcroft, M. M., Khanh Nhu, N. T., Roberts, L. W., Stanton-Cook, M., et al. (2016). Sequential Acquisition of Virulence and Fluoroquinolone Resistance Has Shaped the Evolution of Escherichia coli ST131. *MBio* 7, no pagination. doi:10.1128/mBio.00347-16.

Bennett, J. (2003). Classical enteropathogenic Escherrichia coli or atypical strains? Examination of shigatoxin negative, eaeA positive isolates received in the Enteric Reference Laboratory in 2000. *New Zeal. J. Med. Lab. Sci.* 57, 2–7.

Bergthorsson, U., and Ochman, H. (1998). Distribution of chromosome length variation in natural isolates of Escherichia coli. *Mol. Biol. Evol.* 15, 6–16. doi:10.1093/oxfordjournals.molbev.a025847.

Bevan, E. R., Jones, A. M., and Hawkey, P. M. (2017). Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. *J. Antimicrob. Chemother.* 72, 2145–2155. doi:10.1093/jac/dkx146.

Biran, D., and Ron, E. Z. (2018). "Extraintestinal Pathogenic Escherichia coli," in, 149–161. doi:10.1007/82_2018_108.

Birgy, A., Levy, C., Bidet, P., Thollot, F., Derkx, V., Béchet, S., et al. (2016). ESBL-producing Escherichia coli ST131 versus non-ST131: evolution and risk factors of carriage among French children in the community between 2010 and 2015. *J. Antimicrob. Chemother.* 71, 2949–2956. doi:10.1093/jac/dkw219.

Birgy, A., Madhi, F., Jung, C., Levy, C., Cointe, A., Bidet, P., et al. (2019). Diversity and trends in population structure of ESBL-producing Enterobacteriaceae in febrile urinary tract infections in children in France from 2014 to 2017. *J. Antimicrob. Chemother.* doi:10.1093/jac/dkz423.

Blanc, V., Leflon-Guibout, V., Blanco, J., Haenni, M., Madec, J.-Y., Rafignon, G., et al. (2014). Prevalence of day-care centre children (France) with faecal CTX-M-producing Escherichia coli comprising O25b:H4 and O16:H5 ST131 strains. *J. Antimicrob. Chemother.* 69, 1231–1237. doi:10.1093/jac/dkt519.

Blanco, J., Blanco, J. E., Alonso, M. P., and Blanco, M. (1996). Virulence factors and O groups of Escherichia coli isolates from patients with acute pyelonefritis, cystitis and asymptomatic bacteriuria. *Eur. J. Epidemiol.* 12, 191–198.

Blanco, J., Blanco, M., Blanco, J.-E., Mora, A., and Alonso, P. (2002). "Enterobacterias: características generales. Género Escherichia," in *Manual de Microbiología Veterinaria. McGraw-Hill Interamericana*, 301–325.

Blanco, J., Mora, A., Mamani, R., Lopez, C., Blanco, M., Dahbi, G., et al. (2011). National survey of Escherichia coli causing extraintestinal infections reveals the spread of drug-resistant clonal groups O25b:H4-B2-ST131, O15:H1-D-ST393 and CGA-D-ST69 with high virulence gene content in Spain. *J. Antimicrob. Chemother.* 66, 2011–2021. doi:10.1093/jac/dkr235.

Blanco, M., Alonso, M. P., Nicolas-Chanoine, M.-H., Dahbi, G., Mora, A., Blanco, J. J. E., et al. (2009). Molecular epidemiology of Escherichia coli producing extended-spectrum β-lactamases in Lugo (Spain): dissemination of clone O25b:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* 63, 1135–1141. doi:10.1093/jac/dkp122.

Blanco, M., Blanco, J. E., Dahbi, G., Alonso, M. P., Mora, A., Coira, M. A., et al. (2006). Identification of two new intimin types in atypical enteropathogenic Escherichia coli. *Int. Microbiol.* 9, 103–110.

Blanco, M., Blanco, J. E., Gonzalez, E. A., Mora, A., Jansen, W., Gomes, T. A., et al. (1997). Genes coding for enterotoxins and verotoxins in porcine Escherichia coli strains belonging to different O:K:H serotypes: relationship with toxic phenotypes. *J. Clin. Microbiol.* 35, 2958–63. doi:10.1128/JCM.35.11.2958-2963.1997.

Blango, M. G., Ott, E. M., Erman, A., Veranic, P., and Mulvey, M. A. (2014). Forced Resurgence and Targeting of Intracellular Uropathogenic Escherichia coli Reservoirs. *PLoS One* 9, e93327.

doi:10.1371/journal.pone.0093327.

Blattner, F. R. (1997). The Complete Genome Sequence of Escherichia coli K-12. *Science (80-.).* 277, 1453–1462. doi:10.1126/science.277.5331.1453.

Blondeau, J. M. (2009). *STAT, Steps to antimicrobial therapy*. , ed. Inter-Médica S.A. Port Huron, USA: North American Compendium Inc.

Boehmer, T., Vogler, A. J., Thomas, A., Sauer, S., Hergenroether, M., Straubinger, R. K., et al. (2018). Phenotypic characterization and whole genome analysis of extended-spectrum beta-lactamase-producing bacteria isolated from dogs in Germany. *PLoS One* 13, e0206252. doi:10.1371/journal.pone.0206252.

Bogaerts, P., Huang, T.-D., Bouchahrouf, W., Bauraing, C., Berhin, C., El Garch, F., et al. (2015). Characterization of ESBL- and AmpC-Producing Enterobacteriaceae from Diseased Companion Animals in Europe. *Microb. Drug Resist.* 21, 643–650. doi:10.1089/mdr.2014.0284.

Bogema, D. R., McKinnon, J., Liu, M., Hitchick, N., Miller, N., Venturini, C., et al. (2020). Whole-genome analysis of extraintestinal Escherichia coli sequence type 73 from a single hospital over a 2 year period identified different circulating clonal groups. *Microb. Genomics* 6. doi:10.1099/mgen.0.000255.

Boolchandani, M., D'Souza, A. W., and Dantas, G. (2019). Sequencing-based methods and resources to study antimicrobial resistance. *Nat. Rev. Genet.* doi:10.1038/s41576-019-0108-4.

Borowiak, M., Fischer, J., Hammerl, J. A., Hendriksen, R. S., Szabo, I., and Malorny, B. (2017). Identification of a novel transposon-associated phosphoethanolamine transferase gene, mcr-5, conferring colistin resistance in d-tartrate fermenting Salmonella enterica subsp. enterica serovar Paratyphi B. *J. Antimicrob. Chemother.* 72, 3317–3324. doi:10.1093/jac/dkx327.

Bortolami, A., Zendri, F., Maciuca, E. I., Wattret, A., Ellis, C., Schmidt, V., et al. (2019). Diversity, Virulence, and Clinical Significance of Extended-Spectrum β-Lactamase- and pAmpC-Producing Escherichia coli From Companion Animals. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01260.

Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., et al. (2009). Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* 48, 1–12. doi:10.1086/595011.

Bouillon, J., Snead, E., Caswell, J., Feng, C., Hélie, P., and Lemetayer, J. (2018). Pyelonephritis in Dogs: Retrospective Study of 47 Histologically Diagnosed Cases (2005-2015). *J. Vet. Intern. Med.* 32, 249–259. doi:10.1111/jvim.14836.

Bourne, J. A., Chong, W. L., and Gordon, D. M. (2019). Genetic structure, antimicrobial resistance and frequency of human associated Escherichia coli sequence types among faecal isolates from healthy dogs and cats living in Canberra, Australia. *PLoS One* 14, e0212867. doi:10.1371/journal.pone.0212867.

Boyd, E. F., Hill, C. W., Rich, S. M., and Hartl, D. L. (1996). Mosaic structure of plasmids from natural populations of Escherichia coli. *Genetics* 143, 1091–100.

Brisse, S., Diancourt, L., Laouenan, C., Vigan, M., Caro, V., Arlet, G., et al. (2012). Phylogenetic Distribution of CTX-M- and Non-Extended-Spectrum- -Lactamase-Producing Escherichia coli Isolates: Group B2 Isolates, Except Clone ST131, Rarely Produce CTX-M Enzymes. *J. Clin. Microbiol.* 50, 2974–2981. doi:10.1128/JCM.00919-12.

Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., and MacLean, C. (2019). The Ecology and Evolution of Pangenomes. *Curr. Biol.* 29, R1094–R1103. doi:10.1016/j.cub.2019.08.012.

Brumbaugh, A. R., and Mobley, H. L. (2012). Preventing urinary tract infection: progress toward an effective Escherichia coli vaccine. *Expert Rev. Vaccines* 11, 663–676. doi:10.1586/erv.12.36.

Bubpamala, J., Khuntayaporn, P., Thirapanmethee, K., Montakantikul, P., Santanirand, P., and Chomnawang, M. T. (2018). Phenotypic and genotypic characterizations of extended-spectrum beta-lactamase-producing Escherichia coli in Thailand. *Infect. Drug Resist.* Volume 11, 2151–2157. doi:10.2147/IDR.S174506.

Burrus, V., Pavlovic, G., Decaris, B., and Guédon, G. (2002). Conjugative transposons: the tip of the iceberg. *Mol.*

Microbiol. 46, 601–610. doi:10.1046/j.1365-2958.2002.03191.x.

Busch, A., and Waksman, G. (2012). Chaperone–usher pathways: diversity and pilus assembly mechanism. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1112–1122. doi:10.1098/rstb.2011.0206.

Bush, K. (2013). The ABCD's of β-lactamase nomenclature. *J. Infect. Chemother.* 19, 549–559. doi:10.1007/s10156-013-0640-7.

Bush, K., and Jacoby, G. A. (2010). Updated functional classification of β-lactamases. *Antimicrob. Agents Chemother.* 54, 969–976. doi:10.1128/AAC.01009-09.

Butaye, P. (2003). Mobile genes coding for efflux-mediated antimicrobial resistance in Gram-positive and Gram-negative bacteria. *Int. J. Antimicrob. Agents* 22, 205–210. doi:10.1016/S0924-8579(03)00202-4.

Cabello, F. C. (2006). Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ. Microbiol.* 8, 1137–1144. doi:10.1111/j.1462-2920.2006.01054.x.

Cai, W., Wannemuehler, Y., Dell'Anna, G., Nicholson, B., Barbieri, N. L., Kariyawasam, S., et al. (2013). A Novel Two-Component Signaling System Facilitates Uropathogenic Escherichia coli's Ability to Exploit Abundant Host Metabolites. *PLoS Pathog.* 9, e1003428. doi:10.1371/journal.ppat.1003428.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-2105-10-421.

Campos, J., Cristino, L., Peixe, L., and Antunes, P. (2016). MCR-1 in multidrug-resistant and copper-tolerant clinically relevant Salmonella 1,4,[5],12:i:- and S . Rissen clones in Portugal, 2011 to 2015. *Eurosurveillance* 21, 30270. doi:10.2807/1560-7917.ES.2016.21.26.30270.

Camprubí-Font, C., Ewers, C., Lopez-Siles, M., and Martinez-Medina, M. (2019). Genetic and Phenotypic Features to Screen for Putative Adherent-Invasive Escherichia coli. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.00108.

Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L., and Threlfall, E. J. (2005). Identification of plasmids by PCR-based replicon typing. *J. Microbiol. Methods* 63, 219–228. doi:10.1016/j.mimet.2005.03.018.

Carattoli, A., Villa, L., Feudi, C., Curcio, L., Orsini, S., Luppi, A., et al. (2017). Novel plasmid-mediated colistin resistance mcr-4 g ene in Salmonella and Escherichia coli , Italy 2013, Spain and Belgium, 2015 to 2016. *Eurosurveillance* 22, 30589. doi:10.2807/1560-7917.ES.2017.22.31.30589.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi:10.1128/AAC.02412-14.

Carraro, N., and Burrus, V. (2015). The dualistic nature of integrative and conjugative elements. *Mob. Genet. Elements* 5, 98–102. doi:10.1080/2159256X.2015.1102796.

Carroll, L. M., Gaballa, A., Guldimann, C., Sullivan, G., Henderson, L. O., and Wiedmann, M. (2019). Identification of novel mobilized colistin resistance gene mcr-9 in a multidrug-resistant, colistin-susceptible salmonella enterica serotype typhimurium isolate. *MBio* 10. doi:10.1128/mBio.00853-19.

Carvalho, I., Tejedor-Junco, M. T., González-Martín, M., Corbera, J. A., Suárez-Pérez, A., Silva, V., et al. (2020). Molecular diversity of Extended-spectrum β-lactamase-producing Escherichia coli from vultures in Canary Islands. *Environ. Microbiol. Rep.* 12, 540–547. doi:10.1111/1758-2229.12873.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics* 21, 3422–3423. doi:10.1093/bioinformatics/bti553.

Cash, P. (2014). Proteomic analysis of uropathogenic Escherichia coli. *Expert Rev. Proteomics* 11, 43–58. doi:10.1586/14789450.2014.877845.

Cassini, A., Hõgberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., et al. (2019). Attributable deaths and disability-adjusted life-years caused by

infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* 19, 56–66. doi:10.1016/S1473-3099(18)30605-4.

Chandler, M., de la Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G., and Ton-Hoang, B. (2013). Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.* 11, 525–538. doi:10.1038/nrmicro3067.

Chandler, M., and Siguier, P. (2013). "Insertion Sequences," in *Brenner's Encyclopedia of Genetics: Second Edition* (Elsevier), 86–94. doi:10.1016/B978-0-12-374984-0.00799-3.

Chaudhuri, R. R., and Henderson, I. R. (2012). The evolution of the Escherichia coli phylogeny. *Infect. Genet. Evol.* 12, 214–226. doi:10.1016/j.meegid.2012.01.005.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., et al. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* 33, D325–D328. doi:10.1093/nar/gki008.

Chen, X., Zou, Q., Zhang, W., Wang, R., Yu, F., and Chen, Y. (2019a). Clinical features and microbiological characteristics of hospital- and community-onset Escherichia coli bloodstream infection. *J. Med. Microbiol.* 68, 178–187. doi:10.1099/jmm.0.000904.

Chen, Y., Liu, Z., Zhang, Y., Zhang, Z., Lei, L., and Xia, Z. (2019b). Increasing Prevalence of ESBL-Producing Multidrug Resistance Escherichia coli From Diseased Pets in Beijing, China From 2012 to 2017. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.02852.

Chu, B. C. H., and Vogel, H. J. (2011). A structural and functional analysis of type III periplasmic and substrate binding proteins: their role in bacterial siderophore and heme transport. *Biol. Chem.* 392. doi:10.1515/bc.2011.012.

Chu, S.-M., Hsu, J.-F., Lee, C.-W., Lien, R., Huang, H.-R., Chiang, M.-C., et al. (2014). Neurological Complications after Neonatal Bacteremia: The Clinical Characteristics, Risk Factors, and Outcomes. *PLoS One* 9, e105294. doi:10.1371/journal.pone.0105294.

Clausen, P. T. L. C., Zankari, E., Aarestrup, F. M., and Lund, O. (2016). Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *J. Antimicrob. Chemother.* 71, 2484–2488. doi:10.1093/jac/dkw184.

Clermont, O., Bonacorsi, S., and Bingen, E. (2000). Rapid and Simple Determination of theEscherichia coli Phylogenetic Group. *Appl. Environ. Microbiol.* 66, 4555–4558. doi:10.1128/AEM.66.10.4555-4558.2000.

Clermont, O., Christenson, J. K., Denamur, E., and Gordon, D. M. (2013). The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* 5, 58–65. doi:10.1111/1758-2229.12019.

Clermont, O., Couffignal, C., Blanco, J., Mentré, F., Picard, B., and Denamur, E. (2017). Two levels of specialization in bacteraemic Escherichia coli strains revealed by their comparison with commensal strains. *Epidemiol. Infect.* 145, 872–882. doi:10.1017/S0950268816003010.

Clermont, O., Dixit, O. V. A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., et al. (2019). Characterization and rapid identification of phylogroup G in Escherichia coli, a lineage with high virulence and antibiotic resistance potential. *Environ. Microbiol.* 21, 3107–3117. doi:10.1111/1462-2920.14713.

Clermont, O., Gordon, D. M., Brisse, S., Walk, S. T., and Denamur, E. (2011). Characterization of the cryptic Escherichia lineages: rapid identification and prevalence. *Environ. Microbiol.* 13, 2468–2477. doi:10.1111/j.1462-2920.2011.02519.x.

Clermont, O., Lavollay, M., Vimont, S., Deschamps, C., Forestier, C., Branger, C., et al. (2008). The CTX-M-15-producing Escherichia coli diffusing clone belongs to a highly virulent B2 phylogenetic subgroup. *J. Antimicrob. Chemother.* 61, 1024–1028. doi:10.1093/jac/dkn084.

CLSI (2017). "Performance Standards for Antimicrobial Susceptibility Testing, 27th ed," in (950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA: Clinical and Laboratory Standards Institute), 332.

Codex Alimentarius (2011). Guidelines for risk analysis of foodborne antimicrobial resistance (CAC/GL 77- 2011).

Coelho, A., Mora, A., Mamani, R., Lopez, C.,

Gonzalez-Lopez, J. J., Larrosa, M. N., et al. (2011). Spread of Escherichia coli O25b:H4-B2-ST131 producing CTX-M-15 and SHV-12 with high virulence gene content in Barcelona (Spain). *J. Antimicrob. Chemother.* 66, 517–526. doi:10.1093/jac/dkq491.

Cole, B. K., Ilikj, M., McCloskey, C. B., and Chavez-Bueno, S. (2019). Antibiotic resistance and molecular characterization of bacteremia Escherichia coli isolates from newborns in the United States. *PLoS One* 14, e0219352. doi:10.1371/journal.pone.0219352.

Colpan, A., Johnston, B., Porter, S., Clabots, C., Anway, R., Thao, L., et al. (2013). Escherichia coli Sequence Type 131 (ST131) Subclone H30 as an Emergent Multidrug-Resistant Pathogen Among US Veterans. *Clin. Infect. Dis.* 57, 1256–1265. doi:10.1093/cid/cit503.

Conover, M. S., Ruer, S., Taganna, J., Kalas, V., De Greve, H., Pinkner, J. S., et al. (2016). Inflammation-Induced Adhesin-Receptor Interaction Provides a Fitness Advantage to Uropathogenic E. coli during Chronic Infection. *Cell Host Microbe* 20, 482–492. doi:10.1016/j.chom.2016.08.013.

Conway, T., and Cohen, P. S. (2015). "Commensal and Pathogenic Escherichia coli Metabolism in the Gut," in *Metabolism and Bacterial Pathogenesis* (American Society of Microbiology), 343–362. doi:10.1128/microbiolspec.MBP-0006-2014.

Cortés, P., Blanc, V., Mora, A., Dahbi, G., Blanco, J. E., Blanco, M., et al. (2010). Isolation and Characterization of Potentially Pathogenic Antimicrobial-Resistant Escherichia coli Strains from Chicken and Pig Farms in Spain. *Appl. Environ. Microbiol.* 76, 2799–2805. doi:10.1128/AEM.02421-09.

Costerton, J. W., Geesey, G. G., and Cheng, K.-J. (1978). How Bacteria Stick. *Sci. Am.* 238, 86–95. doi:10.1038/scientificamerican0178-86.

Couturier, M., Bex, F., Bergquist, P. L., and Maas, W. K. (1988). Identification and classification of bacterial plasmids. *Microbiol. Rev.* 52, 375–95.

Cristóvýo, F., Alonso, C. A., Igrejas, G., Sousa, M., Silva, V., Pereira, J. E., et al. (2017). Clonal diversity of extended-spectrum beta-lactamase producing Escherichia coli isolates in fecal samples of wild animals. *FEMS Microbiol. Lett.* 364. doi:10.1093/femsle/fnx039.

Dahbi, G., Mora, A., López, C., Alonso, M. P., Mamani, R., Marzoa, J., et al. (2013). Emergence of new variants of ST131 clonal group among extraintestinal pathogenic Escherichia coli producing extended-spectrum β-lactamases. *Int. J. Antimicrob. Agents* 42, 347–351. doi:10.1016/j.ijantimicag.2013.06.017.

Dahbi, G., Mora, A., Mamani, R., López, C., Alonso, M. P., Marzoa, J., et al. (2014). Molecular epidemiology and virulence of Escherichia coli O16:H5-ST131: Comparison with H30 and H30-Rx subclones of O25b:H4-ST131. *Int. J. Med. Microbiol.* 304, 1247–1257. doi:10.1016/j.ijmm.2014.10.002.

Dale, A. P., and Woodford, N. (2015). Extra-intestinal pathogenic Escherichia coli (ExPEC): Disease, carriage and clones. *J. Infect.* 71, 615–626. doi:10.1016/j.jinf.2015.09.009.

Damborg, P., Morsing, M. K., Petersen, T., Bortolaia, V., and Guardabassi, L. (2015). CTX-M-1 and CTX-M-15-producing Escherichia coli in dog faeces from public gardens. *Acta Vet. Scand.* 57, 83. doi:10.1186/s13028-015-0174-3.

Danzeisen, J. L., Wannemuehler, Y., Nolan, L. K., and Johnson, T. J. (2013). Comparison of Multilocus Sequence Analysis and Virulence Genotyping of Escherichia coli from Live Birds, Retail Poultry Meat, and Human Extraintestinal Infection. *Avian Dis.* 57, 104–108. doi:10.1637/10218-042812-ResNote.1.

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.-L., Barnich, N., et al. (2004). High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease. *Gastroenterology* 127, 412–421. doi:10.1053/j.gastro.2004.04.061.

Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P., et al. (1998). Presence of adherent Escherichia coli strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* 115, 1405–1413. doi:10.1016/S0016-5085(98)70019-8.

Davies, D. (2003). Understanding biofilm resistance to antibacterial agents. *Nat. Rev. Drug Discov.* 2, 114–122.

doi:10.1038/nrd1008.

Day, M. J., Doumith, M., Abernethy, J., Hope, R., Reynolds, R., Wain, J., et al. (2016). Population structure of Escherichia coli causing bacteraemia in the UK and Ireland between 2001 and 2010. *J. Antimicrob. Chemother.* 71, 2139–2142. doi:10.1093/jac/dkw145.

de Been, M., Lanza, V. F., de Toro, M., Scharringa, J., Dohmen, W., Du, Y., et al. (2014). Dissemination of Cephalosporin Resistance Genes between Escherichia coli Strains from Farm Animals and Humans by Specific Plasmid Lineages. *PLoS Genet.* 10, e1004776. doi:10.1371/journal.pgen.1004776.

de Toro, M., Fernández, J., García, V., Mora, A., Blanco, J., de la Cruz, F., et al. (2017). Whole genome sequencing, molecular typing and in vivo virulence of OXA-48-producing Escherichia coli isolates including ST131 H30-Rx, H22 and H41 subclones. *Sci. Rep.* 7, 12103. doi:10.1038/s41598-017-12015-0.

Decano, A. G., and Downing, T. (2019). An Escherichia coli ST131 pangenome atlas reveals population structure and evolution across 4,071 isolates. *Sci. Rep.* 9, 17394. doi:10.1038/s41598-019-54004-5.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., and Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* 62, 434–64.

Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., and van der Meer, J. R. (2017). The hidden life of integrative and conjugative elements. *FEMS Microbiol. Rev.* 41, 512–537. doi:10.1093/femsre/fux008.

Denamur, E., Clermont, O., Bonacorsi, S., and Gordon, D. (2020). The population genetics of pathogenic Escherichia coli. *Nat. Rev. Microbiol.* doi:10.1038/s41579-020-0416-x.

Destoumieux-Garzón, D., Mavingui, P., Boetsch, G., Boissier, J., Darriet, F., Duboz, P., et al. (2018). The One Health Concept: 10 Years Old and a Long Road Ahead. *Front. Vet. Sci.* 5, 1–13. doi:10.3389/fvets.2018.00014.

Desvaux, M., Dalmasso, G., Beyrouthy, R., Barnich, N., Delmas, J., and Bonnet, R. (2020). Pathogenicity Factors of Genomic Islands in Intestinal and Extraintestinal Escherichia coli. *Front. Microbiol.* 11. doi:10.3389/fmicb.2020.02065.

Dhanji, H., Patel, R., Wall, R., Doumith, M., Patel, B., Hope, R., et al. (2011). Variation in the genetic environments of blaCTX-M-15 in Escherichia coli from the faeces of travellers returning to the United Kingdom. *J. Antimicrob. Chemother.* 66, 1005–1012. doi:10.1093/jac/dkr041.

Diaz, M. A., Hernandez-Bello, J. R., Rodriguez-Bano, J., Martinez-Martinez, L., Calvo, J., Blanco, J., et al. (2010). Diversity of Escherichia coli Strains Producing Extended-Spectrum -Lactamases in Spain: Second Nationwide Study. *J. Clin. Microbiol.* 48, 2840–2845. doi:10.1128/JCM.02147-09.

Dickschat, J. S. (2010). Quorum sensing and bacterial biofilms. *Nat. Prod. Rep.* 27, 343. doi:10.1039/b804469b.

Dierikx, C. M., van Duijkeren, E., Schoormans, A. H. W., van Essen-Zandbergen, A., Veldman, K., Kant, A., et al. (2012). Occurrence and characteristics of extended-spectrum-β-lactamase- and AmpC-producing clinical isolates derived from companion animals and horses. *J. Antimicrob. Chemother.* 67, 1368–1374. doi:10.1093/jac/dks049.

Ding, Y., Tang, X., Lu, P., Wu, B., Xu, Z., Liu, W., et al. (2012). Clonal analysis and virulent traits of pathogenic extraintestinal Escherichia coli isolates from swine in China. *BMC Vet. Res.* 8, 140. doi:10.1186/1746-6148-8-140.

Dobrindt, U. (2005). (Patho-)Genomics of Escherichia coli. *Int. J. Med. Microbiol.* 295, 357–371. doi:10.1016/j.ijmm.2005.07.009.

Dobrindt, U., Agerer, F., Michaelis, K., Janka, A., Buchrieser, C., Samuelson, M., et al. (2003). Analysis of Genome Plasticity in Pathogenic and Commensal Escherichia coli Isolates by Use of DNA Arrays. *J. Bacteriol.* 185, 1831–1840. doi:10.1128/JB.185.6.1831-1840.2003.

Dobrindt, U., Blum-Oehler, G., Nagy, G., Schneider, G., Johann, A., Gottschalk, G., et al. (2002). Genetic Structure and Distribution of Four Pathogenicity Islands (PAI I536 to PAI IV536) of Uropathogenic Escherichia coli Strain 536. *Infect. Immun.* 70, 6365–6372. doi:10.1128/IAI.70.11.6365-6372.2002.

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* 2, 414–424. doi:10.1038/nrmicro884.

Dozois, C. M., Dho-Moulin, M., Brée, A., Fairbrother, J. M., Desautels, C., and Curtiss, R. (2000). Relationship between the Tsh Autotransporter and Pathogenicity of Avian Escherichia coli and Localization and Analysis of the tsh Genetic Region. *Infect. Immun.* 68, 4145–4154. doi:10.1128/IAI.68.7.4145-4154.2000.

Du, H., Chen, L., Tang, Y. W., and Kreiswirth, B. N. (2016). Emergence of the mcr-1 colistin resistance gene in carbapenem-resistant Enterobacteriaceae. *Lancet Infect. Dis.* 16, 287–288. doi:10.1016/S1473-3099(16)00056-6.

Dupouy, V., Abdelli, M., Moyano, G., Arpaillange, N., Bibbal, D., Cadiergues, M.-C., et al. (2019). Prevalence of Beta-Lactam and Quinolone/Fluoroquinolone Resistance in Enterobacteriaceae From Dogs in France and Spain—Characterization of ESBL/pAmpC Isolates, Genes, and Conjugative Plasmids. *Front. Vet. Sci.* 6. doi:10.3389/fvets.2019.00279.

Duprilot, M., Baron, A., Blanquart, F., Dion, S., Lettéron, P., Flament-Simon, S.-C., et al. (2019). Success of Escherichia coli O25b:H4 ST131 clade C associated with a decrease in virulence. *bioRxiv*, 786350. doi:10.1101/786350.

Durand, G. A., Raoult, D., and Dubourg, G. (2019). Antibiotic discovery: history, methods and perspectives. *Int. J. Antimicrob. Agents* 53, 371–382. doi:10.1016/j.ijantimicag.2018.11.010.

Dykhuizen, D. (2005). "Species Numbers in Bacteria." Proceedings. *Calif. Acad. Sci. vol. 56,6*, 62–75.

ECDC/EMEA (2009). The bacterial challenge: time to react.

Eguchi, Y., Ishii, E., Hata, K., and Utsumi, R. (2011). Regulation of Acid Resistance by Connectors of Two-Component Signal Transduction Systems in Escherichia coli. *J. Bacteriol.* 193, 1222–1228. doi:10.1128/JB.01124-10.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-Time DNA Sequencing from Single Polymerase Molecules. *Science (80-.).* 323, 133–138. doi:10.1126/science.1162986.

Ellaby, N., Doumith, M., Hopkins, K. L., Woodford, N., and Ellington, M. J. (2019). Emergence of diversity in carbapenemase-producing Escherichia coli ST131, England, January 2014 to June 2016. *Eurosurveillance* 24. doi:10.2807/1560-7917.ES.2019.24.37.1800627.

Ellis, T. N., and Kuehn, M. J. (2010). Virulence and Immunomodulatory Roles of Bacterial Outer Membrane Vesicles. *Microbiol. Mol. Biol. Rev.* 74, 81–94. doi:10.1128/MMBR.00031-09.

Escherich, T. (1885). "Die darmbakterien des neugeborenen und sãuglings." *Fortschr. Med.* 3, 515–522.

Escobar-Páramo, P., Clermont, O., Blanc-Potard, A.-B., Bui, H., Le Bouguénec, C., and Denamur, E. (2004). A Specific Genetic Background Is Required for Acquisition and Expression of Virulence Factors in Escherichia coli. *Mol. Biol. Evol.* 21, 1085–1094. doi:10.1093/molbev/msh118.

EUCAST (2020). The European Committe on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters, Version 10.0.

European Commission (2001). Prudent use of antimicrobial agents in human medicine: third report on implementation of the Council recommendation (2002/77/EC).

European Commission (2015). Guidelines for the prudent use of antimicrobials in veterinary medicine (2015/C 299/04).

European Commission (2017). A European One Health Action Plan against Antimicrobial Resistance (AMR).

European Parliament (2015). safer healthcare in Europe: improving patient safety and fighting antimicrobial resistance (2014/2207(INI).

European Parliament and Council (2001). On the community code relating to veterinary medicinal products (2001/82/EC).

Ewers, C., Bethe, A., Stamm, I., Grobbel, M., Kopp, P. A., Guerra, B., et al. (2014). CTX-M-15-D-ST648 Escherichia coli from companion animals and horses: another pandemic clone combining multiresistance and extraintestinal virulence? *J. Antimicrob. Chemother.* 69, 1224–1230. doi:10.1093/jac/dkt516.

Ewers, C., Gõttig, S., Bũlte, M., Fiedler, S., Tietgen, M., Leidner, U., et al. (2016). Genome Sequence of Avian Escherichia coli Strain IHIT25637, an Extraintestinal Pathogenic E. coli Strain of ST131 Encoding Colistin Resistance Determinant MCR-1. *Genome Announc.* 4, e00863-16. doi:10.1128/genomeA.00863-16.

Ewers, C., Grobbel, M., Stamm, I., Kopp, P. A., Diehl, I., Semmler, T., et al. (2010). Emergence of human pandemic O25:H4-ST131 CTX-M-15 extended-spectrum-β-lactamase-producing Escherichia coli among companion animals. *J. Antimicrob. Chemother.* 65, 651–660. doi:10.1093/jac/dkq004.

Falgenhauer, L., Imirzalioglu, C., Ghosh, H., Gwozdzinski, K., Schmiedel, J., Gentil, K., et al. (2016). Circulation of clonal populations of fluoroquinolone-resistant CTX-M-15-producing Escherichia coli ST410 in humans and animals in Germany. *Int. J. Antimicrob. Agents* 47, 457–465. doi:10.1016/j.ijantimicag.2016.03.019.

FAO-OIE-WHO (2010). The FAO-OIE-WHO Collaboration.

FAO-OIE-WHO (2017). The Tripartite's Commitment.

FAO (2011). Preventing Escherichia coli in food. Rome, Italy: Food Chain Crisis Management Framework.

FAO (2016). THE FAO ACTION PLAN ON ANTIMICROBIAL RESISTANCE 2016-2020.

Fernández-Rodríguez, G., and Prieto-Lago, A. (2017a). "Módulo I. Antibiograma," in *Cómo utilizar e interpretar los antibiogramas para optimizar el tratamiento utilizando como modelo las mamitis bovinas* (Consejo General de Colegios Veterinarios de España).

Fernández-Rodríguez, G., and Prieto-Lago, A. (2017b). "Módulo II. Antibióticos," in *Cómo utilizar e interpretar los antibiogramas para optimizar el tratamiento utilizando como modelo las mamitis bovinas* (Consejo General de Colegios Veterinarios de España).

Fernandez, L., and Hancock, R. E. W. (2012). Adaptive and Mutational Resistance: Role of Porins and Efflux Pumps in Drug Resistance. *Clin. Microbiol. Rev.* 25, 661–681. doi:10.1128/CMR.00043-12.

Ferry, T., Batailler, C., Petitjean, C., Chateau, J., Fevre, C., Forestier, E., et al. (2020). The Potential Innovative Use of Bacteriophages Within the DAC® Hydrogel to Treat Patients With Knee Megaprosthesis Infection Requiring "Debridement Antibiotics and Implant Retention" and Soft Tissue Coverage as Salvage Therapy. *Front. Med.* 7. doi:10.3389/fmed.2020.00342.

Fibke, C. D., Croxen, M. A., Geum, H. M., Glass, M., Wong, E., Avery, B. P., et al. (2019). Genomic Epidemiology of Major Extraintestinal Pathogenic Escherichia coli Lineages Causing Urinary Tract Infections in Young Women Across Canada. *Open Forum Infect. Dis.* 6. doi:10.1093/ofid/ofz431.

Findlay, J., Gould, V. C., North, P., Bowker, K. E., Williams, M. O., MacGowan, A. P., et al. (2020). Characterization of cefotaxime-resistant urinary Escherichia coli from primary care in South-West England 2017–18. *J. Antimicrob. Chemother.* 75, 65–71. doi:10.1093/jac/dkz397.

Finlay, B. B., and McFadden, G. (2006). Anti-Immunology: Evasion of the Host Immune System by Bacterial and Viral Pathogens. *Cell* 124, 767–782. doi:10.1016/j.cell.2006.01.034.

Flament-Simon, S.-C., Duprilot, M., Mayer, N., García, V., Alonso, M. P., Blanco, J., et al. (2019). Association Between Kinetics of Early Biofilm Formation and Clonal Lineage in Escherichia coli. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01183.

Flament-Simon, S.-C., García, V., Duprilot, M., Mayer, N., Alonso, M. P., García-Meniño, I., et al. (2020a). High Prevalence of ST131 Subclades C2-H30Rx and C1-M27 Among Extended-Spectrum β-Lactamase-Producing Escherichia coli Causing Human Extraintestinal Infections in Patients From Two Hospitals of Spain and France During 2015. *Front. Cell. Infect. Microbiol.* 10. doi:10.3389/fcimb.2020.00125.

Flament-Simon, S.-C., Nicolas-Chanoine, M.-H., García, V., Duprilot, M., Mayer, N., Alonso, M. P., et al. (2020b). 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* 9, 161. doi:10.3390/antibiotics9040161.

Fleece, M. E., Pholwat, S., Mathers, A. J., and Houpt, E. R. (2018). Molecular diagnosis of antimicrobial resistance in Escherichia coli. *Expert Rev. Mol. Diagn.* 18, 207–217. doi:10.1080/14737159.2018.1439381.

Flores-Mireles, A. L., Walker, J. N., Caparon, M., and Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* 13, 269–284. doi:10.1038/nrmicro3432.

Fox, S., Goswami, C., Holden, M., Connolly, J. P. R., Mordue, J., O'Boyle, N., et al. (2020). A highly conserved complete accessory Escherichia coli type III secretion system 2 is widespread in bloodstream isolates of the ST69 lineage. *Sci. Rep.* 10, 4135. doi:10.1038/s41598-020-61026-x.

Foxman, B. (2003). Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Disease-a-Month* 113, 5–13. doi:10.1016/S0002-9343(02)01054-9.

Foxman, B. (2014). Urinary Tract Infection Syndromes. *Infect. Dis. Clin. North Am.* 28, 1–13. doi:10.1016/j.idc.2013.09.003.

Foxman, B., Gillespie, B., Koopman, J., Zhang, L., Palin, K., Tallman, P., et al. (2000). Risk Factors for Second Urinary Tract Infection among College Women. *Am. J. Epidemiol.* 151, 1194–1205. doi:10.1093/oxfordjournals.aje.a010170.

Francia, M. V., Varsaki, A., Garcillán-Barcia, M. P., Latorre, A., Drainas, C., and de la Cruz, F. (2004). A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol. Rev.* 28, 79–100. doi:10.1016/j.femsre.2003.09.001.

Frank, C., Werber, D., Cramer, J. P., Askar, M., Faber, M., an der Heiden, M., et al. (2011). Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany. *N. Engl. J. Med.* 365, 1771–80. doi:10.1056/NEJMoa1106483.

Fratamico, P. M., DebRoy, C., Liu, Y., Needleman, D. S., Baranzoni, G. M., and Feng, P. (2016). Advances in Molecular Serotyping and Subtyping of Escherichia coliЍ. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00644.

Freitag, C., Michael, G. B., Kadlec, K., Hassel, M., and Schwarz, S. (2017). Detection of plasmid-borne extended-spectrum β-lactamase (ESBL) genes in Escherichia coli isolates from bovine mastitis. *Vet. Microbiol.* 200, 151–156. doi:10.1016/j.vetmic.2016.08.010.

Garcés, A., Correia, S., Amorim, F., Pereira, J. E., Igrejas, G., and Poeta, P. (2019). First report on extended-spectrum beta-lactamase (ESBL) producing Escherichia coli from European free-tailed bats (Tadarida teniotis) in Portugal: A one-health approach of a hidden contamination problem. *J. Hazard. Mater.* 370, 219–224. doi:10.1016/j.jhazmat.2017.12.053.

García-Meniño, I., Díaz-Jiménez, D., García, V., de Toro, M., Flament-Simon, S. C., Blanco, J., et al. (2019). Genomic Characterization of Prevalent mcr-1, mcr-4, and mcr-5 Escherichia coli Within Swine Enteric Colibacillosis in Spain. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.02469.

García-Meniño, I., García, V., Mora, A., Díaz-Jiménez, D., Flament-Simon, S. C., Alonso, M. P., et al. (2018). Swine Enteric Colibacillosis in Spain: Pathogenic Potential of mcr-1 ST10 and ST131 E. coli Isolates. *Front. Microbiol.* 9, 2659. doi:10.3389/fmicb.2018.02659.

Garcillán-Barcia, M. P., Alvarado, A., and de la Cruz, F. (2011). Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* 35, 936–956. doi:10.1111/j.1574-6976.2011.00291.x.

Garcillán-Barcia, M. P., and de la Cruz, F. (2008). Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* 60, 1–18. doi:10.1016/j.plasmid.2008.03.002.

Garcillán-Barcia, M. P., Francia, M. V., and de La Cruz, F. (2009). The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* 33, 657–687. doi:10.1111/j.1574-6976.2009.00168.x.

Garcillán-Barcia, M. P., Redondo-Salvo, S., Vielva, L., and de la Cruz, F. (2020). "MOBscan: Automated Annotation of MOB Relaxases," in, 295–308. doi:10.1007/978-1-4939-9877-7_21.

Gati, N. S., Middendorf-Bauchart, B., Bletz, S., Dobrindt, U., and Mellmann, A. (2019). Origin and Evolution of Hybrid Shiga

Toxin-Producing and Uropathogenic Escherichia coli Strains of Sequence Type 141. *J. Clin. Microbiol.* 58. doi:10.1128/JCM.01309-19.

Ghigo, J.-M. (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature* 412, 442–445. doi:10.1038/35086581.

Gibreel, T. M., Dodgson, A. R., Cheesbrough, J., Fox, A. J., Bolton, F. J., and Upton, M. (2012). Population structure, virulence potential and antibiotic susceptibility of uropathogenic Escherichia coli from Northwest England. *J. Antimicrob. Chemother.* 67, 346–356. doi:10.1093/jac/dkr451.

Gillings, M. R. (2014). Integrons: Past, Present, and Future. *Microbiol. Mol. Biol. Rev.* 78, 257–277. doi:10.1128/Mmbr.00056-13.

Glasser, A.-L., Boudeau, J., Barnich, N., Perruchot, M.-H., Colombel, J.-F., and Darfeuille-Michaud, A. (2001). Adherent Invasive Escherichia coli Strains from Patients with Crohn's Disease Survive and Replicate within Macrophages without Inducing Host Cell Death. *Infect. Immun.* 69, 5529–5537. doi:10.1128/IAI.69.9.5529-5537.2001.

Goller, C. C., and Seed, P. C. (2010). Revisiting the Escherichia coli polysaccharide capsule as a virulence factor during urinary tract infection: Contribution to intracellular biofilm development. *Virulence* 1, 333–337. doi:10.4161/viru.1.4.12388.

Gómez-Duarte, O. G., Arzuza, O., Urbina, D., Bai, J., Guerra, J., Montes, O., et al. (2010). Detection of Escherichia coli Enteropathogens by Multiplex Polymerase Chain Reaction from Children's Diarrheal Stools in Two Caribbean–Colombian Cities. *Foodborne Pathog. Dis.* 7, 199–206. doi:10.1089/fpd.2009.0355.

Gordon, D. M., and Cowling, A. (2003). The distribution and genetic structure of Escherichia coli in Australian vertebrates: host and geographic effects. *Microbiology* 149, 3575–3586. doi:10.1099/mic.0.26486-0.

Gordon, D. M., Geyik, S., Clermont, O., O'Brien, C. L., Huang, S., Abayasekara, C., et al. (2017a). Fine-Scale Structure Analysis Shows Epidemic Patterns of Clonal Complex 95, a Cosmopolitan Escherichia coli Lineage Responsible for Extraintestinal Infection. *mSphere* 2. doi:10.1128/mSphere.00168-17.

Gordon, D. M., Geyik, S., Clermont, O., O'Brien, C. L., Huang, S., Abayasekara, C., et al. (2017b). Fine-Scale Structure Analysis Shows Epidemic Patterns of Clonal Complex 95, a Cosmopolitan Escherichia coli Lineage Responsible for Extraintestinal Infection. *mSphere* 2. doi:10.1128/mSphere.00168-17.

