

LEBANESE AMERICAN UNIVERSITY

Whole-Genome Comparative Analysis of Carbapenem-
Resistant *Escherichia coli* Isolated from Hospitalized
Patients in Lebanon

By

Christel C. Dagher

A thesis

Submitted in partial fulfillment of the requirements for
the degree of Master of Science in Molecular Biology

School of Arts and Sciences

December 2016

THESIS APPROVAL FORM

Student Name: Christel Dagher I.D. #: 201100165

Thesis Title: Whole-Genome Comparative Analysis of Carbapen-Resistant Escherichia coli Isolated from Hospitalized Patients in Lebanon

Program: M.S. in Molecular Biology

Department: Natural Sciences

School: Arts and Sciences

The undersigned certify that they have examined the final electronic copy of this thesis and approved it in Partial Fulfillment of the requirements for the degree of:

M.Sc. in the major of Molecular Biology

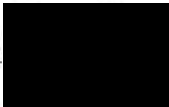
Thesis Advisor's Name	<u>Sima Tikujim</u>	<u>20 / 12 / 2016</u>
		Day Month Year
Committee Member's Name	<u>Tarek Nawas</u>	<u>20 / 12 / 2016</u>
		Day Month Year
Committee Member's Name	<u>Roy Khalaf</u>	<u>20 / 12 / 2016</u>
		Day Month Year

THESIS COPYRIGHT RELEASE FORM

LEBANESE AMERICAN UNIVERSITY NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants the Lebanese American University (LAU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic formats and in any medium, including but not limited to audio or video. You agree that LAU may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that LAU may keep more than one copy of this submission for purposes of security, backup and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant LAU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN LAU, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. LAU will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Name: Christel Dagher

Signature: 

Date: 8-11-2016

PLAGIARISM POLICY COMPLIANCE STATEMENT

I certify that:

1. I have read and understood LAU's Plagiarism Policy.
2. I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.
3. This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: Christel Dagher

Signature: 

Date: 8-12-2016

I dedicate this work to my loving family.

ACKNOWLEDGMENT

To begin with, I would like to thank and express my gratefulness to my supervisor Dr. Sima Tokajian for making this possible. Her dedication and hard work has helped in completing this work and I am fortunate to have had the opportunity of working with her.

I also want to thank my committee members Dr. Roy Khalaf and Dr. Tarek Na'was for their contribution to my work. Not to mention my gratitude to Ms. Maya Farah and Mrs. Helena Bou Farah for their help and assistance in the labs.

A special thanks goes to my dearest friends and colleagues Mr. Melhem Bilen, Ms. Cynthia Harb, Ms. Tina Panossian, Ms. Natalia Abou Zaki and Ms. Tamara Salloum. Thank you for your support, motivation and love.

Last but not least, to my family members, saying 'thank you' is insufficient to express my gratitude for being my eternal constant. I am beyond blessed to have you as my loving and supporting family.

Whole-Genome Comparative Analysis of Carbapenem-Resistant *Escherichia coli* Isolated from Hospitalized Patients in Lebanon

Christel C. Dagher

ABSTRACT

Antibiotic resistant bacteria that are challenging and nearly impossible to treat are becoming more prevalent and this is resulting in an urgent public health concern. The horizontal gene transfer of antibiotic resistance genes between bacteria as well as the complex antibiotic resistance mechanisms involved have aided in the production of an extensive spectrum of bacterial species with multi-drug resistant patterns. Resistance mechanisms of Gram-negative bacteria, such as extraintestinal *Escherichia coli* (ExPEC), vary extensively and the emergence of extended spectrum β -lactamases (ESBLs) as well as carbapenemases have compromised the effectiveness of the majority of antibiotics. ExPEC is one of the major causes of community- and nosocomial-acquired infections. Distinct from intestinal pathogens and commensals, ExPEC causes infections of the bloodstream, respiratory tract, urinary tract, peritoneum and cerebrospinal fluid. The universal burden of diseases caused by such a microorganism is staggering and hundreds of thousands of people are annually affected. In this study, 27 carbapenem resistant ExPEC isolates were recovered from hospitalized patients at the AUBMC (American University of Beirut Medical Center) and were used for identification, phylogenetic typing

and molecular detection of associated antibiotic resistance genes (OXA, SHV, TEM, CTX-M and CTX-M-15). High-throughput sequencing was performed on ten isolates and bioinformatic tools were utilized to identify resistance and virulence determinants, plasmids, and phages. Overall, out of the 27 isolates, 59.6 %, 51.9 %, 7.4 %, 88.9 % and 63 % carried the *bla*_{OXA}, *bla*_{TEM}, *bla*_{SHV}, *bla*_{CTX-M} and *bla*_{CTX-M-15} genes, respectively. The distribution of the isolates among the different phylogenetic groups was as follows: 5 (18.5 %) belonged to group A, 3 (11.1 %) group B1, 6 (22.2 %) group B2 and 13 (48.2 %) group D. The average genome was 5.2 Mb with an average of 224–342 contigs and a G+C % content of 50.49 % – 50.79 %. The isolates were derived from 7 different lineages: ST-405 (ECC153, ECC157, ECC173 and ECC188), ST-205 (EC 174), ST-410 (ECC194), ST-448 (ECC161), ST-617 (ECC202), ST-1284 (ECC167) and ST-648 (ECC149). Analysis revealed that these isolates harbored different β -lactamase genes including *bla*_{OXA-1}, *bla*_{OXA-10}, *bla*_{OXA-181}, *bla*_{TEM-1b}, *bla*_{CTX-M-15}, *bla*_{CMY-2}, *bla*_{CMY-42} in addition to *aac(6')Ib-cr* gene which confers tetracycline and aminoglycoside resistance among others. To the best of our knowledge, this is the first detection of *bla*_{OXA-181} encoding *E. coli* from Lebanon in addition to it being the first comprehensive genome-wide comparative analysis of carbapenem resistant ExPEC isolates. The pandemic potential of these ExPEC strains merits further large-scale comparative and functional genomic studies to better understand the biology of this clinically important bacterium.

Key words: *Escherichia coli*, ExPEC, ESBL, Carbapenem Resistance, High Throughput Sequencing, Lebanon, Virulence, CTX-M-15, OXA-181.

TABLE OF CONTENTS

I-	INTRODUCTION.....	1
II-	LITERATURE REVIEW.....	4
	2.1. Overview of <i>Escherichia coli</i>	4
	2.2. Evolution of Pathogens.....	5
	2.3. The Eight Pathovars.....	6
	2.3.1. ExPEC Pathovar.....	7
	2.4. Extended Spectrum Beta Lactamases (ESBLs).....	9
	2.4.1. Types of ESBLs.....	9
	2.4.1.1. SHV.....	10
	2.4.1.2. TEM.....	10
	2.4.1.3. CTX-M.....	11
	2.4.1.4. OXA.....	11
	2.5. Carbapenems.....	12
	2.6. <i>E. coli</i> Resistance Mechanisms.....	13
	2.7. ESBLs and Carbapenem Resistance in Lebanon.....	15
	2.8. High Throughput Sequencing (HTS).....	16
III-	MATERIALS AND METHODS.....	19
	3.1. Ethical Approval.....	19
	3.2. Bacterial Isolates.....	19
	3.3. Antimicrobial Susceptibility Testing.....	19
	3.4. DNA Extraction.....	19
	3.5. PCR Assays.....	20
	3.5.1. Phylogenetic Group Determination.....	20
	3.5.2. β -lactamase Gene Detection.....	21
	3.6. Genome Sequencing.....	22
	3.7. Data Analysis.....	22
	3.7.1. Genome Assembly.....	22
	3.7.2. RAST Genome Annotation.....	22

3.7.3. Determination of Virulence Genes, Resistance Genes and Pathogenicity..	23
3.7.4. Plasmid Detection, MLST and Serotyping.....	23
3.7.5. Phage Detection.....	23
3.7.6. Circular Visualization.....	23
IV- RESULTS.....	25
4.1. Genome Statistics.....	25
4.2. Representative and Comparative Circular Genomes.....	26
4.3. Phylogenetic Grouping.....	28
4.4. MLST and Serotyping.....	29
4.5. Antibiotic Resistance.....	31
4.6. Virulence Factors.....	35
4.7. Plasmids.....	37
4.8. Phages.....	38
4.9. Insertion Sequences.....	39
4.10. Phylogenetic Analysis.....	39
V- DISCUSSION.....	41
5.1. Antibiotic Resistance.....	41
5.2. Virulence.....	46
5.3. Plasmids.....	50
5.4. Phages and Insertion Sequences.....	51
5.5. Phylogenetic Analysis.....	51
5.6. Future Work.....	52
VI- CONCLUSION.....	53
VII- BIBLIOGRAPHY.....	55
VIII- ANNEX I.....	74

LIST OF TABLES

Table	Page
1. Primers used in phylogenetic grouping.....	20
2. Primer sequences and PCR conditions used to identify β -lactamase genes.....	21
3. General features of the sequenced isolates.....	25
4. Demography and phylogenetic grouping.....	28
5. Correlation between phylogenetic groups of sequenced isolates and their MLST types, serotypes and clonal complexes.....	30
6. Comparative analysis of β -lactams and the antibiograms of the sequenced isolates.....	32
7. Comparative analysis of fluoroquinolone, phenicol and tetracycline resistance of sequenced isolates.....	33
8. Comparative analysis of trimethoprim, aminoglycoside, macrolide and sulphonamide antibiotic resistance of sequenced isolates.....	33
9. Comparative analysis of rifampicin, fosfomycin and quinolone antibiotic resistance of sequenced isolates.....	34
10. Antimicrobial resistance gene patterns of all isolates.....	34
11. Genes encoding the most important virulence factors on the <i>E. coli</i> genomes....	36
12. Comparative analysis of plasmids in sequenced isolates.....	38

LIST OF FIGURES

Figure	Page
1. Genome atlas of <i>E. coli</i> ECC 149.....	26
2. Comparative circular representation of the <i>E. coli</i> ECC 149 isolate genome BLASTed against the genomes of two reference strains.....	27
3. Bar diagram showing the percentage resistance of all <i>E. coli</i> isolates to the tested antibiotics.....	35
4. A pie chart demonstrating the different percentages of major insertion sequence families for ECC 149.....	39
5. Maximum likelihood tree generated using PhyloSift.....	40