Grozdanov, L., Raasch, C., Schulze, J., Sonnenborn, U., Gottschalk, G., Hacker, J., et al. (2004). Analysis of the Genome Structure of the Nonpathogenic Probiotic Escherichia coli Strain Nissle 1917. *J. Bacteriol.* 186, 5432–5441. doi:10.1128/JB.186.16.5432-5441.2004.

Guglielmini, J., Néron, B., Abby, S. S., Garcillán-Barcia, M. P., la Cruz, F. de, and Rocha, E. P. C. (2014). Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res.* 42, 5715–5727. doi:10.1093/nar/gku194.

Guinée, P. A. M., Jansen, W. H., Wadstrõm, T., and Sellwood, R. (1981). "Escherichia Coli Associated with Neonatal Diarrhoea in Piglets and Calves," in *Laboratory Diagnosis in Neoanatal Calf and Pig Diarrhoea*, eds. P. W. de Leeuw and P. A. M. Guinée (Dordrecht: Springer Netherlands), 126–162. doi:10.1007/978-94-009-8328-1_18.

Gupta, N., Limbago, B. M., Patel, J. B., and Kallen, A. J. (2011). Carbapenem-Resistant Enterobacteriaceae: Epidemiology and Prevention. *Clin. Infect. Dis.* 53, 60–67. doi:10.1093/cid/cir202.

Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., et al. (2014). ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in Bacterial Genomes. *Antimicrob. Agents Chemother.* 58, 212–220. doi:10.1128/AAC.01310-13.

Hacker, J., and Carniel, E. (2001). Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO Rep.* 2, 376–381. doi:10.1093/embo-reports/kve097.

Hacker, J., and Kaper, J. B. (2000). Pathogenicity Islands and the Evolution of Microbes. *Annu. Rev. Microbiol.* 54, 641–679. doi:10.1146/annurev.micro.54.1.641.

Hadjadj, L., Riziki, T., Zhu, Y., Li, J., Diene, S., and Rolain, J.-M. (2017). Study of mcr-1

Gene-Mediated Colistin Resistance in Enterobacteriaceae Isolated from Humans and Animals in Different Countries. *Genes (Basel).* 8, 394. doi:10.3390/genes8120394.

Haenni, M., Poirel, L., Kieffer, N., Châtre, P., Saras, E., Métayer, V., et al. (2016). Co-occurrence of extended spectrum β lactamase and MCR-1 encoding genes on plasmids. *Lancet Infect. Dis.* 16, 281–282. doi:10.1016/S1473-3099(16)00007-4.

Haenni, M., Saras, E., Métayer, V., Médaille, C., and Madec, J.-Y. (2014). High Prevalence of bla CTX-M-1 /IncI1/ST3 and bla CMY-2 /IncI1/ST2 Plasmids in Healthy Urban Dogs in France. *Antimicrob. Agents Chemother.* 58, 5358–5362. doi:10.1128/AAC.02545-14.

Hagan, E. C., and Mobley, H. L. T. (2009). Haem acquisition is facilitated by a novel receptor Hma and required by uropathogenic Escherichia coli for kidney infection. *Mol. Microbiol.* 71, 79–91. doi:10.1111/j.1365-2958.2008.06509.x.

Hall-Stoodley, L., Costerton, J. W., and Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. *Nat. Rev. Microbiol.* 2, 95–108. doi:10.1038/nrmicro821.

Hammerl, J. A., Borowiak, M., Schmoger, S., Shamoun, D., Grobbel, M., Malorny, B., et al. (2018). mcr-5 and a novel mcr-5.2 variant in Escherichia coli isolates from food and food-producing animals, Germany, 2010 to 2017. *J. Antimicrob. Chemother.*, 1433–1435. doi:10.1093/jac/dky020.

Han, Z., Pinkner, J. S., Ford, B., Chorell, E., Crowley, J. M., Cusumano, C. K., et al. (2012). Lead Optimization Studies on FimH Antagonists: Discovery of Potent and Orally Bioavailable Ortho-Substituted Biphenyl Mannosides. *J. Med. Chem.* 55, 3945–3959. doi:10.1021/jm300165m.

Han, Z., Pinkner, J. S., Ford, B., Obermann, R., Nolan, W., Wildman, S. A., et al. (2010). Structure-Based Drug Design and Optimization of Mannoside Bacterial FimH Antagonists. *J. Med. Chem.* 53, 4779–4792. doi:10.1021/jm100438s.

Hannan, T. J., Totsika, M., Mansfield, K. J., Moore, K. H., Schembri, M. A., and Hultgren, S. J. (2012). Host–pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic Escherichia coli bladder infection. *FEMS Microbiol. Rev.* 36, 616–648. doi:10.1111/j.1574-6976.2012.00339.x.

Hansen, K. H., Bortolaia, V., Nielsen, C. A., Nielsen, J. B., Schønning, K., Agersø, Y., et al. (2016). Host-Specific Patterns of Genetic Diversity among IncI1-Iγ and IncK Plasmids Encoding CMY-2 β-Lactamase in Escherichia coli Isolates from Humans, Poultry Meat, Poultry, and Dogs in Denmark. *Appl. Environ. Microbiol.* 82, 4705–4714. doi:10.1128/AEM.00495-16.

Harada, K., Nakai, Y., and Kataoka, Y. (2012). Mechanisms of resistance to cephalosporin and emergence of O25b-ST131 clone harboring CTX-M-27 β-lactamase in extraintestinal pathogenic Escherichia coli from dogs and cats in Japan. *Microbiol. Immunol.* 56, 480–485. doi:10.1111/j.1348-0421.2012.00463.x.

Harding, G. K. M., and Ronald, A. R. (1994). The management of urinary infections; what have we learned in the past decade? *Int. J. Antimicrob. Agents* 4, 83–88. doi:10.1016/0924-8579(94)90038-8.

Hasman, H., Hammerum, A. M., Hansen, F., Hendriksen, R. S., Olesen, B., Agersø, Y., et al. (2015). Detection of mcr-1 encoding plasmid-mediated colistin-resistant Escherichia coli isolates from human bloodstream infection and imported chicken meat, Denmark 2015. *Eurosurveillance* 20, 30085. doi:10.2807/1560-7917.ES.2015.20.49.30085.

Hawkey, J., Hamidian, M., Wick, R. R., Edwards, D. J., Billman-Jacobe, H., Hall, R. M., et al. (2015). ISMapper: Identifying transposase insertion sites in bacterial genomes from short read sequence data. *BMC Genomics* 16, 16–667. doi:10.1186/s12864-015-1860-2.

Hayes, F. (2001). The Horizontal Gene Pool — Bacterial Plasmids and Gene Spread. *Heredity (Edinb).* 86, 251–252. doi:10.1046/j.1365-2540.2001.0902b.x.

Hernández, M., Iglesias, M. R., Rodríguez-Lázaro, D., Gallardo, A., Quijada, N., Miguela-Villoldo, P., et al. (2017). Co-occurrence of colistin-resistance genes mcr-1 and mcr-3 among multidrug-resistant Escherichia coli isolated from cattle, Spain, September 2015. *Euro surveillance. Eur. J. Infect. Dis. Epidemiol. Prev. Control since 1996* 22, 30586.

doi:10.2807/1560-7917.ES.2017.22.31.30586.

Herrera Estévez, A. (2015). El papel de los alimentos en la transmisión de Escherichia coli potencialmente patógenas para el hombre: prevalencia y caracterización de cepas diarreagénicas y productoras de β-lactamasas de espectro extendido. doi:http://hdl.handle.net/10347/13687.

Hochhut, B., Wilde, C., Balling, G., Middendorf, B., Dobrindt, U., Brzuszkiewicz, E., et al. (2006). Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic Escherichia coli strain 536. *Mol. Microbiol.* 61, 584–595. doi:10.1111/j.1365-2958.2006.05255.x.

Høiby, N., Ciofu, O., and Bjarnsholt, T. (2010). Pseudomonas aeruginosa biofilms in cystic fibrosis. *Future Microbiol.* 5, 1663–1674. doi:10.2217/fmb.10.125.

Høiby, N., Ciofu, O., Johansen, H. K., Song, Z., Moser, C., Jensen, P. Ø., et al. (2011). The clinical impact of bacterial biofilms. *Int. J. Oral Sci.* 3, 55–65. doi:10.4248/IJOS11026.

Hojabri, Z., Darabi, N., Arab, M., Saffari, F., and Pajand, O. (2019). Clonal diversity, virulence genes content and subclone status of Escherichia coli sequence type 131: comparative analysis of E. coli ST131 and non-ST131 isolates from Iran. *BMC Microbiol.* 19, 117. doi:10.1186/s12866-019-1493-8.

Holden, N. J., and Gally, D. L. (2004). Switches, cross-talk and memory in Escherichia coli adherence. *J. Med. Microbiol.* 53, 585–593. doi:10.1099/jmm.0.05491-0.

Holmes, A. H., Moore, L. S. P., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., et al. (2016). Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet* 387, 176–187. doi:10.1016/S0140-6736(15)00473-0.

Hombach, M., Bõttger, E. C., and Roos, M. (2013). The critical influence of the intermediate category on interpretation errors in revised EUCAST and CLSI antimicrobial susceptibility testing guidelines. *Clin. Microbiol. Infect.* 19, E59–E71. doi:10.1111/1469-0691.12090.

Honardoost, M., Rajabpour, A., and Vakil, L. (2018). Molecular epidemiology; New but impressive. *Med. J. Islam. Repub. Iran* 32, 312–316. doi:10.14196/mjiri.32.53.

Hong, J. S., Song, W., Park, H.-M., Oh, J.-Y., Chae, J.-C., Shin, S., et al. (2019). Clonal Spread of Extended-Spectrum Cephalosporin-Resistant Enterobacteriaceae Between Companion Animals and Humans in South Korea. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01371.

Hordijk, J., Schoormans, A., Kwakernaak, M., Duim, B., Broens, E., Dierikx, C., et al. (2013). High prevalence of fecal carriage of extended spectrum β-lactamase/AmpC-producing Enterobacteriaceae in cats and dogs. *Front. Microbiol.* 4. doi:10.3389/fmicb.2013.00242.

Horner, C., Fawley, W., Morris, K., Parnell, P., Denton, M., and Wilcox, M. (2014). Escherichia coli bacteraemia: 2 years of prospective regional surveillance (2010-12). *J. Antimicrob. Chemother.* 69, 91–100. doi:10.1093/jac/dkt333.

Hussain, A., Ranjan, A., Nandanwar, N., Babbar, A., Jadhav, S., and Ahmed, N. (2014). Genotypic and Phenotypic Profiles of Escherichia coli Isolates Belonging to Clinical Sequence Type 131 (ST131), Clinical Non-ST131, and Fecal Non-ST131 Lineages from India. *Antimicrob. Agents Chemother.* 58, 7240–7249. doi:10.1128/AAC.03320-14.

Hutton, T. A., Innes, G. K., Harel, J., Garneau, P., Cucchiara, A., Schifferli, D. M., et al. (2018). Phylogroup and virulence gene association with clinical characteristics of Escherichia coli urinary tract infections from dogs and cats. *J. Vet. Diagnostic Investig.* 30, 64–70. doi:10.1177/1040638717729395.

Inouye, M., Dashnow, H., Raven, L.-A., Schultz, M. B., Pope, B. J., Tomita, T., et al. (2014). SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 6, 90. doi:10.1186/s13073-014-0090-6.

Izdebski, R., Baraniak, A., Fiett, J., Adler, A., Kazma, M., Salomon, J., et al. (2013). Clonal Structure, Extended-Spectrum β-Lactamases, and Acquired AmpC-Type Cephalosporinases of Escherichia coli Populations Colonizing Patients in Rehabilitation Centers in Four Countries. *Antimicrob. Agents Chemother.* 57, 309–316. doi:10.1128/AAC.01656-12.

Jacoby, G. A. (2009). AmpC β-Lactamases. *Clin.*

Microbiol. Rev. 22, 161–182. doi:10.1128/CMR.00036-08.

Jamborova, I., Johnston, B. D., Papousek, I., Kachlikova, K., Micenkova, L., Clabots, C., et al. (2018). Extensive Genetic Commonality among Wildlife, Wastewater, Community, and Nosocomial Isolates of Escherichia coli Sequence Type 131 (H 30R1 and H 30Rx Subclones) That Carry bla CTX-M-27 or bla CTX-M-15. *Antimicrob. Agents Chemother.* 62. doi:10.1128/AAC.00519-18.

Jantunen, M. E., Saxén, H., Salo, E., and Siitonen, A. (2002). Recurrent Urinary Tract Infections in Infancy: Relapses or Reinfections? *J. Infect. Dis.* 185, 375–9. doi:10.1086/338771.

Jarlier, V., Nicolas, M.-H., Fournier, G., and Philippon, A. (1988). Extended Broad-Spectrum -Lactamases Conferring Transferable Resistance to Newer -Lactam Agents in Enterobacteriaceae: Hospital Prevalence and Susceptibility Patterns. *Clin. Infect. Dis.* 10, 867–878. doi:10.1093/clinids/10.4.867.

Jaureguy, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., et al. (2008). Phylogenetic and genomic diversity of human bacteremic Escherichia coli strains. *BMC Genomics* 9, 560. doi:10.1186/1471-2164-9-560.

Jin, L., Wang, R., Wang, X., Wang, Q., Zhang, Y., Yin, Y., et al. (2018). Emergence of mcr-1 and carbapenemase genes in hospital sewage water in Beijing, China. *J. Antimicrob. Chemother.* 73, 84–87. doi:10.1093/jac/dkx355.

Joensen, K. G., Tetzschner, A. M. M. M., Iguchi, A., Aarestrup, F. M., and Scheutz, F. (2015). Rapid and Easy In Silico Serotyping of Escherichia coli Isolates by Use of Whole-Genome Sequencing Data. *J. Clin. Microbiol.* 53, 2410–2426. doi:10.1128/JCM.00008-15.

Johnson, A. P., and Woodford, N. (2013). Global spread of antibiotic resistance: the example of New Delhi metallo-β-lactamase (NDM)-mediated carbapenem resistance. *J. Med. Microbiol.* 62, 499–513. doi:10.1099/jmm.0.052555-0.

Johnson, J. R. (1991). Virulence factors in Escherichia coli urinary tract infection. *Clin. Microbiol. Rev.* 4, 80–128. doi:10.1128/CMR.4.1.80.

Johnson, J. R., Clabots, C., and Kuskowski, M. A. (2008a). Multiple-Host Sharing, Long-Term Persistence, and Virulence of Escherichia coli Clones from Human and Animal Household Members. *J. Clin. Microbiol.* 46, 4078–4082. doi:10.1128/JCM.00980-08.

Johnson, J. R., Davis, G., Clabots, C., Johnston, B. D., Porter, S., DebRoy, C., et al. (2016a). Household Clustering of Escherichia coli Sequence Type 131 Clinical and Fecal Isolates According to Whole Genome Sequence Analysis. *Open Forum Infect. Dis.* 3. doi:10.1093/ofid/ofw129.

Johnson, J. R., Jelacic, S., Schoening, L. M., Clabots, C., Shaikh, N., Mobley, H. L. T., et al. (2005a). The IrgA Homologue Adhesin Iha Is an Escherichia coli Virulence Factor in Murine Urinary Tract Infection. *Infect. Immun.* 73, 965–971. doi:10.1128/IAI.73.2.965-971.2005.

Johnson, J. R., Johnston, B., Clabots, C. R., Kuskowski, M. A., Roberts, E., and DebRoy, C. (2008b). Virulence Genotypes and Phylogenetic Background of Escherichia coli Serogroup O6 Isolates from Humans, Dogs, and Cats. *J. Clin. Microbiol.* 46, 417–422. doi:10.1128/JCM.00674-07.

Johnson, J. R., Johnston, B. D., Porter, S. B., Clabots, C., Bender, T. L., Thuras, P., et al. (2019). Rapid Emergence, Subsidence, and Molecular Detection of Escherichia coli Sequence Type 1193- fimH64 , a New Disseminated Multidrug-Resistant Commensal and Extraintestinal Pathogen. *J. Clin. Microbiol.* 57. doi:10.1128/JCM.01664-18.

Johnson, J. R., Kaster, N., Kuskowski, M. A., and Ling, G. V. (2003). Identification of Urovirulence Traits in Escherichia coli by Comparison of Urinary and Rectal E. coli Isolates from Dogs with Urinary Tract Infection. *J. Clin. Microbiol.* 41, 337–345. doi:10.1128/JCM.41.1.337-345.2003.

Johnson, J. R., Kuskowski, M. A., Owens, K., Clabots, C., and Singer, R. S. (2009). Virulence genotypes and phylogenetic background of fluoroquinolone-resistant and susceptible Escherichia coli urine isolates from dogs with urinary tract infection. *Vet. Microbiol.* 136, 108–114. doi:10.1016/j.vetmic.2008.10.006.

Johnson, J. R., Moseley, S. L., Roberts, P. L., and Stamm, W. E. (1988). Aerobactin and

other virulence factor genes among strains of Escherichia coli causing urosepsis: association with patient characteristics. *Infect. Immun.* 56, 405–412. doi:10.1128/IAI.56.2.405-412.1988.

Johnson, J. R., and O'Bryan, T. T. (2004). Detection of the Escherichia coli Group 2 Polysaccharide Capsule Synthesis Gene kpsM by a Rapid and Specific PCR-Based Assay. *J. Clin. Microbiol.* 42, 1773–1776. doi:10.1128/JCM.42.4.1773-1776.2004.

Johnson, J. R., O'Bryan, T. T., Kuskowski, M., and Maslow, J. N. (2001). Ongoing Horizontal and Vertical Transmission of Virulence Genes and papA Alleles among Escherichia coliBlood Isolates from Patients with Diverse-Source Bacteremia. *Infect. Immun.* 69, 5363–5374. doi:10.1128/IAI.69.9.5363-5374.2001.

Johnson, J. R., Oswald, E., O'Bryan, T. T., Kuskowski, M. A., and Spanjaard, L. (2002). Phylogenetic Distribution of Virulence-Associated Genes among Escherichia coli Isolates Associated with Neonatal Bacterial Meningitis in The Netherlands. *J. Infect. Dis.* 185, 774–784. doi:10.1086/339343.

Johnson, J. R., Owens, K., Gajewski, A., and Clabots, C. (2008c). Escherichia coli Colonization Patterns among Human Household Members and Pets, with Attention to Acute Urinary Tract Infection. *J. Infect. Dis.* 197, 218–224. doi:10.1086/524844.

Johnson, J. R., Porter, S., Johnston, B., Kuskowski, M. A., Spurbeck, R. R., Mobley, H. L. T., et al. (2015). Host Characteristics and Bacterial Traits Predict Experimental Virulence for Escherichia coli Bloodstream Isolates From Patients With Urosepsis. *Open Forum Infect. Dis.* 2. doi:10.1093/ofid/ofv083.

Johnson, J. R., and Russo, T. A. (2018). Molecular Epidemiology of Extraintestinal Pathogenic Escherichia coli. *EcoSal Plus* 8. doi:10.1128/ecosalplus.ESP-0004-2017.

Johnson, J. R., Russo, T. A., Tarr, P. I., Carlino, U., Bilge, S. S., Vary, J. C., et al. (2000). Molecular Epidemiological and Phylogenetic Associations of Two Novel Putative Virulence Genes, iha and iroNE. coli, among Escherichia coli Isolates from Patients with Urosepsis. *Infect. Immun.* 68, 3040–3047. doi:10.1128/IAI.68.5.3040-3047.2000.

Johnson, J. R., Scheutz, F., Ulleryd, P., Kuskowski, M. A., O'Bryan, T. T., and Sandberg, T. (2005b). Host-Pathogen Relationships among Escherichia coli Isolates Recovered from Men with Febrile Urinary Tract Infection. *Clin. Infect. Dis.* 40, 813–822. doi:10.1086/428048.

Johnson, J. R., Scheutz, F., Ulleryd, P., Kuskowski, M. A., O'Bryan, T. T., and Sandberg, T. (2005c). Phylogenetic and Pathotypic Comparison of Concurrent Urine and Rectal Escherichia coli Isolates from Men with Febrile Urinary Tract Infection. *J. Clin. Microbiol.* 43, 3895–3900. doi:10.1128/JCM.43.8.3895-3900.2005.

Johnson, J. R., Stapleton, A. E., Russo, T. A., Scheutz, F., Brown, J. J., and Maslow, J. N. (1997). Characteristics and prevalence within serogroup O4 of a J96-like clonal group of uropathogenic Escherichia coli O4:H5 containing the class I and class III alleles of papG. *Infect. Immun.* 65, 2153–2159. doi:10.1128/IAI.65.6.2153-2159.1997.

Johnson, J. R., and Stell, A. L. (2000). Extended Virulence Genotypes of Escherichia coli Strains from Patients with Urosepsis in Relation to Phylogeny and Host Compromise. *J. Infect. Dis.* 181, 261–272. doi:10.1086/315217.

Johnson, J. R., Tchesnokova, V., Johnston, B., Clabots, C., Roberts, P. L., Billig, M., et al. (2013). Abrupt Emergence of a Single Dominant Multidrug-Resistant Strain of Escherichia coli. *J. Infect. Dis.* 207, 919–928. doi:10.1093/infdis/jis933.

Johnson, T. J., Danzeisen, J. L., Youmans, B., Case, K., Llop, K., Munoz-Aguayo, J., et al. (2016b). Separate F-Type Plasmids Have Shaped the Evolution of the H 30 Subclone of Escherichia coli Sequence Type 131. *mSphere* 1. doi:10.1128/msphere.00121-16.

Johnson, T. J., Wannemuehler, Y., Doetkott, C., Johnson, S. J., Rosenberger, S. C., and Nolan, L. K. (2008d). Identification of Minimal Predictors of Avian Pathogenic Escherichia coli Virulence for Use as a Rapid Diagnostic Tool. *J. Clin. Microbiol.* 46, 3987–3996. doi:10.1128/JCM.00816-08.

Johnson, T. J., Wannemuehler, Y. M., and Nolan, L. K. (2008e). Evolution of the iss Gene in Escherichia coli. *Appl. Environ. Microbiol.* 74, 2360–2369. doi:10.1128/AEM.02634-

07.

Jorgensen, J. H., and Ferraro, M. J. (2009). Antimicrobial Susceptibility Testing: A Review of General Principles and Contemporary Practices. *Clin. Infect. Dis.* 49, 1749–1755. doi:10.1086/647952.

Justice, S. S., Hung, C., Theriot, J. A., Fletcher, D. A., Anderson, G. G., Footer, M. J., et al. (2004). Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis. *Proc. Natl. Acad. Sci.* 101, 1333–1338. doi:10.1073/pnas.0308125100.

Kallonen, T., Brodrick, H. J., Harris, S. R., Corander, J., Brown, N. M., Martin, V., et al. (2017). Systematic longitudinal survey of invasive Escherichia coli in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res.* 27, 1437–1449. doi:10.1101/gr.216606.116.

Kanamaru, S., Kurazono, H., Ishitoya, S., Terai, A., Habuchi, T., Nakano, M., et al. (2003). Distribution and Genetic Association of Putative Uropathogenic Virulence Factors Iron, iha, kpsMT, ompT and usp in Escherichia Coli Isolated From Urinary Tract Infections in Japan. *J. Urol.* 170, 2490–2493. doi:10.1097/01.ju.0000094185.48467.dc.

Kaper, J. B., Nataro, J. P., and Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nat. Rev. Microbiol.* 2, 123–140. doi:10.1038/nrmicro818.

Karkaba, A., Hill, K., Benschop, J., Pleydell, E., and Grinberg, A. (2019). Carriage and population genetics of extended spectrum β-lactamase-producing Escherichia coli in cats and dogs in New Zealand. *Vet. Microbiol.* 233, 61–67. doi:10.1016/j.vetmic.2019.04.015.

Kasman, L. M., and Porter, L. D. (2020). *Bacteriophages*.

Kauffmann, F., and Ørskov, F. (1956). Die Bakteriologie derEscherichia coli-enteritis. *A. Adam (ed.), Sauglings-Enteritis*, 1–41.

Kawamura, K., Sugawara, T., Matsuo, N., Hayashi, K., Norizuki, C., Tamai, K., et al. (2017). Spread of CTX-Type Extended-Spectrum β-Lactamase-Producing Escherichia coli Isolates of Epidemic Clone B2-O25-ST131 Among Dogs and Cats in Japan. *Microb. Drug Resist.* 23, 1059–1066. doi:10.1089/mdr.2016.0246.

Kempf, I., Jouy, E., and Chauvin, C. (2016). Colistin use and colistin resistance in bacteria from animals. *Int. J. Antimicrob. Agents* 48, 598–606. doi:10.1016/j.ijantimicag.2016.09.016.

Kidsley, A. K., O'Dea, M., Ebrahimie, E., Mohammadi-Dehcheshmeh, M., Saputra, S., Jordan, D., et al. (2020a). Genomic analysis of fluoroquinolone-susceptible phylogenetic group B2 extraintestinal pathogenic Escherichia coli causing infections in cats. *Vet. Microbiol.* 245, 108685. doi:10.1016/j.vetmic.2020.108685.

Kidsley, A. K., O'Dea, M., Saputra, S., Jordan, D., Johnson, J. R., Gordon, D. M., et al. (2020b). Genomic analysis of phylogenetic group B2 extraintestinal pathogenic E. coli causing infections in dogs in Australia. *Vet. Microbiol.* 248, 108783. doi:10.1016/j.vetmic.2020.108783.

Kline, K. A., Fãlker, S., Dahlberg, S., Normark, S., and Henriques-Normark, B. (2009). Bacterial Adhesins in Host-Microbe Interactions. *Cell Host Microbe* 5, 580–592. doi:10.1016/j.chom.2009.05.011.

Klinger, G., Chin, C.-N., Beyene, J., and Perlman, M. (2000). Predicting the Outcome of Neonatal Bacterial Meningitis. *Pediatrics* 106, 477–482. doi:10.1542/peds.106.3.477.

Kobayashi, R. K. T., Gaziri, L. C. J., and Vidotto, M. C. (2010). Functional activities of the Tsh protein from avian pathogenic Escherichia coli (APEC) strains. *J. Vet. Sci.* 11, 315. doi:10.4142/jvs.2010.11.4.315.

Koh, E.-I., Hung, C. S., Parker, K. S., Crowley, J. R., Giblin, D. E., and Henderson, J. P. (2015). Metal selectivity by the virulence-associated yersiniabactin metallophore system. *Metallomics* 7, 1011–1022. doi:10.1039/C4MT00341A.

Kondratyeva, K., Salmon-Divon, M., and Navon-Venezia, S. (2020). Meta-analysis of Pandemic Escherichia coli ST131 Plasmidome Proves Restricted Plasmid-clade Associations. *Sci. Rep.* 10, 36. doi:10.1038/s41598-019-56763-7.

Kosek, M., Bern, C., and Guerrant, R. L. (2003). The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull. World Health Organ.* 81, 197–204.

Kovac, J., Bakker, H. den, Carroll, L. M., and Wiedmann, M. (2017). Precision food safety: A systems approach to food safety facilitated by genomics tools. *TrAC Trends Anal. Chem.* 96, 52–61. doi:10.1016/j.trac.2017.06.001.

Krawczyk, P. S., Lipinski, L., and Dziembowski, A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* 46, e35–e35. doi:10.1093/nar/gkx1321.

Kudinha, T., Johnson, J. R., Andrew, S. D., Kong, F., Anderson, P., and Gilbert, G. L. (2013). Distribution of phylogenetic groups, sequence type ST131, and virulence-associated traits among Escherichia coli isolates from men with pyelonephritis or cystitis and healthy controls. *Clin. Microbiol. Infect.* 19, E173–E180. doi:10.1111/1469-0691.12123.

Kundu, P., Blacher, E., Elinav, E., and Pettersson, S. (2017). Our Gut Microbiome: The Evolving Inner Self. *Cell* 171, 1481–1493. doi:10.1016/j.cell.2017.11.024.

Kuo, S.-C., Huang, W.-C., Wang, H.-Y., Shiau, Y.-R., Cheng, M.-F., and Lauderdale, T.-L. (2016). Colistin resistance gene mcr -1 in Escherichia coli isolates from humans and retail meats, Taiwan. *J. Antimicrob. Chemother.* 71, 2327–2329. doi:10.1093/jac/dkw122.

La Combe, B., Clermont, O., Messika, J., Eveillard, M., Kouatchet, A., Lasocki, S., et al. (2019). Pneumonia-Specific Escherichia coli with Distinct Phylogenetic and Virulence Profiles, France, 2012–2014. *Emerg. Infect. Dis.* 25, 710–718. doi:10.3201/eid2504.180944.

Lane, M. C., Alteri, C. J., Smith, S. N., and Mobley, H. L. T. (2007). Expression of flagella is coincident with uropathogenic Escherichia coli ascension to the upper urinary tract. *Proc. Natl. Acad. Sci.* 104, 16669–16674. doi:10.1073/pnas.0607898104.

Lanza, V. F., de Toro, M., Garcillán-Barcia, M. P., Mora, A., Blanco, J., Coque, T. M., et al. (2014). Plasmid Flux in Escherichia coli ST131 Sublineages, Analyzed by Plasmid Constellation Network (PLACNET), a New Method for Plasmid Reconstruction from Whole Genome Sequences. *PLoS Genet.* 10, e1004766. doi:10.1371/journal.pgen.1004766.

Lara, F. B. M., Nery, D. R., de Oliveira, P. M., Araujo, M. L., Carvalho, F. R. Q., Messias-Silva, L. C. F., et al. (2017). Virulence Markers and Phylogenetic Analysis of Escherichia coli Strains with Hybrid EAEC/UPEC Genotypes Recovered from Sporadic Cases of Extraintestinal Infections. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.00146.

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi:10.1128/JCM.06094-11.

Laxminarayan, R., Matsoso, P., Pant, S., Brower, C., Røttingen, J.-A., Klugman, K., et al. (2016). Access to effective antimicrobials: a worldwide challenge. *Lancet* 387, 168–175. doi:10.1016/S0140-6736(15)00474-2.

Le Bouguenec, C., Archambaud, M., and Labigne, A. (1992). Rapid and specific detection of the pap, afa, and sfa adhesin-encoding operons in uropathogenic Escherichia coli strains by polymerase chain reaction. *J. Clin. Microbiol.* 30, 1189–93. doi:10.1128/JCM.30.5.1189-1193.1992.

Le Gall, T., Clermont, O., Gouriou, S., Picard, B., Nassif, X., Denamur, E., et al. (2007). Extraintestinal Virulence Is a Coincidental By-Product of Commensalism in B2 Phylogenetic Group Escherichia coli Strains. *Mol. Biol. Evol.* 24, 2373–2384. doi:10.1093/molbev/msm172.

Lebeaux, D., Ghigo, J.-M., and Beloin, C. (2014). Biofilm-Related Infections: Bridging the Gap between Clinical Management and Fundamental Aspects of Recalcitrance toward Antibiotics. *Microbiol. Mol. Biol. Rev.* 78, 510–543. doi:10.1128/MMBR.00013-14.

LeCuyer, T. E., Byrne, B. A., Daniels, J. B., Diaz-Campos, D. V., Hammac, G. K., Miller, C. B., et al. (2018). Population Structure and Antimicrobial Resistance of Canine Uropathogenic Escherichia coli. *J. Clin. Microbiol.* 56. doi:10.1128/JCM.00788-18.

Ledeboer, N. A., and Hodinka, R. L. (2011). Molecular Detection of Resistance Determinants. *J. Clin. Microbiol.* 49, S20–S24. doi:10.1128/JCM.00771-11.

Leflon-Guibout, V., Blanco, J., Amaqdouf, K.,

Mora, A., Guize, L., and Nicolas-Chanoine, M.-H. (2008). Absence of CTX-M Enzymes but High Prevalence of Clones, Including Clone ST131, among Fecal Escherichia coli Isolates from Healthy Subjects Living in the Area of Paris, France. *J. Clin. Microbiol.* 46, 3900–3905. doi:10.1128/JCM.00734-08.

Leflon-Guibout, V., Jurand, C., Bonacorsi, S., Espinasse, F., Guelfi, M. C., Duportail, F., et al. (2004a). Emergence and Spread of Three Clonally Related Virulent Isolates of CTX-M-15-Producing Escherichia coli with Variable Resistance to Aminoglycosides and Tetracycline in a French Geriatric Hospital. *Antimicrob. Agents Chemother.* 48, 3736–3742. doi:10.1128/AAC.48.10.3736-3742.2004.

Leflon-Guibout, V., Jurand, C., Bonacorsi, S., Espinasse, F., Guelfi, M. C., Duportail, F., et al. (2004b). Emergence and Spread of Three Clonally Related Virulent Isolates of CTX-M-15-Producing Escherichia coli with Variable Resistance to Aminoglycosides and Tetracycline in a French Geriatric Hospital. *Antimicrob. Agents Chemother.* 48, 3736–3742. doi:10.1128/AAC.48.10.3736-3742.2004.

Lefort, A., Panhard, X., Clermont, O., Woerther, P.-L., Branger, C., Mentre, F., et al. (2011). Host Factors and Portal of Entry Outweigh Bacterial Determinants To Predict the Severity of Escherichia coli Bacteremia. *J. Clin. Microbiol.* 49, 777–783. doi:10.1128/JCM.01902-10.

Lescat, M., Clermont, O., Woerther, P. L., Glodt, J., Dion, S., Skurnik, D., et al. (2013). Commensal Escherichia coli strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environ. Microbiol. Rep.* 5, 49–57. doi:10.1111/j.1758-2229.2012.00374.x.

Levy, M., Kolodziejczyk, A. A., Thaiss, C. A., and Elinav, E. (2017). Dysbiosis and the immune system. *Nat. Rev. Immunol.* 17, 219–232. doi:10.1038/nri.2017.7.

Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nat. Rev. Microbiol.* 5, 48–56. doi:10.1038/nrmicro1557.

Lewis, K. (2008). "Multidrug Tolerance of Biofilms and Persister Cells," in, 107–131. doi:10.1007/978-3-540-75418-3_6.

Li, A., Yang, Y., Miao, M., Chavda, K. D., Mediavilla, J. R., Xie, X., et al. (2016). Complete Sequences of mcr-1 -Harboring Plasmids from Extended-Spectrum-β-Lactamase- and Carbapenemase-Producing Enterobacteriaceae. *Antimicrob. Agents Chemother.* 60, 4351–4354. doi:10.1128/AAC.00550-16.

Li, R., Xie, M., Lv, J., Wai-Chi Chan, E., and Chen, S. (2017a). Complete genetic analysis of plasmids carrying mcr-1 and other resistance genes in an Escherichia coli isolate of animal origin. *J. Antimicrob. Chemother.* 72, 696–699. doi:10.1093/jac/dkw509.

Li, R., Xie, M., Zhang, J., Yang, Z., Liu, L., Liu, X., et al. (2017b). Genetic characterization of mcr-1 -bearing plasmids to depict molecular mechanisms underlying dissemination of the colistin resistance determinant. *J. Antimicrob. Chemother.* 72, 393–401. doi:10.1093/jac/dkw411.

Li, R., Yu, H., Xie, M., Chen, K., Dong, N., Lin, D., et al. (2018). Genetic basis of chromosomally-encoded mcr-1 gene. *Int. J. Antimicrob. Agents* 51, 578–585. doi:10.1016/j.ijantimicag.2017.11.015.

Li, X.-Z., Plésiat, P., and Nikaido, H. (2015). The Challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria. *Clin. Microbiol. Rev.* 28, 337–418. doi:10.1128/CMR.00117-14.

Li, X. P., Fang, L. X., Jiang, P., Pan, D., Xia, J., Liao, X. P., et al. (2017c). Emergence of the colistin resistance gene mcr-1 in Citrobacter freundii. *Int. J. Antimicrob. Agents* 49, 786–787. doi:10.1016/j.ijantimicag.2017.04.004.

Liakopoulos, A., Betts, J., La Ragione, R., van Essen-Zandbergen, A., Ceccarelli, D., Petinaki, E., et al. (2018). Occurrence and characterization of extended-spectrum cephalosporin-resistant Enterobacteriaceae in healthy household dogs in Greece. *J. Med. Microbiol.* 67, 931–935. doi:10.1099/jmm.0.000768.

Lindsey, R. L., Garcia-Toledo, L., Fasulo, D., Gladney, L. M., and Strockbine, N. (2017). Multiplex polymerase chain reaction for identification of Escherichia coli , Escherichia albertii and Escherichia fergusonii. *J. Microbiol. Methods* 140, 1–4. doi:10.1016/j.mimet.2017.06.005.

Liu, B.-T., Song, F.-J., Zou, M., Hao, Z.-H., and Shan, H. (2016a). Emergence of colistin resistance gene mcr-1 in Cronobacter sakazakii producing NDM-9 and Escherichia coli from the same animal.

Antimicrob. Agents Chemother. 61, AAC.01444-16. doi:10.1128/AAC.01444-16.

Liu, C. M., Stegger, M., Aziz, M., Johnson, T. J., Waits, K., Nordstrom, L., et al. (2018a). Escherichia coli ST131- H 22 as a Foodborne Uropathogen. *MBio* 9, e00470-18. doi:10.1128/mBio.00470-18.

Liu, H., Zhou, H., Li, Q., Peng, Q., Zhao, Q., Wang, J., et al. (2018b). Molecular characteristics of extended-spectrum β-lactamase-producing Escherichia coli isolated from the rivers and lakes in Northwest China. *BMC Microbiol.* 18, 125. doi:10.1186/s12866-018-1270-0.

Liu, J., Keelan, P., Bennett, P. M., and Enne, V. I. (2009). Characterization of a novel macrolide efflux gene, mef(B), found linked to sul3 in porcine Escherichia coli. *J. Antimicrob. Chemother.* 63, 423–426. doi:10.1093/jac/dkn523.

Liu, X., Liu, H., Li, Y., and Hao, C. (2016b). High Prevalence of β-lactamase and Plasmid-Mediated Quinolone Resistance Genes in Extended-Spectrum Cephalosporin-Resistant Escherichia coli from Dogs in Shaanxi, China. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.01843.

Liu, X., Liu, H., Li, Y., and Hao, C. (2017). Association between virulence profile and fluoroquinolone resistance in Escherichia coli isolated from dogs and cats in China. *J. Infect. Dev. Ctries.* 11, 306–313. doi:10.3855/jidc.8583.

Liu, X., Liu, H., Wang, L., Peng, Q., Li, Y., Zhou, H., et al. (2018c). Molecular Characterization of Extended-Spectrum β-Lactamase-Producing Multidrug Resistant Escherichia coli From Swine in Northwest China. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.01756.

Liu, X., Thungrat, K., and Boothe, D. M. (2016c). Occurrence of OXA-48 Carbapenemase and Other β-Lactamase Genes in ESBL-Producing Multidrug Resistant Escherichia coli from Dogs and Cats in the United States, 2009–2013. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.01057.

Liu, Y.-Y., Wang, Y., Walsh, T. R., Yi, L.-X., Zhang, R., Spencer, J., et al. (2016d). Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* 16, 161–168. doi:10.1016/S1473-3099(15)00424-7.

Lo, A. W., Moriel, D. G., Phan, M.-D., Schulz, B. L., Kidd, T. J., Beatson, S. A., et al. (2017). 'Omic' Approaches to Study Uropathogenic Escherichia coli Virulence. *Trends Microbiol.* 25, 729–740. doi:10.1016/j.tim.2017.04.006.

Loncaric, I., Misic, D., Szostak, M. P., Kũnzel, F., Schãfer-Somi, S., and Spergser, J. (2020). Broad-Spectrum Cephalosporin-Resistant and/or Fluoroquinolone-Resistant Enterobacterales Associated with Canine and Feline Urogenital Infections. *Antibiotics* 9, 387. doi:10.3390/antibiotics9070387.

Looft, T., Johnson, T. A., Allen, H. K., Bayles, D. O., Alt, D. P., Stedtfeld, R. D., et al. (2012). In-feed antibiotic effects on the swine intestinal microbiome. *Proc. Natl. Acad. Sci.* 109, 1691–1696. doi:10.1073/pnas.1120238109.

Lu, H., Giordano, F., and Ning, Z. (2016a). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics. Proteomics Bioinformatics* 14, 265–279. doi:10.1016/j.gpb.2016.05.004.

Lu, S., Jin, D., Wu, S., Yang, J., Lan, R., Bai, X., et al. (2016b). Insights into the evolution of pathogenicity of Escherichia coli from genomic analysis of intestinal E. coli of Marmota himalayana in Qinghai–Tibet plateau of China. *Emerg. Microbes Infect.* 5, 1–9. doi:10.1038/emi.2016.122.

Luo, C., Walk, S. T., Gordon, D. M., Feldgarden, M., Tiedje, J. M., and Konstantinidis, K. T. (2011). Genome sequencing of environmental Escherichia coli expands understanding of the ecology and speciation of the model bacterial species. *Proc. Natl. Acad. Sci.* 108, 7200–7205. doi:10.1073/pnas.1015622108.

Luo, J., Yao, X., Lv, L., Doi, Y., Huang, X., Huang, S., et al. (2017). Emergence of mcr-1 in Raoultella ornithinolytica and Escherichia coli Isolates from Retail Vegetables in China. *Antimicrob. Agents Chemother.* 61, e01139-17. doi:10.1128/AAC.01139-17.

Ma, J., Cheng, Z., Bai, Q., Zhao, K., Pan, Z., and Yao, H. (2020). Screening virulence factors of porcine extraintestinal pathogenic Escherichia coli (an emerging pathotype) required for optimal growth in

swine blood. *Transbound. Emerg. Dis.*, tbed.13848. doi:10.1111/tbed.13848.

Machado, J., Grimont, F., and Grimont, P. A. D. (2000). Identification of Escherichia coli flagellar types by restrictionof the amplified fliC gene. *Res. Microbiol.* 151, 535–546. doi:10.1016/S0923-2508(00)00223-0.

Macnab, R. M. (1996). "Flagella and motility," in *Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), Escherichia and Salmonella: cellular and molecular biology, 2nd ed., vol. 1. American Society for Microbi*, 123–145.

Madec, J. Y., and Haenni, M. (2018). Antimicrobial resistance plasmid reservoir in food and food-producing animals. *Plasmid* 99, 72–81. doi:10.1016/j.plasmid.2018.09.001.

Madsen, J. S., Burmølle, M., Hansen, L. H., and Sørensen, S. J. (2012). The interconnection between biofilm formation and horizontal gene transfer. *FEMS Immunol. Med. Microbiol.* 65, 183–195. doi:10.1111/j.1574-695X.2012.00960.x.