LIST OF ABBREVIATIONS

air: Enteroaggregative Immunoglobulin Repeat Protein

AMK: amikacin

AKT: serine/threonine – specific protein kinase

astA: EAST-1 Heat-Stable Toxin

AUBMC: American University of Beirut Medical Center

capU: Hexosyltransferase Homolog

CIPX: ciprofloxacin

CLSI: Clinical and Laboratory Standards Institute

CTX-M: resistance to cefotaxime

DAEC: diffusely adherent *E. coli*

DTA: deep tracheal aspirate

EAEC: enteroaggregative *E. coli*

EDTA: ethylenediaminetetraacetic

EHEC: enterohaemorrhagic *E. coli*

EIEC: enteroinvasive *E. coli*

EilA: *Salmonella* HilA Homolog

EPEC: enteropathogenic *E. coli*

ESBL: Extended Spectrum Beta Lactamases

ETEC: enterotoxigenic *E. coli*

ETP: ertapenem

ExPEC: extraintestinal *E. coli*

FimH: type 1 fimbriae adhesin

gad: Glutamate Decarboxylase

GEN: gentamicin

GES: Guiana extended spectrum

GIM: German imipenemase

HGT: horizontal gene transfer

HlyA: haemolysin A toxin

HTS: high throughput genome sequencing

IBCs: intracellular bacterial communities

IMI: imipenem hydrolyzing carbapenemase

IMP: a metallo- β -lactamase in *Pseudomonas aeruginosa*

IPM: imipenem

iss: Increased Serum Survival

KPC: *Klebsiella pneumonia* carbapenemase

lpfA: Long Polar Fimbriae

MLST: Multi-Locus Sequence Typing

NCCLS: National Committee for Clinical Laboratory Standards

NDM: New Delhi metallo- β -lactamase

NGS: next generation sequencing

NMC-A: not metalloenzyme carbapenemase class A

NMEC: neonatal meningitis *E. coli*

OMP: outer membrane protein

OXA: oxacillin-hydrolyzing β -lactamases

PAIs: pathogenicity islands

PCR: polymerase chain reactions

PFGE: pulsed-field gel electrophoresis

PMNs: polymorphonuclear neutrophils

QRDR: quinolone resistance determining region

RAST: Rapid Annotation using Subsystem Technology

Sat: secreted autotransporter toxin

SenB: Plasmid-Encoded Enterotoxin

SHV: sulfhydryl variable

SIM: Seoul imipenemase

SME: *Serratia marsescens* enzyme

SPM: Sao Paulo metallo- β -lactamases

ST: sequence type

TEM: Temoneria β -lactamase

TMP/SMX: trimethoprim/sulfamethoxazole

***tra*:** transferase gene

TZB: tazobactam β -lactamase inhibitor

UPEC: uropathogenic *E. coli*

UTI: urinary tract infections

Vat: vacuolating autotransporter toxin

VIM: Verona integron–encoded metallo- β -lactamase

WHO: World Health Organization

Chapter One

INTRODUCTION

Escherichia coli is one of the most common member of the family *Enterobacteriaceae* that is frequently associated with both hospital- and community-acquired infections (Doumith et al., 2009). Besides being a harmless intestinal inhabitant, it is highly versatile and most often a deadly pathogen. It can cause intestinal and extraintestinal diseases with the help of virulence factors that alter several cellular processes (Kaper, 2004). *E. coli* is the main causative agent of urinary tract infections (UTI) and it is associated with more than 80% of such infections. UTIs are one of the most frequent human bacterial infections constituting around 25-40% of all nosocomial infections, making them a significant financial and medical burden on the healthcare system (Munk Vejborg et al., 2011). Over the past decade, there has been an increase in extended spectrum β -lactamase producing and carbapenem resistant *E. coli* in Lebanon (Moubareck et al., 2005). Although worldwide cases of carbapenem resistant *E. coli* have been reported, carbapenems are still the main choice of antibiotics in the treatment of multidrug resistant or ESBL producing *E. coli* (Lartigue et al., 2007). A varying number of different mechanisms are thought to be involved in the resistance to carbapenems. Primarily, the process includes the production of carbapenemases like class A KPC, class B metallo- β -lactamases (IMP, VIM and NDM-1) as well as class D OXA-type enzymes (OXA-48) (Matar et al., 2008). Moreover, resistance may be due to AmpC type enzymes along with impermeability in the membrane or the production of class A ESBLs (Moubareck et al., 2005). Membrane impermeability is usually linked to mutations in the porin channels and this in turn leads

to porins not functioning properly. Additionally, impermeability may be the outcome of complete absence of OmpC and/or OmpF porin proteins (Livermore et al., 2006), or associated with active drug effluxing pumps (Gröbner et al., 2009).

The alarming increase of carbapenem resistance has triggered the need of quickly and accurately identifying such bacterial strains. Traditional techniques such as pulsed-field gel electrophoresis (PFGE), multi-locus sequence typing (MLST) and polymerase chain reactions (PCR) among others are limited, time consuming, tedious and costly. High-throughput genome sequencing on the other hand, makes the entire content of the bacterial genome accessible enabling detailed analysis and facilitating the identification of virulence and antimicrobial resistance genes (Hasman et al., 2014). High-throughput genome sequencing studies performed on *E. coli* and numerous other microorganisms have shown great value in giving insight about the diverse drug resistance elements and in describing bacterial transmission, evolution, and outbreaks (Zhu et al., 2013).

In this study, 27 carbapenem resistant *E. coli* isolates were recovered from hospitalized patients at the AUBMC and were used for the identification, molecular characterization and phylogenetic typing. High-throughput sequencing was performed on ten isolates to enable the identification of genes responsible for resistance mechanisms as well as the ones involved in pathogenicity. The overall objectives of the study were to:

- Investigate the population dynamics of pathogenic carbapenem resistant *E. coli*.
- Identify genes involved in host adaptation and investigate variation in virulence potential.

- Identify antimicrobial resistance gene reservoirs in the collected isolates and characterize the fitness and virulence of these resistant pathogens isolated from Lebanon.
- Determine the phylogenetic groups.
- Examine and compare the genes present in the different phylogenetic groups.
- Evaluate the phylogenetic distribution of antibiotic-resistant *E. coli*.
- Detect the presence of different resistance encoding genes (TEM, SHV, OXA, CTX-M and CTX-M-15) through individual PCR assays, correlate to genes detected using high-throughput sequencing, and estimate variation within the genes.
- Determine the susceptibility or resistance of collected isolates towards the most common antimicrobial agents used to treat *E. coli* associated infections.
- Use genome sequence data to construct a genome-based phylogeny.

Chapter Two

LITERATURE REVIEW

2.1. Overview of *Escherichia coli*

E. coli is a ubiquitous Gram-negative organism and is a significant member of the intestinal microflora. This usually harmless commensal organism requires only the acquisition of mobile genetic elements to turn into a highly versatile pathogen capable of causing diseases including: gastroenteritis and extraintestinal infections of the bloodstream, central nervous system and urinary tract (Croxen et al., 2010). It is one of the major causes of community- and nosocomial-acquired infections. The universal burden of diseases caused by such a microorganism is staggering as hundreds of thousands of people are annually affected (Croxen et al., 2010). To date, eight *E. coli* pathovars have been identified and each utilizes a large set of virulence factors to overwhelm the host's cellular functions (Croxen et al., 2010).

The worldwide increase of bacterial resistance to antimicrobial agents has become worrisome. The emergence of a wide range of strains and species with multi-drug resistant patterns are the outcome of complex resistant mechanisms (El-Herte et al., 2012). There has been a remarkable increase in the number of ESBL and carbapenem producing microorganisms, in particular *E. coli* (Baroud et al., 2012). ESBLs enable Gram-negative bacteria to become resistant to all β -lactam antibiotics excluding carbapenems and cephamycins (Ben-Ami et al., 2009). The production of ESBLs is a noteworthy resistance tool that interferes with the treatment of infectious diseases caused by members of the

family *Enterobacteriaceae* and is a serious threat to the presently available antibiotics (Shaikh et al., 2015). Similarly, the detection of carbapenem resistance in *E. coli* is a major concern since carbapenems are the last resort drugs used to treat patients with severe infections (Peirano et al., 2014). Therefore, there is a substantial need for the development of cost-effective, quicker and trustworthy diagnostic tools as well as improved therapies (Dhillon et al., 2012).

2.2. Evolution of pathogens:

The gain or loss of mobile genetic elements plays a crucial role in influencing the genome of a pathogen. A rapid way for a microorganism to acquire new traits is through horizontal gene transfer (HGT), which could enhance the survival and fitness of the pathogen as it coevolves with its host. Furthermore, a large collection of genes coding for virulence known as pathogenicity islands (PAIs) can be either integrated into the chromosome or located on plasmids in pathogenic bacteria. Notably, non-pathogenic bacteria do not possess PAIs (Shames et al., 2010).

Generally, mobile genetic elements such as transposons and bacteriophages often flank PAIs. Hence, it is expected to find most of the virulence genes of *E. coli* located on PAIs in addition to prophages and plasmids. Prophages are commonly defective, but some do have the ability to form infectious elements. When a bacterium acquires new traits via HGT, it gains advantages such as becoming capable of inhabiting new niches. A bacterium undergoing several HGT events will become more exposed to selective pressures, which will favor the survival of more virulent variants that can become epidemic (Asadulghani et al., 2009).

The genome of pathogenic *E. coli* is diverse and can differ from the commensal ones by up to 1 MB as a result of gaining and losing PAIs beside other mobile genetic elements. The core genome of *E. coli* is made up of nearly 2,200 genes and its pan-genome is composed of 13,000 genes. It is interesting to note that even though the genomes of most pathogenic *E. coli* code for approximately 5,000 genes, less than 50% are actually part of the core genome. This permits plasticity and extensive genetic diversity in the pathogenic isolates (Touchon et al., 2009).

2.3. The Eight Pathovars:

The different pathovars were extensively studied and grouped as either extraintestinal *E. coli* (ExPEC) or diarrhoeagenic *E. coli*. The two most common ExPEC pathovars include neonatal meningitis *E. coli* (NMEC) and uropathogenic *E. coli* (UPEC). The rest 6 diarrhoeagenic pathovars are enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC), diffusely adherent *E. coli* (DAEC) and enteroaggregative *E. coli* (EAEC).

Most of the used virulence strategies are shared among the pathogenic strains. All pathovars except EIEC are able to adhere to host cells with the help of pili or fimbriae. After attachment, *E. coli* has to take over the host's cellular processes often by utilizing secreted proteins. This way, the pathogen can manipulate the host cell's signaling pathways and enable its invasion, successful colonization, escape from host's immune response consequently leading to diseases. Every pathovar has its own set of mechanisms to invade and colonize, but yet they all seem to target almost the same processes (Croxen et al., 2010).

2.3.1. ExPEC Pathovar:

ExPECs are part of the normally asymptomatic microflora residing in the intestine and the gut, however when these isolates access other niches, they cause diseases in humans including urinary tract infections, septicemia or meningitis in newborns. Even though several specific virulence factors have been associated with ExPEC infections, the majority of ExPECs cannot be discriminated based on the set of virulence factors that they possess as most often they utilize a combination of multiple virulence factors. This significant genetic heterogeneity and genomic plasticity within *E. coli* species are the outcome of loss and acquisition of genomic information in addition to the high recombination rates (Köhler and Dobrindt, 2011). Therefore, ExPECs have the potential to invade varying tissues and to cause infections in all age groups. The most common ExPEC infections are bacteremia and UTIs, but it has also been linked to infections in skin, soft tissue and respiratory tract. As previously mentioned, several variants such as neonatal meningitis *E. coli* (NMEC) and uropathogenic *E. coli* (UPEC) fall under the ExPEC pathovar (Poolman et al., 2016).