Maeyama, Y., Taniguchi, Y., Hayashi, W., Ohsaki, Y., Osaka, S., Koide, S., et al. (2018). Prevalence of ESBL/AmpC genes and specific clones among the third-generation cephalosporin-resistant Enterobacteriaceae from canine and feline clinical specimens in Japan. *Vet. Microbiol.* 216, 183–189. doi:10.1016/j.vetmic.2018.02.020.

Magiorakos, A.-P. P., Srinivasan, A., Carey, R. B. B., Carmeli, Y., Falagas, M. E. E., Giske, C. G. G., et al. (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* 18, 268–281. doi:10.1111/j.1469-0691.2011.03570.x.

Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., et al. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci.* 95, 3140–3145. doi:10.1073/pnas.95.6.3140.

Maiden, M. C. J., van Rensburg, M. J. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., et al. (2013). MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol.* 11, 728–736. doi:10.1038/nrmicro3093.

Mamani, R., Flament-Simon, S. C., García, V., Mora, A., Alonso, M. P., López, C., et al. (2019). Sequence Types, Clonotypes, Serotypes, and Virotypes of Extended-Spectrum β-Lactamase-Producing Escherichia coli Causing Bacteraemia in a Spanish Hospital Over a 12-Year Period (2000 to 2011). *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01530.

Manges, A. R., Geum, H. M., Guo, A., Edens, T. J., Fibke, C. D., and Pitout, J. D. D. (2019). Global Extraintestinal Pathogenic Escherichia coli (ExPEC) Lineages. *Clin. Microbiol. Rev.* 32. doi:10.1128/CMR.00135-18.

Manges, A. R., and Johnson, J. R. (2012). Food-Borne Origins of Escherichia coli Causing Extraintestinal Infections. *Clin. Infect. Dis.* 55, 712–719. doi:10.1093/cid/cis502.

Mann, R., Mediati, D. G., Duggin, I. G., Harry, E. J., and Bottomley, A. L. (2017). Metabolic Adaptations of Uropathogenic E. coli in the Urinary Tract. *Front. Cell. Infect. Microbiol.* 7. doi:10.3389/fcimb.2017.00241.

Marc, D., and Dho-Moulin, M. (1996). Analysis of the fim cluster of an avian O2 strain of Escherichia coli: serogroup-specific sites within fimA and nucleotide sequence of fimI. *J. Med. Microbiol.* 44, 444–452. doi:10.1099/00222615-44-6-444.

Marques, C., Belas, A., Franco, A., Aboim, C., Gama, L. T., and Pomba, C. (2018). Increase in antimicrobial resistance and emergence of major international high-risk clonal lineages in dogs and cats with urinary tract infection: 16 year retrospective study. *J. Antimicrob. Chemother.* 73, 377–384. doi:10.1093/jac/dkx401.

Martin Bland, J., and Altman, D. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *Lancet* 327, 307–310. doi:10.1016/S0140-6736(86)90837-8.

Martin, P., Marcq, I., Magistro, G., Penary, M., Garcie, C., Payros, D., et al. (2013). Interplay between Siderophores and Colibactin Genotoxin Biosynthetic Pathways in Escherichia coli. *PLoS Pathog.* 9, e1003437.

doi:10.1371/journal.ppat.1003437.

Martinez-Medina, M., Mora, A., Blanco, M., Lopez, C., Alonso, M. P., Bonacorsi, S., et al. (2009a). Similarity and Divergence among Adherent-Invasive Escherichia coli and Extraintestinal Pathogenic E. coli Strains. *J. Clin. Microbiol.* 47, 3968–3979. doi:10.1128/JCM.01484-09.

Martinez-Medina, M., Naves, P., Blanco, J., Aldeguer, X., Blanco, J. E., Blanco, M., et al. (2009b). Biofilm formation as a novel phenotypic feature of adherent-invasive Escherichia coli (AIEC). *BMC Microbiol.* 9, 202. doi:10.1186/1471-2180-9-202.

Martínez, J. L. (2012). Natural Antibiotic Resistance and Contamination by Antibiotic Resistance Determinants: The Two Ages in the Evolution of Resistance to Antimicrobials. *Front. Microbiol.* 3. doi:10.3389/fmicb.2012.00001.

Martínez, J. L. (2014). General principles of antibiotic resistance in bacteria. *Drug Discov. Today Technol.* 11, 33–39. doi:10.1016/j.ddtec.2014.02.001.

Martins, M. T., Rivera, I. G., Clark, D. L., Stewart, M. H., Wolfe, R. L., and Olson, B. H. (1993). Distribution of uidA gene sequences in Escherichia coli isolates in water sources and comparison with the expression of beta-glucuronidase activity in 4-methylumbelliferyl-beta-D-glucuronide media. *Appl. Environ. Microbiol.* 59, 2271–6. doi:10.1128/AEM.59.7.2271-2276.1993.

Martinson, J. N. V., and Walk, S. T. (2020). Escherichia coli Residency in the Gut of Healthy Human Adults. *EcoSal Plus* 9. doi:10.1128/ecosalplus.ESP-0003-2020.

Massot, M., Daubié, A.-S., Clermont, O., Jauréguy, F., Couffignal, C., Dahbi, G., et al. (2016). Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of Escherichia coli from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology* 162, 642–650. doi:10.1099/mic.0.000242.

Matamoros, S., van Hattem, J. M., Arcilla, M. S., Willemse, N., Melles, D. C., Penders, J., et al. (2017). Global phylogenetic analysis of Escherichia coli and plasmids carrying the mcr-1 gene indicates bacterial diversity but plasmid restriction. *Sci. Rep.* 7, 15364. doi:10.1038/s41598-017-15539-7.

Mathers, A. J., Peirano, G., and Pitout, J. D. D. (2015). The Role of Epidemic Resistance Plasmids and International High-Risk Clones in the Spread of Multidrug-Resistant Enterobacteriaceae. *Clin. Microbiol. Rev.* 28, 565–591. doi:10.1128/CMR.00116-14.

Matsumura, Y., Noguchi, T., Tanaka, M., Kanahashi, T., Yamamoto, M., Nagao, M., et al. (2017a). Population structure of Japanese extraintestinal pathogenic Escherichia coli and its relationship with antimicrobial resistance. *J. Antimicrob. Chemother.*, dkw530. doi:10.1093/jac/dkw530.

Matsumura, Y., Pitout, J. D. D., Gomi, R., Matsuda, T., Noguchi, T., Yamamoto, M., et al. (2016). Global Escherichia coli Sequence Type 131 Clade with bla CTX-M-27 Gene. *Emerg. Infect. Dis.* 22, 1900–1907. doi:10.3201/eid2211.160519.

Matsumura, Y., Pitout, J. D. D., Peirano, G., DeVinney, R., Noguchi, T., Yamamoto, M., et al. (2017b). Rapid Identification of Different Escherichia coli Sequence Type 131 Clades. *Antimicrob. Agents Chemother.* 61, 1–9. doi:10.1128/AAC.00179-17.

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., et al. (2013). The Comprehensive Antibiotic Resistance Database. *Antimicrob. Agents Chemother.* 57, 3348–3357. doi:10.1128/AAC.00419-13.

McDermott, P. F., Tyson, G. H., Kabera, C., Chen, Y., Li, C., Folster, J. P., et al. (2016). Whole-Genome Sequencing for Detecting Antimicrobial Resistance in Nontyphoidal Salmonella. *Antimicrob. Agents Chemother.* 60, 5515–5520. doi:10.1128/AAC.01030-16.

McGann, P., Snesrud, E., Maybank, R., Corey, B., Ong, A. C., Clifford, R., et al. (2016). Escherichia coli Harboring mcr-1 and bla CTX-M on a Novel IncF Plasmid: First Report of mcr-1 in the United States. *Antimicrob. Agents Chemother.* 60, 4420–4421. doi:10.1128/AAC.01103-16.

Meek, R. W., Vyas, H., and Piddock, L. J. V. (2015). Nonmedical Uses of Antibiotics: Time to Restrict Their Use? *PLOS Biol.* 13, e1002266. doi:10.1371/journal.pbio.1002266.

Melo, L. C., Boisson, M. N. G., Saras, E., Médaille, C., Boulouis, H.-J., Madec, J.-

Y., et al. (2016). OXA-48-producing ST372 Escherichia coli in a French dog. *J. Antimicrob. Chemother.*, dkw531. doi:10.1093/jac/dkw531.

Melo, L. C., Haenni, M., Saras, E., Duprilot, M., Nicolas-Chanoine, M.-H., and Madec, J.-Y. (2019). Emergence of the C1-M27 cluster in ST131 Escherichia coli from companion animals in France. *J. Antimicrob. Chemother.* 74, 3111–3113. doi:10.1093/jac/dkz304.

Melo, L. C., Oresco, C., Leigue, L., Netto, H. M., Melville, P. A., Benites, N. R., et al. (2018). Prevalence and molecular features of ESBL/pAmpC-producing Enterobacteriaceae in healthy and diseased companion animals in Brazil. *Vet. Microbiol.* 221, 59–66. doi:10.1016/j.vetmic.2018.05.017.

Méric, G., Kemsley, E. K., Falush, D., Saggers, E. J., and Lucchini, S. (2013). Phylogenetic distribution of traits associated with plant colonization in Escherichia coli. *Environ. Microbiol.* 15, 487–501. doi:10.1111/j.1462-2920.2012.02852.x.

Merino, I., Hernández-García, M., Turrientes, M.-C., Pérez-Viso, B., López-Fresneña, N., Diaz-Agero, C., et al. (2018). Emergence of ESBL-producing Escherichia coli ST131-C1-M27 clade colonizing patients in Europe. *J. Antimicrob. Chemother.* 73, 2973–2980. doi:10.1093/jac/dky296.

Merino, I., Shaw, E., Horcajada, J. P., Cercenado, E., Mirelis, B., Pallarés, M. A., et al. (2016). CTX-M-15- H 30Rx-ST131 subclone is one of the main causes of healthcare-associated ESBL-producing Escherichia coli bacteraemia of urinary origin in Spain. *J. Antimicrob. Chemother.* 71, 2125–2130. doi:10.1093/jac/dkw133.

Micali, S., Isgro, G., Bianchi, G., Miceli, N., Calapai, G., and Navarra, M. (2014). Cranberry and Recurrent Cystitis: More than Marketing? *Crit. Rev. Food Sci. Nutr.* 54, 1063–1075. doi:10.1080/10408398.2011.625574.

Middendorf, B., Hochhut, B., Leipold, K., Dobrindt, U., Blum-Oehler, G., and Hacker, J. (2004). Instability of Pathogenicity Islands in Uropathogenic Escherichia coli 536. *J. Bacteriol.* 186, 3086–3096. doi:10.1128/JB.186.10.3086-3096.2004.

Milkman, R., and Stoltzfus, A. (1988). Molecular evolution of the Escherichia coli chromosome. II. Clonal segments. *Genetics* 120, 359–66.

Mora, A. Herrera, A. López, C. Dahbi, G. Mamani, R. Pita, J-M. Alonso, M-P. Llovo, J. Bernárdez, M-I. Blanco, J-E. Blanco, M. Blanco, J. (2011). Characteristics of the Shiga-toxin-producing enteroaggregative Escherichia coli O104:H4 German outbreak strain and of STEC strains isolated in Spain. *Int. Microbiol.* 14, 121–141. doi:10.2436/20.1501.01.142.

Mora, A., García-Peña, F. J., Alonso, M. P., Pedraza-Diaz, S., Ortega-Mora, L. M., Garcia-Parraga, D., et al. (2018). Impact of human-associated Escherichia coli clonal groups in Antarctic pinnipeds: presence of ST73, ST95, ST141 and ST131. *Sci. Rep.* 8, 4678. doi:10.1038/s41598-018-22943-0.

Mora, A., Herrrera, A., López, C., Dahbi, G., Mamani, R., Pita, J. M., et al. (2011). Characteristics of the Shiga-toxin-producing enteroaggregative Escherichia coli O104:H4 German outbreak strain and of STEC strains isolated in Spain. *Int. Microbiol.* 14, 121–141. doi:10.2436/20.1501.01.142.

Mora, A., Viso, S., López, C., Alonso, M. P., García-Garrote, F., Dabhi, G., et al. (2013). Poultry as reservoir for extraintestinal pathogenic Escherichia coli O45:K1:H7-B2-ST95 in humans. *Vet. Microbiol.* 167, 506–512. doi:10.1016/j.vetmic.2013.08.007.

Morabia, A. (2020). Pandemics and methodological developments in epidemiology history. *J. Clin. Epidemiol.* 125, 164–169. doi:10.1016/j.jclinepi.2020.06.008.

Morales, C., Lee, M. D., Hofacre, C., and Maurer, J. J. (2004). Detection of a Novel Virulence Gene and a Salmonella Virulence Homologue Among Escherichia coli Isolated from Broiler Chickens. *Foodborne Pathog. Dis.* 1, 160–165. doi:10.1089/fpd.2004.1.160.

Moreno, E., Andreu, A., Pigrau, C., Kuskowski, M. A., Johnson, J. R., and Prats, G. (2008). Relationship between Escherichia coli Strains Causing Acute Cystitis in Women and the Fecal E. coli Population of the Host. *J. Clin. Microbiol.* 46, 2529–2534. doi:10.1128/JCM.00813-08.

Moriel, D. G., Tan, L., Goh, K. G. K., Phan, M.-D., Ipe, D. S., Lo, A. W., et al. (2016). A Novel Protective Vaccine Antigen from the Core Escherichia coli Genome. *mSphere* 1. doi:10.1128/mSphere.00326-16.

Moulin-Schouleur, M., Reperant, M., Laurent, S., Bree, A., Mignon-Grasteau, S., Germon, P., et al. (2007). Extraintestinal Pathogenic Escherichia coli Strains of Avian and Human Origin: Link between Phylogenetic Relationships and Common Virulence Patterns. *J. Clin. Microbiol.* 45, 3366–3376. doi:10.1128/JCM.00037-07.

Moulin-Schouleur, M., Schouler, C., Tailliez, P., Kao, M.-R., Bree, A., Germon, P., et al. (2006). Common Virulence Factors and Genetic Relationships between O18:K1:H7 Escherichia coli Isolates of Human and Avian Origin. *J. Clin. Microbiol.* 44, 3484–3492. doi:10.1128/JCM.00548-06.

Mulder, C. J. J., van Alphen, L., and Zanen, H. C. (1984). Neonatal Meningitis Caused by Escherichia coli in The Netherlands. *J. Infect. Dis.* 150, 935–940. doi:10.1093/infdis/150.6.935.

Mulder, C. J. J., and Zanen, H. C. (1984). A study of 280 cases of neonatal meningitis in The Netherlands. *J. Infect.* 9, 177–184. doi:10.1016/S0163-4453(84)91351-3.

Mulvey, M. A. (2002). Adhesion and entry of uropathogenic Escherichia coli. *Cell. Microbiol.* 4, 257–271. doi:10.1046/j.1462-5822.2002.00193.x.

Munita, J. M., and Arias, C. A. (2016). Mechanisms of Antibiotic Resistance. *Microbiol. Spectr.* 4. doi:10.1128/microbiolspec.VMBF-0016-2015.

Murray, P., Baron, E., Pfaller, M., Tenover, F., and Yolken, R. (1999). "Manual of clinical microbiology, 7th ed," in *ASM Press* (Washington, DC).

Mysorekar, I. U., and Hultgren, S. J. (2006). Mechanisms of uropathogenic Escherichia coli persistence and eradication from the urinary tract. *Proc. Natl. Acad. Sci.* 103, 14170–14175. doi:10.1073/pnas.0602136103.

Naderi, G., Haghi, F., Zeighami, H., Hemati, F., and Masoumian, N. (2016). Distribution of pathogenicity island (PAI) markers and phylogenetic groups in diarrheagenic and commensal Escherichia coli from young children. *Gastroenterol. Hepatol. from bed to bench* 9, 316–324.

Nagy, G., Dobrindt, U., Kupfer, M., Emödy, L., Karch, H., and Hacker, J. (2001). Expression of Hemin Receptor Molecule ChuA Is Influenced by RfaH in Uropathogenic Escherichia coliStrain 536. *Infect. Immun.* 69, 1924–1928. doi:10.1128/IAI.69.3.1924-1928.2001.

Naseer, U., Olsson-Liljequist, B. E., Woodford, N., Dhanji, H., Cantón, R., Sundsfjord, A., et al. (2012). Multi-Locus Variable Number of Tandem Repeat Analysis for Rapid and Accurate Typing of Virulent Multidrug Resistant Escherichia coli Clones. *PLoS One* 7, e41232. doi:10.1371/journal.pone.0041232.

Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic Escherichia coli. *Clin. Microbiol. Rev.* 11, 142–201.

Naves, P., del Prado, G., Huelves, L., Gracia, M., Ruiz, V., Blanco, J., et al. (2008). Correlation between virulence factors and in vitro biofilm formation by Escherichia coli strains. *Microb. Pathog.* 45, 86–91. doi:10.1016/j.micpath.2008.03.003.

Nesta, B., Spraggon, G., Alteri, C., Gomes Moriel, D., Rosini, R., Veggi, D., et al. (2012). FdeC, a Novel Broadly Conserved Escherichia coli Adhesin Eliciting Protection against Urinary Tract Infections. *MBio* 3. doi:10.1128/mBio.00010-12.

Nicolas-Chanoine, M.-H., Bertrand, X., and Madec, J.-Y. (2014). Escherichia coli ST131, an Intriguing Clonal Group. *Clin. Microbiol. Rev.* 27, 543–574. doi:10.1128/CMR.00125-13.

Nicolas-Chanoine, M.-H., Blanco, J., Leflon-Guibout, V., Demarty, R., Alonso, M. P., Canica, M. M., et al. (2007a). Intercontinental emergence of Escherichia coli clone O25:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* 61, 273–281. doi:10.1093/jac/dkm464.

Nicolas-Chanoine, M.-H., Gruson, C., Bialek-Davenet, S., Bertrand, X., Thomas-Jean, F., Bert, F., et al. (2013). 10-Fold increase (2006-11) in the rate of healthy subjects with extended-spectrum -lactamase-producing Escherichia coli faecal carriage in a Parisian check-up centre. *J. Antimicrob. Chemother.* 68, 562–568. doi:10.1093/jac/dks429.

Nicolas-Chanoine, M.-H. M. H., Blanco, J., Leflon-Guibout, V., Demarty, R., Alonso, M. P., Canica, M. M., et al. (2007b). Intercontinental emergence of Escherichia coli clone O25:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* 61, 273–281. doi:10.1093/jac/dkm464.

Nicolas-Chanoine, M.-H., Petitjean, M., Mora, A., Mayer, N., Lavigne, J.-P., Boulet, O., et al. (2017). The ST131 Escherichia coli H22 subclone from human intestinal microbiota: Comparison of genomic and phenotypic traits with those of the globally successful H30 subclone. *BMC Microbiol.* 17, 71. doi:10.1186/s12866-017-0984-8.

Nielsen, K. L., Dynesen, P., Larsen, P., and Frimodt-Møller, N. (2014). Faecal Escherichia coli from patients with E. coli urinary tract infection and healthy controls who have never had a urinary tract infection. *J. Med. Microbiol.* 63, 582–589. doi:10.1099/jmm.0.068783-0.

Nielubowicz, G. R., and Mobley, H. L. T. (2010). Host–pathogen interactions in urinary tract infection. *Nat. Rev. Urol.* 7, 430–441. doi:10.1038/nrurol.2010.101.

Nikaido, H., and Pagès, J.-M. (2012). Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria. *FEMS Microbiol. Rev.* 36, 340–363. doi:10.1111/j.1574-6976.2011.00290.x.

Nordmann, P., Naas, T., and Poirel, L. (2011). Global Spread of Carbapenemase-producing Enterobacteriaceae. *Emerg. Infect. Dis.* 17, 1791–1798. doi:10.3201/eid1710.110655.

Novais, Á., Pires, J., Ferreira, H., Costa, L., Montenegro, C., Vuotto, C., et al. (2012). Characterization of Globally Spread Escherichia coli ST131 Isolates (1991 to 2010). *Antimicrob. Agents Chemother.* 56, 3973–3976. doi:10.1128/AAC.00475-12.

Novick, R. P. (1989). Staphylococcal Plasmids and their Replication. *Annu. Rev. Microbiol.* 43, 537–563. doi:10.1146/annurev.mi.43.100189.002541.

Novick, R. P., Christie, G. E., and Penadés, J. R. (2010). The phage-related chromosomal islands of Gram-positive bacteria. *Nat. Rev. Microbiol.* 8, 541–551. doi:10.1038/nrmicro2393.

Nowak, K., Fahr, J., Weber, N., Lũbke-Becker, A., Semmler, T., Weiss, S., et al. (2017). Highly diverse and antimicrobial susceptible Escherichia coli display a naïve bacterial population in fruit bats from the Republic of Congo. *PLoS One* 12, e0178146. doi:10.1371/journal.pone.0178146.

Nowicki, B., Moulds, J., Hull, R., and Hull, S. (1988). A hemagglutinin of uropathogenic Escherichia coli recognizes the Dr blood group antigen. *Infect. Immun.* 56, 1057–1060. doi:10.1128/IAI.56.5.1057-1060.1988.

Nowicki, B., Selvarangan, R., and Nowicki, S. (2001). Family of Escherichia coli Dr Adhesins: Decay-Accelerating Factor Receptor Recognition and Invasiveness. *J. Infect. Dis.* 183, S24–S27. doi:10.1086/318846.

Nowrouzian, F., Adlerberth, I., and Wold, A. E. (2001). P fimbriae, capsule and aerobactin characterize colonic resident Escherichia coli. *Epidemiol. Infect.* 126, 11–18. doi:10.1017/S0950268801005118.

Nowrouzian, F., Hesselmar, B., Saalman, R., Strannegård, I.-L., Åberg, N., Wold, A. E., et al. (2003). Escherichia coli in Infants' Intestinal Microflora: Colonization Rate, Strain Turnover, and Virulence Gene Carriage. *Pediatr. Res.* 54, 8–14. doi:10.1203/01.PDR.0000069843.20655.EE.

Nũesch-Inderbinen, M. T., Baschera, M., Zurfluh, K., Hãchler, H., Nũesch, H., and Stephan, R. (2017). Clonal Diversity, Virulence Potential and Antimicrobial Resistance of Escherichia coli Causing Community Acquired Urinary Tract Infection in Switzerland. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.02334.

O'Brien, V. P., Hannan, T. J., Nielsen, H. V., and Hultgren, S. J. (2016). Drug and Vaccine Development for the Treatment and Prevention of Urinary Tract Infections. *Microbiol. Spectr.* 4. doi:10.1128/microbiolspec.UTI-0013-2012.

O'Neill, J. (2016). TACKLING DRUG-RESISTANT INFECTIONS GLOBALLY: FINAL REPORT AND RECOMMENDATIONS.

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi:10.1038/35012500.

OIE (2003). Chapter 6.9. Monitoring of the quantities and usage patterns of antimicrobial agents used in food-producing animals.

OIE (2016). The OIE Strategy on Antimicrobial Resistance and the Prudent Use of Antimicrobials.

OIE (2019). OIE list of antimicrobial agents of veterinary importance.

Olaitan, A. O., Chabou, S., Okdah, L., Morand, S., and Rolain, J.-M. (2016). Dissemination of the mcr-1 colistin resistance gene. *Lancet Infect. Dis.* 16, 147. doi:10.1016/S1473-3099(15)00540-X.

Olaitan, A. O., Morand, S., and Rolain, J.-M. (2014). Mechanisms of polymyxin resistance: acquired and intrinsic resistance in bacteria. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00643.

Olesen, B., Scheutz, F., Andersen, R. L., Menard, M., Boisen, N., Johnston, B., et al. (2012). Enteroaggregative Escherichia coli O78:H10, the Cause of an Outbreak of Urinary Tract Infection. *J. Clin. Microbiol.* 50, 3703–3711. doi:10.1128/JCM.01909-12.

Olivares, E., Badel-Berchoux, S., Provot, C., Jaulhac, B., Prévost, G., Bernardi, T., et al. (2016). The BioFilm Ring Test: a Rapid Method for Routine Analysis of Pseudomonas aeruginosa Biofilm Formation Kinetics. *J. Clin. Microbiol.* 54, 657–661. doi:10.1128/JCM.02938-15.

Orlek, A., Stoesser, N., Anjum, M. F., Doumith, M., Ellington, M. J., Peto, T., et al. (2017). Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. *Front. Microbiol.* 8, 182. doi:10.3389/fmicb.2017.00182.

Orskov, F., and Orskov, I. (1992). Escherichia coli serotyping and disease in man and animals. *Can. J. Microbiol.* 38, 699–704.

Orskov, I., Orskov, F., Jann, B., and Jann, K. (1977). Serology, chemistry, and genetics of O and K antigens of Escherichia coli. *Bacteriol. Rev.* 41, 667–710.

Ortiz de la Tabla, V., Ortega, A., Buñuel, F., Pérez-Vázquez, M., Marcos, B., and Oteo, J. (2017). Detection of the high-risk clone ST131 of Escherichia coli carrying the colistin resistance gene mcr-1 and causing acute peritonitis. *Int. J. Antimicrob. Agents* 49, 115–116. doi:10.1016/j.ijantimicag.2016.10.003.

Oteo, J., González-López, J. J., Ortega, A., Quintero-Zárate, J. N., Bou, G., Cercenado, E., et al. (2014). Inhibitor-Resistant TEM- and OXA-1-Producing Escherichia coli Isolates Resistant to Amoxicillin-Clavulanate Are More Clonal and Possess Lower Virulence Gene Content than Susceptible Clinical Isolates. *Antimicrob. Agents Chemother.* 58, 3874–3881. doi:10.1128/AAC.02738-13.

Pagès, J.-M., James, C. E., and Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat. Rev. Microbiol.* 6, 893–903. doi:10.1038/nrmicro1994.

Pal, C., Asiani, K., Arya, S., Rensing, C., Stekel, D. J., Larsson, D. G. J., et al. (2017). "Metal Resistance and Its Association With Antibiotic Resistance," in, 261–313. doi:10.1016/bs.ampbs.2017.02.001.

Partridge, S. R., Di Pilato, V., Doi, Y., Feldgarden, M., Haft, D. H., Klimke, W., et al. (2018a). Proposal for assignment of allele numbers for mobile colistin resistance (mcr) genes. *J. Antimicrob. Chemother.* 73, 2625–2630. doi:10.1093/jac/dky262.

Partridge, S. R., Kwong, S. M., Firth, N., and Jensen, S. O. (2018b). Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* 31, 1–61. doi:10.1128/cmr.00088-17.

Peirano, G., Bradford, P. A., Kazmierczak, K. M., Badal, R. E., Hackel, M., Hoban, D. J., et al. (2014a). Global Incidence of Carbapenemase-Producing Escherichia coli ST131. *Emerg. Infect. Dis.* 20, 1928–1931. doi:10.3201/eid2011.141388.

Peirano, G., and Pitout, J. D. D. (2019). Extended-Spectrum β-Lactamase-Producing Enterobacteriaceae: Update on Molecular Epidemiology and Treatment Options. *Drugs* 79, 1529–1541. doi:10.1007/s40265-019-01180-3.

Peirano, G., van der Bij, A. K., Freeman, J. L., Poirel, L., Nordmann, P., Costello, M., et al. (2014b). Characteristics of Escherichia coli Sequence Type 131 Isolates That Produce Extended-Spectrum β-Lactamases: Global Distribution of the H 30-Rx Sublineage. *Antimicrob. Agents*

Chemother. 58, 3762–3767. doi:10.1128/AAC.02428-14.

Peirano, G., van der Bij, A. K., Gregson, D. B., and Pitout, J. D. D. (2012). Molecular Epidemiology over an 11-Year Period (2000 to 2010) of Extended-Spectrum -Lactamase-Producing Escherichia coli Causing Bacteremia in a Centralized Canadian Region. *J. Clin. Microbiol.* 50, 294–299. doi:10.1128/JCM.06025-11.

Penteado, A. ., Ugrinovich, L. ., Blanco, J., Blanco, M., Blanco, J. ., Mora, A., et al. (2002). Serobiotypes and virulence genes of Escherichia coli strains isolated from diarrheic and healthy rabbits in Brazil. *Vet. Microbiol.* 89, 41–51. doi:10.1016/S0378-1135(02)00148-7.

Pepin-Puget, L., El Garch, F., Bertrand, X., Valot, B., and Hocquet, D. (2020). Genome analysis of enterobacteriaceae with non-wild type susceptibility to third-generation cephalosporins recovered from diseased dogs and cats in Europe. *Vet. Microbiol.* 242, 108601. doi:10.1016/j.vetmic.2020.108601.

Perez-Perez, F. J., and Hanson, N. D. (2002). Detection of Plasmid-Mediated AmpC -Lactamase Genes in Clinical Isolates by Using Multiplex PCR. *J. Clin. Microbiol.* 40, 2153–2162. doi:10.1128/JCM.40.6.2153-2162.2002.

Perna, N. T., Plunkett, G., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., et al. (2001). Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. *Nature* 409, 529–533. doi:10.1038/35054089.

Perry, J. A., Westman, E. L., and Wright, G. D. (2014). The antibiotic resistome: what's new? *Curr. Opin. Microbiol.* 21, 45–50. doi:10.1016/j.mib.2014.09.002.

Petty, N. K., Zakour, N. L. B., Stanton-Cook, M., Skippington, E., Totsika, M., Forde, B. M., et al. (2014). Global dissemination of a multidrug resistant Escherichia coli clone. *Proc. Natl. Acad. Sci. U. S. A.* 111, 5694–5699. doi:10.1073/pnas.1322678111.

Pham Thanh, D., Thanh Tuyen, H., Nguyen Thi Nguyen, T., Chung The, H., Wick, R. R., Thwaites, G. E., et al. (2016). Inducible colistin resistance via a disrupted plasmid-borne mcr-1 gene in a 2008 Vietnamese Shigella sonnei isolate. *J. Antimicrob. Chemother.* 71, 2314–2317. doi:10.1093/jac/dkw173.

Phan, M.-D., Peters, K. M., Sarkar, S., Lukowski, S. W., Allsopp, L. P., Moriel, D. G., et al. (2013). The Serum Resistome of a Globally Disseminated Multidrug Resistant Uropathogenic Escherichia coli Clone. *PLoS Genet.* 9, e1003834. doi:10.1371/journal.pgen.1003834.

Philippon, A., Arlet, G., and Jacoby, G. A. (2002). Plasmid-Determined AmpC-Type β-Lactamases. *Antimicrob. Agents Chemother.* 46, 1–11. doi:10.1128/AAC.46.1.1-11.2002.

Picard, B., Garcia, J. S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., et al. (1999). The Link between Phylogeny and Virulence inEscherichia coli Extraintestinal Infection. *Infect. Immun.* 67, 546–553. doi:10.1128/IAI.67.2.546-553.1999.

Piercey, M. J., Hingston, P. A., and Truelstrup Hansen, L. (2016). Genes involved in Listeria monocytogenes biofilm formation at a simulated food processing plant temperature of 15°C. *Int. J. Food Microbiol.* 223, 63–74. doi:10.1016/j.ijfoodmicro.2016.02.009.

Pires, J., Bernasconi, O. J., Kasraian, S., Hilty, M., Perreten, V., and Endimiani, A. (2016). Intestinal colonisation with extended-spectrum cephalosporin-resistant Escherichia coli in Swiss pets: molecular features, risk factors and transmission with owners. *Int. J. Antimicrob. Agents* 48, 759–760. doi:10.1016/j.ijantimicag.2016.08.007.

Pitout, J. D. D., and DeVinney, R. (2017). Escherichia coli ST131: a multidrug-resistant clone primed for global domination. *F1000Research* 6, 195. doi:10.12688/f1000research.10609.1.

Platell, J. L., Cobbold, R. N., Johnson, J. R., Heisig, A., Heisig, P., Clabots, C., et al. (2011). Commonality among Fluoroquinolone-Resistant Sequence Type ST131 Extraintestinal Escherichia coli Isolates from Humans and Companion Animals in Australia. *Antimicrob. Agents Chemother.* 55, 3782–3787. doi:10.1128/AAC.00306-11.

Platell, J. L., Trott, D. J., Johnson, J. R., Heisig, P., Heisig, A., Clabots, C. R., et al. (2012). Prominence of an O75 Clonal Group (Clonal Complex 14) among Non-ST131 Fluoroquinolone-Resistant Escherichia coli Causing Extraintestinal Infections in Humans and Dogs in Australia. *Antimicrob. Agents Chemother.* 56, 3898–

3904. doi:10.1128/AAC.06120-11.

Poirel, L., Kieffer, N., Brink, A., Coetze, J., Jayol, A., and Nordmann, P. (2016). Genetic Features of MCR-1-Producing Colistin-Resistant Escherichia coli Isolates in South Africa. *Antimicrob. Agents Chemother.* 60, 4394–4397. doi:10.1128/AAC.00444-16.

Poirel, L., Madec, J., Lupo, A., Schink, A., Kieffer, N., Nordmann, P., et al. (2018). Antimicrobial Resistance in Escherichia coli. *Microbiol. Spectr.* 6, 1–27. doi:10.1128/microbiolspec.ARBA-0026-2017.

Pomba, C., Rantala, M., Greko, C., Baptiste, K. E., Catry, B., van Duijkeren, E., et al. (2016). Public health risk of antimicrobial resistance transfer from companion animals. *J. Antimicrob. Chemother.*, dkw481. doi:10.1093/jac/dkw481.

Ponnusamy, P., Natarajan, V., and Sevanan, M. (2012). In vitro biofilm formation by uropathogenic Escherichia coli and their antimicrobial susceptibility pattern. *Asian Pac. J. Trop. Med.* 5, 210–213. doi:10.1016/S1995-7645(12)60026-1.

Pratt, L. A., and Kolter, R. (1998). Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol. Microbiol.* 30, 285–293. doi:10.1046/j.1365-2958.1998.01061.x.

Price, L. B., Johnson, J. R., Aziz, M., Clabots, C., Johnston, B., Tchesnokova, V., et al. (2013). The Epidemic of Extended-Spectrum-β-Lactamase-Producing Escherichia coli ST131 Is Driven by a Single Highly Pathogenic Subclone, H30-Rx. *MBio* 4. doi:10.1128/mBio.00377-13.

Pulss, S., Stolle, I., Stamm, I., Leidner, U., Heydel, C., Semmler, T., et al. (2018). Multispecies and Clonal Dissemination of OXA-48 Carbapenemase in Enterobacteriaceae From Companion Animals in Germany, 2009—2016. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.01265.

Quainoo, S., Coolen, J. P. M., van Hijum, S. A. F. T., Huynen, M. A., Melchers, W. J. G., van Schaik, W., et al. (2017). Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. *Clin. Microbiol. Rev.* 30, 1015–1063. doi:10.1128/CMR.00016-17.

Radhouani, H., Silva, N., Poeta, P., Torres, C., Correia, S., and Igrejas, G. (2014). Potential impact of antimicrobial resistance in wildlife, environment and human health. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00023.

Ramsay, J. P., and Firth, N. (2017). Diverse mobilization strategies facilitate transfer of non-conjugative mobile genetic elements. *Curr. Opin. Microbiol.* 38, 1–9. doi:10.1016/j.mib.2017.03.003.

Randall, L. P., Horton, R. A., Lemma, F., Martelli, F., Duggett, N. A. D., Smith, R. P., et al. (2018). Longitudinal study on the occurrence in pigs of colistin-resistant Escherichia coli carrying mcr-1 following the cessation of use of colistin. *J. Appl. Microbiol.* 125, 596–608. doi:10.1111/jam.13907.

Rasheed, J. K., Jay, C., Metchock, B., Berkowitz, F., Weigel, L., Crellin, J., et al. (1997). Evolution of extended-spectrum beta-lactam resistance (SHV-8) in a strain of Escherichia coli during multiple episodes of bacteremia. *Antimicrob. Agents Chemother.* 41, 647–53. doi:10.1128/AAC.41.3.647.

Ratiner, Y. A., Sihvonen, L. M., Liu, Y., Wang, L., and Siitonen, A. (2010). Alteration of flagellar phenotype of Escherichia coli strain P12b, the standard type strain for flagellar antigen H17, possessing a new non-fliC flagellin gene flnA, and possible loss of original flagellar phenotype and genotype in the course of subcultur. *Arch. Microbiol.* 192, 267–278. doi:10.1007/s00203-010-0556-x.

Rebelo, A. R., Bortolaia, V., Kjeldgaard, J. S., Pedersen, S. K., Leekitcharoenphon, P., Hansen, I. M., et al. (2018). Multiplex PCR for detection of plasmid-mediated colistin resistance determinants, mcr-1, mcr-2, mcr-3, mcr-4 and mcr-5 for surveillance purposes. *Eurosurveillance* 23. doi:10.2807/1560-7917.ES.2018.23.6.17-00672.

Reid, C. J., McKinnon, J., and Djordjevic, S. P. (2019). Clonal ST131-H22 Escherichia coli strains from a healthy pig and a human urinary tract infection carry highly similar resistance and virulence plasmids. *Microb. Genomics* 5. doi:10.1099/mgen.0.000295.

Reid, S. D., Herbelin, C. J., Bumbaugh, A. C., Selander, R. K., and Whittam, T. S. (2000). Parallel evolution of virulence in

pathogenic Escherichia coli. *Nature* 406, 64–67. doi:10.1038/35017546.

Ren, J., Ahlgren, N. A., Lu, Y. Y., Fuhrman, J. A., and Sun, F. (2017). VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* 5, 69. doi:10.1186/s40168-017-0283-5.

Riley, L. W. (2014). Pandemic lineages of extraintestinal pathogenic Escherichia coli. *Clin. Microbiol. Infect.* 20, 380–390. doi:10.1111/1469-0691.12646.

Roberts, M. C. (2008). Update on macrolide-lincosamide-streptogramin, ketolide, and oxazolidinone resistance genes. *FEMS Microbiol. Lett.* 282, 147–159. doi:10.1111/j.1574-6968.2008.01145.x.

Robertson, J., and Nash, J. H. E. (2018). MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb. Genomics* 4. doi:10.1099/mgen.0.000206.

Robin, F., Beyrouthy, R., Bonacorsi, S., Aissa, N., Bret, L., Brieu, N., et al. (2017). Inventory of Extended-Spectrum-β-Lactamase-Producing Enterobacteriaceae in France as Assessed by a Multicenter Study. *Antimicrob. Agents Chemother.* 61. doi:10.1128/AAC.01911-16.

Robinson, A. E., Heffernan, J. R., and Henderson, J. P. (2018). The iron hand of uropathogenic Escherichia coli : the role of transition metal control in virulence. *Future Microbiol.* 13, 745–756. doi:10.2217/fmb-2017-0295.

Rocha-Gracia, R. C., Cortés-Cortés, G., Lozano-Zarain, P., Bello, F., Martínez-Laguna, Y., and Torres, C. (2015). Faecal Escherichia coli isolates from healthy dogs harbour CTX-M-15 and CMY-2 β-lactamases. *Vet. J.* 203, 315–319. doi:10.1016/j.tvjl.2014.12.026.

Rodloff, A., Bauer, T., Ewig, S., Kujath, P., and Mũller, E. (2008). Susceptible, Intermediate, and Resistant – The Intensity of Antibiotic Action. *Dtsch. Aerzteblatt Online*. doi:10.3238/arztebl.2008.0657.

Rodriguez-Siek, K. E., Giddings, C. W., Doetkott, C., Johnson, T. J., and Nolan, L. K. (2005). Characterizing the APEC pathotype. *Vet. Res.* 36, 241–256. doi:10.1051/vetres:2004057.

Roer, L., Hansen, F., Thomsen, M. C. F., Knudsen, J. D., Hansen, D. S., Wang, M., et al. (2017). WGS-based surveillance of third-generation cephalosporin-resistant Escherichia coli from bloodstream infections in Denmark. *J. Antimicrob. Chemother.* 72, 1922–1929. doi:10.1093/jac/dkx092.

Roer, L., Johannesen, T. B., Hansen, F., Stegger, M., Tchesnokova, V., Sokurenko, E., et al. (2018). CHTyper, a Web Tool for Subtyping of Extraintestinal Pathogenic Escherichia coli Based on the fumC and fimH Alleles. *J. Clin. Microbiol.* 56, e00063-18. doi:10.1128/JCM.00063-18.

Ron, E. Z. (2006). Host specificity of septicemic Escherichia coli: human and avian pathogens. *Curr. Opin. Microbiol.* 9, 28–32. doi:10.1016/j.mib.2005.12.001.

Ronald, A. (2003). The etiology of urinary tract infection: Traditional and emerging pathogens. *Disease-a-Month* 49, 71–82. doi:10.1067/mda.2003.8.

Rosen, D. A., Hooton, T. M., Stamm, W. E., Humphrey, P. A., and Hultgren, S. J. (2007). Detection of Intracellular Bacterial Communities in Human Urinary Tract Infection. *PLoS Med.* 4, e329. doi:10.1371/journal.pmed.0040329.

Rozov, R., Brown Kav, A., Bogumil, D., Shterzer, N., Halperin, E., Mizrahi, I., et al. (2016). Recycler: an algorithm for detecting plasmids from de novo assembly graphs. *Bioinformatics*, btw651. doi:10.1093/bioinformatics/btw651.

Rozwandowicz, M., Brouwer, M. S. M. M., Fischer, J., Wagenaar, J. A., Gonzalez-Zorn, B., Guerra, B., et al. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *J. Antimicrob. Chemother.* 73, 1121–1137. doi:10.1093/jac/dkx488.

Russo, T. (2003). Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. *Microbes Infect.* 5, 449–456. doi:10.1016/S1286-4579(03)00049-2.

Russo, T. A., Beanan, J. M., Olson, R., MacDonald, U., and Cope, J. J. (2009). Capsular polysaccharide and the O-specific antigen impede antibody binding: A potential obstacle for the successful development of an extraintestinal pathogenic Escherichia coli vaccine. *Vaccine* 27, 388–395.

doi:10.1016/j.vaccine.2008.10.082.

Russo, T. A., and Johnson, J. R. (2000). Proposal for a New Inclusive Designation for Extraintestinal Pathogenic Isolates of Escherichia coli: ExPEC. *J. Infect. Dis.* 181, 1753–1754. doi:10.1086/315418.

Sabaté Brescó, M., Harris, L. G., Thompson, K., Stanic, B., Morgenstern, M., O'Mahony, L., et al. (2017). Pathogenic Mechanisms and Host Interactions in Staphylococcus epidermidis Device-Related Infection. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01401.

Saha, S., Tariq, R., Tosh, P. K., Pardi, D. S., and Khanna, S. (2019). Faecal microbiota transplantation for eradicating carriage of multidrug-resistant organisms: a systematic review. *Clin. Microbiol. Infect.* 25, 958–963. doi:10.1016/j.cmi.2019.04.006.

Saïdani, N., Lagier, J.-C., Cassir, N., Million, M., Baron, S., Dubourg, G., et al. (2019). Faecal microbiota transplantation shortens the colonisation period and allows re-entry of patients carrying carbapenamase-producing bacteria into medical care facilities. *Int. J. Antimicrob. Agents* 53, 355–361. doi:10.1016/j.ijantimicag.2018.11.014.

Saladin, M., Cao, V. T. B., Lambert, T., Donay, J.-L., Herrmann, J.-L., Ould-Hocine, Z., et al. (2002). Diversity of CTX-M beta-lactamases and their promoter regions from Enterobacteriaceae isolated in three Parisian hospitals. *FEMS Microbiol. Lett.* 209, 161–168. doi:10.1111/j.1574-6968.2002.tb11126.x.

Samei, A., Haghi, F., and Zeighami, H. (2016). Distribution of pathogenicity island markers in commensal and uropathogenic Escherichia coli isolates. *Folia Microbiol. (Praha).* 61, 261–268. doi:10.1007/s12223-015-0433-8.

San Millan, A. (2018). Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends Microbiol.* 26, 978–985. doi:10.1016/j.tim.2018.06.007.

Sarkar, S., Ulett, G. C., Totsika, M., Phan, M.-D., and Schembri, M. A. (2014). Role of Capsule and O Antigen in the Virulence of Uropathogenic Escherichia coli. *PLoS One* 9, e94786. doi:10.1371/journal.pone.0094786.

Sarkar, S., Vagenas, D., Schembri, M. A., and Totsika, M. (2016). Biofilm formation by multidrug resistant Escherichia coli ST131 is dependent on type 1 fimbriae and assay conditions. *Pathog. Dis.* 74, ftw013. doi:10.1093/femspd/ftw013.

Sartor, B. (2018). Intestinal bacterial biofilms modulate mucosal immune responses. *J. Immunol. Sci.* 2, 13–18. doi:10.29245/2578-3009/2018/2.1122.

Sauget, M., Cholley, P., Vannier, A., Thouverez, M., Nicolas-Chanoine, M.-H., Hocquet, D., et al. (2016). Trends of extended-spectrum β-lactamase-producing Escherichia coli sequence type 131 and its H 30 subclone in a French hospital over a 15-year period. *Int. J. Antimicrob. Agents* 48, 744–747. doi:10.1016/j.ijantimicag.2016.09.023.

Savageau, M. A. (1983). Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. *Am. Nat.* 122, 732–744. doi:10.1086/284168.

SCENIHR (2010). Research strategy to address the knowledge gaps on the antimicrobial resistance effects of biocides.

Schaufler, K., Bethe, A., Lũbke-Becker, A., Ewers, C., Kohn, B., Wieler, L. H., et al. (2015). Putative connection between zoonotic multiresistant extended-spectrum beta-lactamase (ESBL)-producing Escherichia coli in dog feces from a veterinary campus and clinical isolates from dogs. *Infect. Ecol. Epidemiol.* 5, 25334. doi:10.3402/iee.v5.25334.

Schiebel, J., Bõhm, A., Nitschke, J., Burdukiewicz, M., Weinreich, J., Ali, A., et al. (2017). Genotypic and Phenotypic Characteristics Associated with Biofilm Formation by Human Clinical Escherichia coli Isolates of Different Pathotypes. *Appl. Environ. Microbiol.* 83. doi:10.1128/AEM.01660-17.

Schink, A.-K., Kadlec, K., Kaspar, H., Mankertz, J., and Schwarz, S. (2013). Analysis of extended-spectrum- -lactamase-producing Escherichia coli isolates collected in the GERM-Vet monitoring programme. *J. Antimicrob. Chemother.* 68, 1741–1749. doi:10.1093/jac/dkt123.

Schmidt, C. (2019). Phage therapy's latest makeover. *Nat. Biotechnol.* 37, 581–586. doi:10.1038/s41587-019-0133-z.

Schmidt, H., Knop, C., Franke, S., Aleksic, S., Heesemann, J., and Karch, H. (1995).

Development of PCR for screening of enteroaggregative Escherichia coli. *J. Clin. Microbiol.* 33, 701–705. doi:10.1128/JCM.33.3.701-705.1995.

Schreiber, H. L., Conover, M. S., Chou, W.-C., Hibbing, M. E., Manson, A. L., Dodson, K. W., et al. (2017). Bacterial virulence phenotypes of Escherichia coli and host susceptibility determine risk for urinary tract infections. *Sci. Transl. Med.* 9, eaaf1283. doi:10.1126/scitranslmed.aaf1283.

Schubert, S., Rakin, A., Karch, H., Carniel, E., and Heesemann, J. (1998). Prevalence of the "High-Pathogenicity Island" of Yersinia Species among Escherichia coliStrains That Are Pathogenic to Humans. *Infect. Immun.* 66, 480–485. doi:10.1128/IAI.66.2.480-485.1998.

Schuldiner, S. (2018). The Escherichia coli effluxome. *Res. Microbiol.* 169, 357–362. doi:10.1016/j.resmic.2018.02.006.

Schultsz, C., Pool, G. J., van Ketel, R., de Wever, B., Speelman, P., and Dankert, J. (1994). Detection of enterotoxigenic Escherichia coli in stool samples by using nonradioactively labeled oligonucleotide DNA probes and PCR. *J. Clin. Microbiol.* 32, 2393–7. doi:10.1128/JCM.32.10.2393-2397.1994.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi:10.1093/bioinformatics/btu153.

Sekse, C., Holst-Jensen, A., Dobrindt, U., Johannessen, G. S., Li, W., Spilsberg, B., et al. (2017). High Throughput Sequencing for Detection of Foodborne Pathogens. *Front. Microbiol.* 8, 1–26. doi:10.3389/fmicb.2017.02029.

Sengupta, M., and Austin, S. (2011). Prevalence and Significance of Plasmid Maintenance Functions in the Virulence Plasmids of Pathogenic Bacteria. *Infect. Immun.* 79, 2502–2509. doi:10.1128/IAI.00127-11.

Seni, J., Falgenhauer, L., Simeo, N., Mirambo, M. M., Imirzalioglu, C., Matee, M., et al. (2016). Multiple ESBL-Producing Escherichia coli Sequence Types Carrying Quinolone and Aminoglycoside Resistance Genes Circulating in Companion and Domestic Farm Animals in Mwanza, Tanzania, Harbor Commonly Occurring Plasmids. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00142.

Servin, A. L. (2005). Pathogenesis of Afa/Dr Diffusely Adhering Escherichia coli. *Clin. Microbiol. Rev.* 18, 264–292. doi:10.1128/CMR.18.2.264-292.2005.

Servin, A. L. (2014). Pathogenesis of Human Diffusely Adhering Escherichia coli Expressing Afa/Dr Adhesins (Afa/Dr DAEC): Current Insights and Future Challenges. *Clin. Microbiol. Rev.* 27, 823–869. doi:10.1128/CMR.00036-14.

Sevilla, E., Mainar-Jaime, R. C., Moreno, B., Martín-Burriel, I., Morales, M., Andrés-Lasheras, S., et al. (2020). Antimicrobial resistance among canine enteric Escherichia coli isolates and prevalence of attaching–effacing and extraintestinal pathogenic virulence factors in Spain. *Acta Vet. Hung.* 68, 1–7. doi:10.1556/004.2020.00013.

Sheikh, J., Hicks, S., Dall'Agnol, M., Phillips, A. D., and Nataro, J. P. (2008). Roles for Fis and YafK in biofilm formation by enteroaggregative Escherichia coli. *Mol. Microbiol.* 41, 983–997. doi:10.1046/j.1365-2958.2001.02512.x.

Sherlock, O., Schembri, M. A., Reisner, A., and Klemm, P. (2004). Novel Roles for the AIDA Adhesin from Diarrheagenic Escherichia coli: Cell Aggregation and Biofilm Formation. *J. Bacteriol.* 186, 8058–8065. doi:10.1128/JB.186.23.8058-8065.2004.

Shlaes, D. M., Sahm, D., Opiela, C., and Spellberg, B. (2013). The FDA Reboot of Antibiotic Development. *Antimicrob. Agents Chemother.* 57, 4605–4607. doi:10.1128/AAC.01277-13.

Shokouhi Mostafavi, S. K., Najar Peerayeh, S., Mohabbati Mobarez, A., and Parizi, M. K. (2018). Detection of virulence factors, phylogroups, serogroups and biofilm formation among CTX-M-1 positive Escherichia coli isolated from patients with pyelonephritis. *Biomed. Res.* 29. doi:10.4066/biomedicalresearch.29-18-270.

Shpigel, N. Y., Elazar, S., and Rosenshine, I. (2008). Mammary pathogenic Escherichia coli. *Curr. Opin. Microbiol.* 11, 60–65. doi:10.1016/j.mib.2008.01.004.

Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865–891. doi:10.1111/1574-6976.12067.

Simarro, E., Navarro, F., Ruiz, J., Miró, E., Gómez, J., and Mirelis, B. (2000). Salmonella enterica Serovar Virchow with CTX-M-Like β-Lactamase in Spain. *J. Clin. Microbiol.* 38, 4676–4678. doi:10.1128/JCM.38.12.4676-4678.2000.

Singer, R. S. (2015). Urinary tract infections attributed to diverse ExPEC strains in food animals: evidence and data gaps. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00028.

Skov, R. L., and Monnet, D. L. (2016). Plasmid-mediated colistin resistance (mcr-1 gene): Three months later, the story unfolds. *Eurosurveillance* 21, 30155. doi:10.2807/1560-7917.ES.2016.21.9.30155.

Smelov, V., Naber, K., and Bjerklund Johansen, T. E. (2016). Letter to the Editor: Diagnostic Criteria in Urological Diseases do not Always Match with Findings by Extended Culture Techniques and Metagenomic Sequencing of 16S rDNA. *Open Microbiol. J.* 10, 23–26. doi:10.2174/1874285801610010023.

Smith, J. L., Fratamico, P. M., and Gunther, N. W. (2007). Extraintestinal Pathogenic Escherichia coli. *Foodborne Pathog. Dis.* 4, 134–163. doi:10.1089/fpd.2007.0087.

Snesrud, E., He, S., Chandler, M., Dekker, J. P., Hickman, A. B., McGann, P., et al. (2016). A Model for Transposition of the Colistin Resistance Gene mcr-1 by IS Apl1. *Antimicrob. Agents Chemother.* 60, 6973–6976. doi:10.1128/AAC.01457-16.

Snesrud, E., McGann, P., and Chandler, M. (2018). The Birth and Demise of the ISApl1-mcr-1-ISApl1 Composite Transposon: the Vehicle for Transferable Colistin Resistance. *MBio* 9, e02381-17. doi:10.1128/mBio.02381-17.

Sonnevend, Á., Ghazawi, A., Alqahtani, M., Shibl, A., Jamal, W., Hashmey, R., et al. (2016). Plasmid-mediated colistin resistance in Escherichia coli from the Arabian Peninsula. *Int. J. Infect. Dis.* 50, 85–90. doi:10.1016/j.ijid.2016.07.007.

Spratt, B. G. (2004). "Exploring the Concept of Clonality in Bacteria," in *Genomics, Proteomics, and Clinical Bacteriology* (Totowa, NJ: Humana Press), 323–352. doi:10.1385/1-59259-763-7:323.

Spurbeck, R. R., Dinh, P. C., Walk, S. T., Stapleton, A. E., Hooton, T. M., Nolan, L. K., et al. (2012). Escherichia coli Isolates That Carry vat , fyuA , chuA , and yfcV Efficiently Colonize the Urinary Tract. *Infect. Immun.* 80, 4115–4122. doi:10.1128/IAI.00752-12.

Stenutz, R., Weintraub, A., and Widmalm, G. (2006). The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol. Rev.* 30, 382–403. doi:10.1111/j.1574-6976.2006.00016.x.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-Component Signal Transduction. *Annu. Rev. Biochem.* 69, 183–215. doi:10.1146/annurev.biochem.69.1.183.

Stoesser, N., Batty, E. M., Eyre, D. W., Morgan, M., Wyllie, D. H., Del Ojo Elias, C., et al. (2013). Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. *J. Antimicrob. Chemother.* 68, 2234–2244. doi:10.1093/jac/dkt180.

Stoesser, N., Mathers, A. J., Moore, C. E., Day, N. P., and Crook, D. W. (2016a). Colistin resistance gene mcr-1 and pHNSHP45 plasmid in human isolates of Escherichia coli and Klebsiella pneumoniae. *Lancet Infect. Dis.* 16, 285–286. doi:10.1016/S1473-3099(16)00010-4.

Stoesser, N., Sheppard, A. E., Pankhurst, L., De Maio, N., Moore, C. E., Sebra, R., et al. (2016b). Evolutionary History of the Global Emergence of the Escherichia coli Epidemic Clone ST131. *MBio* 7, no pagination. doi:10.1128/mBio.02162-15.

Stoesser, N., Sheppard, A. E., Peirano, G., Sebra, R., Lynch, T., Anson, L., et al. (2016c). Complete Sequencing of Plasmids Containing bla OXA-163 and bla OXA-48 in Escherichia coli Sequence Type 131. *Antimicrob. Agents Chemother.* 60, 6948–6951. doi:10.1128/AAC.01130-16.

Stoll, B. J., Hansen, N. I., Sanchez, P. J., Faix, R. G., Poindexter, B. B., Van Meurs, K. P., et al. (2011). Early Onset Neonatal Sepsis: The Burden of Group B Streptococcal and E. coli Disease Continues. *Pediatrics* 127, 817–826. doi:10.1542/peds.2010-2217.

Suay-García, B., Galán, F., Rodríguez-Iglesias, M. A., and Pérez-Gracia, M. T. (2019). Detection and Characterization of Extended-Spectrum Beta-Lactamases-Producing Escherichia coli in Animals. *Vector-Borne Zoonotic Dis.* 19, 115–120.

doi:10.1089/vbz.2018.2333.

Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011a). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi:10.1093/bioinformatics/btr039.

Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011b). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi:10.1093/bioinformatics/btr039.

Sun, J., Fang, L.-X., Wu, Z., Deng, H., Yang, R.-S., Li, X.-P., et al. (2017). Genetic Analysis of the IncX4 Plasmids: Implications for a Unique Pattern in the mcr-1 Acquisition. *Sci. Rep.* 7, 424. doi:10.1038/s41598-017-00095-x.

Sun, J., Zhang, H., Liu, Y.-H., and Feng, Y. (2018). Towards Understanding MCR-like Colistin Resistance. *Trends Microbiol.* 26, 794–808. doi:10.1016/j.tim.2018.02.006.

Surgers, L., Boersma, P., Girard, P.-M., Homor, A., Geneste, D., Arlet, G., et al. (2018). Molecular epidemiology of ESBL-producing E. coli and K. pneumoniae: establishing virulence clusters. *Infect. Drug Resist.* Volume 12, 119–127. doi:10.2147/IDR.S179134.

Surgers, L., Boyd, A., Girard, P.-M., Arlet, G., and Decré, D. (2019). Biofilm formation by ESBL-producing strains of Escherichia coli and Klebsiella pneumoniae. *Int. J. Med. Microbiol.* 309, 13–18. doi:10.1016/j.ijmm.2018.10.008.

Tamadonfar, K. O., Omattage, N. S., Spaulding, C. N., and Hultgren, S. J. (2019). "Reaching the End of the Line: Urinary Tract Infections," in *Bacteria and Intracellularity* (American Society of Microbiology), 83–99. doi:10.1128/microbiolspec.BAI-0014-2019.

Tan, C., Tang, X., Zhang, X., Ding, Y., Zhao, Z., Wu, B., et al. (2012). Serotypes and virulence genes of extraintestinal pathogenic Escherichia coli isolates from diseased pigs in China. *Vet. J.* 192, 483–488. doi:10.1016/j.tvjl.2011.06.038.

Tapiainen, T., Hanni, A.-M., Salo, J., Ikãheimo, I., and Uhari, M. (2014). Escherichia coli biofilm formation and recurrences of urinary tract infections in children. *Eur. J. Clin. Microbiol. Infect. Dis.* 33, 111–115. doi:10.1007/s10096-013-1935-4.

Tarr, P. I., Bilge, S. S., Vary, J. C., Jelacic, S., Habeeb, R. L., Ward, T. R., et al. (2000). Iha: a Novel Escherichia coli O157:H7 Adherence-Conferring Molecule Encoded on a Recently Acquired Chromosomal Island of Conserved Structure. *Infect. Immun.* 68, 1400–1407. doi:10.1128/IAI.68.3.1400-1407.2000.

Tasse, J., Trouillet-Assant, S., Josse, J., Martins-Simjes, P., Valour, F., Langlois-Jacques, C., et al. (2018). Association between biofilm formation phenotype and clonal lineage in Staphylococcus aureus strains from bone and joint infections. *PLoS One* 13, e0200064. doi:10.1371/journal.pone.0200064.

Tchesnokova, V., Billig, M., Chattopadhyay, S., Linardopoulou, E., Aprikian, P., Roberts, P. L., et al. (2013). Predictive Diagnostics for Escherichia coli Infections Based on the Clonal Association of Antimicrobial Resistance and Clinical Outcome. *J. Clin. Microbiol.* 51, 2991–2999. doi:10.1128/JCM.00984-13.

Tchesnokova, V. L., Rechkina, E., Larson, L., Ferrier, K., Weaver, J. L., Schroeder, D. W., et al. (2019). Rapid and Extensive Expansion in the United States of a New Multidrug-resistant Escherichia coli Clonal Group, Sequence Type 1193. *Clin. Infect. Dis.* 68, 334–337. doi:10.1093/cid/ciy525.

Tegetmeyer, H. E., Jones, S. C. P., Langford, P. R., and Baltes, N. (2008). ISApl1, a novel insertion element of Actinobacillus pleuropneumoniae, prevents ApxIV-based serological detection of serotype 7 strain AP76. *Vet. Microbiol.* 128, 342–353. doi:10.1016/j.vetmic.2007.10.025.

Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010). The population genetics of commensal Escherichia coli. *Nat. Rev. Microbiol.* 8, 207–217. doi:10.1038/nrmicro2298.

Terlizzi, M. E., Gribaudo, G., and Maffei, M. E. (2017). UroPathogenic Escherichia coli (UPEC) Infections: Virulence Factors, Bladder Responses, Antibiotic, and Non-antibiotic Antimicrobial Strategies. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01566.

Thankavel, K., Madison, B., Ikeda, T., Malaviya, R., Shah, A. H., Arumugam, P. M., et al. (1997). Localization of a domain in the FimH adhesin of Escherichia coli type 1 fimbriae capable of receptor recognition and use of a domain-specific antibody to

confer protection against experimental urinary tract infection. *J. Clin. Invest.* 100, 1123–1136. doi:10.1172/JCI119623.

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nat. Rev. Microbiol.* 3, 711–721. doi:10.1038/nrmicro1234.

Thompson, M. F., Litster, A. L., Platell, J. L., and Trott, D. J. (2011). Canine bacterial urinary tract infections: New developments in old pathogens. *Vet. J.* 190, 22–27. doi:10.1016/j.tvjl.2010.11.013.

Tomenius, H., Pernestig, A.-K., Jonas, K., Georgellis, D., Mõllby, R., Normark, S., et al. (2006). The Escherichia coli BarA-UvrY two-component system is a virulence determinant in the urinary tract. *BMC Microbiol.* 6, 27. doi:10.1186/1471-2180-6-27.

Tornieporth, N. G., John, J., Salgado, K., de Jesus, P., Latham, E., Melo, M. C., et al. (1995). Differentiation of pathogenic Escherichia coli strains in Brazilian children by PCR. *J. Clin. Microbiol.* 33, 1371–1374. doi:10.1128/JCM.33.5.1371-1374.1995.

Tóth, I., Hérault, F., Beutin, L., and Oswald, E. (2003). Production of cytolethal distending toxins by pathogenic Escherichia coli strains isolated from human and animal sources: establishment of the existence of a new cdt variant (Type IV). *J. Clin. Microbiol.* 41, 4285–91. doi:10.1128/jcm.41.9.4285-4291.2003.

Totsika, M., Beatson, S. A., Sarkar, S., Phan, M.-D., Petty, N. K., Bachmann, N., et al. (2011). Insights into a Multidrug Resistant Escherichia coli Pathogen of the Globally Disseminated ST131 Lineage: Genome Analysis and Virulence Mechanisms. *PLoS One* 6, e26578. doi:10.1371/journal.pone.0026578.

Touchon, M., Perrin, A., de Sousa, J. A. M., Vangchhia, B., Burn, S., O'Brien, C. L., et al. (2020). Phylogenetic background and habitat drive the genetic diversification of Escherichia coli. *PLOS Genet.* 16, e1008866. doi:10.1371/journal.pgen.1008866.

Toval, F., Kohler, C.-D., Vogel, U., Wagenlehner, F., Mellmann, A., Fruth, A., et al. (2014). Characterization of Escherichia coli Isolates from Hospital Inpatients or Outpatients with Urinary Tract Infection. *J. Clin. Microbiol.* 52, 407–418. doi:10.1128/JCM.02069-13.

Trobos, M., Christensen, H., Sunde, M., Nordentoft, S., Agerso, Y., Simonsen, G. S., et al. (2009). Characterization of sulphonamide-resistant Escherichia coli using comparison of sul2 gene sequences and multilocus sequence typing. *Microbiology* 155, 831–836. doi:10.1099/mic.0.024190-0.

Turrientes, M.-C., González-Alba, J.-M., del Campo, R., Baquero, M.-R., Cantón, R., Baquero, F., et al. (2014). Recombination Blurs Phylogenetic Groups Routine Assignment in Escherichia coli: Setting the Record Straight. *PLoS One* 9, e105395. doi:10.1371/journal.pone.0105395.

Tyson, G. H., McDermott, P. F., Li, C., Chen, Y., Tadesse, D. A., Mukherjee, S., et al. (2015). WGS accurately predicts antimicrobial resistance in Escherichia coli. *J. Antimicrob. Chemother.* 70, 2763–2769. doi:10.1093/jac/dkv186.

Ukah, U. V., Glass, M., Avery, B., Daignault, D., Mulvey, M. R., Reid-smith, R. J., et al. (2018). Risk factors for acquisition of multidrug-resistant Escherichia coli and development of community-acquired urinary tract infections. *Epidemiol. Infect.* 146, 46–57. doi:10.1017/S0950268817002680.

Ulett, G. C., Valle, J., Beloin, C., Sherlock, O., Ghigo, J.-M., and Schembri, M. A. (2007). Functional Analysis of Antigen 43 in Uropathogenic Escherichia coli Reveals a Role in Long-Term Persistence in the Urinary Tract. *Infect. Immun.* 75, 3233–3244. doi:10.1128/IAI.01952-06.

Ulla-Maija Kãrkkãinen, R. I. (2000). Recurrence of Urinary Tract Infections in Adult Patients with Community-acquired Pyelonephritis Caused by E. coli: A 1-year Follow-up. *Scand. J. Infect. Dis.* 32, 495–499. doi:10.1080/003655400458767.

Ulleryd, P., Sandberg, T., Scheutz, F., Clabots, C., Johnston, B. D., Thuras, P., et al. (2015). Colonization with Escherichia coli Strains among Female Sex Partners of Men with Febrile Urinary Tract Infection. *J. Clin. Microbiol.* 53, 1947–1950. doi:10.1128/JCM.00579-15.

Umeda, K., Hase, A., Matsuo, M., Horimoto, T., and Ogasawara, J. (2019). Prevalence and genetic characterization of cephalosporin-resistant Enterobacteriaceae among dogs

and cats in an animal shelter. *J. Med. Microbiol.* 68, 339–345. doi:10.1099/jmm.0.000933.

Valat, C., Drapeau, A., Beurlet, S., Bachy, V., Boulouis, H.-J., Pin, R., et al. (2020). Pathogenic Escherichia coli in Dogs Reveals the Predominance of ST372 and the Human-Associated ST73 Extra-Intestinal Lineages. *Front. Microbiol.* 11. doi:10.3389/fmicb.2020.00580.

Valenza, G., Werner, M., Eisenberger, D., Nickel, S., Lehner-Reindl, V., Hõller, C., et al. (2019). First report of the new emerging global clone ST1193 among clinical isolates of extended-spectrum β-lactamase (ESBL)-producing Escherichia coli from Germany. *J. Glob. Antimicrob. Resist.* 17, 305–308. doi:10.1016/j.jgar.2019.01.014.

Van Boeckel, T. P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B. T., Levin, S. A., et al. (2014). Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. *Lancet Infect. Dis.* 14, 742–750. doi:10.1016/S1473-3099(14)70780-7.

Van Boeckel, T. P., Glennon, E. E., Chen, D., Gilbert, M., Robinson, T. P., Grenfell, B. T., et al. (2017). Reducing antimicrobial use in food animals. *Science (80-.).* 357, 1350–1352. doi:10.1126/science.aao1495.

van der Mee-Marquet, N. L., Blanc, D. S., Gbaguidi-Haore, H., Dos Santos Borges, S., Viboud, Q., Bertrand, X., et al. (2015). Marked increase in incidence for bloodstream infections due to Escherichia coli, a side effect of previous antibiotic therapy in the elderly. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00646.

van Dijk, E. L., Auger, H., Jaszczyszyn, Y., and Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends Genet.* 30, 418–426. doi:10.1016/j.tig.2014.07.001.

van Hoek, A. H. A. M., Veenman, C., Florijn, A., Huijbers, P. M. C., Graat, E. A. M., de Greeff, S., et al. (2018). Longitudinal study of ESBL Escherichia coli carriage on an organic broiler farm. *J. Antimicrob. Chemother.* doi:10.1093/jac/dky362.

Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P T* 40, 277–83.

Vielva, L., de Toro, M., Lanza, V. F., and de la Cruz, F. (2017). PLACNETw: a web-based tool for plasmid reconstruction from bacterial genomes. *Bioinformatics* 33, 3796–3798. doi:10.1093/bioinformatics/btx462.

Vigil, P. D., Stapleton, A. E., Johnson, J. R., Hooton, T. M., Hodges, A. P., He, Y., et al. (2011). Presence of Putative Repeat-in-Toxin Gene tosA in Escherichia coli Predicts Successful Colonization of the Urinary Tract. *MBio* 2. doi:10.1128/mBio.00066-11.

Villa, L., García-Fernández, A., Fortini, D., and Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *J. Antimicrob. Chemother.* 65, 2518–2529. doi:10.1093/jac/dkq347.

Vincent, C., Boerlin, P., Daignault, D., Dozois, C. M., Dutil, L., Galanakis, C., et al. (2010). Food Reservoir for Escherichia coli Causing Urinary Tract Infections. *Emerg. Infect. Dis.* 16, 88–95. doi:10.3201/eid1601.091118.

Wagner, S., Gally, D. L., and Argyle, S. A. (2014). Multidrug-resistant Escherichia coli from canine urinary tract infections tend to have commensal phylotypes, lower prevalence of virulence determinants and ampC-replicons. *Vet. Microbiol.* 169, 171–178. doi:10.1016/j.vetmic.2014.01.003.

Walk, S. T., Alm, E. W., Gordon, D. M., Ram, J. L., Toranzos, G. A., Tiedje, J. M., et al. (2009). Cryptic Lineages of the Genus Escherichia. *Appl. Environ. Microbiol.* 75, 6534–6544. doi:10.1128/AEM.01262-09.

Wang, J., Huang, X.-Y., Xia, Y.-B., Guo, Z.-W., Ma, Z.-B., Yi, M.-Y., et al. (2018a). Clonal Spread of Escherichia coli ST93 Carrying mcr-1-Harboring IncN1-IncHI2/ST3 Plasmid Among Companion Animals, China. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.02989.

Wang, L., Rothemund, D., Curd, H., and Reeves, P. R. (2003). Species-Wide Variation in the Escherichia coli Flagellin (H-Antigen) Gene. *J. Bacteriol.* 185, 2936–2943. doi:10.1128/JB.185.9.2936-2943.2003.

Wang, S., Zhao, S.-Y., Xiao, S.-Z., Gu, F.-F., Liu, Q.-Z., Tang, J., et al. (2016). Antimicrobial Resistance and Molecular Epidemiology of Escherichia coli Causing Bloodstream Infections in Three Hospitals in Shanghai, China. *PLoS One* 11, e0147740.

doi:10.1371/journal.pone.0147740.

Wang, X., Wang, Y., Zhou, Y., Li, J., Yin, W., Wang, S., et al. (2018b). Emergence of a novel mobile colistin resistance gene, mcr-8 , in NDM-producing Klebsiella pneumoniae. *Emerg. Microbes Infect.* 7, 1–9. doi:10.1038/s41426-018-0124-z.

Watts, R. E., Hancock, V., Ong, C.-L. Y., Vejborg, R. M., Mabbett, A. N., Totsika, M., et al. (2010). Escherichia coli Isolates Causing Asymptomatic Bacteriuria in Catheterized and Noncatheterized Individuals Possess Similar Virulence Properties. *J. Clin. Microbiol.* 48, 2449–2458. doi:10.1128/JCM.01611-09.

Wedley, A. L., Dawson, S., Maddox, T. W., Coyne, K. P., Pinchbeck, G. L., Clegg, P., et al. (2017). Carriage of antimicrobial resistant Escherichia coli in dogs: Prevalence, associated risk factors and molecular characteristics. *Vet. Microbiol.* 199, 23–30. doi:10.1016/j.vetmic.2016.11.017.

Weissman, S. J., Johnson, J. R., Tchesnokova, V., Billig, M., Dykhuizen, D., Riddell, K., et al. (2012). High-Resolution Two-Locus Clonal Typing of Extraintestinal Pathogenic Escherichia coli. *Appl. Environ. Microbiol.* 78, 1353–1360. doi:10.1128/AEM.06663-11.

Welch, R. A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., et al. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. *Proc. Natl. Acad. Sci.* 99, 17020–17024. doi:10.1073/pnas.252529799.

Wells, T. J., Totsika, M., and Schembri, M. A. (2010). Autotransporters of Escherichia coli: a sequence-based characterization. *Microbiology* 156, 2459–2469. doi:10.1099/mic.0.039024-0.

Whittam, T. S., Wolfe, M. L., and Wilson, R. A. (1989). Genetic relationships among Escherichia coli isolates causing urinary tract infections in humans and animals. *Epidemiol. Infect.* 102, 37–46. doi:10.1017/S0950268800029666.

WHO (2011). Takling antybiotic resistance from a food safety perspectif in europe.

WHO (2014). Antimicrobial resistance: global report on surveillance 2014.

WHO (2015). Global action plan on antimicrobial resistance.

WHO (2017). Executive summary: the selection and use of essential medicines. Report of the 21st WHO Expert Committee on the Selection and Use of Essential Medicines.

Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of de novo genome assemblies: Fig. 1. *Bioinformatics* 31, 3350–3352. doi:10.1093/bioinformatics/btv383.

Wiles, T. J., Kulesus, R. R., and Mulvey, M. A. (2008). Origins and virulence mechanisms of uropathogenic Escherichia coli. *Exp. Mol. Pathol.* 85, 11–19. doi:10.1016/j.yexmp.2008.03.007.

Wiles, T. J., and Mulvey, M. A. (2013). The RTX pore-forming toxin α-hemolysin of uropathogenic Escherichia coli : progress and perspectives. *Future Microbiol.* 8, 73–84. doi:10.2217/fmb.12.131.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L. H., et al. (2006). Sex and virulence in Escherichia coli: an evolutionary perspective. *Mol. Microbiol.* 60, 1136–1151. doi:10.1111/j.1365-2958.2006.05172.x.

Wu, J., Lan, F., Lu, Y., He, Q., and Li, B. (2017). Molecular Characteristics of ST1193 Clone among Phylogenetic Group B2 Non-ST131 Fluoroquinolone-Resistant Escherichia coli. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.02294.

Wu, R., Yi, L., Yu, L., Wang, J., Liu, Y., Chen, X., et al. (2018). Fitness Advantage of mcr-1–Bearing IncI2 and IncX4 Plasmids in Vitro. *Front. Microbiol.* 9, 331. doi:10.3389/fmicb.2018.00331.

Xia, S., Fan, X., Huang, Z., Xia, L., Xiao, M., Chen, R., et al. (2014). Dominance of CTX-M-Type Extended-Spectrum β-Lactamase (ESBL)-Producing Escherichia coli Isolated from Patients with Community-Onset and Hospital-Onset Infection in China. *PLoS One* 9, e100707. doi:10.1371/journal.pone.0100707.

Xicohtencatl-Cortes, J., Cruz-Córdova, A., Cázares-Domínguez, V., Escalona-Venegas, G., Zavala-Vega, S., Arellano-Galindo, J., et al. (2019). Uropathogenic Escherichia coli strains harboring tosA gene were associated to high virulence genes and a multidrug-resistant profile. *Microb. Pathog.* 134, 103593.

doi:10.1016/j.micpath.2019.103593.

Yamaji, R., Rubin, J., Thys, E., Friedman, C. R., and Riley, L. W. (2018). Persistent Pandemic Lineages of Uropathogenic Escherichia coli in a College Community from 1999 to 2017. *J. Clin. Microbiol.* 56, e01834-17. doi:10.1128/JCM.01834-17.

Yamamoto, S., Terai, A., Yuri, K., Kurazono, H., Takeda, Y., and Yoshida, O. (1995). Detection of urovirulence factors in Escherichia coli by multiplex polymerase chain reaction. *FEMS Immunol. Med. Microbiol.* 12, 85–90. doi:10.1111/j.1574-695X.1995.tb00179.x.

Yamamoto, S., Tsukamoto, T., Terai, A., Kurazono, H., Takeda, Y., and Yoshida, O. (1997). Genetic evidence supporting the fecal-perineal-urethral hypothesis in cystitis caused by Escherichia coli. *J. Urol.* 157, 1127–9.

Yang, R.-S., Feng, Y., Lv, X.-Y., Duan, J.-H., Chen, J., Fang, L.-X., et al. (2016). Emergence of NDM-5- and MCR-1-Producing Escherichia coli Clones ST648 and ST156 from a Single Muscovy Duck (Cairina moschata). *Antimicrob. Agents Chemother.* 60, 6899–6902. doi:10.1128/AAC.01365-16.

Yang, Y.-Q., Li, Y.-X., Lei, C.-W., Zhang, A.-Y., and Wang, H.-N. (2018). Novel plasmid-mediated colistin resistance gene mcr-7.1 in Klebsiella pneumoniae. *J. Antimicrob. Chemother.* 73, 1791–1795. doi:10.1093/jac/dky111.

Yano, H., Rogers, L. M., Knox, M. G., Heuer, H., Smalla, K., Brown, C. J., et al. (2013). Host range diversification within the IncP-1 plasmid group. *Microbiology* 159, 2303–2315. doi:10.1099/mic.0.068387-0.

Yao, X., Doi, Y., Zeng, L., Lv, L., and Liu, J.-H. (2016). Carbapenem-resistant and colistin-resistant Escherichia coli co-producing NDM-9 and MCR-1. *Lancet Infect. Dis.* 16, 288–289. doi:10.1016/S1473-3099(16)00057-8.

Yin, W., Li, H., Shen, Y., Liu, Z., Wang, S., Shen, Z., et al. (2017). Erratum to: Yin et al., "Novel plasmid-mediated colistin resistance Gene mcr-3 in Escherichia coli" (mBio, (2017) 8(3), e00543-17, 10.1128/mBio.00543-17). *MBio* 8. doi:10.1128/mBio.01166-17.

Zacchè, M. M., and Giarenis, I. (2016). Therapies in early development for the treatment of urinary tract inflammation. *Expert Opin. Investig. Drugs* 25, 531–540. doi:10.1517/13543784.2016.1161024.

Zankari, E., Allesøe, R., Joensen, K. G., Cavaco, L. M., Lund, O., and Aarestrup, F. M. (2017). PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J. Antimicrob. Chemother.* 72, 2764–2768. doi:10.1093/jac/dkx217.

Zankari, E. E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012a). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi:10.1093/jac/dks261.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012b). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi:10.1093/jac/dks261.

Zankari, E., Hasman, H., Kaas, R. S., Seyfarth, A. M., Agerso, Y., Lund, O., et al. (2013). Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J. Antimicrob. Chemother.* 68, 771–777. doi:10.1093/jac/dks496.

Zeng, K., Doi, Y., Patil, S., Huang, X., and Tian, G.-B. (2016). Emergence of the Plasmid-Mediated mcr-1 Gene in Colistin-Resistant Enterobacter aerogenes and Enterobacter cloacae: TABLE 1. *Antimicrob. Agents Chemother.* 60, 3862–3863. doi:10.1128/AAC.00345-16.

Zhang, P. L. C., Shen, X., Chalmers, G., Reid-Smith, R. J., Slavic, D., Dick, H., et al. (2018). Prevalence and mechanisms of extended-spectrum cephalosporin resistance in clinical and fecal Enterobacteriaceae isolates from dogs in Ontario, Canada. *Vet. Microbiol.* 213, 82–88. doi:10.1016/j.vetmic.2017.11.020.

Zhao, F., and Zong, Z. (2016). Kluyvera ascorbata carrying the mcr-1 colistin resistance gene from hospital sewage. *Antimicrob. Agents Chemother.* 60, AAC.01165-16. doi:10.1128/AAC.01165-16.

Zhong, L.-L., Phan, H. T. T. T., Shen, C., Vihta, K.-D., Sheppard, A. E., Huang, X., et al. (2018). High Rates of Human Fecal

Carriage of mcr-1–Positive Multidrug-Resistant Enterobacteriaceae Emerge in China in Association With Successful Plasmid Families. *Clin. Infect. Dis.* 66, 676–685. doi:10.1093/cid/cix885.

Zhou, F., and Xu, Y. (2010). cBar: a computer program to distinguish plasmid-derived from chromosome-derived sequence fragments in metagenomics data. *Bioinformatics* 26, 2051–2052. doi:10.1093/bioinformatics/btq299.

Zhou, Z., Alikhan, N.-F., Mohamed, K., Fan, Y., and Achtman, M. (2020). The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. *Genome Res.* 30, 138–152. doi:10.1101/gr.251678.119.

Zhu, Y., Dong, W., Ma, J., Yuan, L., Hejair, H. M. A., Pan, Z., et al. (2017). Characterization and virulence clustering analysis of extraintestinal pathogenic Escherichia coli isolated from swine in China. *BMC Vet. Res.* 13, 94. doi:10.1186/s12917-017-0975-x.

Zogg, A. L., Zurfluh, K., Schmitt, S., Nüesch-Inderbinen, M., and Stephan, R. (2018). Antimicrobial resistance, multilocus sequence types and virulence profiles of ESBL producing and non-ESBL producing uropathogenic Escherichia coli isolated from cats and dogs in Switzerland. *Vet. Microbiol.* 216, 79–84. doi:10.1016/j.vetmic.2018.02.011.

Zurfluh, K., Klumpp, J., Nüesch-Inderbinen, M., and Stephan, R. (2016a). Full-Length Nucleotide Sequences of *mcr-1* -Harboring Plasmids Isolated from Extended-Spectrum-β-Lactamase-Producing Escherichia coli Isolates of Different Origins. *Antimicrob. Agents Chemother.* 60, 5589–5591. doi:10.1128/AAC.00935-16.

Zurfluh, K., Tasara, T., Poirel, L., Nordmann, P., and Stephan, R. (2016b). Draft Genome Sequence of Escherichia coli S51, a Chicken Isolate Harboring a Chromosomally Encoded mcr-1 Gene. *Genome Announc.* 4, 4–5. doi:10.1128/genomeA.00796-16.

8 ANNEXES

8.1 ANNEX 1: MATERIAL AND METHODS SUPPLEMENTARY MATERIAL

Table S1. Assembly information of genomes.