UPEC pathovar is responsible for approximately 80% of all UTIs resulting in acute cystitis in the urinary bladder beside pyelonephritis in kidneys. A major issue encountered by the UPEC (uropathogenic *E. coli*) is to move from the intestinal into the urinary tract, causing an infection there, while utilizing amino acids and peptides as main carbon sources for survival (Alteri et al., 2009). UPEC has an incomparable mechanism for organ tropism, where it moves up the urinary tract from the urethra to the bladder and then to the kidneys. Not to mention its outstanding ability to avoid clearance by urination and to evade the host's innate immune response. A group of strictly regulated virulence factors including:

secreted toxins [vacuolating autotransporter toxin (Vat) and secreted autotransporter toxin (Sat)], polysaccharide capsule, iron harvesting systems and multiple fimbriae, collectively contribute to the enhanced pathogenesis of UPEC (Flores-Mireles et al., 2015).

Upon entry, UPEC adheres to the uroepithelium facilitated by FimH adhesin, and binds to the glycosylated uroplakin Ia in the bladder covering the terminally differentiated superficial capping cells (Wiles et al., 2008). It has been shown that the interaction between uroplakin IIIa (transmembrane protein) and FimH adhesin leads to several phosphorylation events that in turn stimulate signaling pathways involved in apoptosis and invasion (Saldaña et al., 2009). Moreover, the invasion of UPEC is also affected by the destabilization of microtubules and the interaction of FimH with $\alpha 3$ and $\beta 1$ integrins, all of which are located along with actin at the invasion site (Eto et al., 2007). These interactions result in the rearrangement of actin locally by means of activating Rho-GTPases and kinases, leading to the envelopment and internalization of the bacteria. Once UPEC invades, it now replicates quickly and synthesizes complexes called intracellular bacterial communities (IBCs) that provide a temporary safe environment (Hannan et al., 2010). UPEC can escape from the IBCs via a fluxing mechanism, and the motile *E. coli* leave the epithelial cells moving into the lumen of the bladder (Justice et al., 2004).

In the course of an infection, the influx of polymorphonuclear neutrophils (PMNs) results in damage to the tissue, and the attachment and invasion of UPEC causes bladder cells' apoptosis and exfoliation. Likewise, low concentration of pore-forming haemolysin A (HlyA) toxin can hinder AKT triggering and leads to apoptosis and exfoliation of the host cell (Flores-Mireles et al., 2015). Breaching the superficial capping cells momentarily exposes the concealed transitional cells to invasion and propagation of UPEC. The

replication of invading UPEC is restricted when they are trafficked inside endocytic vesicles entangled with actin fibers (Mulvey et al., 2001). However, replication is activated again once the bacteria disrupt the host's actin leading to the formation of IBCs in the cytosol. This dormant state acts as a reservoir as it protects the pathogen from the host's immune response and subsequently allows for a long-term colonization of the bladder (Mysorekar et al., 2009).

Untreated UTIs can lead to the spread of the pathogen to the kidneys and result in the progression of the disease. The capability of bacteria to move up to the kidney is facilitated by the reciprocal management of motility and type 1 pili. Bacteria expressing the type 1 pili have fewer flagella than those that don't; consequently down regulating type 1 pili will result in more motile UPEC. This motility is needed for the bacteria to ascend from the bladder to the kidney (Lane et al., 2007). UTIs affect 150 million individuals each year worldwide, the detected increase in antibiotic resistance along with the high recurrence rates are major threats (Flores-Mireles et al., 2015).

2.4. Extended Spectrum Beta Lactamases (ESBLs):

The ongoing exposure of bacterial isolates to a wide variety of antibiotics in particular β -lactams has resulted in the spread of resistance against these drugs due to selective pressure. The first β -lactamase detected in *E. coli* was in the year 1940 (Turner, 2005). β -lactamases are enzymes capable of hydrolyzing β -lactams and are classified according to two methods: the Ambler and Bush-Jacoby-Mederos. The Ambler scheme is based on molecular classification and protein homology, whereas the Bush-Jacoby-Mederos scheme is based on functional similarities, which makes it more relevant for a

microbiologist or a physician (Rasmussen et al., 1997). The ESBLs belong to the group 2be of the Bush-Jacoby-Mederos scheme (Bush et al., 1995). Group 2b β -lactamases such as SHV-1, TEM-1 and TEM-2 are precursors of group 2be, which are β -lactamases with an extended spectrum.

ESBLs on the other hand, are β -lactamases that enable bacterial resistance by hydrolyzing: penicillins, first-, second-, third-generation cephalosporins and aztreonam, but are inhibited by clavulanic acid. AmpC type β -lactamases is an alternative collection of enzymes that are isolated from extended spectrum cephalosporin-resistant Gram-negative bacteria, such as *E. coli*. AmpC type β -lactamases can be located on a plasmid or on a chromosome and differ from ESBLs by resisting inhibition by clavulanic acid as well as other inhibitors (Rupp et al., 2003).

2.4.1. Types of ESBLs:

2.4.1.1. SHV:

The term refers to sulfhydryl variable and it is mostly associated with *Klebsiella* sp. SHV-1 β -lactamase evolved as being part of *Klebsiella*'s genome, but was later integrated into a plasmid facilitating its spread to other *Enterobacteriaceae*. SHV-1 is linked to resistance to broad-spectrum penicillins, but not oximino substituted cephalosporins. In 1983, a *K. ozaenae* isolated from Germany showed efficient cefotaxime hydrolysis and to a minor degree ceftazidime (Knothe et al., 1983). After sequencing, it was found that in comparison to SHV-1, the β -lactamase differed at the 238 position, glycine being replaced by serine, and hence was named SHV-2. Currently, more than 36 SHV associated ESBLs have been identified (Rupp et al., 2003).

2.4.1.2.TEM:

The first reported TEM-1 was isolated from *E. coli* in 1965 from a patient named Temoneria (hence the term TEM) in Athens, Greece. TEM ESBLs are derived from TEM-1 and TEM-2. TEM-1 is the main plasmid borne β -lactamase of ampicillin resistant *E. coli*. TEM-1 is capable of hydrolyzing ampicillin, carbenicillin, oxacillin, cephalothin and is inhibited by clavulanic acid. While TEM-2 has the same properties as TEM-1, it differs by having a better native promoter and different isoelectric point. TEM-1, TEM-2 and TEM-13 are not ESBLs. The first TEM able to hydrolyze extended spectrum cephalosporins was TEM-3, and to date, more than 100 TEM types have been identified (Paterson et al., 2005).

2.4.1.3.CTX-M:

The designation refers to its ability to break down cefotaxime, but they are also able to hydrolyze ceftazidime, cefepime. CTX-M are inhibited by clavulanic acid, but are more efficiently inhibited by tazobactam. The number of identified CTX-M has been expanding drastically and being detected on every populated continent. While TEM and SHV ESBLs were the consequence of amino acid substitution, CTX-M-ESBLs were obtained via HGT (transposons or conjugative plasmids) (Paterson et al., 2005).

2.4.1.4.OXA:

These Class D β -lactamases were primarily penicillinases, but have the ability to hydrolyze oxacillin as well and thus the name oxacillinases (OXA) (Evans et al., 2014). OXA is predominantly found in *Pseudomonas aeruginosa*, but has been identified in several other Gram-negative bacteria. OXA-1, which is the most common OXA β -

lactamase has been detected in one to ten percent of *E. coli* isolates (Shaikh et al., 2015). Studies suggest that class D β -lactamase genes are not just acquired resistance genes, but are also naturally synthesized in environmental species and clinically significant pathogens (Poirel et al., 2010). These enzymes are thought to be embedded into class 1 integrons and recent discoveries suggest that other genetic structures such as transposons and insertion sequences may also be associated with such genes (Poirel et al., 2010). Generally, this class of β -lactamases is not inhibited by sulbactam, tazobactam and clavulanic acid, but their activities may be inhibited by sodium chloride *in vitro*, which is a unique characteristic and is useful for *in vitro* identification (Evans et al., 2014).

Moreover, some class D β -lactamases that have acquired their resistance genes were shown to also hydrolyze carbapenems. The first class D β -lactamase with carbapenemase activity was OXA-23, which was detected in *Acinetobacter baumannii* isolate from Scotland and was plasmid mediated as it was transferred to *Acinetobacter junii* (Poirel et al., 2010). Since then, several carbapenem hydrolyzing class D β -lactamases have been identified including OXA-48, OXA-149 and OXA-181 (Evans et al., 2014).

2.5. Carbapenems

Carbapenems, and based on the homology of amino acid sequences, are classified into three groups: A, B and D. Groups A and D are serine carbapenemases, while the group B includes the metallo- β -lactamases. The genes coding for carbapenemase are generally located on mobile genetic elements hence playing a major role in its rapid transfer and spread (El-Herte et al., 2012).

The first detection of class A carbapenemase was in 1982 and it included all the following: not metalloenzyme carbapenemase class A (NMC-A), Guiana extended spectrum (GES), imipenem hydrolyzing (IMI), *Serratia marsescens* enzyme (SME) and *Klebsiella pneumoniae* carbapenemase (KPC) (El-Herte et al., 2012). Enzymes belonging to class A carbapenemase are able to hydrolyze all β -lactams, meropenem, fluoroquinolones and aminoglycosides, but they are inhibited by clavulanic acid and tazobactam. The most clinically significant member of this class is KPC. It was detected in North Carolina in 1996 and currently there are nine different gene variants coding for the plasmid borne *bla*_{KPC} gene (Queenan et al., 2007).

Class B carbapenemases include VIM, IMP, NDM-1, Sao Paulo metallo- β -lactamases (SPM), German imipenemase (GIM) and Seoul imipenemase (SIM). These enzymes have the ability to break down all β -lactams and carbapenems, but they are inhibited by ethylene diamine triacetic acid (EDTA). The genes responsible for this phenotype can be chromosomal or plasmid-mediated (Carrer et al., 2010).

On the other hand, class D also known as oxacillin-hydrolyzing β -lactamases (OXA), can hydrolyze penicillin, meropenem, imipenem, but not aztreonam or extended spectrum cephalosporins. The majority of OXA carbapenemases seem to be chromosomally associated. The first *E. coli* OXA-48 carbapenemase producing isolate reported in Lebanon was in 2008. To date, approximately 121 different variants have been identified (Matar et al., 2010).