Genome	kmer	No. of contigs	No. of contigs > 1kb	N50	Longest contig (bp)	Total bp
Human origin genomes						
LREC_263	93	113	72	203374	350007	5140392
LREC_264	115	103	61	243728	693027	5105112
LREC_265	117	102	50	222698	624019	5088133
LREC_266	113	102	58	196351	623442	5203023
LREC_268	115	151	77	290559	512299	5399129
LREC_269	117	159	68	221057	515993	5140621
LREC_270	117	102	63	217209	957382	5087242
LREC_271	89	114	66	194949	928860	5128192
LREC_272	117	147	71	323644	757085	5269464
LREC_273	117	152	71	214599	696746	5267138
LREC_274	117	103	60	338527	562364	5173361
LREC_275	117	94	46	297751	512897	5194217
LREC_276	113	101	49	204926	673059	5341151
LREC_277	117	172	74	191857	414832	5306873
LREC_278	113	123	62	193172	659355	5296741
LREC_279	113	165	67	192014	431183	5406813
LREC_280	115	130	64	204363	658862	5403286
LREC_281	115	81	46	288560	774020	5346311
LREC_282	113	86	48	265392	619346	5265015
LREC_283	113	113	49	342281	658956	5163448
LREC_284	113	92	47	282391	720386	4895377
LREC_285	117	86	56	234793	849655	4967927
LREC_286	117	149	58	342415	626222	5211994
LREC_287	117	125	65	258141	660190	5181373
LREC_288	117	113	55	379986	776221	5170343
LREC_289	113	81	38	472249	926038	5077629
LREC_290	115	107	54	261454	484687	5094876
LREC_291	113	120	53	319777	884132	5138992
LREC_292	117	122	51	251466	884253	5114696
LREC_293	113	168	81	184646	829973	5402287
LREC_294	113	178	91	125770	358142	4992032
LREC_295	113	142	64	184135	361914	4987057
LREC_296	117	194	104	112670	263666	5184662
LREC_298	89	115	55	269939	896901	5176306
LREC_299	117	104	38	388083	844183	5167647
LREC_300	117	106	37	384528	809990	5089070
LREC_301	117	173	85	182793	467306	5314537
LREC_302	115	133	67	157126	357708	4870632
LREC_303	117	129	64	235849	488907	5068117
LREC_304	117	133	63	175873	360474	4936865
LREC_305	109	199	58	366005	758657	5377588
LREC_306	115	223	73	359018	759064	5323487
LREC_307	113	141	53	296449	540602	5238001
LREC_308	117	97	43	247620	1147271	5092158
LREC_309	117	86	39	447425	738469	5066850
LREC_310	117	112	46	410939	526915	5296446
LREC_311	113	118	61	332451	638235	5133186
LREC_312	115	110	57	332455	589442	5048514
LREC_313	113	155	73	198707	638574	5313931
LREC_314	113	189	77	110784	295290	4871776
LREC_315	117	215	100	144552	448963	5197349
LREC_316	113	140	67	173221	408695	4800578
LREC_317	113	136	68	176315	382221	4715420

Table S1. *Cont.*

Genome	kmer	No. of contigs	No. of contigs > 1kb	N50	Longest contig (bp)	Total bp
Human origin genomes						
LREC_318	117	154	77	180225	375767	4896108
LREC_319	113	129	71	131563	263942	4877624
LREC_320	111	155	84	117598	232285	5023956
LREC_321	89	131	69	150213	392450	4758372
LREC_322	87	220	113	119222	292155	5307554
LREC_323	89	187	83	130886	434550	4944756
LREC_324	85	169	79	135677	605768	5247965
LREC_325	89	162	81	113152	344562	5053174
LREC_326	93	123	53	324350	765552	5380491
LREC_327	91	131	54	398957	765417	5387169
LREC_328	89	207	97	172902	432860	5754113
LREC_329	89	173	78	204185	599010	5774177
LREC_331	89	95	52	159131	431279	5054900
LREC_332	89	123	74	160816	435673	5141499
LREC_333	89	227	81	169280	411679	5164249
LREC_334	93	211	90	201968	619956	5256962
LREC_335	89	151	77	212378	564133	5220032
LREC_336	89	156	80	212147	372413	5266493
LREC_337	89	277	127	116032	457865	5585838
LREC_338	89	128	50	260986	680138	5246463
LREC_339	89	195	78	303701	686628	5330377
LREC_340	89	124	50	305907	680041	5244820
Canine origin genomes						
LREC_341	89	72	44	326460	672653	4840787
LREC_342	93	58	32	673054	720458	4789379
LREC_343	89	170	52	402644	823930	5173983
LREC_344	89	107	32	628916	1207514	5110297
LREC_345	91	180	52	391453	824252	5272043
LREC_346	93	116	35	598033	823983	5217811
LREC_347	98	118	32	592244	1702793	5104975
LREC_348	89	181	52	409123	824017	5158712
LREC_349	93	129	37	520829	824077	5150651
LREC_350	91	133	38	463354	824208	5150646
LREC_351	93	136	39	583831	824105	5159107
LREC_352	89	118	38	399516	657523	5082876
LREC_353	89	121	38	399704	657637	5083364
LREC_354	91	127	40	399799	657586	5080519
LREC_355	91	179	52	402920	824131	5159806
LREC_356	93	98	37	397327	2540863	5355508
LREC_357	91	85	25	864233	1677509	5045168
LREC_358	93	99	28	659764	777701	5099113
LREC_359	89	11	29	628732	1317107	5093206
LREC_360	93	109	34	506708	825202	5088464
LREC_361	93	112	29	628853	1317062	5093489
LREC_362	93	105	31	592465	743574	5056878
Porcine origin genomes						
LREC_153	101	101	34	400,612	1,112,169	5,358,477
LREC_154	101	140	59	202,640	547,553	5,133,266
LREC_155	89	129	69	274,127	719,927	5,482,745
LREC_157	101	93	40	325,448	633,164	5,094,477
LREC_158	89	64	32	392,821	674,978	5,001,845
LREC_159	89	119	69	165,826	712,711	5,375,593
LREC_160	101	136	64	220,432	696,874	5,252,736
LREC_161	89	116	54	370,583	601,486	5,456,371
LREC_162	93	75	34	468,418	1,208,336	4,940,669
LREC_168	93	124	50	412,870	801,182	5,265,150
LREC_176	117	171	78	169,181	511,118	5,451,627

Table S2. Database sources and input parameters.

DATA BASE	Minimum Threshold for %ID	Minimum length for %ID	Software version	Database version	Reference DOI
Clermont Typing Web	DEFAULT	DEFAULT		20.03	DOI: 10.1099/mgen.0.000192; DOI: 10.11111/1462-2920.14713
SerotypeFinder 2.0	85	60	2.0.1 (2019-01-28)	1.0.0 (2019-02-27)	DOI: 10.1128/JCM.00008-15
PlasmidFinder 2.1	95	60	2.0.1 (2020-02-07)	2.0.1 (2020-02-07)	DOI: 10.1007/978-1-4939-9877-7_20
pMLST 2.0	DEFAULT	DEFAULT	0.1.0 (2018-12-20)	(2019-07-29)	DOI: 10.1007/978-1-4939-9877-7_20
ResFinder 3.2	90	60	3.2 (2020-03-31)	(2020-02-11)	doi: 10.1093/jac/dks261
PointFinder	90	60	3.1.0 (2019-02-27)	(2019-07-02)	doi: 10.1093/jac/dks261
VirulenceFinder 2.0	90	60	2.0 (2020-02-06)	(2019-09-30)	doi:10.1128/JCM.03617-13
CARD db	90	80	March 2020 release	3.0.8	DOI: 10.1093/nar/gkz935
MLST 2.0 (Multi-Locus Sequence Typing)	DEFAULT	DEFAULT	2.0.4 (2019-05-08)	2.0.0 (2020-04-13)	doi: 10.12.0/JCM.06094-11
MyDbFinder 2.0 (Upload your own database)	80	60	2.0		
CSI Phylogeny 1.4	DEFAULT	DEFAULT	1.4	(2019-05-26)	https://doi.org/10.1371/journal.pone.0104984
ICEberg (ICEfinder and VRprofile)	DEFAULT	DEFAULT	ICEberg 2.0		https://doi.org/10.1093/nar/gkr846

Table S3. Plasmid references.

Plasmid genome	Isolate	Species	Assembly accesion	size (bp)	Isolation sources	*Escherichia coli* (ST)	Country	Plasmid type	mcr-1 environment	Reference (DOI)
pMR0516mcr	MRSN388634	*Escherichia coli*	KX276657.1	225069	Human urine	Yes (ST457)	USA (Pennsylvania)	IncF (F18:A-:B1)	ISApl1-mcr-1	McGann, P. *et al.* (2016) (10.1128/AAC.01103-16)
pKP81-BE	KP81	*Escherichia coli*	KU994859.1	91041	Pig feaces	Yes (ST10)	Belgium	IncFII (F2:A-B)	ISApl1-mcr-1	Xavier, B.B. *et al.* (2016) (https://doi.org/10.1093/jac/dkw191)
pECJS-59-244	EC59	*Escherichia coli*	KX084394.1	243572	Pig feaces	Yes	China	IncHI2	Tn6330	Ruichao, Li. *et al.* (2017) (https://doi.org/10.1093/jac/dkw411)
pHNSHP45-2	SHP45	*Escherichia coli*	KU341381.1	251493	Pig feaces	Yes	China	IncHI2	ISApl1-mcr-1-orf	Zhi C. *et al.* (2016) (10.1016/S1473-3099(16)00063-3)
pMCR_WCHEC050613	WCHEC050613	*Escherichia coli*	CP019214.2	289112	Wastewater hospital	Yes	China	IncHI2	ISApl1-mcr-1-orf	Zhao,F. *et al.* Unpublished
pEC2474_MCR1	Z247	*Escherichia coli*	CP021209	223982	Human blood	Yes	China	IncHI2	ISApl1-mcr-1-orf	Zheng, B. *et al.* (2016) (https://doi.org/10.1093/cid/ciw553)
pICBEC7Pmcr	ICBEC7P	*Escherichia coli*	NZ_CP017246.1	34992	Magellanic penguins pododermatitis	Yes	Brazil	IncX4	mcr-1-orf	Sellera, FP. *et al.* (2016) (https://doi.org/10.1093/jac/dkw543)
pMCR_WCHEC1618	WCHEC1618	*Escherichia coli*	KY463454.1	33309	Wastewater hospital	Yes	China	IncX4	mcr-1-orf	Zhao, F. *et al.* (2017) (10.3389/fmicb.2017.02094)
pESTMCR	ENV-187	*Escherichia coli*	KU743383.1	33311	Pig feaces	Yes	Estonia	IncX4	mcr-1-orf	Brauer, A. *et al.* (2017) (10.1128/AAC.00443-16)
pMCR1_IncX4	SZ04	*Klebsiella*	KU761327.1	33287	human	no (ST25)	China	IncX4	mcr-1-orf	Du,H. *et al.* (2016) (10.1016/S1473-3099(16)00056-6)
pECGD-8-33	GD-8	*Escherichia coli*	KX254343.1	33307	Pig feaces	Yes	China	IncX4	mcr-1-orf	Ruichao, Li. *et al.* (2017) (https://doi.org/10.1093/jac/dkw411)
pECSF1	SE15	*Escherichia coli*	NC_013655.1	122345	human commensal	Yes (ST131)	-	IncF	Negative	Toh,H. *et al.* (2010)
plasmid F	K-12	*Escherichia coli*	AP001918.1	99159	human commensal	Yes	-	IncF	Negative	Shimizu,H. *et al.* Unpublished

Table S3. Cont.

Plasmid genome	Isolate	Species	Assembly accesion	size (bp)	Isolation sources	*Escherichia coli* (ST)	Country	Plasmid type	mcr-1 environment	Reference (DOI)
R478	-	*Serratia marcescens*	U62007.2	12792	-	no	-	IncHI2	Negative	Gabant,P. *et al.* (1993)
pSH146_32	-	*Salmonella enterica*	JX258655.1	32447	-	no	-	IncX4	Negative	Gokulan,K. *et al.* (2013)
pJIE143	JIE143	*Escherichia coli*	JN194214.1	34345	human-ESBL	Yes (ST131)	-	IncX4	Negative	Partridge,S.R. *et al.* (2011)
pHS30-1	HS30-1	*Escherichia coli*	CP029493.1	179444	-	Yes	-	Composite transposon-Tn6330		Li,R. *et al.* Unpublished
pJIE186_2	JIE186	*Escherichia coli*	NC_020271	137728	Human urine	yes (ST131)	Sydney, Australia	IncF	Negative	Zong Z. (2013) (DOI: 10.1128/AAC.01081-12)
ColE1	A745	*Escherichia coli*	J01566.1	6646	-	yes	-	ColE1	Negative	Tomizawa,J.I. *et al.* (1977) (DOI: 10.1073/pnas.74.5.1865)
p2ESCUM	UMN026	*Escherichia coli*	NC_011739.1	33809	Human urine	yes (ST69)	Minnesota, UK.	IncX1	Negative	Lescat *et al.* (2009) (DOI: 10.1128/AAC.00123-09)
pSH696_34	-	*Salmonella enterica*	JX258654.1	33765	-	no	-	IncX4	Negative	Gokulan,K. *et al.* (2013) (DOI: 10.1371/journal.pone.0077866)
pEK204	C	*Escherichia coli*	EU935740	93732	Human	yes (ST131)	Belfast, UK.	IncI	Negative	Woodford N. *et al.* (2009) (DOI: 10.1128/AAC.00688-09)
pEC743-OXA48	Ecol_743	*Escherichia coli*	CP015071.1	69471	Human urine	yes (ST131)	Dubai, United Arab Emirates	IncL/M	Negative	Stoesser,N. *et al.* (2016) (doi:10.1128/AAC.01130-16.)
pCE10D	CE10	*Escherichia coli*	NC_017650.1	1549	neonate with meningitis	no	North America	-	Negative	Lu,S. *et al.* (2011) (DOI: 10.1128/JB.06284-11)
pEC0674	PEC15-AB00674	*Escherichia coli*	MF684783.1	6268	Pig feaces	no	Germany	-	Negative	Hammerl JA. *Et al.* (2018) (DOI: 10.1093/jac/dky020)

8.2. ANNEX 2: STUDIES SUPPLEMENTARY MATERIAL

8.2.1 STUDY 1

Supplementary material in Flament-Simon SC *et al.* 2020 Antibiotics 9 (4) 161 doi: 10.3390/antibiotics9040161.

Table S4. Phylogenetic groups of the 196 studied *E. coli* isolates and dominant sequence types with further characterization for ST131 (clades, subclades and virotypes).

Characteristic	No. (%) of isolates		*P-Value*
	Lugo, Spain (n= 100)	Clichy, France (n= 96)	
Phylogenetic group			
A	14	15	
B1	10	8	
B2	48	56	
C	11	4	
D	9	5	
E	5	5	
F	3	3	
ST131			
Total	12	11	
Clade A	5	1	
Clade B	2	2	
Cluster C1-M27	0	1	
Non-C1-M27 subclade C1	3	3	
Subclade C2 (*H*30-Rx)	2	3	
Clade-NT[a]	0	1	
Clonotype CH40-22	1	2	
Clonotype CH40-30	5	8	
Clonotype CH40-41	5	1	
Clonotype CH40-298	1	0	
Virotype A	3	0	
Virotype A-like	1	0	
Virotype C2	1	2	
Virotype C3	3	0	
Virotype D2	1	1	
Virotype D3	0	1	
Virotype D5-like	1	0	
Virotype E	0	2	
Virotype F	0	1	
Virotype NT	2	4	
Non-ST131 ST[b]			
ST10	3	6	
ST12	4	4	
ST58	4	1	
ST69	8	5	
ST73	9	12	
ST88	9	3	
ST95	6	4	
ST127	3	4	
ST141	3	11	0.0265

[a] NT = Not typeable; [b] Represented by at least 5 isolates.

Table S5. Characteristics of the 13 ESBL-producing *E. coli* isolates.

Isolate	Country	ESBL enzyme	Clone	Serotype	ExPEC status	UPEC status	ST131 clade, subclade and cluster
A24	Spain	CTX-M-1	A-CH11-54-ST34	O9:H1	-	-	
C6	France	CTX-M-1	A-CH11-43-ST3596	O8:HNM	-	-	
A46	Spain	CTX-M-1	A-CH27-23-ST new 2-437like	O98:HNM	-	-	
A7	Spain	CTX-M-14	B1-CH4-27-ST58	O9:H25	-	-	
A10	Spain	CTX-M-14	B1-CH6-35-ST448	O11:H10	-	-	
C22	France	CTX-M-15	B2-CH40-30-ST131	O25:H4	+	+	C2
C27	France	CTX-M-15	B2-CH40-30-ST131	O25:H4	-	+	NT[a]
C44	France	CTX-M-15	B2-CH40-30-ST131	O25:H4	+	+	C2
C3	France	CTX-M-27	B2-CH40-30-ST131	O25:H4	+	+	C1-M27
A82	Spain	CTX-M-1	B2-CH40-41-ST131	ONT:H5	+	-	A
C80	France	CTX-M-55	B2-CH14-64-ST1193	O75:HNM	+	+	
A62	Spain	CTX-M-14	E-CH37-27-ST405	O18:H6	-	-	
C93	France	CTX-M-32	F-CH4-58-ST648	ONT:H42	+	-	

[a] NT = Not typeable

Table S6. Antimicrobial resistance and Virulence factor(VF)-encoding genes.

VF gene	Number (%) isolates		*P-Value*
	MDR (n=73)	non-MDR (n= 123)	MDR vs. non-MDR
Adhesins			
fimH	71 (97.3)	122 (99.2)	
fimAv$_{MT78}$	10 (13.7)	17 (13.8)	
papAH	24 (32.9)	**60 (48.8)**[a]	0.03661
papC	26 (35.6)	**61 (49.6)**	0.07414
papEF	29 (39.7)	63 (51.2)	
sfa/focDE	7 (29.6)	**51 (41.5)**	0.000001
afa/draBC	7 (9.6)	4 (3.3)	
yfcV	28 (38.4)	**78 (63.4)**	0.00104
Toxins			
sat	20 (27.4)	27 (22.0)	
cnf1	8 (11.0)	**41 (33.3)**	0.00054
hlyA	7 (9.6)	**48 (39.0)**	0.000014
hlyF	20 (27.4)	24 (19.5)	
cdtB	3 (4.1)	9 (7.3)	
tsh	2 (2.7)	5 (4.1)	
vat	12 (16.4)	**68 (55.3)**	$8.5x10^{-8}$
Iron uptake			
iucD	42 (57.5)	57 (46.3)	
iutA	42 (57.5)	57 (46.3)	
iroN	24 (32.9)	**69 (56.1)**	0.00190
fyuA	59 (80.8)	103 (83.7)	
chuA	43 (58.9)	**92 (74.8)**	0.02546
Capsule			
kpsM II	34 (46.6)	**85 (69.1)**	0.00241
neuC-K1	5 (6.8)	**30 (24.4)**	0.00180
kpsM II-K2	6 (8.2)	11 (8.9)	
kpsM II-K5	23 (31.5)	44 (35.8)	
kpsM III	2 (2.7)	4 (3.3)	
Miscellaneous			
cvaC	13 (17.8)	22 (17.9)	
iss	21 (28.8)	25 (20.3)	
traT	**51 (69.9)**	60 (48.8)	0.00463
ibeA	4 (5.5)	**19 (15.4)**	0.03993
malX	30 (41.1)	**79 (64.2)**	0.00185
usp	29 (39.7)	**78 (63.4)**	0.00176
ompT	50 (68.5)	**101 (82.1)**	0.03517
ExPEC status	38 (52.1)	**83 (67.5)**	0.03459
UPEC status	27 (37.0)	**79 (64.2)**	0.00034
Mean of VF	9.90	12.67	

[a] Statistically significant differences are highlighted in bold.

8.2.2. STUDY 2

Supplementary material in Flament-Simon SC *et al.* 2020 Front Cell Infect Microbiol 10:125 doi: 10.3389/fcimb.2020.00125.

8.2.3. STUDY 3

Supplementary material in Flament-Simon SC *et al.* 2020 Front Microbiol 10:1183 doi: 10.3389/fmicb.2019.01183.

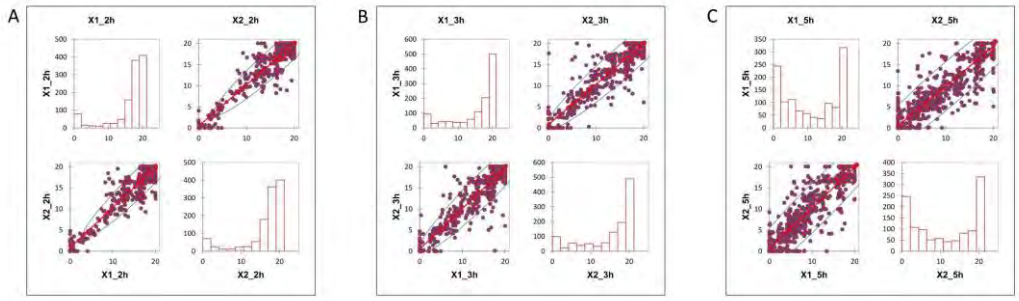


Figure S1A. Pearson´s correlation between the BFI values obtained for each of the 394 *E. coli* isolates. A: 0.96 at 2h, B: 0.96 at 3h and C: 0.95 at 5h.

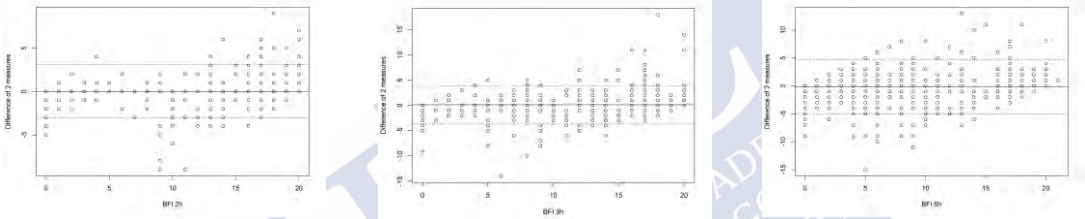


Figure S1B. Assessment of repeatability between the BFI values obtained for each of the 394 *E. coli* isolates using the Bland and Altman statistical method (Bland and Altman, 1986). The mean of individual BFI measures (continuous line) and standard deviation (dashed line) of all representing ±2SD from mean line were calculated at 2, 3 and 5h. Only 3.4% of values at 2h, 5.5% at 3h and 6.8% at 5h exceed the established standard deviation.

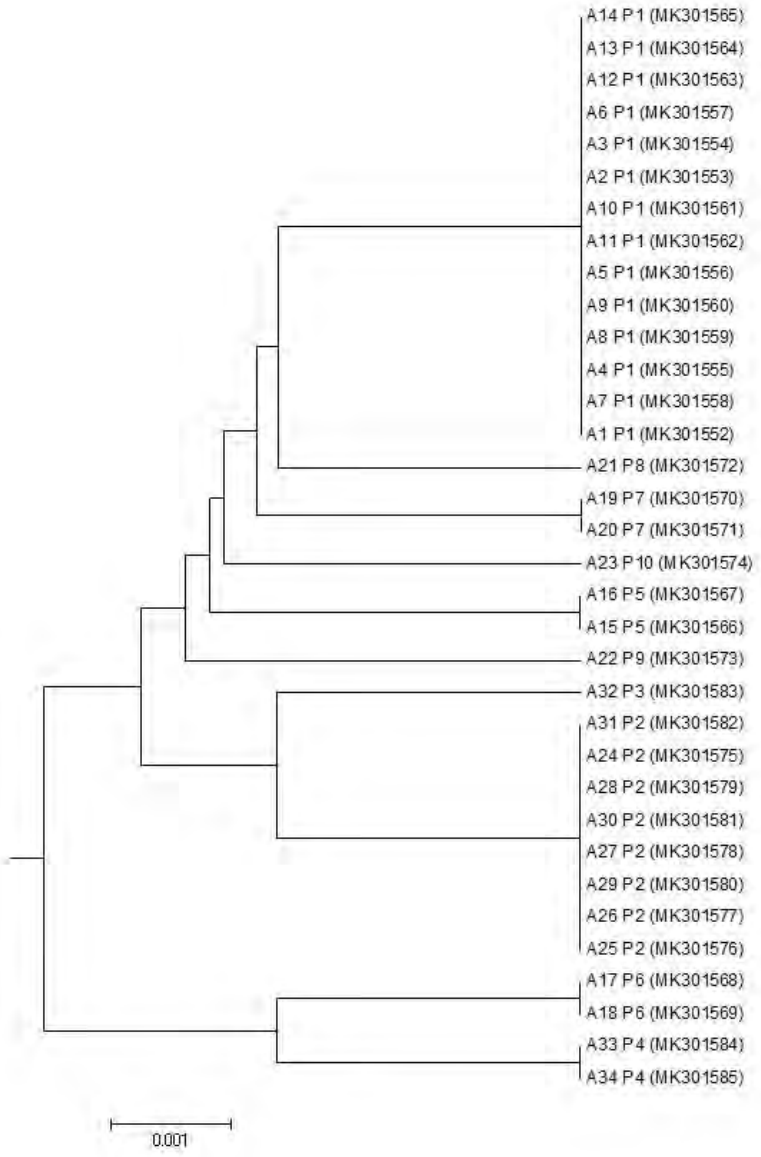


Figure S2. Phylogenetic tree of FimB. In parenthesis are indicated the access numbers of the sequences registered in GenBank database.

Table S7. Distribution of phylogroups A, B1, B2, C, D, E and F among G1, G2 and G3 isolates.

Phylogroup	Number (%) of isolates				*P-Value*[a]
	Total (n=394)	G1 (n=36)	G2 (n=28)	G3 (n=330)	G1 vs. G3
A	59 (15.0)	0 (0)	6 (21.4)	53 (16.1)	0.0098
B1	38 (9.6)	3 (8.3)	1 (3.6)	34 (10.3)	
B2	202 (51.3)	28 (77.8)	17 (60.7)	157 (47.6)	0.0006
C	38 (9.6)	0 (0)	0 (0)	38 (11.5)	0.0419
D	17 (4.3)	0 (0)	4 (11.1)	13 (3.9)	
E	20 (5.1)	2 (5.6)	0 (0)	18 (5.4)	
F	20 (5.1)	3 (8.3)	0 (0)	17 (5.2)	

[a] Two-tailed *P-Value*s by Fisher´s exact test are shown where *P*<0.05.

Table S8. Distribution of the most frequent sequence types (STs) among G1, G2 and G3 isolates.

Sequence Type	Number (%) of isolates				*P-Value*[a]
	Total (n=394)	G1 (n=36)	G2 (n=28)	G3 (n=330)	G1 vs. G3
ST10	17 (4.3)	0 (0)	2 (7.1)	15 (4.5)	
ST69	15 (3.8)	0 (0)	3 (10.7)	12 (3.6)	
ST73	22 (5.6)	4 (11.1)	5 (17.9)	13 (3.9)	
ST88	17 (4.3)	0 (0)	0 (0)	17 (5.2)	
ST95	14 (3.6)	0 (0)	0 (0)	14 (4.2)	
ST131	104 (26.4)	4 (11.1)	3 (10.7)	97 (29.4)	0.0289
ST141	16 (4.1)	9 (25)	2 (7.1)	5 (1.5)	<0.0001
Remaining-ST	189 (48.0)	19 (57.8)	13 (46.4)	157 (47.6)	

[a] Two-tailed *P-Value*s by Fisher´s exact test are shown where *P*<0.05

Table S11 *fim*B gene status in G1, G2 and G3 isolates.

***fim*B gene status**	**Number (%) of isolates**				***P-Value*[a]**
	Total (n=127)	**G1 (n=14)**	**G2 (n=12)**	**G3 (n=101)**	**G1 vs. G3**
Intact	89 (70.1)	13 (92.9)	10 (83.3)	66 (65.3)	0.0301
Disrupted	28 (22.0)	1 (7.1)	1 (8.3)	26 (25.7)	
Absent	10 (7.9)	0	1 (8.3)	9 (8.9)	

[a]One-tailed *P-Value*s by Fisher´s exact test are shown *P*<0.05.

Table S13. Distribution of FimB variants in G1, G2 and G3 isolates.

FimB variant	Number (%) of isolates				*P-Value*[a]
	Total (n=89)	G1 (n=13)	G2 (n=10)	G3 (n=66)	G1 vs. G3
P1	52 (58.4)	3 (23.1)	3 (30)	46 (69.7)	0.0023
P2	17 (19.1)	7 (53.8)	3 (30)	7 (10.6)	0.0011
P3	4 (4.5)	2 (15.4)	2 (20)	0	0.0235
P2 or P3	21 (23.6)	9 (69.2)	5 (50)	7 (10.6)	<0.0001
Others	16 (18.0)	1 (7.7)	2 (20)	13 (19.7)	

[a]Two-tailed *P-Value*s by Fisher´s exact test are shown where *P*<0.05.

Table S9. Distribution (%) of the 31 clones including at least three isolates into G1, G2 and G3.

Clone	Number (%) of isolates			*P-Value*[a]
	G1 (n=36)	G2 (n=28)	G3 (n=330)	G1 vs. G3
A-CH11-54-ST10	0	1 (3.6)	9 (2.7)	
A-CH11-0-ST93	0	0	4 (1.2)	
A-CH11-0-ST167	0	1 (3.6)	2 (0.6)	
A-CH99-54-ST361	0	0	5 (1.5)	
A-CH11-54-ST744	0	0	5 (1.5)	
B1-CH4-27-ST58	0	0	4 (1.2)	
B1-CH6-31-ST453	0	0	3 (0.9)	
B2-CH13-106-ST12	1 (2.8)	1 (3.6)	2 (0.6)	
B2-CH24-10-ST73	1 (2.8)	2 (7.1)	3 (0.9)	
B2-CH24-103-ST73	0	1 (3.6)	5 (1.5)	
B2-CH24-12-ST73	0	1 (3.6)	2 (0.6)	
B2-CH24-30-ST73	1 (2.8)	1 (3.6)	2 (0.6)	
B2-CH38-15-ST95	0	0	6 (1.8)	
B2-CH14-2-ST127	2 (5.6)	1 (3.6)	1 (0.3)	0.0265
B2-CH40-22-ST131	3 (8.3)	0	1 (0.3)	0.0033
B2-CH40-30-ST131	0	3 (10.7)	84 (25.5)	<0.0001
B2-CH40-41-ST131	0	0	10 (3)	
B2-CH52-5-ST141	6 (16.7)	2 (7.1)	5 (1.5)	0.0002
B2-CH14-27-ST404	0	0	3 (0.9)	
B2-CH14-64-ST1193	0	0	4 (1.2)	
C-CH4-35-ST23	0	0	3 (0.9)	
C-CH4-39-ST88	0	0	11 (3.3)	
C-CH4-24-ST410	0	0	7 (2.1)	
C-CH263-32-ST1615	0	0	6 (1.8)	
D-CH35-27-ST69	0	3 (10.7)	11 (3.3)	
E-CH26-0-ST38	0	0	3 (0.9)	
E-CH100-96-ST362	2 (5.6)	0	1 (0.3)	0.0265
E-CH37-27-ST405	0	0	4 (1.2)	
F-CH32-41-ST59	0	0	4 (1.2)	
F-CH45-97-ST117	0	0	3 (0.9)	
F-CH88-58-ST354	0	0	5 (1.5)	

[a] Two-tailed *P-Value*s by Fisher´s exact test are shown where $P<0.05$.

Table S10. Distribution of the 26 G1-detected clones among G1, G2 and G3 isolates

Clone	Number (%) of isolates			*P-Value*[a]	
	G1 (n=36)	G2 (n=28)	G3 (n=330)	G1 vs. G2	G1 vs. G3
B1-CH23-54-ST205	1 (2.8)	0	0		
B1-CH6-31-ST1196	1 (2.8)	0	0		
B1-CH4-31-ST2766	1 (2.8)	0	0		
B2-CH13-106-ST12	1 (2.8)	1 (3.6)	2 (0.6)		
B2-CH13-223-ST12	1 (2.8)	0	0		
B2-CH24-10-ST73	1 (2.8)	2 (7.1)	3		
B2-CH24-13-ST73	1 (2.8)	0	0		
B2-CH24-30-ST73	1 (2.8)	1 (3.6)	2 (0.6)		
B2-CH24-32-ST73	1 (2.8)	0	0		
B2-CH24-2-ST104	1 (2.8)	0	0		
B2-CH14-2-ST127	2 (5.6)	1 (3.6)	1 (0.3)		0.0265
B2-CH40-22-ST131	3 (8.3)	0	1 (0.3)		0.0033
B2-CH40-298-ST131	1 (2.8)	0	0		
B2-CH40-22-ST2556	1 (2.8)	0	0		
B2-CH52-5-ST141	6 (16.7)	2 (7.1)	5 (1.5)		0.0002
B2-CH52-14-ST141	2 (5.6)	0	0		0.0094
B2-CH52-76-ST141	1 (2.8)	0	0		
B2-CH38-5-ST569	1 (2.8)	0	0		
B2-CH108-75-ST636	1 (2.8)	0	0		
B2-CH24-12-ST1154	1 (2.8)	0	0		
B2-CH319-197-ST2015	1 (2.8)	0	0		
B2-CH43-0-ST2558	1 (2.8)	0	0		
E-CH100-96-ST362	2 (5.6)	0	1 (0.3)		0.0265
F-CH4-171-ST648	1 (2.8)	0	1 (0.3)		
F-CH4-58-ST648	1 (2.8)	0	1 (0.3)		
F-CH231-58-ST1485	1 (2.8)	0	0		
	36 (100)	7 (25)	17 (5.2)	<0.0001	<0.0001

[a] Two-tailed *P-Value*s by Fisher´s exact test are shown where $P<0.05$.

Table S12. Nucleotide and amino acid differences in 34 *fimB* alleles detected in 84 *Escherichia coli* clinical isolates.

Strain/FimB/accession n°	Nucleotide position																																							Amino acid position							
	39	57	72	80	93	103	114	117	156	198	237	246	249	256	285	294	297	315	319	324	330	339	345	384	396	402	485	495	498	519	537	540	556	570	573	579	580	#		27	35	86	107	162	180	186	195
A1 P1 (MK301552)	G	A	A	C	A	T	T	G	T	A	C	G	T	G	C	G	C	G	G	T	A	G	G	C	C	G	A	T	G	T	C	T	G	A	C	A	C	C		T	Y	V	V	Q	N	G	Q
A22 P9 (MK301573)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	I	.	.	.	.
A7 P1 (MK301558)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	A	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A16 P5 (MK301567)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	A	.	.	G	A	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.		.	.	.	.	.	K	.	.
A4 P1 (MK301555)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A8 P1 (MK301559)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A9 P1 (MK301560)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A23 P10 (MK301574)	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		N	.	.	.	.	.	.	.
A5 P1 (MK301556)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A15 P5 (MK301566)	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.		.	.	.	.	.	K	.	.
A17 P6 (MK301568)	.	.	.	.	.	C	.	.	.	.	T	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	H	.	.	.	.	.	.
A11 P1 (MK301562)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A10 P1 (MK301561)	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A2 P1 (MK301553)	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A19 P7 (MK301570)	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	.		.	.	.	.	.	.	S	.
A3 P1 (MK301554)	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.		.	.	.	.	.	.	.	.
A18 P6 (MK301569)	.	.	.	.	.	C	.	.	.	.	.	.	C	.	.	T	T	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	H	.	.	.	.	.	.
A6 P1 (MK301557)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	G	A	.		.	.	.	.	.	.	.	.
A12 P1 (MK301563)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
A20 P7 (MK301571)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	.	.	.	.		.	.	.	.	.	.	S	.
A21 P8 (MK301572)	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	T	T	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	R	.	.	.
A31 P2 (MK301582)	T	.	.	.	T	.	.	.	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A24 P2 (MK301575)	.	.	T	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A33 P4 (MK301584)	.	.	T	.	T	C	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	H	I	.	.	.	.	.
A34 P4 (MK301585)	.	.	.	.	T	C	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	H	I	.	.	.	.	.
A32 P3 (MK301583)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	G		.	.	I	.	.	.	.	E
A28 P2 (MK301579)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	C	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A30 P2 (MK301581)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	C	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A27 P2 (MK301578)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A29 P2 (MK301580)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A26 P2 (MK301577)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A25 P2 (MK301576)	.	.	.	.	T	.	A	A	C	.	.	.	C	A	.	T	T	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	I	.	.	.	.	.
A14 P1 (MK301565)	T	G	.	.	.	.	.	.	.	.	.	.	C	.	A	T	T	.	.	.	.	.	.	T	T	A	.	.	.	.	T	.	.	.	.	G	A	.		.	.	.	.	.	.	.	.
A13 P1 (MK301564)	.	G	.	.	.	.	.	.	.	.	.	.	C	.	.	.	T	.	.	.	.	.	.	T	T	A	.	.	.	A	.	.	.	.	T	G	A	.		.	.	.	.	.	.	.	.

8.2.4. STUDY 4

PLACNETw is a graph-based tool for reconstruction of plasmids. However, manual pruning of the graphs is a necessary requirement, and not all plasmids could always be successfully separate without assuming some risk of error. For this reason, some plasmids remain together and the plamidome remains presumptive.

The presumptive plamsidome has been described following structure: plasmid name (size), plasmid typing by relaxase (MOB) and/or Incompatibility group (Inc.), number of copies based on coverage (plasmid coverage/chromosome coverage) and the carried virulence and resistance genes. Transferability was presumptively deduced based on identification of relaxases (mobilizable plasmids) and mating pair formation (Mpf) system proteins (conjugative plasmids).

Plasmidome analysis of STC131 genomes (n=30)

Genome LREC263 (B2-CH40-30-ST131) includes five plasmids:
pLREC263_1+2 (126 kb). MOB_{F12}/IncF [F1:A6:B66] and a MOB_{P}/ColRNAI plasmid that could not be separated. Conjugative (Mpf presence).
pLREC263_3 (6 kb). ColRNAI (>10 copies/genome).
pLREC263_4 (4kb). MOB_{Q}/ColRNAI (7-10 copies/genome).
pLREC263_5 (4 kb). MOB_{Q12}/ColRNAI (2-6 copies/genome).

Genome LREC264 (B2-CH40-30-ST131) includes six plasmids:
pLREC264_1 (99 kb). MOB_{F12}/IncF [F1:A2:B20]. Harbouring resistance genes that confer resistance to beta-lactams ($bla_{CTX\text{-}M\text{-}27}$), aminoglycosides (*aad*A5, *aph*(3")-Ib, *aph*(6)-Id), sulfonamides (*sul1*, *sul2*), trimethoprim (*dfr*A17) and *tet*racycline (*tet(*A)).
pLREC264_2 (37 kb). MOB_{P3}/IncX4 with a ColRNAI presumably cointegrated (2-6 copies/genome). Conjugative (Mpf presence).
pLREC264_3 (16 Kb). MOB_{Q12} (2-6 copies/genome). Harbouring the virulence gen *sen*B.
pLREC264_4 (5 kb). ColRNAI (2-6 copies/genome).
pLREC264_5 (2 kb). Col(BS512) and Inc13 (2-6 copies/genome).
pLREC264_6 (2 kb). ColRNAI (2-6 copies/genome).

Genome LREC265 (B2-CH40-30-ST131) includes three plasmids:
pLREC265_1 (150 kb). MOB_{F}/IncF [F4:A2:B20] with a ColRNAI presumably cointegrated. Harbouring the virulence genes *sen*B and *iha*. Conjugative (Mpf presence).
pLREC265_2 (5 kb). MOB_{Q12}/ColRNAI (2-6 copies/genome).
pLREC265_3 (3 Kb). ColRNAI (2-6 copies/genome).

Genome LREC266 (B2-CH40-30-ST131) includes six plasmids:
pLREC266_1 (132 kb). MOB_{F12}/IncF [F1:A2:B20] with a ColRNAI presumably cointegrated. Harbouring the virulence gene *sen*B and resistance genes to beta-lactams ($bla_{CTX\text{-}M\text{-}27}$), aminoglycosides (*aph*(3")-Ib, *aph*(6)-Id) sulfonamides (*sul2*) and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).
pLREC266_2 (32 kb). MOB_{P3}/IncX4. Conjugative (Mpf presence).
pLREC266_3 (7 kb). MOB_{P51}/ColRNAI (2-6 copies/genome).
pLREC266_4 (6 kb). MOB_{P51}/ColRNAI (2-6 copies/genome).
pLREC266_5 (5 kb). MOB_{Q12}/ColRNAI (7-10 copies/genome).
pLREC266_6 (2 kb). ColRNAI (7-10 copies/genome).

Genome LREC268 (B2-CH40-30-ST131) includes six plasmids:
pLREC268_1 (125 kb). MOB_{F}/IncF [F2:A1:B-]. Harbouring resistance genes to quinolones and aminoglycosides (*acc*(6')-Ib-cr), macrolides (*mph*(A)), sulfonamides (*sul1*), trimethoprim (*dfr*A14) and phenicols (*cat*B3). Conjugative (Mpf presence).
pLREC268_2 (65 kb). MOB_{P}/IncK (IncB/O/K/Z). Conjugative (Mpf presence).
pLREC268_3 (36 kb). MOB_{P3}/IncX.
pLREC268_4 (19 kb). MOB_{F11}/IncN. Conjugative (Mpf presence).
pLREC268_5 (7 kb). MOB_{P51}/ColRNAI.
pLREC268_6 (4 kb). MOB_{Q}/ColRNAI (2-6 copies/genome).

Genome LREC269 (B2-CH40-30-ST131) includes five plasmids:
pLREC269_1 (87 kb). MOB_{P12}/ Incк (IncB/O/K/Z). Conjugative (Mpf presence).
pLREC269_2 (46 kb). IncF [F2:A1:B-] with a ColRNAI and an IncN presumably cointegrated. Conjugative (Mpf presence).
pLREC269_3 (5 kb). MOB_{Q12}/ColRNAI (2-6 copies/genome).
pLREC269_4 (2 kb). MOB_{V2} (2-6 copies/genome).
pLREC269_5 (2 kb). ColRNAI.

Genome LREC270 (B2-CH40-30-ST131) includes four plasmids:
pLREC270_1 (99 kb). IncF [F2:A1:B-]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$) macrolides (*mph*(A)), aminoglycosides (*aac*(6')-Ib-cr, *aad*A5), quinolones (*aac*(6')-Ib-cr), sulfonamides (*sul1*), trimethoprim (*dfr*A17), phenicols (*cat*B3) and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).
pLREC270_2 (4 kb). MOB_{Q}/ColRNAI (2-6 copies/genome). Conjugative (Mpf presence).
pLREC270_3 (2 kb). MOB_{V2} (2-6 copies/genome).
pLREC270_4 (2 kb). ColRNAI (>10 copies/genome).

Genome LREC271 (B2-CH40-30-ST131) includes four plasmids:
pLREC271_1 (95 kb). IncF [F2:A1:B-]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$), macrolides (*mph*(A)), aminoglycosides (*aac*(6')-Ib-cr, aadA5), quinolones (*aac*(6')-Ib-cr), sulfonamides (*sul1*), trimethoprim (*dfr*A17), phenicols (*cat*B3) and *tet*racycline (*tet*(A)).
pLREC271_2 (5 kb). MOB_{Q12}/ColRNAI (2-6 copies/genome).
pLREC271_3 (2 kb). MOB_{V2} (2-6 copies/genome).
pLREC271_4 (2 kb). ColRNAI (7-10 copies/genome).

Genome LREC272 (B2-CH40-30-ST131) includes four plasmids:
pLREC272_1 (211 kb). MOB_{F12}/IncF [F2-F10:A1:B-]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$), macrolides (*mph*(A)), aminoglycosides (*aac*(6')-Ib-cr, aadA5), quinolones (*aac*(6')-Ib-cr, *qnr*S1), sulfonamides (*sul1*), trimethoprim (*dfr*A14, *dfr*A17) and phenicols (*cat*B3).
pLREC272_2 (20 kb). MOB_{F11}/IncN.
pLREC272_3 (7 kb). MOB_{P51}/ColRNAI.
pLREC272_4 (4 kb). MOB_{Qu}/ColRNAI

Genome LREC273 (B2-CH40-30-ST131) includes four plasmids:
pLREC273_1 (109 kb). MOB_{F}/IncF [F2:A1:B-]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$), quinolones-aminoglycosides (*aac*(6')-Ib-cr), phenicols (*cat*B3) and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).
pLREC273_2 (36 kb). MOB_{P3}/IncX4. Conjugative (Mpf presence).
pLREC273_3 (4 kb). MOB_{Qu}/ColRNAI (2-6 copies/genome).
pLREC273_4 (2 kb). Col(MG828) (2-6 copies/genome).

Genome LREC274 (B2-CH40-30-ST131) includes four plasmids:
pLREC274_1 (74 kb). IncF [F-:A1:B-]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$), aminoglycosides-quinolones (*aac*(6')-Ib-cr), phenicols (*cat*B3) and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).
pLREC274_2 (36 kb). MOB_{P3}/IncX4. Conjugative (Mpf presence).
pLREC274_3 (4 kb). MOB_{Qu}/ColRNAI (2-6 copies/genome).
pLREC274_4 (2 kb). Col(MG828) (2-6 copies/genome).

Genome LREC275 (B2-CH40-30-ST131) includes two plasmids:
pLREC275_1 (152 kb). MOB_{F12}/IncF [F22:A1:B1]. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$), sulfonamides (*sul1*), trimethoprim (*dfr*A7) and *tet*racycline (*tet*(B)). Conjugative (Mpf presence).
pLREC275_2 (2kb). ColRNAI (7-10 copies/genome).

Genome LREC276 (B2-CH40-30-ST131) includes one plasmid:
pLREC276_1 (220 kb). MOB_{F}/IncF [F36:A4:B1] with a ColRNAI presumably cointegrated. Harbouring resistance genes to beta-lactams (*bla*$_{CTX-M-15}$), macrolides (*mph*(A)), aminoglycosides (aadA5), sulfonamides (*sul1*) and trimethoprim (*dfr*A17). Conjugative (Mpf presence).

Genome LREC277 (B2-CH40-30-ST131) includes one plasmid:
pLREC277_1 (92 kb). MOB_{F12}/IncF [F2:A2:B-]. Harbouring resistance genes to beta-lactams (bla_{TEM-1B}), aminoglycosides (*aph*(3")-Ib, *aph*(6)-Id), sulfonamides (*sul2*) and trimethoprim (*dfr*A17, *dfr*A8). Conjugative (Mpf presence).

Genome LREC278 (B2-CH40-30-ST131) includes one plasmid:
pLREC278_1 (616 kb). MOB_F/IncF [F36:A4:B-]. Harbouring resistance genes to beta-lactams ($bla_{CTX-M-15}$, bla_{OXA-1}), aminoglycosides and quinolones (*aac*(6')-Ib-cr) and phenicols (*cat*B3). Conjugative (Mpf presence).

Genome LREC279 (B2-CH40-30-ST131) includes two plasmids:
pLREC279_1 (87 kb). MOB_F/IncF [F10:A-:B-]. Conjugative (Mpf presence).
pLREC279_2 (49 kb). MOB_F/IncN-ST9. Conjugative (Mpf presence).

Genome LREC280 (B2-CH40-30-ST131) includes two plasmids:
pLREC280_1 (145 kb). MOB_F/IncF [F31:A4:B1] with a ColRNAI presumably cointegrated. Harbouring resistance genes to beta-lactams ($bla_{CTX-M-15}$), aminoglycosides (aadA5), sulfonamides (*sul1*), macrolides (*mph*(A)), and trimethoprim (*dfr*A17). Conjugative (Mpf presence).
pLREC280_2 (2 kb). ColRNAI and Inc13 (2-6 copies/genome).

Genome LREC281 (B2-CH40-30-ST131) includes three plasmids:
pLREC281_1 (374 kb). MOB_{F12}/IncF [F36:A4:B1]. Harbouring the virulence genes *sat* and *iha* and resistance genes to beta-lactams ($bla_{CTX-M-15}$, bla_{OXA-1}), aminoglycosides and quinolones (*aac*(6')-Ib-cr), phenicols (*cat*B3) and tetracycline (*tet* (A)). Conjugative (Mpf presence).
pLREC281_2 (4 kb). MOB_Q.
pLREC281_3 (2 kb). Col(BS512) and Inc13 (2-6 copies/genome).

Genome LREC282 (B2-CH40-30-ST131) includes four plasmids:
pLREC282_1+2 (246 kb). MOB_{F12}/IncF [F31-F36:A4:B1] and a MOB_{P51}/ColRNAI that could not be separated. Harbouring the virulence gene *iss* and resistance genes to beta-lactams ($bla_{CTX-M-15}$, bla_{OXA-1}), aminoglycosides and quinolones (*aac*(6')-Ib-cr), phenicols (*cat*B3) and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).
pLREC282_3 (4 kb). MOB_{Qu}
pLREC282_4 (2 kb). Col(BS512) and Inc13 (2-6 copies/genome).

Genome LREC283 (B2-CH40-30-ST131) did not present any plasmids.

Genome LREC284 (B2-CH40-30-ST131) includes three plasmids:
pLREC284_1 (68 kb). IncF [F2:A1:B-]. Harboring resistance genes to beta-lactams ($bla_{CTX-M-15}$), macrolides (*mph*(A)), aminoglycosides (*aad*A5), sulfonamides (*sul1*) and trimethoprim (*dfr*A17).
pLREC284_2 (34 kb). MOB_{P3}/IncX1.
pLREC284_3 (2kb) ColRNAI (>10 copies/genome).

Genome LREC285 (B2-CH40-30-ST131) includes two plasmids:
pLREC285_1 (137 kb). MOB_F/IncF [F1-F35:A-:B41]. Harbouring resistance genes to aminoglycosides (*aad*A5), sulfonamides (*sul2*, *sul1*) and trimethoprim (*dfr*A17). Conjugative (Mpf presence).
pLREC285_2 (4 kb). MOB_{Qu} (2-6 copies/genome).

Genome LREC286 (B2-CH40-30-ST131) includes two plasmids:
pLREC286_1+2 (113 kb). MOB_F/IncF [F10:A6:B20] and a MOB_P/ColRNAI that could not be separated. Harbouring resistance genes to beta-lactams (bla_{TEM-1B}), aminoglycosides (*aad*A5, *aac*(3)-IId, *aph*(6)-Id), trimethoprim (*dfr*A17) and sulfoamidas (*sul2*). Conjugative (Mpf presence).

Genome LREC287 (B2-CH40-35-ST131) includes three plasmids:
pLREC287_1 (118 kb). MOB_F/IncF [F2:A1:B-]. Harbouring resistance genes to beta-lactams ($bla_{CTX-M-15}$, bla_{OXA-1}), aminoglycosides (*aac*(6')-Ib-cr, *aad*A5), quinolones (*aac*(6')-Ib-cr), sulfonamides (*sul1*), trimethoprim (*dfr*A17) and phenicols (*cat*B3). Conjugative (Mpf presence).
pLREC287_2 (5 kb). MOB_{Q12}/ColRNAI (2-6 copies/genome).
pLREC287_3 (2 kb). ColRNAI (2-6 copies/genome).

Genome LREC288 (B2-CH40-41-ST131) includes a MOB_H in the chromosome (ICE) and two plasmids:
pLREC288_1 (294 kb). MOB_F/IncF [F-:A-:B10] with a ColRNAI presumably cointegrated. Harbouring the virulence gen *sen*B and resistance genes to beta-lactams (bla_{TEM-1B}), aminoglycosides (*aad*A5, *aph*(3'')-Ib , *aph*(6)-Id), sulfonamides (*sul1*, *sul2*) and trimethoprim (*dfr*A17). Conjugative (Mpf presence).
pLREC288_2 (4 kb). MOB_Q/ColRNAI.

Genome LREC289 (B2-CH40-41-ST131) includes a MOB_H in the chromosome (ICE) and two plasmids:
pLREC289_1 (550 kb). MOB_{F12}/IncF [F29:A-:B10]. Harbouring the virulence gen *iha* and resistance genes to beta-lactams ($bla_{CTX-M-1}$) and macrolides (*mph*(A)). Conjugative (Mpf presence).
pLREC289_2 (109 kb). IncFIB (H89-PhagePlasmid).

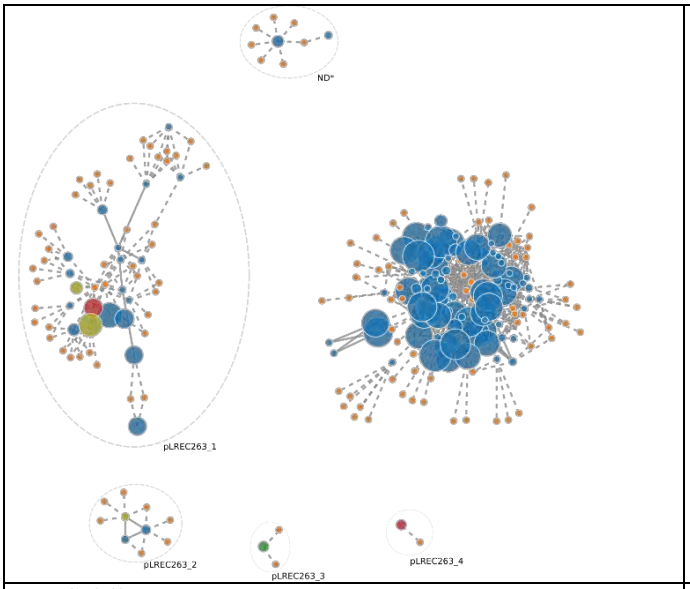
Genome LREC290 (B2-CH40-41-ST131) includes a MOB_Q in the chromosome (ICE) and two plasmids:
pLREC290_1 (143 kb). MOB_F.
pLREC290_2 (121 kb). MOB_{F12}/IncF [F29:A-:B10] with a ColRNAI presumably cointegrated. Harbouring resistance genes to beta-lactams (bla_{TEM-1B}) macrolides (*mph*(A)), aminoglycosides (*aac*(3)-IId, *aad*A5, *aph*(3")-Ib, *aph*(6)-Id), sulfonamides (*sul1*, *sul2*), trimethoprim (*dfr*A17) and tertaciclinas (*tet*(A)). Conjugative (Mpf presence).

Genome LREC291 (B2-CH40-41-ST131) includes a MOB_H in the chromosome (ICE) and one plasmid:
pLREC291_1 (150 kb). MOB_{F12}/IncF [F29:A-:B10] with a ColRNAI presumably cointegrated. Harbouring the virulence gen *sen*B and resistance genes to beta-lactams (bla_{TEM-1B}), macrolides (*mph*(A)), aminoglycosides (*aad*A5, *aph*(3")-Ib, *aph*(6)-Id), sulfonamides (*sul1*, *sul2*), trimethoprim (*dfr*A17)and *tet*racycline (*tet*(A)). Conjugative (Mpf presence).

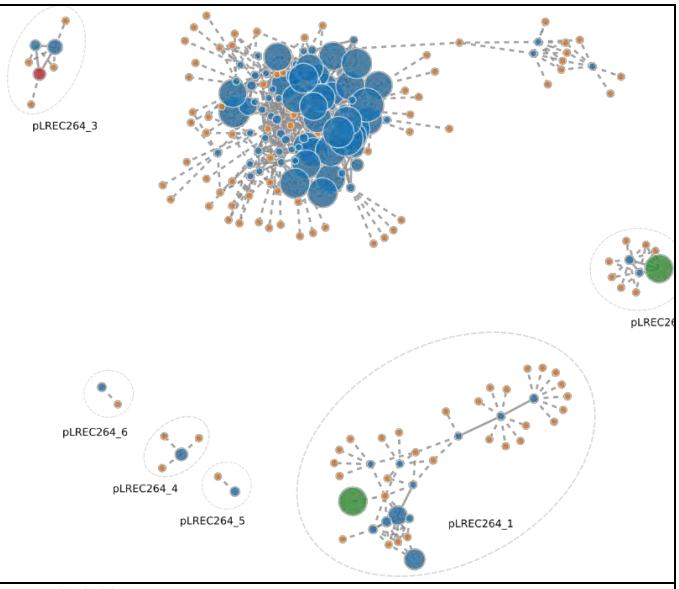
Genome LREC292 (B2-CH40-41-ST131) includes a MOB_H in the chromosome (ICE) and one plasmid:
pLREC292_1 (152 kb). MOB_F/IncF [F29:A-:B10] with a ColRNAI presumably cointegrated. Harbouring the virulence gen *sen*B and resistance genes to beta-lactams (bla_{TEM-1B}), aminoglycosides (*aad*A5, *aph*(3")-Ib, *aph*(6)-Id), sulfonamides (*sul1*), trimethoprim (*dfr*A17), *tet*racycline (*tet*(A)) and macrolides (*mph*(A)). Conjugative (Mpf presence).

Genome LREC293 (B2-CH40-5-ST2279) includes one plasmid:
pLREC293_1 (172 kb). MOB_{F12}/IncF [F31:A4:B1] with a ColRNAI presumably cointegrated. Harbouring the virulence genes *sen*B and *iha* and resistance genes to *tet*racycline (*tet*(A)). Conjugative (Mpf presence).

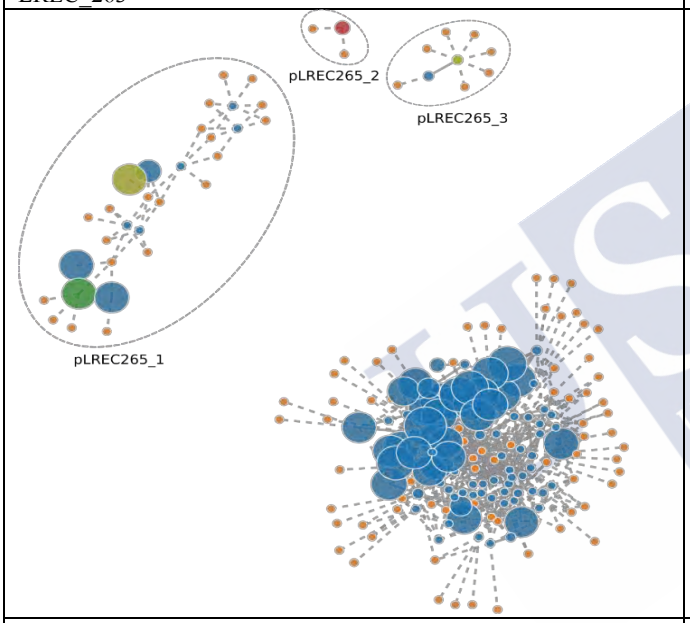
PLACNETw reconstruction of the 30 ESBLEC STC131 group of genomes.



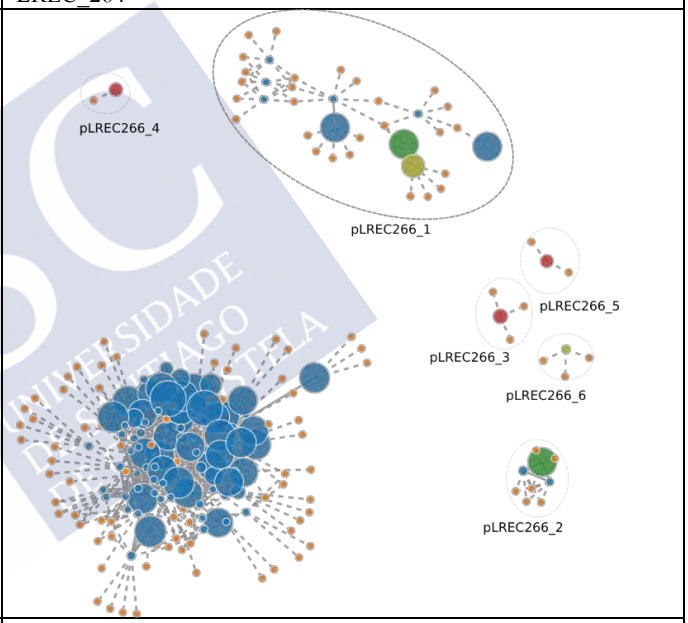
LREC_263



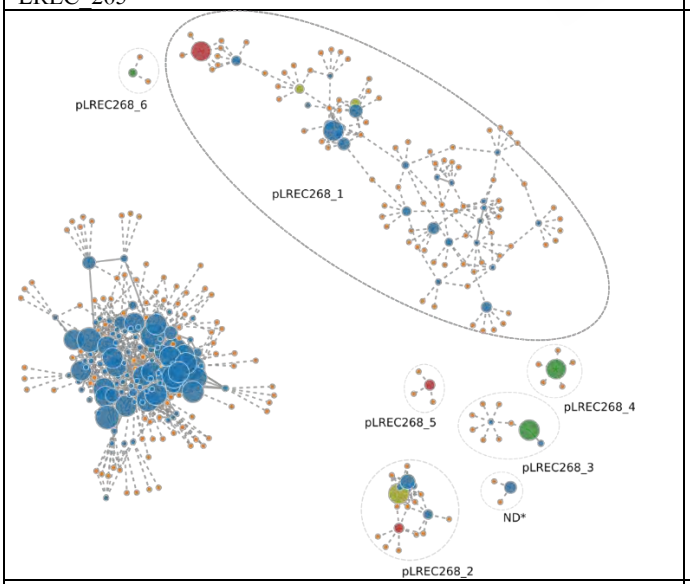
LREC_264



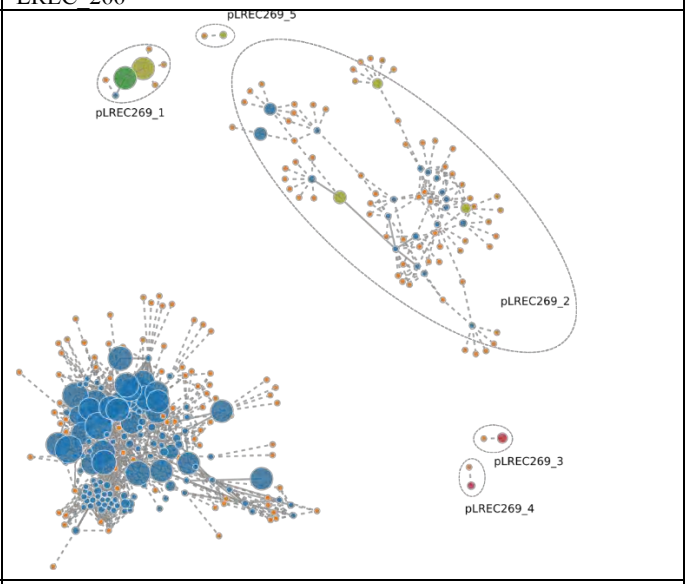
LREC_265



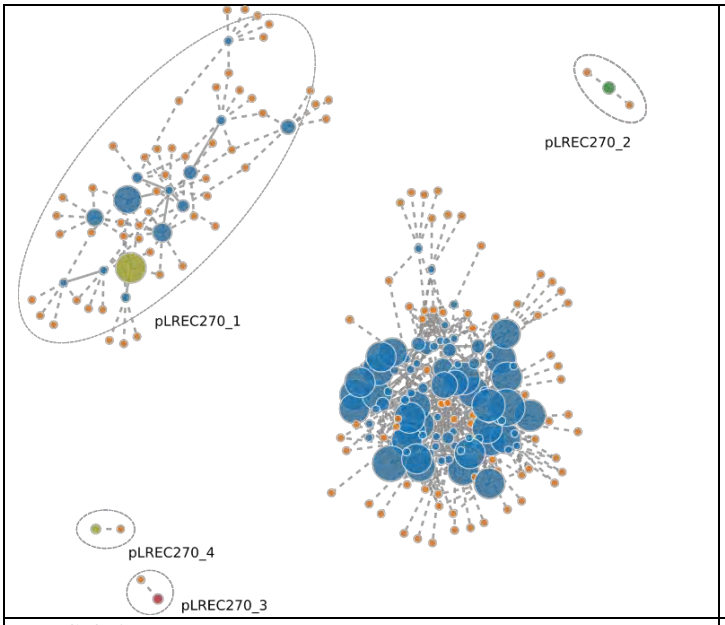
LREC_266



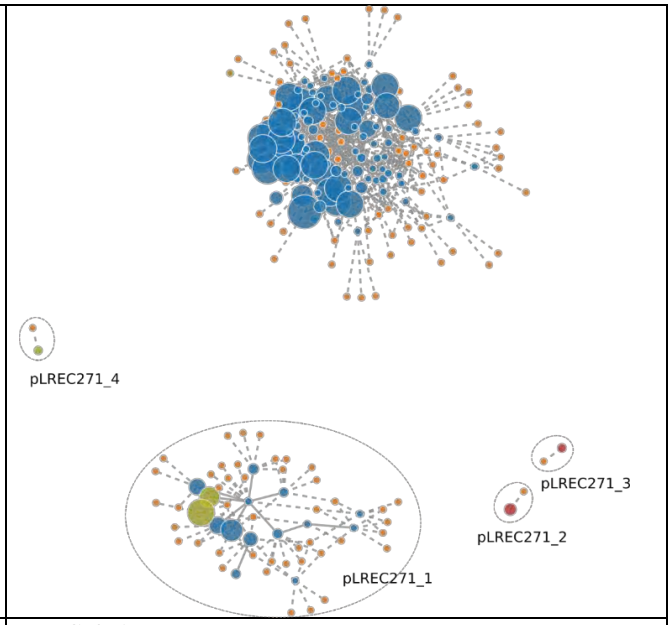
LREC_268



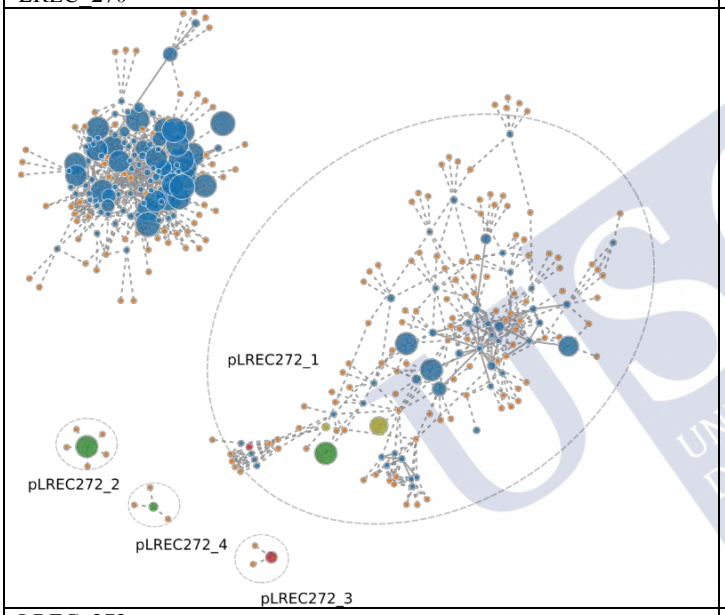
LREC_269



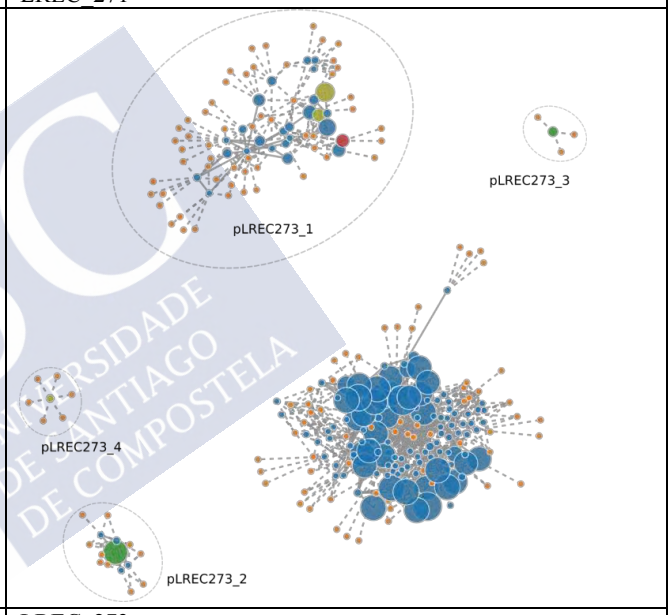
LREC_270



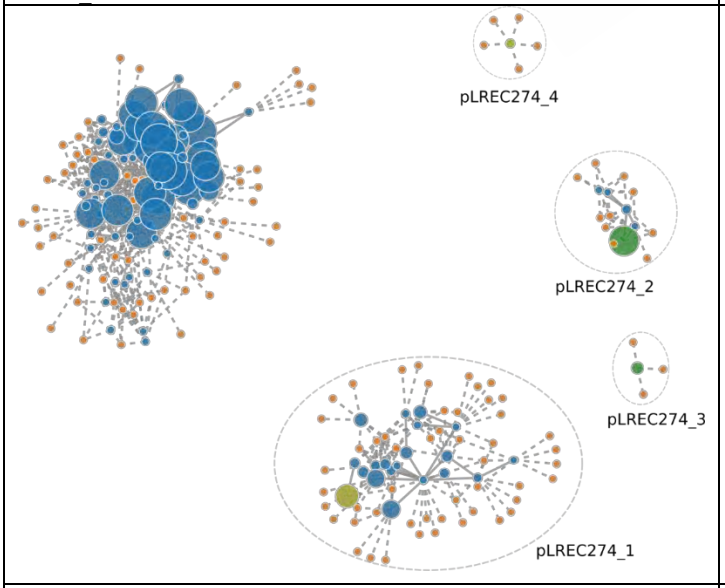
LREC_271



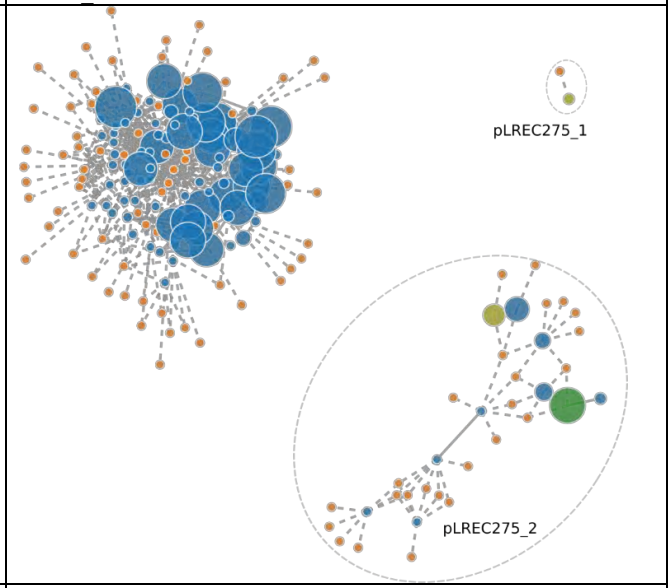
LREC_272



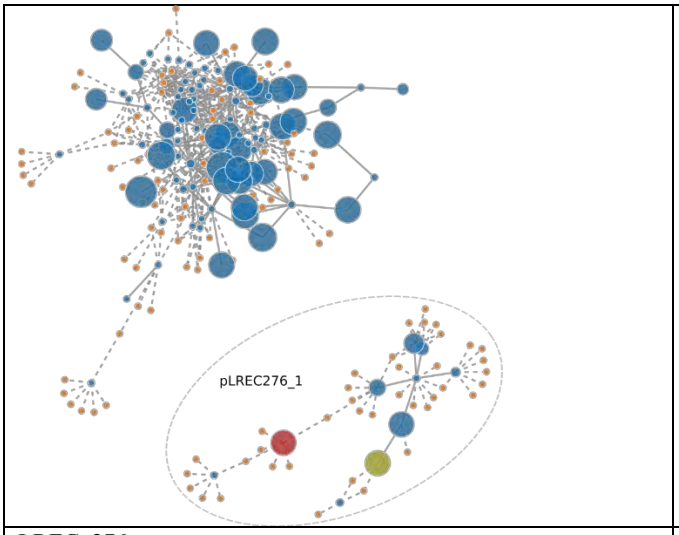
LREC_273



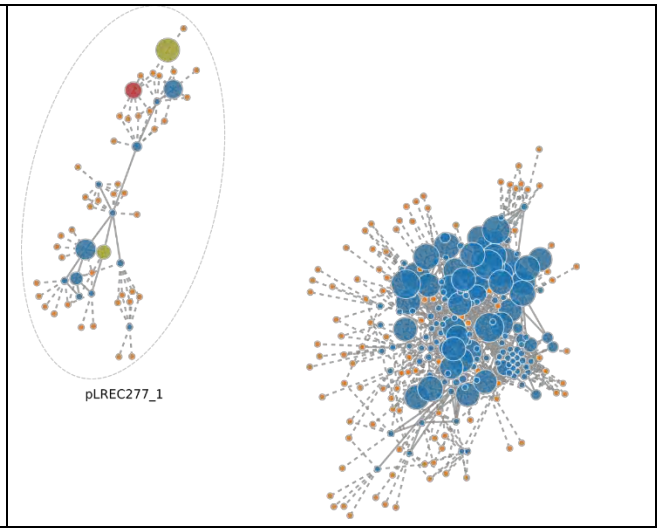
LREC_274



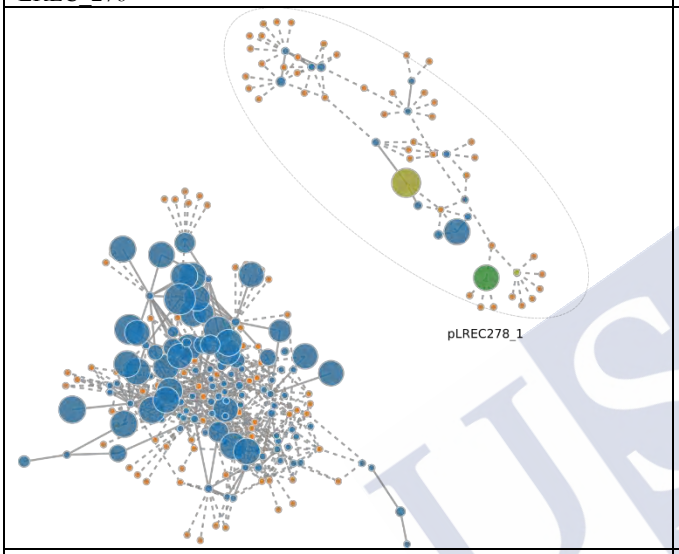
LREC_275



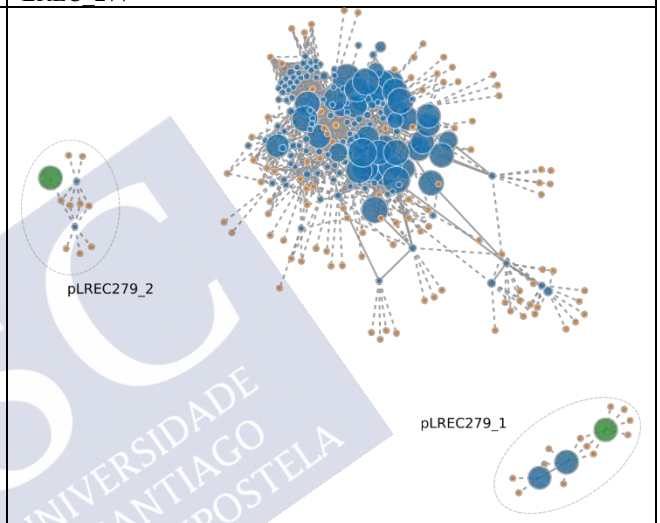
LREC_276



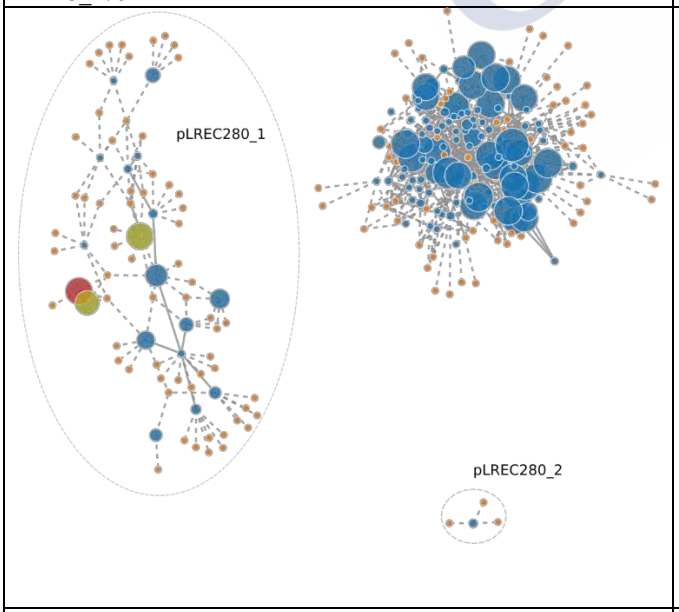
LREC_277



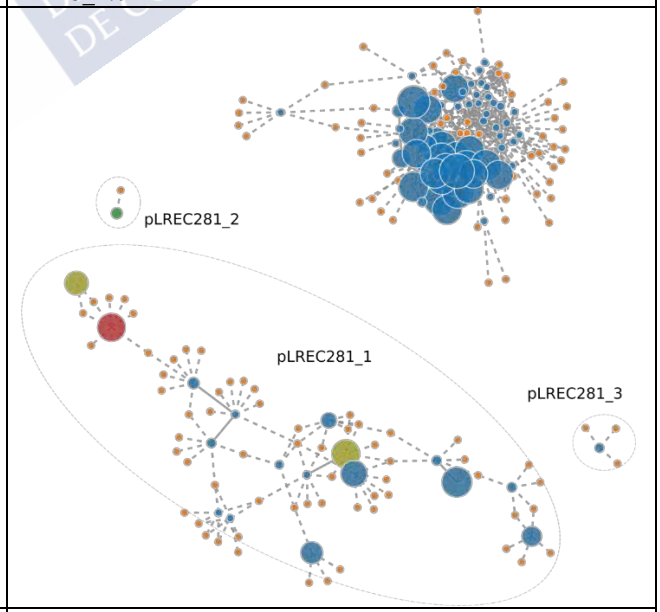
LREC_278



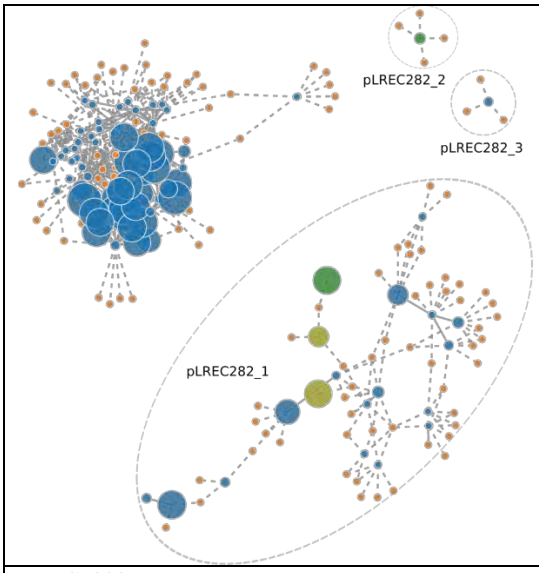
LREC_279



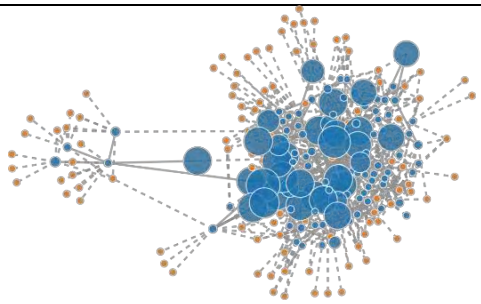
LREC_280



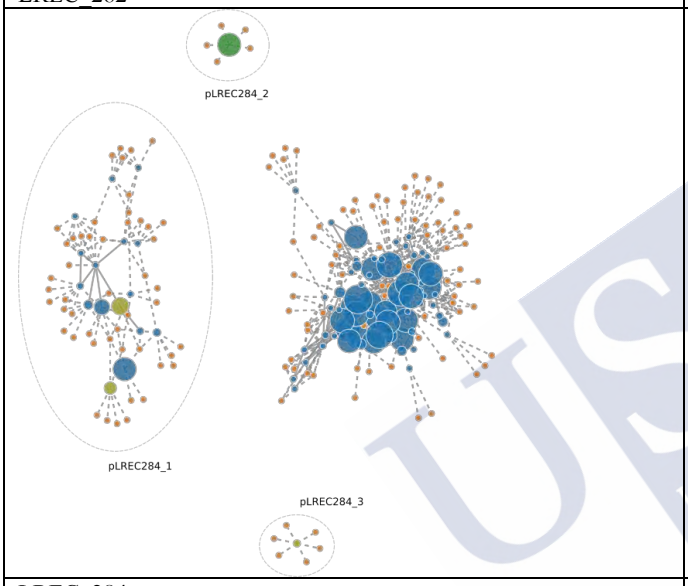
LREC_281



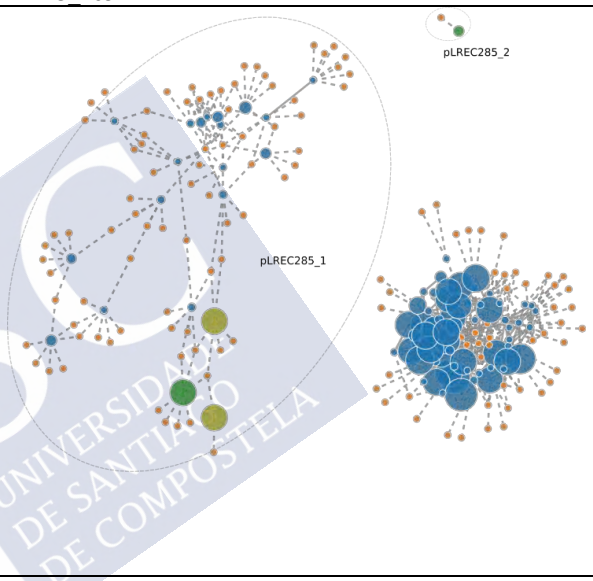
LREC_282



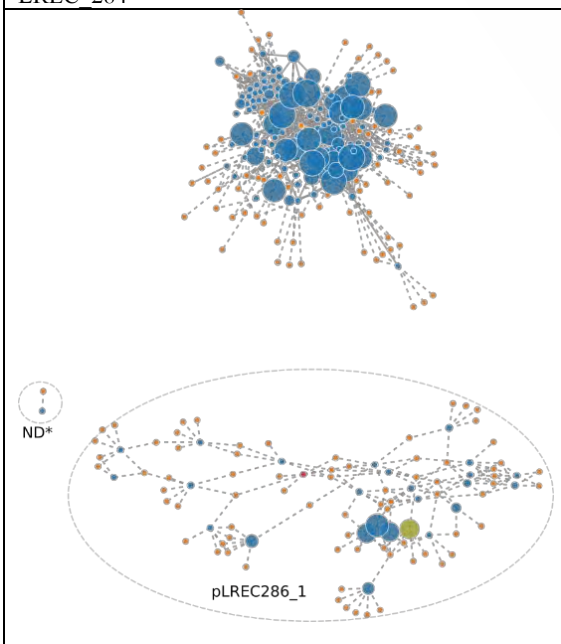
LREC_283



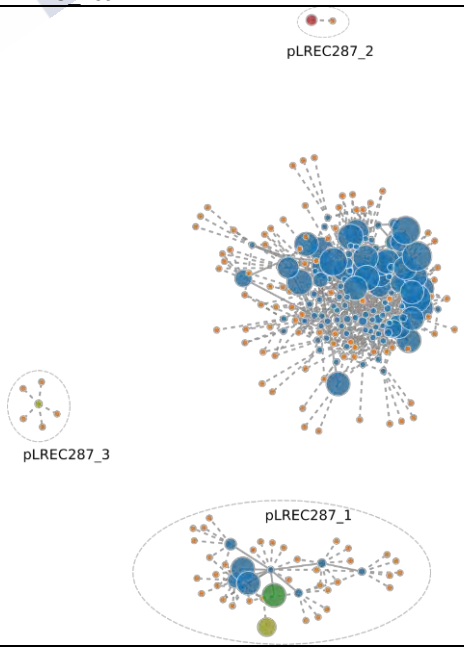
LREC_284



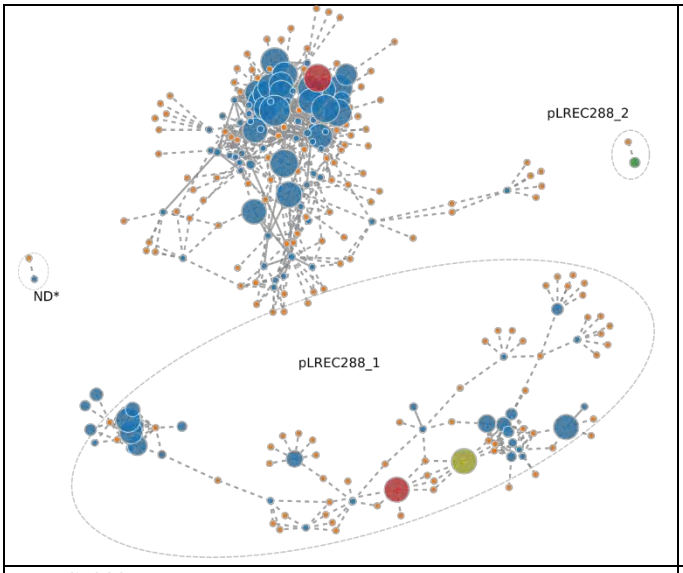
LREC_285



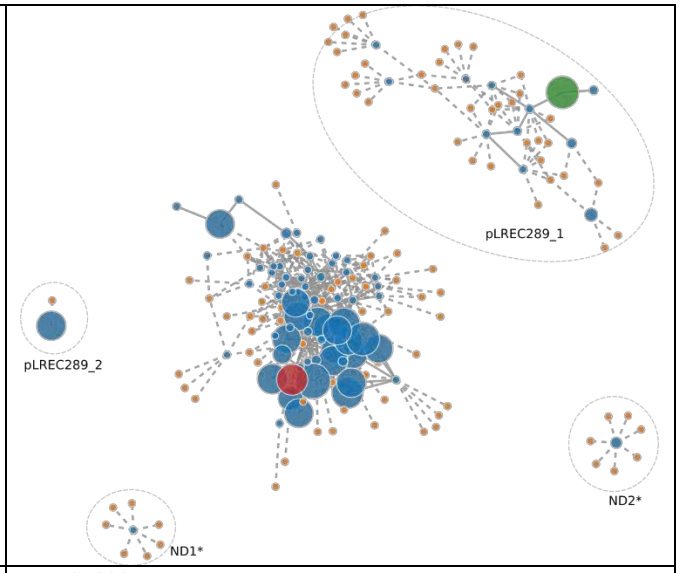
LREC_286



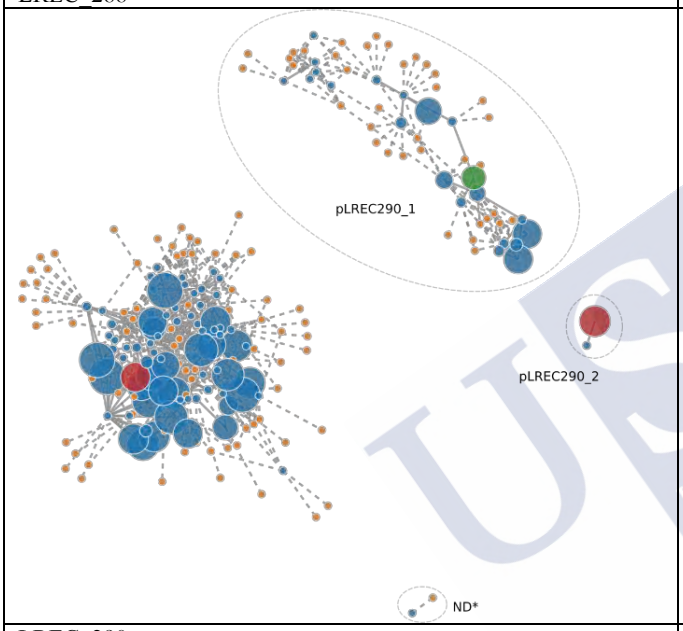
LREC_287



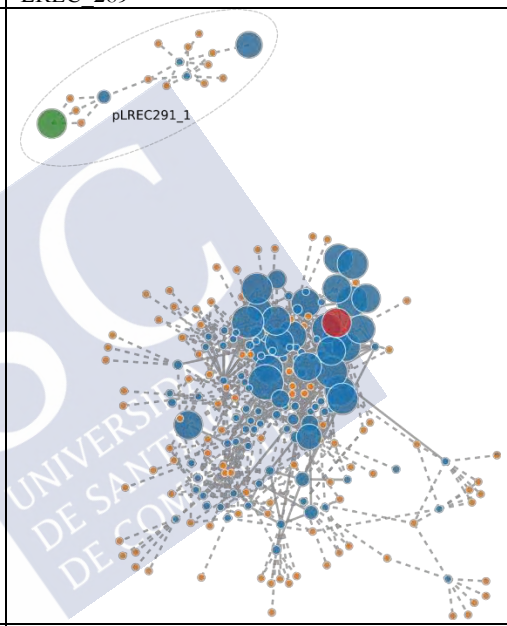
LREC_288



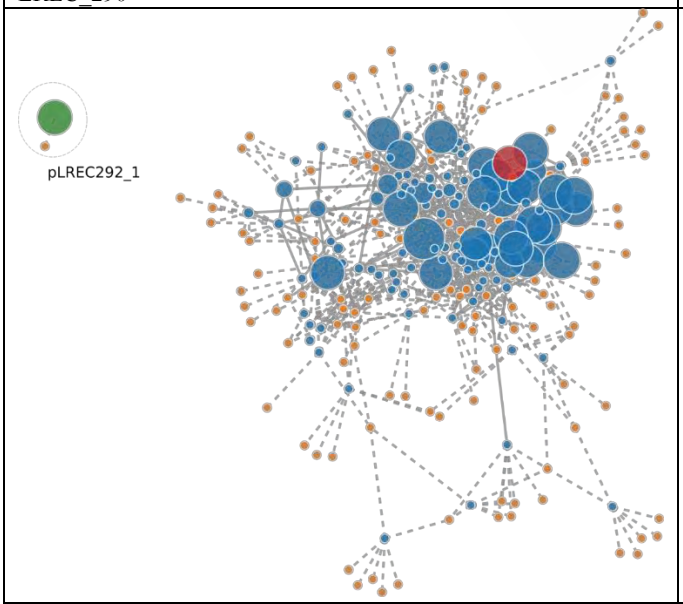
LREC_289



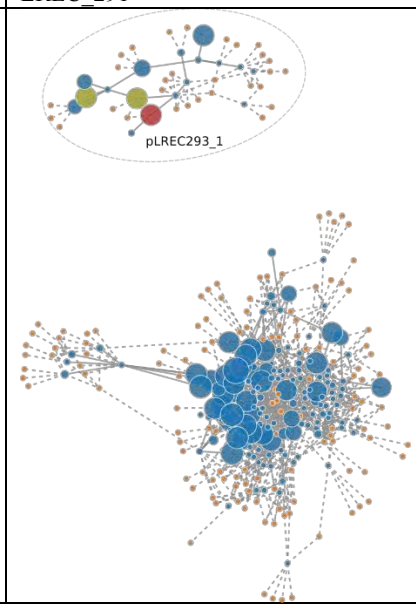
LREC_290



LREC_291



LREC_292



LREC_293

8.2.5. STUDY 5

Supplementary materials in Flament-Simon SC *et al.* 2020 Microorganims 8 (11): E1712 doi: 10.3390/microorganisms8111712 (Table S2: SNP matrix and VF-encoding genes of 197 ST372 genomes and Table S8: *In silico* determination of VF-encoding genes, antibiotic-resistance encoding genes (ARGs) and point mutations in 23 ST372 *E. coli* genomes).