2.6. *E. coli* Resistance Mechanisms:

Several risk factors predispose the patient to ESBL infections including: sickness severity, duration of hospitalization, time spent in the intensive care unit, urinary and arterial catheterization, prior exposure to antibiotics, intubation and mechanical ventilation (Bradford et al., 2001). Usually, SHV, TEM and OXA β -lactamases are correlated with ampicillin resistance in *E. coli* (Domínguez et al., 2002). Moreover, the dominant resistance mechanism to quinolone in *E. coli* involves the alteration of the target of quinolone, which is linked to the quinolone resistance determining region (QRDR), *par* and *gyrA* genes. On the other hand, aminoglycoside resistant *E. coli* express enzymes that modify aminoglycosides and involve several genetic determinants [(*ant(2'')*), *aac(6')-I*, *aph(3')-I*, *aac(3)-I*, *aac(3)-II*, *aac(3)-III*, *aac(3)-IV*], with tetracycline resistance being encoded by *tet* genes (*tetA*, *tetB*, *tetC*, *tetD*, *tetE* and *tetI*) (Domínguez et al., 2002).

Carbapenem resistance is due to efflux pump activity, porin impermeability as well as the possession of carbapenemase encoding genes (Baroud et al., 2012). The most important elements involved in sustaining carbapenem resistance are the acquisition of class A (KPC), class B (NDM, VIM, IMP) or class D (OXA-181, OXA-48) carbapenemases. Such genes are usually located on plasmids and are linked to mobile genetic elements such as transposons, insertion sequences and integrons which facilitate their spread (Nordmann et al., 2012).

There are also other mechanisms involved in resistance to carbapenem including: the expression of *AmpC* and alteration of outer membrane protein (OMP) resulting in decreased carbapenem permeability. *AmpC* is a class C lactamase and is plasmid mediated (Baroud et al., 2011). It enables the resistance to quinolones, penicillins, cephalosporins, oxyimiocephalosporins, cephamycins, ertapenem, tetracycline, sulfonamide,

chloramphenicol, aminoglycosides and trimethoprim. Detection, prevention, antimicrobial management and necessary infection control measures are all critical in limiting the spread of resistant isolates (El-Herte et al., 2012).

2.7. ESBLs and Carbapenem Resistance in Lebanon:

E. coli strains with antibiotic resistance are increasingly being documented worldwide, with the rate being higher in isolates recovered from the Mediterranean region than in ones from various European countries (Stedt et al., 2014). The frequency of infections caused by ESBL producing strains is drastically increasing, particularly in Lebanon (Moubareck et al., 2005). In 1994, a study done in Lebanon at the American University of Beirut Medical Center (AUBMC) showed that as much as 65% of clinical *E. coli* isolates were ampicillin resistant (Araj et al., 1994). Approximately ten years later, the same institute showed a rise in that percentage to up to 72% (Araj et al., 2008). The percentage of ESBL producing *E. coli* rose from 1.3% in 1997 to 4% in 2001 (Daoud et al., 2003). Another study done at AUBMC showed that between 1998 and 2002, the percentage of ESBL producing *E. coli* increased from 3% to 5% (Samaha-Kfoury et al., 2003). Additionally, in 2006 when Daoud et al., investigated the intestinal carriage of ESBL producing bacteria in the intensive care unit of five different Lebanese hospitals, they discovered that out of 118 isolated strains, 95 (80.5%) of them were indeed *E. coli* (Daoud et al., 2006). Moreover, a study done at Saint George Hospital in Beirut, aimed to examine the bacterial etiology of UTIs over a ten years period. They revealed that 61% of all clinically significant urinary isolates were in fact *E. coli* and that the percentage of ESBL production increased from 2.3% in 2000 to 16.8% in 2009 (Daoud et al., 2011).

Furthermore, another study aimed to investigate urinary *E. coli* isolated from Lebanese patients between 2005 and 2012. They discovered via molecular analysis that among the urinary *E. coli* isolated from Lebanese patients, CTX-M-15 is the most frequent ESBL as it was found in 83% of resistant strains. Also, throughout the course of the study, the percentage of ESBLs rose from 12.6% to 29.4% in all *E. coli* isolates whereas it rose from 11.6% to 25.3% when considering urinary isolates (Daoud et al., 2015).

OXA-48 was the first *E. coli* linked carbapenemase detected in Lebanon (Matar et al., 2010), with Beyrouthy et al. (2014) showing an increase in carbapenem resistant *E. coli* from 0.4% (2008–2010) to 1.6% (2012). This was primarily associated with the emergence of OXA-48 carbapenemase. In fact, *E. coli* constituted 10% of the OXA-48 clinical isolates in 2008-2010 and 73% in 2012 (Beyrouthy et al., 2014). As a result, like other countries, Lebanon is now facing a significant threat with the emergence of carbapenem resistant *Enterobacteriaceae* (El-Herte et al., 2012).

2.8. High Throughput Sequencing (HTS):

HTS has been developed for relatively inexpensive and rapid DNA sequencing. Both the rapid turnaround time and low cost will mean that pathogen HTS can overcome the gap between the practice of diagnostic microbiology and microbial research. This will modify our understanding of pathogens' evolution and global spread of antimicrobial resistance, which is identified by the World Health Organization (WHO) as one of the three greatest threats to human health (Köser et al., 2012).