Table S14. Prevalence of the phylogenetic groups in the 197 canine *E. coli* isolates.

Phylogenetic group	Number of isolates
A	32 (16.2%)
B1	26 (13.2%)
B2	84 (42.6%)
C	10 (5.1%)
D	6 (3.0%)
E	14 (7.1%)
F	18 (9.1%)
Clade V	1 (0.5%)
Not typeable	6 (3.0%)

Table S15. Comparison of the distribution of the phylogenetic groups among the 197 canine isolates according to the strain ExPEC and UPEC status.

Phylogenetic group	Number of isolates		P-value[1]
	ExPEC and/or UPEC (n= 91)	non-ExPEC and non-UPEC (n=106)	
A	2 (2.2%)	30 (28.3%)	<0.00001
B1	2 (2.2%)	24 (22.6%)	0.00001
B2	78 (85.7%)	6 (5.7%)	<0.00001
C	0	10 (9.4%)	0.00199
D	1 (1.1%)	5 (4.7%)	
E	1 (1.1%)	13 (12.3%)	0.00371
F	7 (7.7%)	11 (10.4%)	
Clade V	0	1 (0.9%)	
Not typeable	0	6 (5.7%)	0.03155

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

Table S16. Comparison of the distribution of the phylogenetic groups among the 197 canine isolates according to the strain multidrug resistant (MDR) status.

Phylogenetic group	Number of isolates		P-value[1]
	MDR (n= 28)	non-MDR (n=169)	
A	4 (14.3%)	28 (16.6%)	
B1	8 (28.6%)	18 (10.7%)	0.01597
B2	2 (7.1%)	82 (48.5%)	0.00004
C	1 (3.6%)	9 (5.3%)	
D	5 (17.9%)	1 (0.6%)	0.00022
E	5 (17.9%)	9 (5.3%)	0.03238
F	3 (10.7%)	15 (8.9%)	
Clade V	0	1 (0.6%)	
Not typeable	0	6 (3.6%)	

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

Table S17. Comparison of the strain ExPEC and UPEC status among canine multidrug resistant (MDR) and non-MDR isolates.

Status	Number of isolates (%)		P-value[1]
	MDR (n= 28)	non-MDR (n=169)	
ExPEC	8 (28.6%)	66 (39.1%)	
UPEC	4 (14.3%)	78 (46.2%)	0.00156
ExPEC and/or UPEC	8 (28.6%)	83 (49.1%)	
Non ExPEC and non-UPEC	20 (71.4%)	86 (50.9%)	

[1]Two-tailed *P-Value*s by Fisher´s exact probability test are shown where $P < 0.05$.

Table S18. New sequence types observed in 18 canine *E. coli* isolates.

ST New	Allele numbers						
	adk	*fum*C	*gyr*B	*icd*	*mdh*	*pur*A	rec**A**
1	10	11	4	666	560	8	2
2	13	13	9	13	16	[2]126-like	9
3	13	13	9	13	439	10	9
4	88	103	19	18	23	44	26
5	88	103	[1]19-like	36	23	44	26
6	6	11	4	8	235	8	2
7	13	40	9	13	16	126	9
8	13	363	302	97	23	94	93
9	13	24	9	912	17	11	25
10	36	24	10	912	17	10	25
11	9	23	33	628	11	8	6
12	13	40	459	912	23	28	109
13	13	40	13	912	23	25	66
14	21	35	27	6	286	5	4
15	6	4	4	16	474	8	14
16	6	4	5	18	11	319	6
17	6	29	33	402	11	7	2
18	63	30	3	628	43	350	6

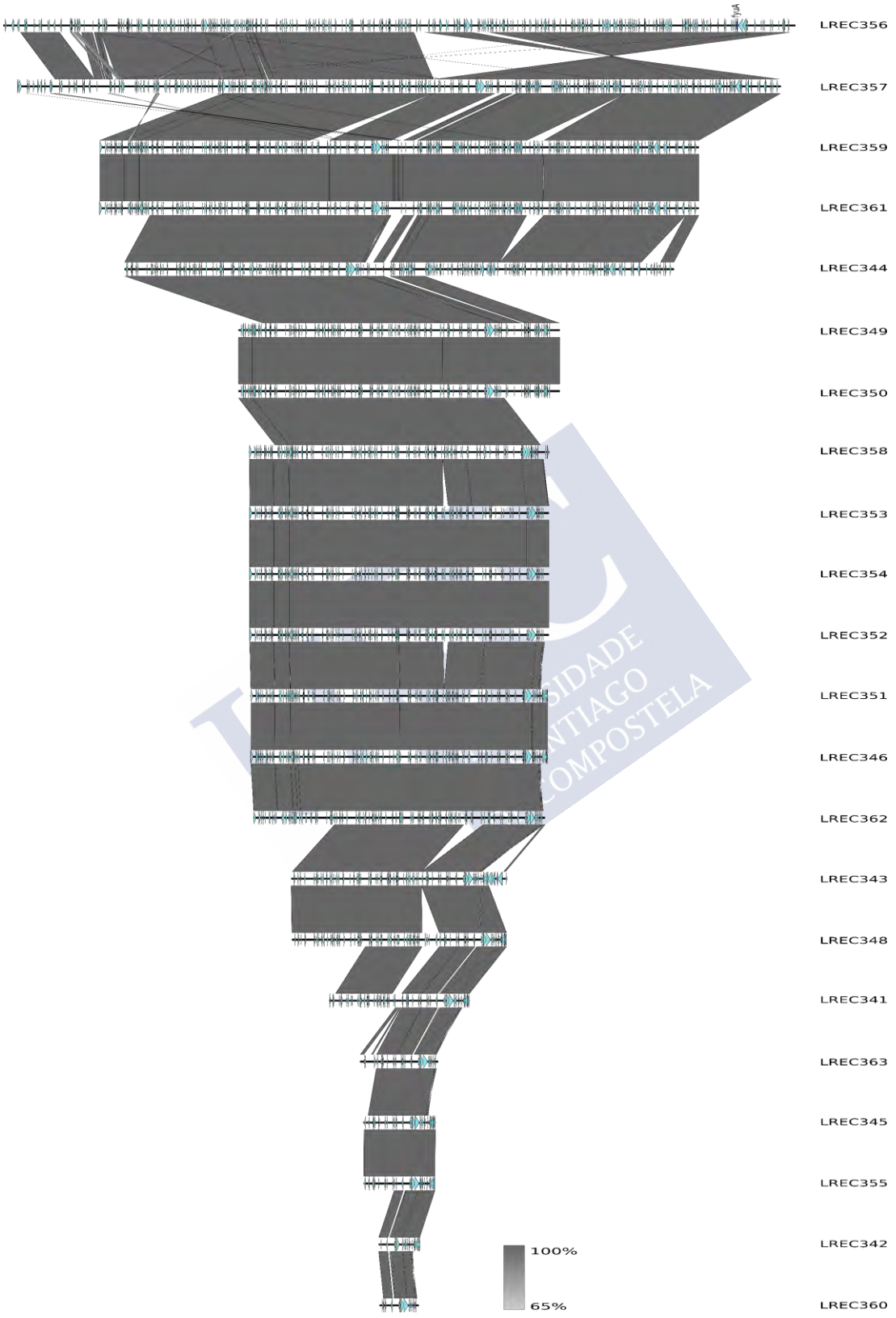
[1]Nucleotide sequence of *gyrB* 19-like, closest match: *Escherichia coli* Achtman MLST *gyrB*: 19 (1 difference found: [49]C→[49]T)

GGTCTGCACGGCGTTGGTGTTTCGGTAGTAAACGCCCTGTCGCAAAAATTGGAGCTGGTTATCCAGCGCGAGGGTAAAATTCACCGTCAGATCTACGAACA
CGGTGTACCGCAGGCCCCGCTGGCGGTTACCGGCGAGACTGAAAAAACCGGCACTATGGTGCGTTTCTGGCCAAGCCTTGAAACCTTCACCAATGTGACCG
AGTTCGAATATGACATTCTGGCGAAACGTCTGCGTGAGTTGTCGTTCCTCAACTCCGGCGTTTCCATTCGTCTGCGCGACAAGCGCGACGGCAAAGAAGAC
CACTTCCACTATGAAGGCGGCATCAAGGCGTTCGTTGAATATCTGAACAAGAACAAAACGCCGATCCACCCGAATATCTTCTACTTCTCCACCGAAAAAGA
CGGTATTGGCGTCGAAGTGGCGTTGCAGTGGAACGATGGCTTCCAGGAAAACATCT

[2]Nucleotide sequence of *purA* 126-like, closest match: *Escherichia coli* Achtman MLST *purA:* 126 (1 difference found: [118]A →[118]G)

ATAACGCGCGTGAGAAAGCGCGTGGCGCGAAAGCGATCGGCACCACCGGTCGAGGTATCGGGCCTGCTTATGAAGATAAAGTGGCACGTCGCGGTCTGCGT
GTTGGCGACCTTTTCGGCAAAGAAACCTTCGCTGAAAAACTGAAAGAAGTGATGGAATATCACAACTTCCAGTTGGTTAACTACTACAAAGCTGAAGCGGT
TGATTACCAGAAAGTTCTGGATGATACGATGGCTGTTGCCGACATCCTGACTTCTATGGTTGTTGACGTTTCTGATCTGCTCGACCAAGCGCGTCAGCGTG
GCGATTTCGTCATGTTTGAAGGTGCGCAGGGTACGCTGCTGGATATCGACCACGGCACTTATCCGTACGTAACTTCTTCCAACACTACTGCTGGTGGCGTG
GCGACCGGTTCCGGCCTGGGCCCGCGTTATGTTGATTACGTTTTGGGTATCCTCAAAGCTTACTCCACTCGTGT

Figure S3. Comparison of contigs harbouring integrative conjugative elements (ICEs) from 22 ST372 *E. coli* genomes.

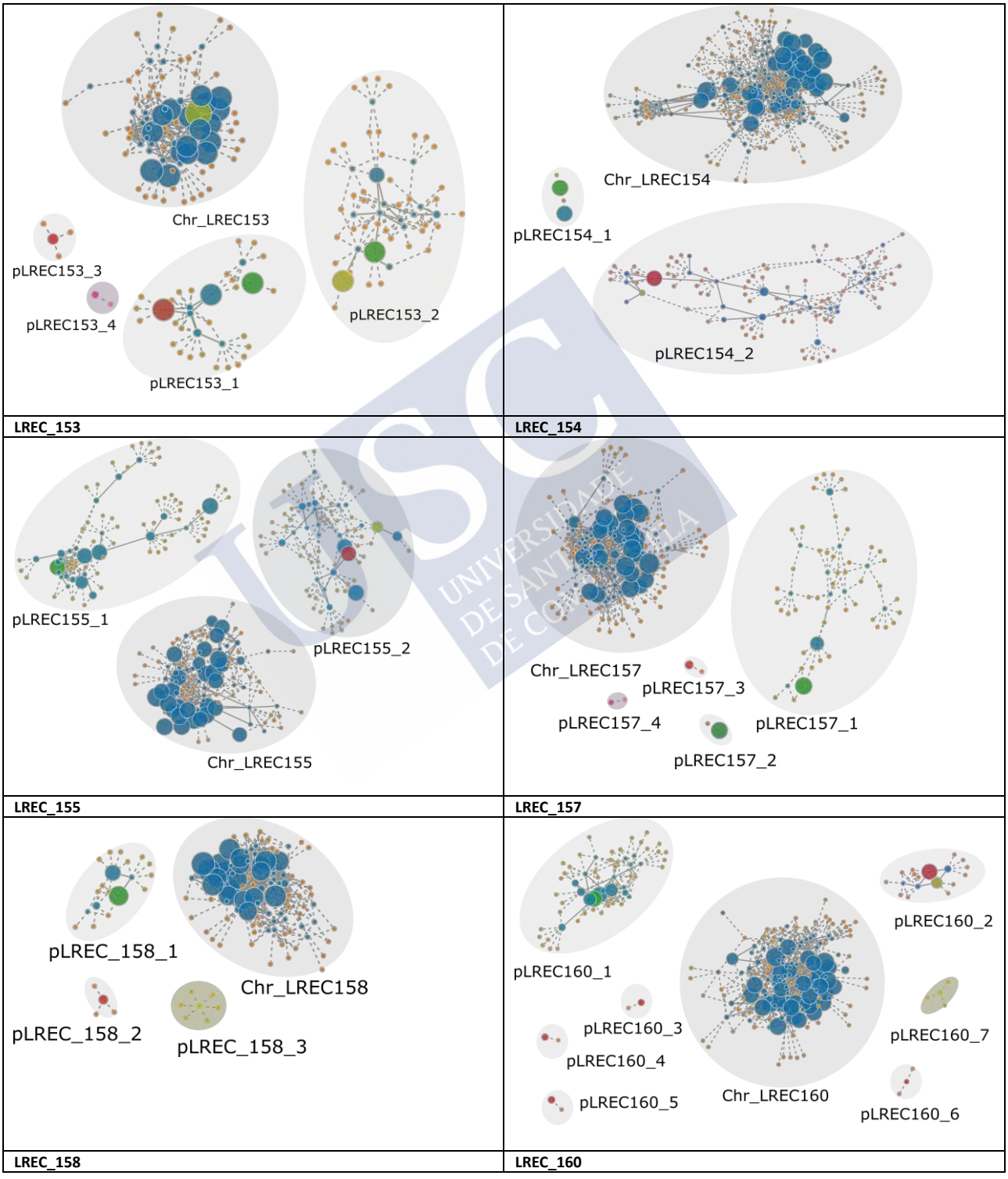


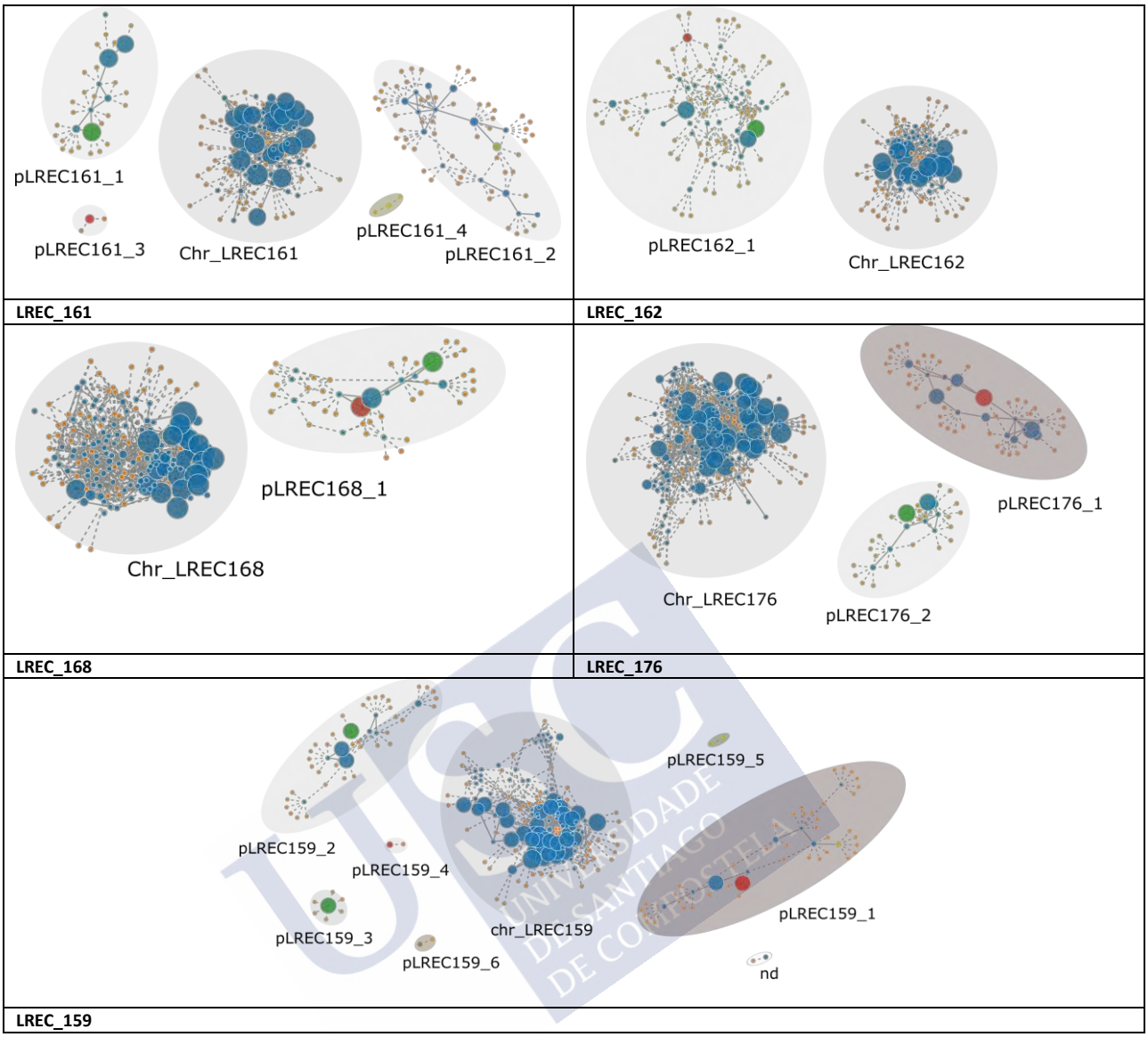
The Figure includes 22 complete contigs with a length ranging from 2,540,863 to 86,362 (pb). The shadow parallelograms denote genetic regions that exhibit sequence homology among different segments. Light shadow denotes regions with a lower level of sequence identity (99%) by BLAST. The image was generated using EasyFig (default parameters).

8.2.6. STUDY 6

Supplementary materials in Flament-Simon SC *et al.* 2020 Front Microbiol 2020 11:387 doi: 10.3389/fmicb.2020.00387 (Supplementary Table S1. Primers used for PCR amplification of virulence genes and virulence genes of 11 ST131 porcine isolates.and Supplementary Table S2**.** Core-genome analysis isolates, SNPs and references).

Figure S4. PLACNETw reconstruction of the 11 ST131 genomes from porcine origin.





The network contains nodes of different colours (blue for contigs, orange for references, yellow for replication proteins, red for relaxases and green for replication proteins and relaxases). Nodes are linked by scaffold links (solid grey lines). The size of the contig nodes is proportional to the contig length (Lanza *et al.*, 2014). *nd, is a 2,014 pb contig (transposases + aminoglycoside resistance protein + LinF) which could not be assigned due of lack of scaffold links to both chromosome and plasmids.

Figure S5. Structural comparison of IncF plasmids from the 11 ST131 clade B porcine isolates.

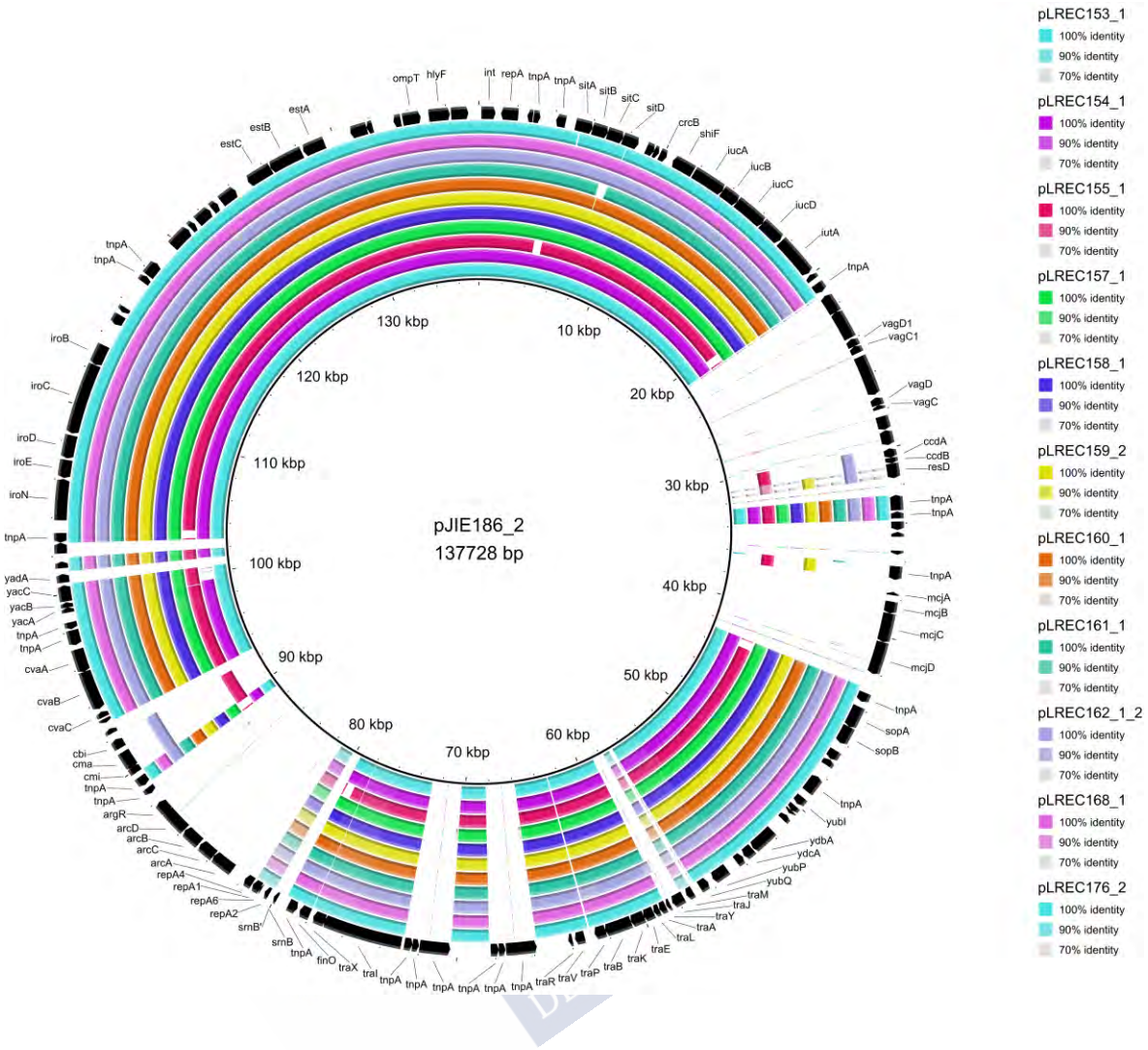


Figure S6. Structural comparison of ColE1-like plasmids from ST131 clade B porcine isolates.

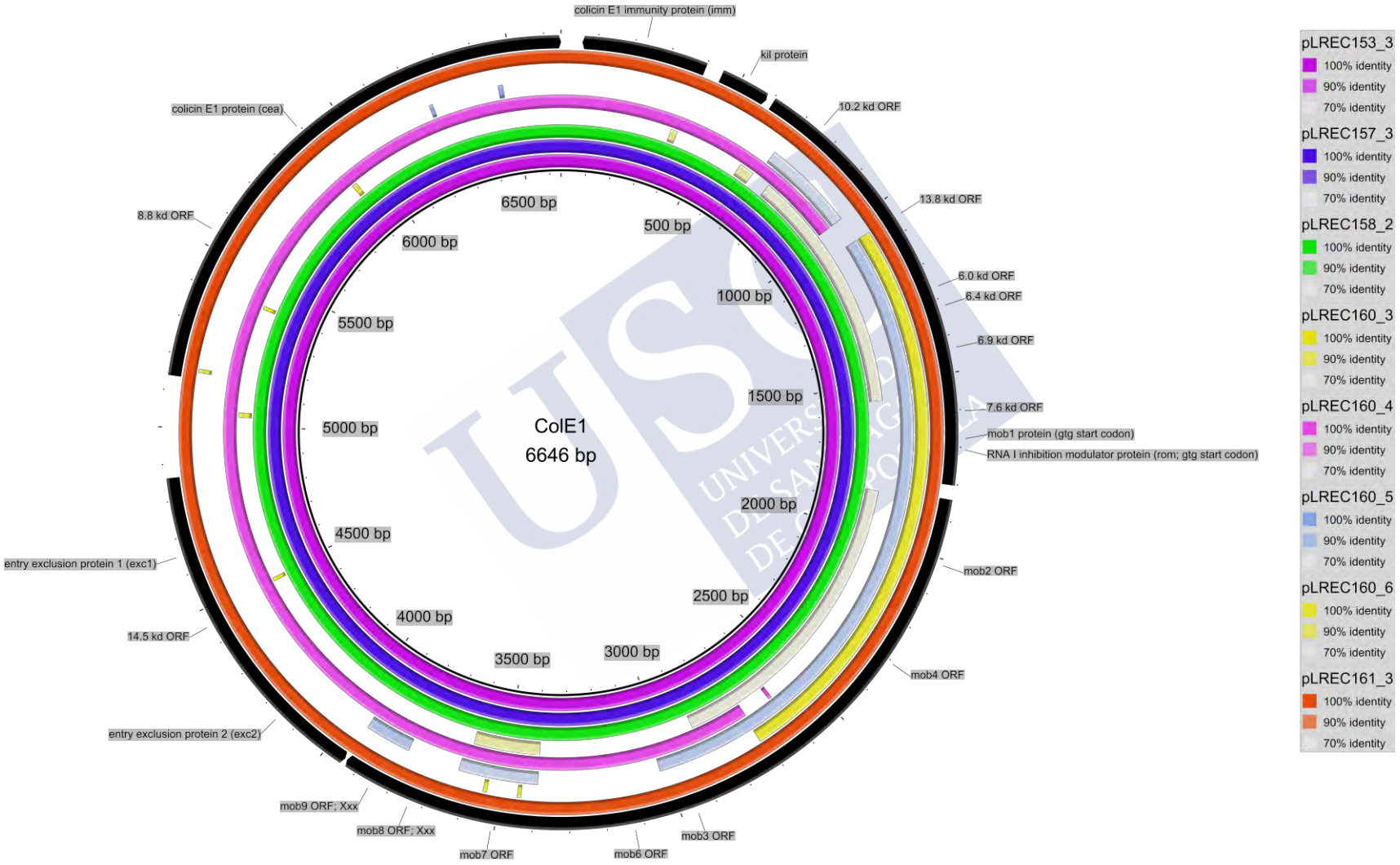


Figure S7. Structural comparison of an IncX1 plasmid from a ST131 clade B porcine isolate with pESCUM as a reference.

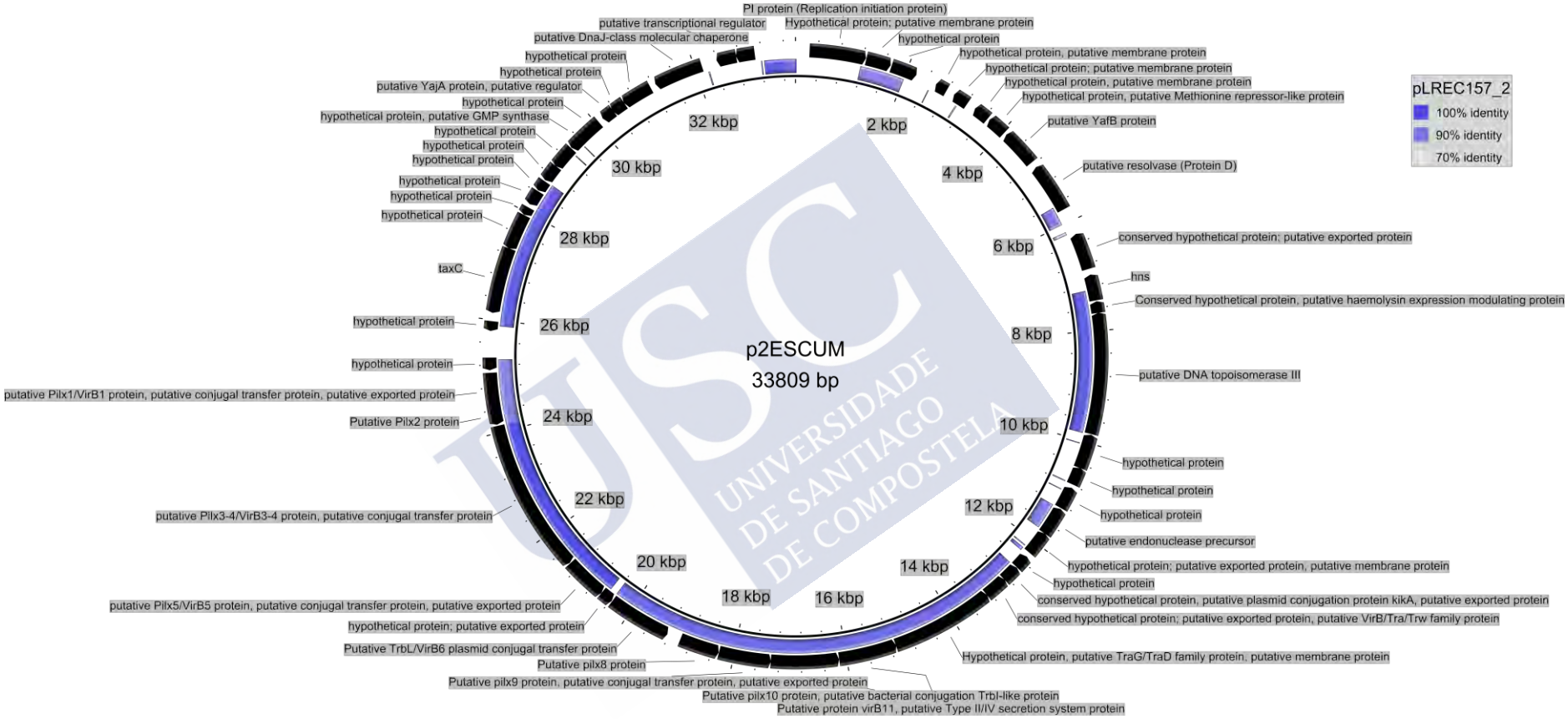


Figure S8. Structural comparison of IncX4 plasmids from ST131 clade B isolates.

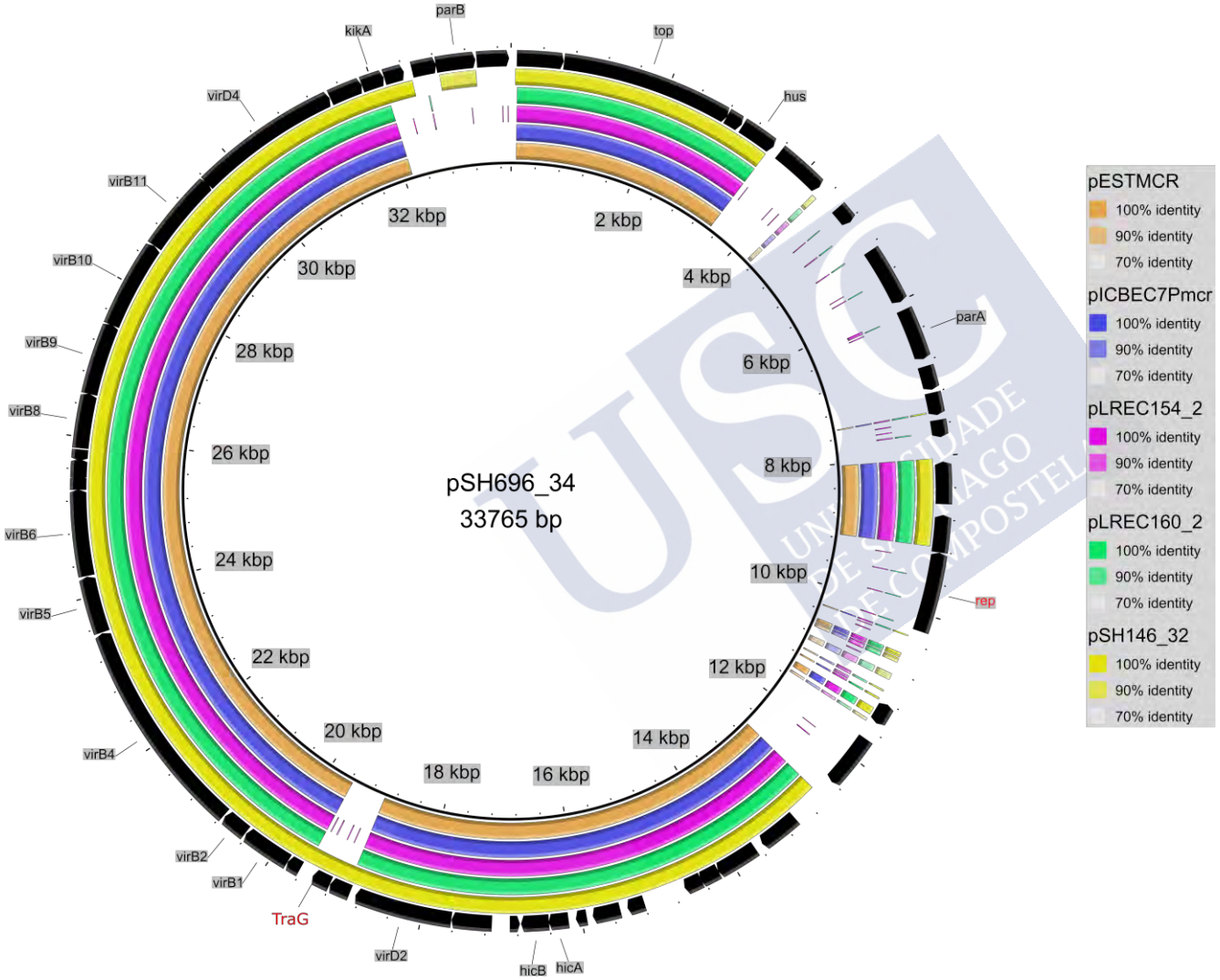


Figure S9. Structural comparison of IncI1 plasmids from the ST131 clade B porcine isolates with pEK204 as a reference.

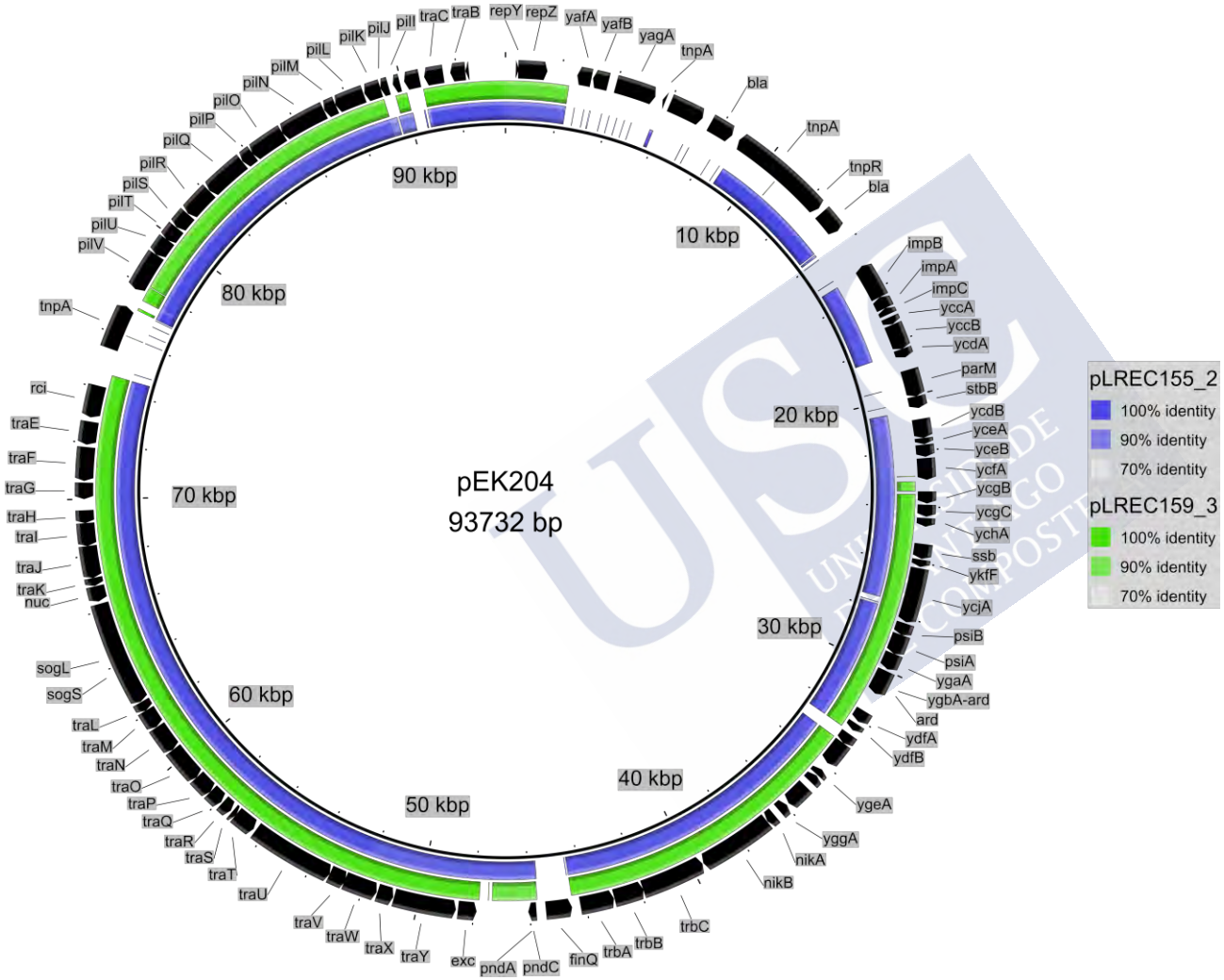


Figure S10. Structural comparison of the IncL/M plasmid from a ST131 clade B porcine isolate with pEC743-OXA48 as a reference.

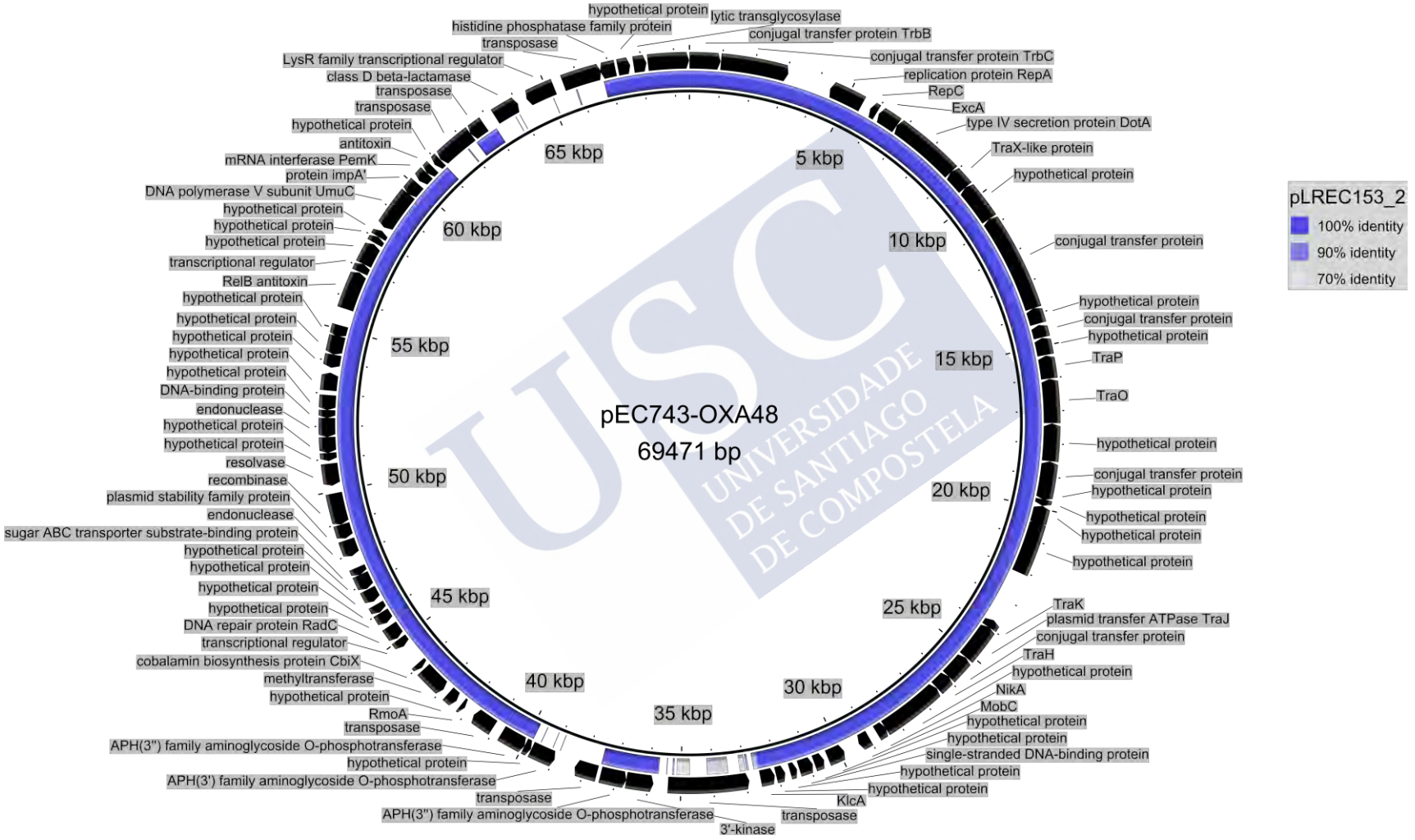


Figure S11. Structural comparison of the cryptic no-MOB plasmids from ST131 clade B porcine isolates with pCE10D as a reference.

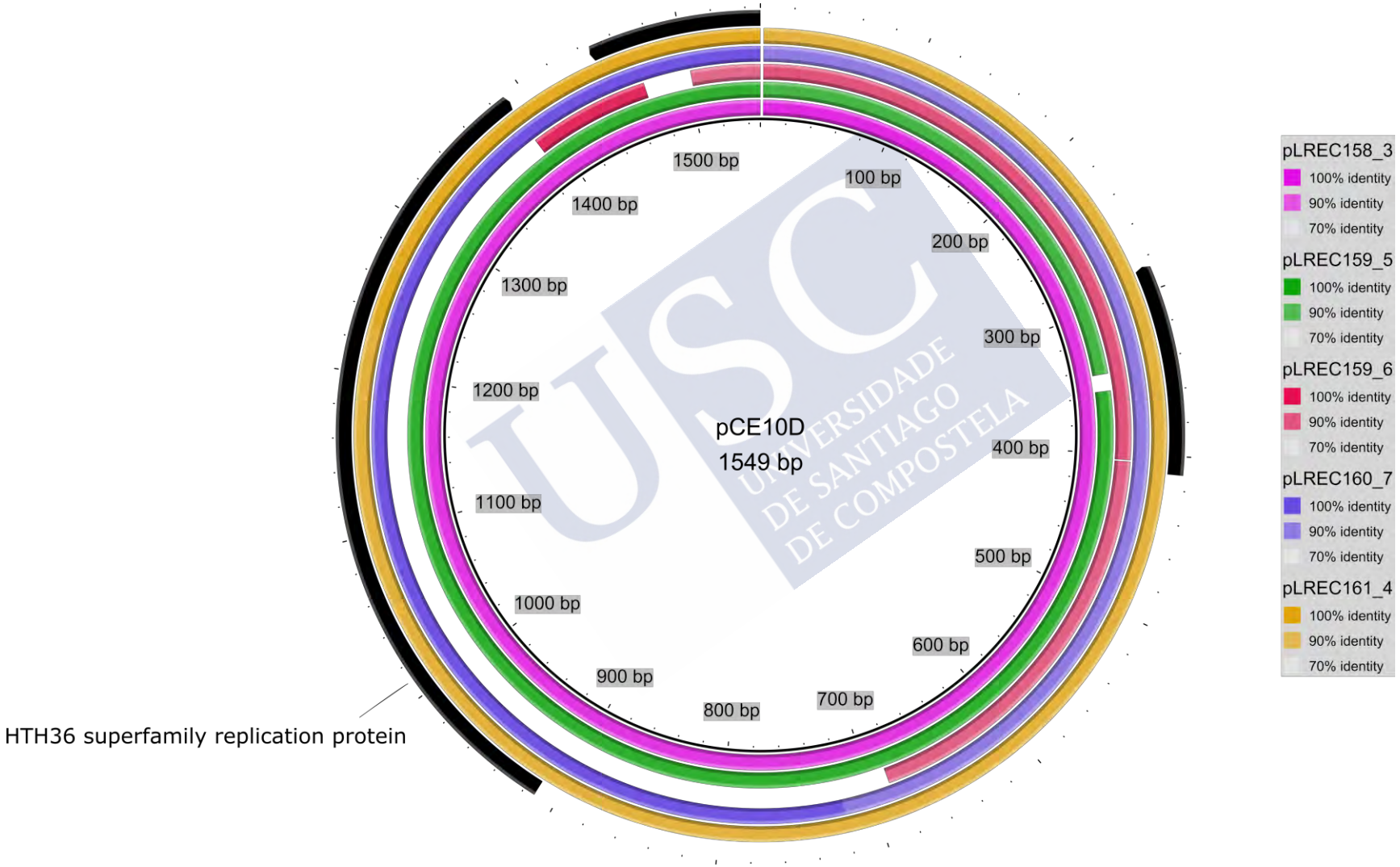


Figure S12. Structural comparison of small MOBv2 plasmids, ST131 clade B porcine plasmids with pEC0674 as a reference.

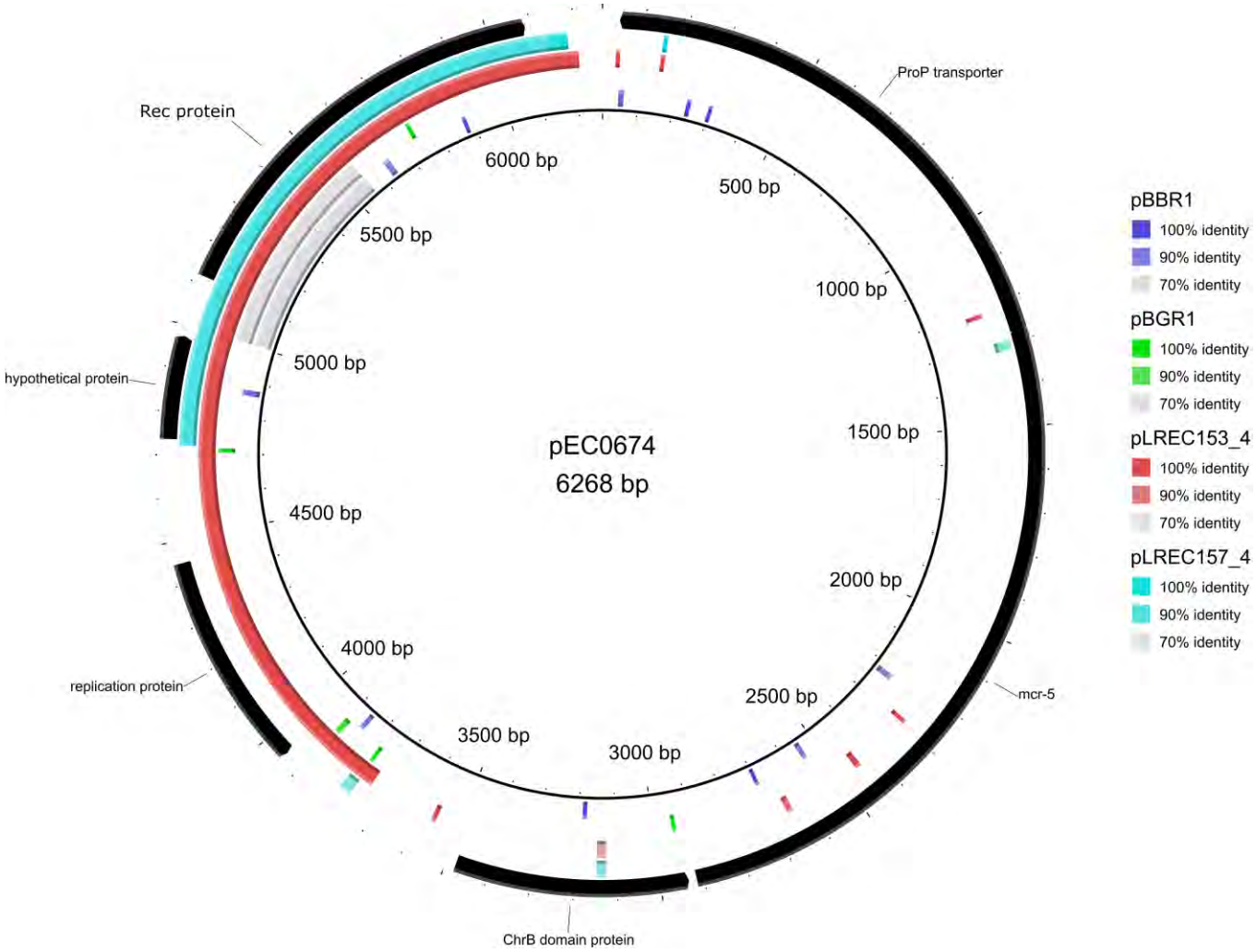


Figure S13. Structural comparison between IncX4 *mcr-1* and non-*mcr*-1 plasmids with pESTMCR as reference.

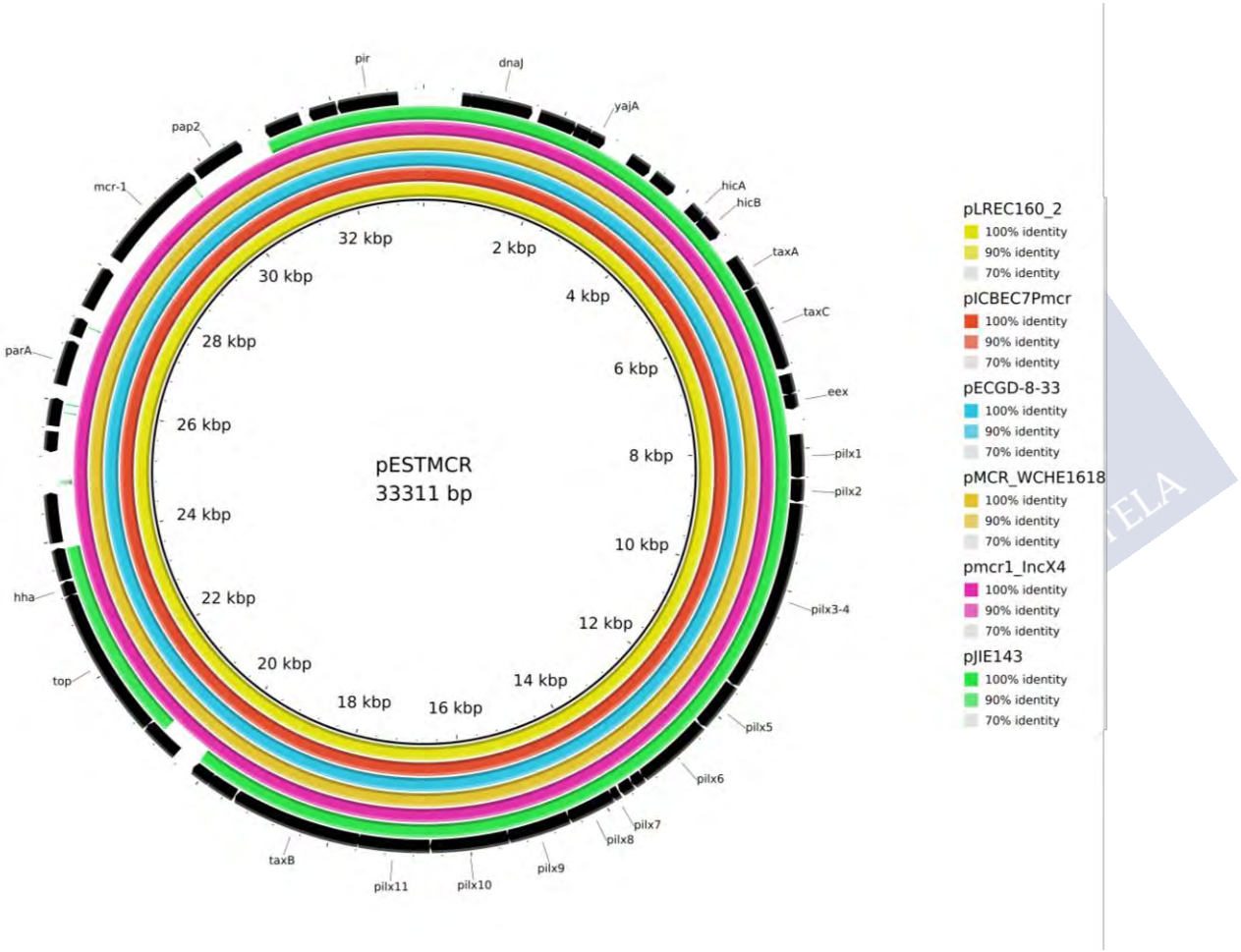


Figure S14. Structural comparison between IncX4 *mcr-1* and non-*mcr*-1 plasmids with pMCR_WCHEC1618 as reference.

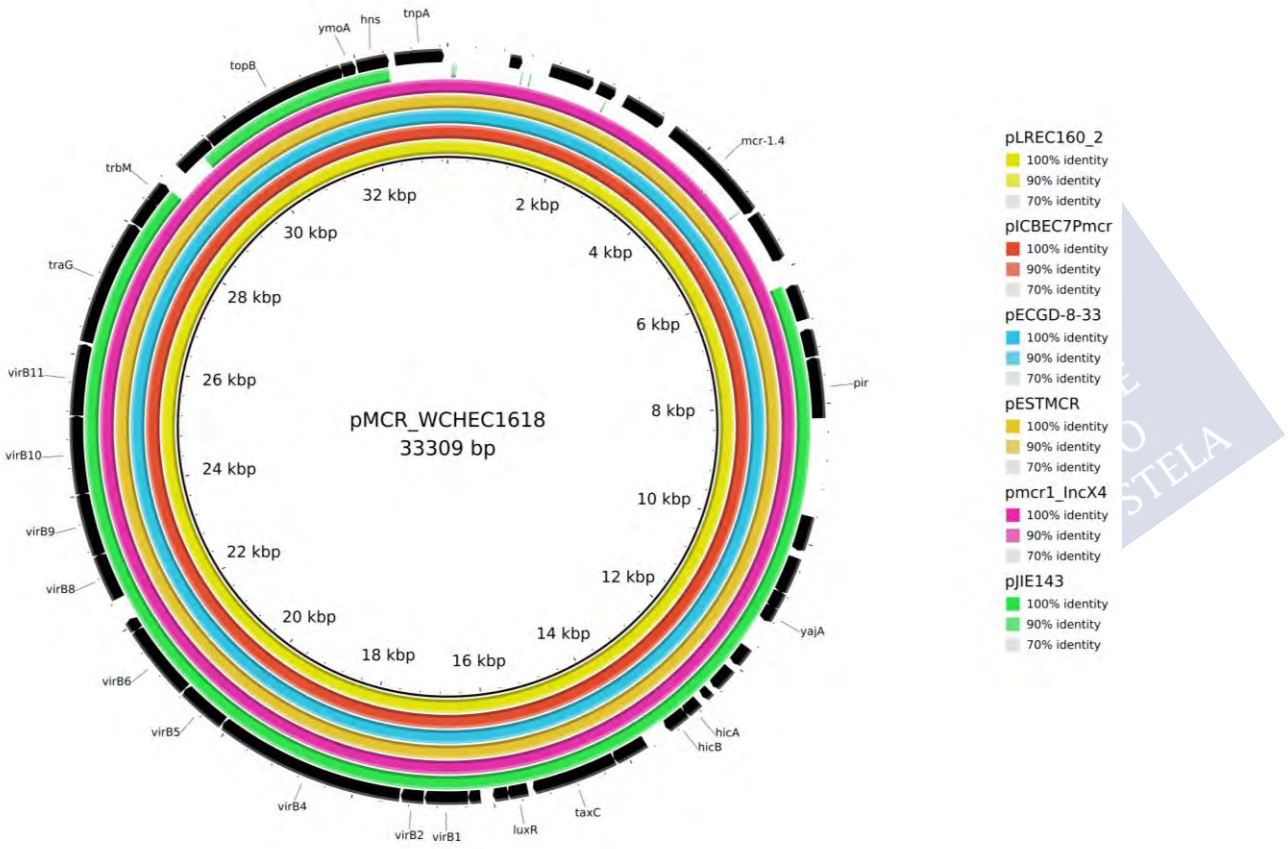


Figure S15. Structural comparison between IncX4 *mcr-1* and non-*mcr*-1 plasmids with pECGD-8-33 as reference.

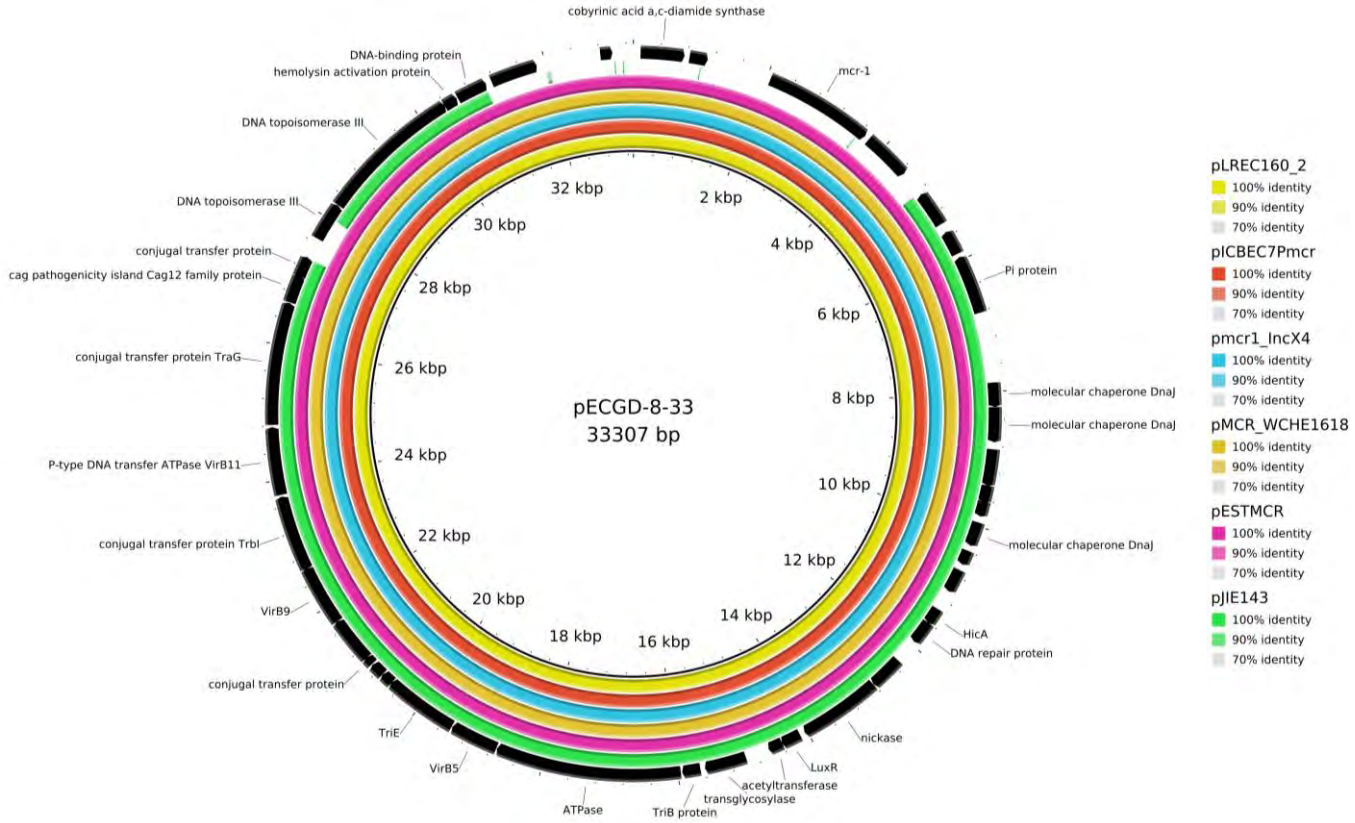


Figure S16. Structural comparison between IncX4 *mcr-1* and non-*mcr*-1 plasmids with pmcr1_IncX4 as reference.

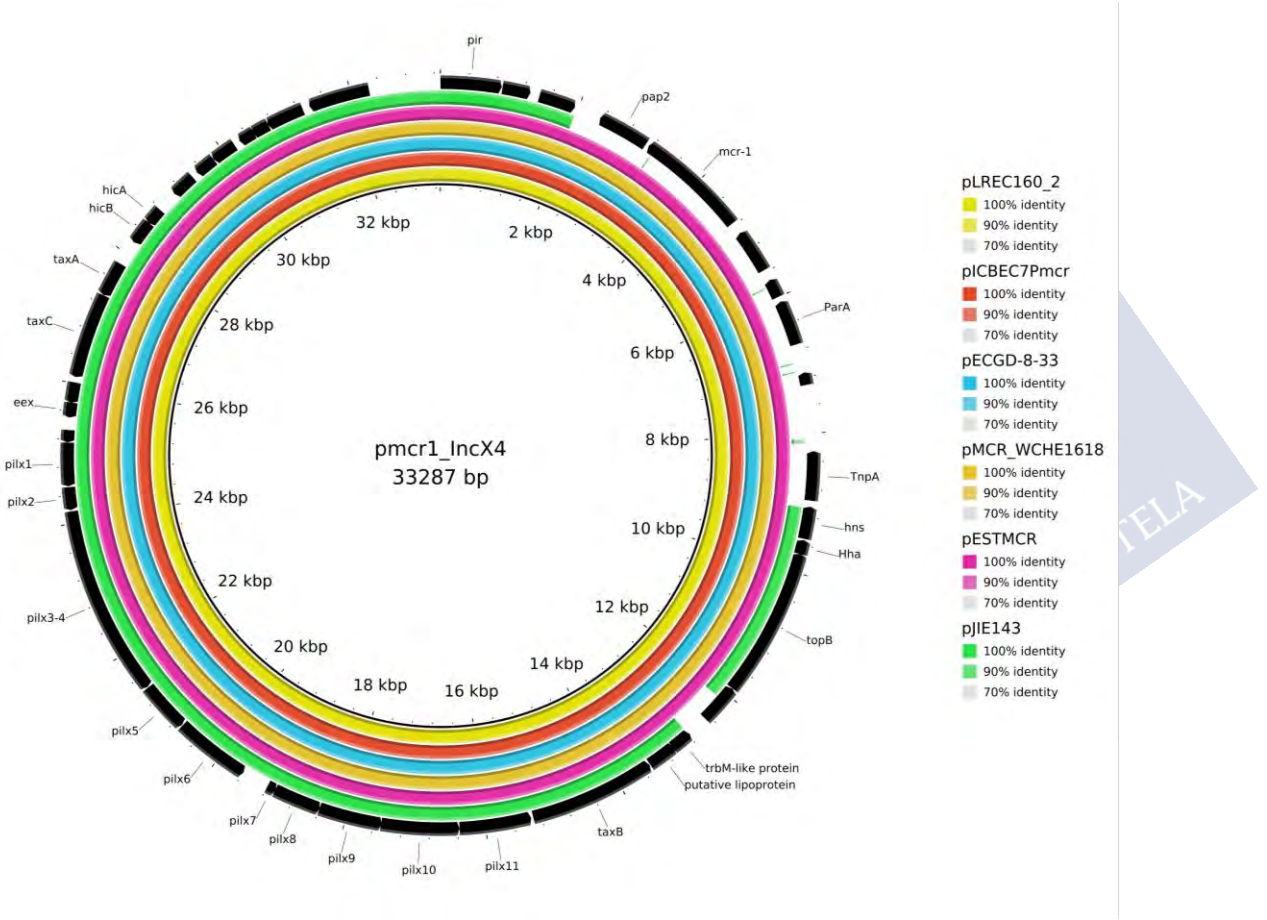


Figure S17. Structural comparison between IncF *mcr-1* and non-*mcr*-1 plasmids with pKP81_BE as reference.

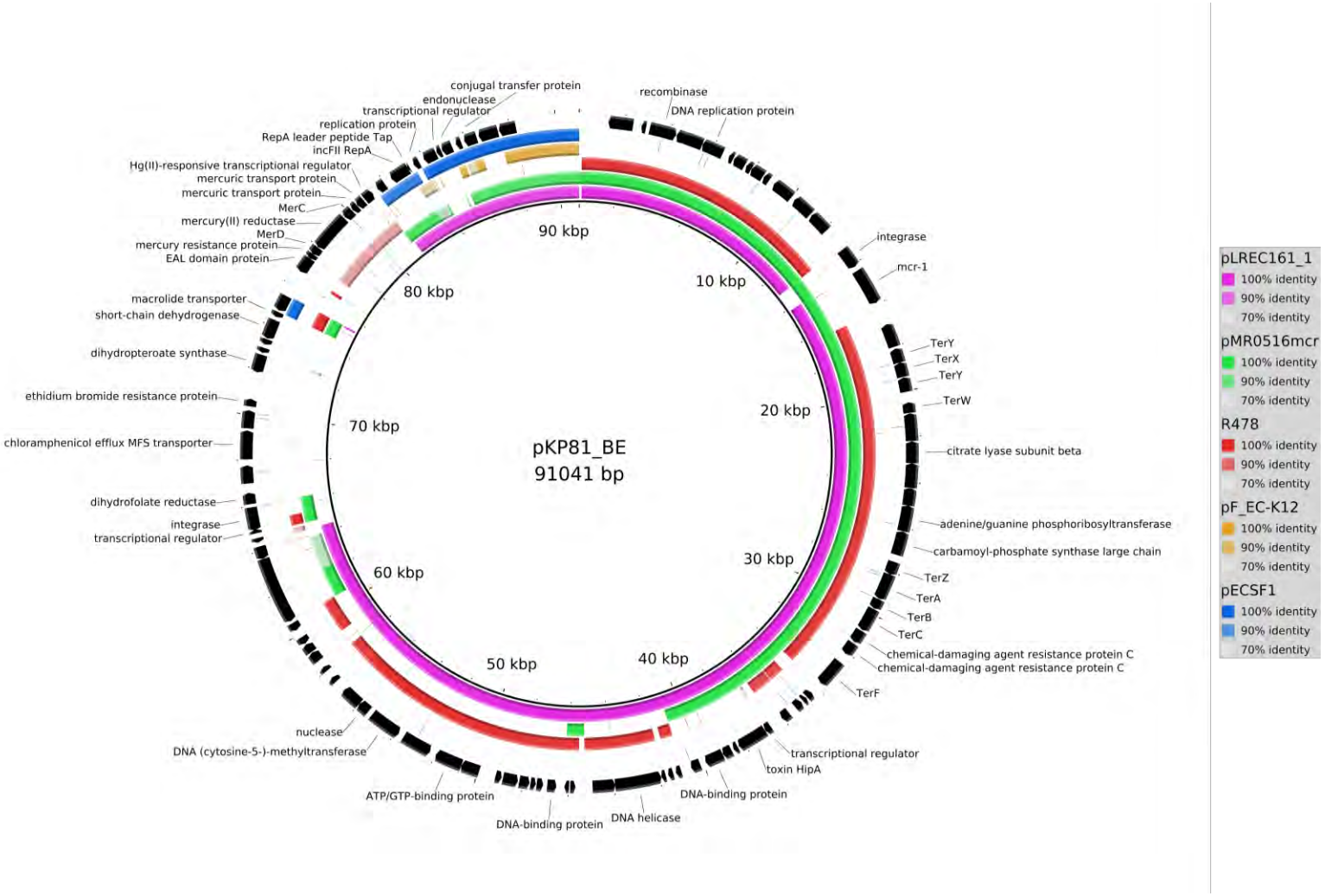


Figure S18. Structural comparison between IncF *mcr-1* and non-*mcr*-1 plasmids with plasmid F from *Escherichia coli* K-12 (pF_EC-K12) as reference.

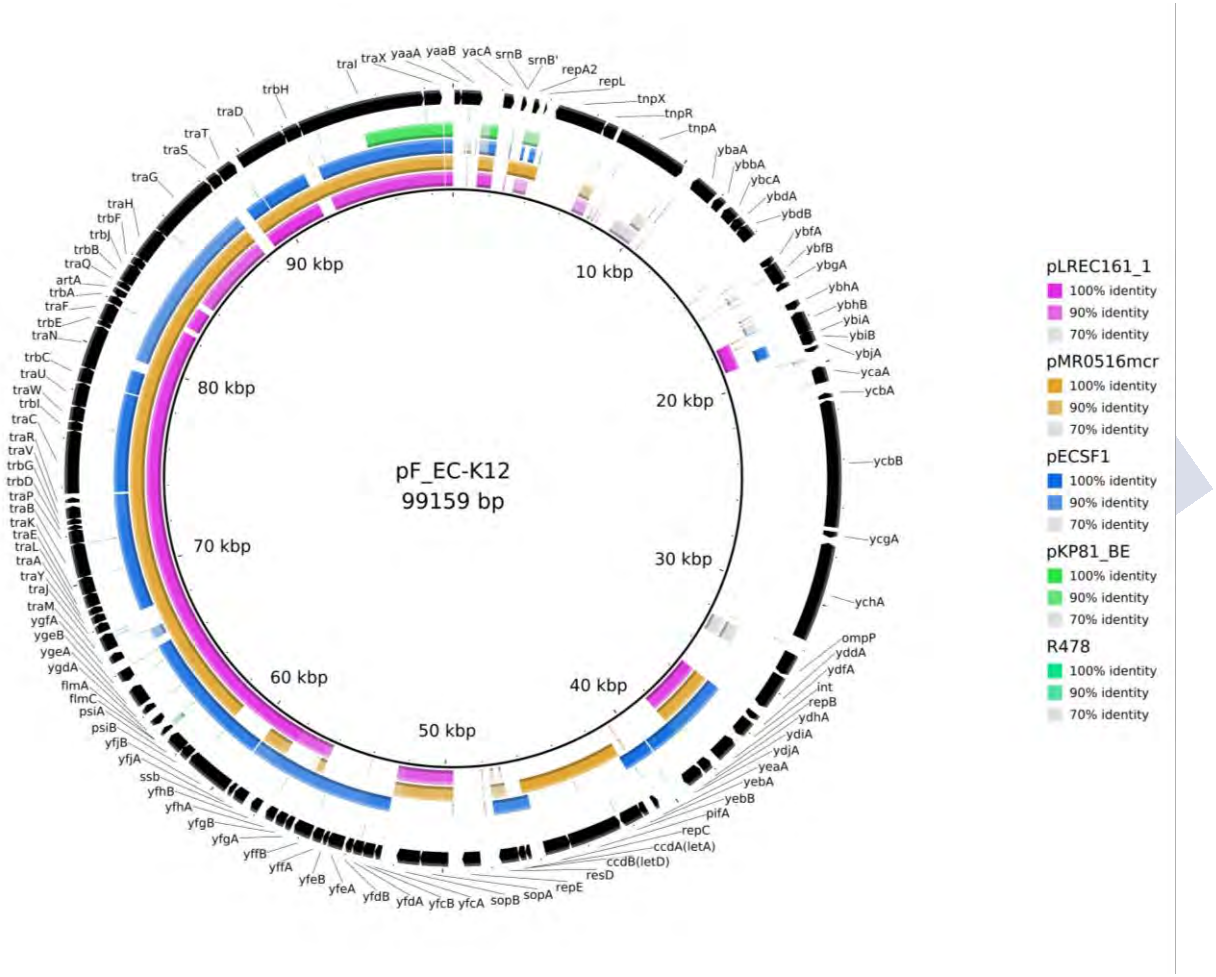


Figure S19. Structural comparison between plasmids from ST131 mcr-1 porcine isolates, other mcr-1 plasmids and non-mcr-1 plasmids with pHNSHP45-2 as reference.

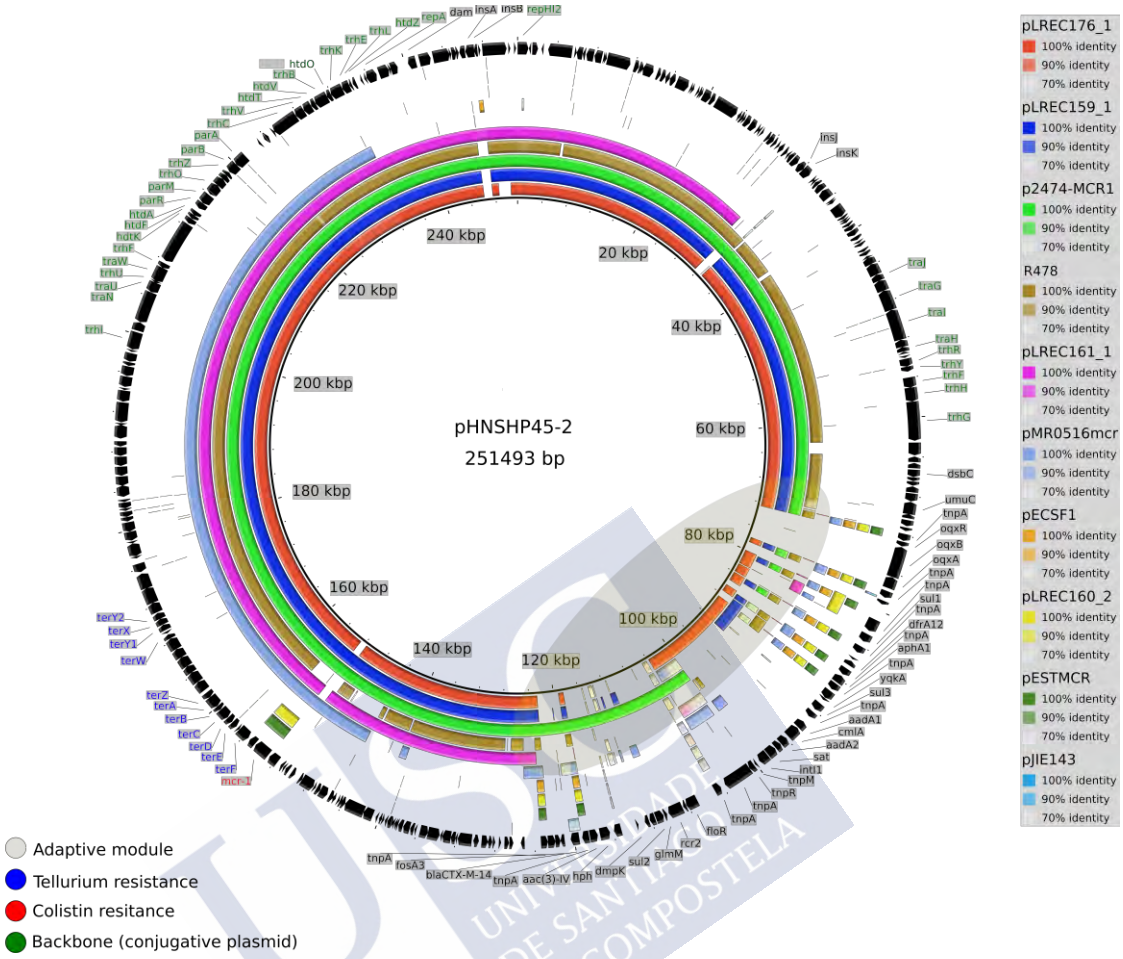


Figure S20. Genetic environment of the *mcr*-1 gene found in the studied genomes. The *mcr*-1 gene is colored in red, the *pap*2 gene in blue and the insertion elements in green. Hyp: hypothetical protein.

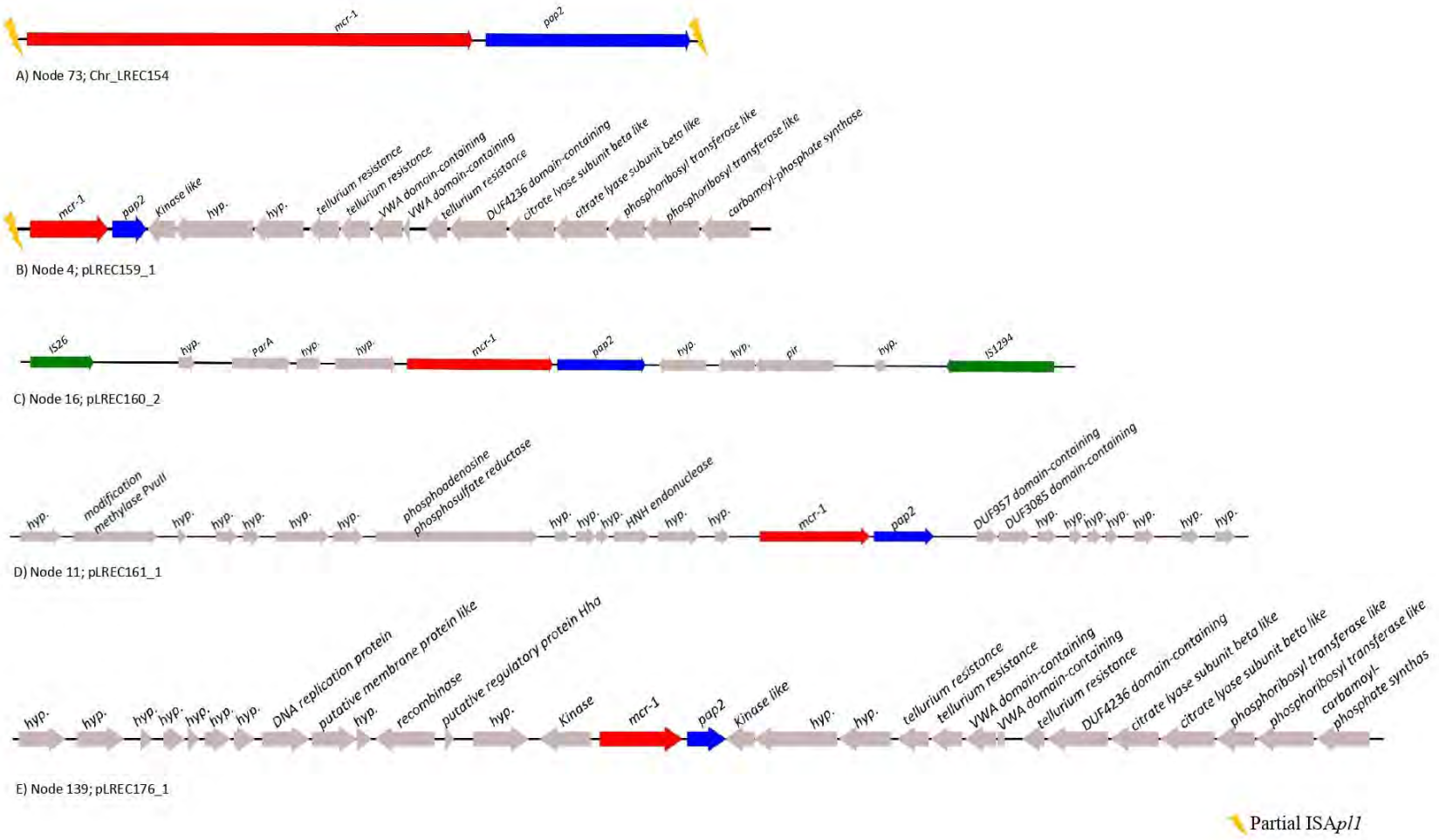
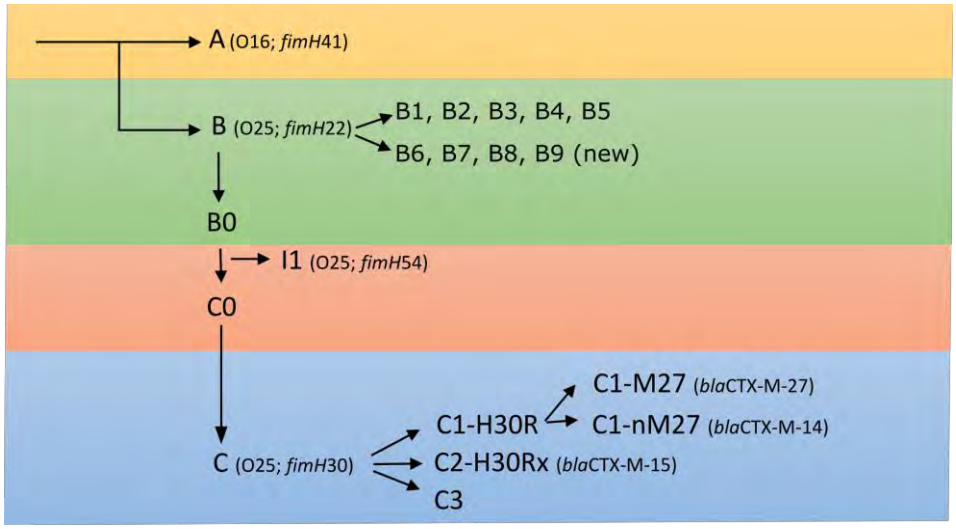


Figure S21. Graphic representation of the ST131 lineage.



USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Escherichia coli is the leading cause of urinary tract and bloodstream infections in humans and animals. The treatment of these infections has been seriously complicated by the appearance of multidrug-resistant isolates and especially by the rapid dissemination of extended-spectrum P-lactamase-producing *E. coli* (ESBLEC).

The present doctoral thesis includes six studies in which we evaluated the prevalence, the phenotypic and the genotypic characteristics of sequence type 131 (ST131) strains and other high-risk clones among *E. coli* strains isolated from patients with extraintestinal infections. Further, we study animals (pigs and dogs) as reservoirs for ST131 strains and other high-risk clones.