HTS is being applied in numerous of ways to address concerns of infectious diseases. One of many applications is epidemiological typing which is used to identify laboratory cross contaminations, outline transmissions pathways and facilitate outbreak investigations

(McAdam et al., 2014). Current techniques used for bacterial genotyping have limited resolution since they only cover small regions of microbial genomes, but sequencing entire genomes results in the ultimate resolution for studies of epidemiology as established by several studies including *E. coli* O104:H4 in Germany (Köser et al., 2012). For instance, the extent of molecular epidemiology revealing information about the transmission patterns and sources of an outbreak depends on the samples used and the resolution of the technology utilized. In 2011, Germany experienced an outbreak of *E. coli* O104:H4; however, the isolates from Germany were indistinguishable by conventional tests from *E. coli* O104:H4 isolates from an outbreak in France that same year. Therefore, the epidemiological analysis of these isolates was only possible by using HTS which in turn established the two outbreaks as different (Grad et al., 2012).

Additionally, the high resolution of HTS enables the identification of molecular mechanisms involved in the occurrence of pathogenic clones, the interpretation of transmission pathways during localized outbreaks and global pandemics as well as facilitates the evolutionary analysis of bacterial populations in a patient during infection (McAdam et al., 2014). HTS also aids in accurately defining phylogenetic relationships between different clades within a species. This improvement in phylogenetic analysis will facilitate studies done on how bacterial lineages are distributed geographically as well as simplify the detection of emerging strains with unique genotypes. It is now possible to track the evolution and transmission of bacteria locally, globally and even within a single host (Klemm et al., 2016). Moreover, HTS facilitates transcriptomic analysis with several advantages over conventional hybridization approaches such as accurate quantification, single nucleotide resolution and genome wide coverage. Nevertheless, combining

transposon mutagenesis with HTS provides a dynamic approach to identify bacterial elements required for the survival *in vivo*. This increase in applications of the new technology offers extensive insight into bacterial pathogenesis, epidemiology and evolution (McAdam et al., 2014).

Chapter Three

MATERIALS AND METHODS

3.1. Ethical Approval

Ethical approval was not needed since the clinical isolates were gathered and stored as part of routine clinical care. Clinical isolates and patient records/information remained anonymous.

3.2. Bacterial Isolates

Samples we collected from the American University of Beirut Medical Center (AUBMC). All isolates used in this study were screened for carbapenem resistance. Samples were cultured overnight on Tryptone Soy Agar (Bio-Rad, USA) medium for subsequent experimental work.

3.3. Antimicrobial Susceptibility Testing:

Antimicrobial susceptibility test by the disk agar diffusion technique was performed to establish the resistance patterns of the isolates to: amikacin, ciprofloxacin, gentamicin, tazobactam, trimethoprim/sulfamethoxazole, ertapenem, imipenem and meropenem (Biorad). The test is performed by applying a bacterial inoculum of 0.5 McFarland to the surface of a Mueller-Hinton agar plate. The zone diameters of each drug are interpreted using the criteria published by the Clinical and Laboratory Standards Institute (CLSI, formerly the National Committee for Clinical Laboratory Standards or NCCLS) (Wayne, 2009).

3.4. DNA Extraction

Bacterial DNA was extracted using the Nucleospin Tissue kit (Macherey-Nagel, Germany) following manufacturer's instructions.

3.5. PCR Assays

All PCR assays were performed on PerkinElmer GeneAmp 9700 (PerkinElmer, Wellesly, Massachusetts). All PCR assay runs incorporated a negative (one reagent control without template DNA) and a positive control (reference strain used for the gene amplified). Product size was determined by comparison with a 500 bp and 100 bp molecular weight markers (Fermentas, Vilnius, Lithuania).

3.5.1. Phylogenetic Group Determination:

The final volume of the reaction was 20 µl mixture containing: 1 X PCR buffer, 2.5 mM MgCl₂, 20 pmol of each primer (Table 1), 2.5 U of Platinum Taq DNA polymerase (Invitrogen Inc, Carlsbad, CA), 2 µM of each deoxynucleoside triphosphate (dNTP) and 3 µl DNA. The following reaction parameters were used: initial denaturation at 94°C for 4 min; denaturation at 94°C for 10 s, annealing at 59°C for 10 s, and elongation at 72°C for 1 min, repeated for 30 cycles; final extension at 72°C for 5 min (Clermont et al., 2000). PCR products were separated by electrophoresis using 10 µl of the PCR product using 2.5% agarose gel stained with 0.5 µg/mL ethidium bromide, and visualized on a UV Bioimaging system (GeneSnap system from Syngene).

Table 1. Primers used in phylogenetic grouping (Clermont et al., 2000).

Primer	Primer Sequence (5'-3')	Size
ChuA-F	GACGAACCAACGGTCAGGAT	279 bp
ChuA-R	TGCCGCCAGTACCAAAGACA	
YjaA-F	TGAAGTGTCAGGAGACGCTG	211 bp
YjaA-R	ATGGAGAATGCGTTCCTCAAC	

TspE4C2-F	GAGTAATGTCGGGGCATTCA	152 bp
TspE4C2-R	CGCGCCAACAAAGTATTACG	

3.5.2. β - lactamase Genes Detection:

Table 2. Primer sequences and PCR conditions used to identify β - lactamase genes.

Gene	Primers 5'-3'	Size (bp)	PCR mix	PCR conditions
SHV-F SHV-R	GGTTATGCGTTATATTCGCC TTAGCGTTGCCAGTGCTC	865	25 μ l containing: 1.5 U <i>Taq</i> DNA polymerase, 1X buffer, 1.5 mM MgCl ₂ , 200 μ M dNTPs, 0.2 μ M primer and 2 μ l DNA template ¹	one cycle of 5 min at 96 °C, followed by 35 cycles of 1 min at 96 °C, 1 min at 60 °C and 1 min at 72 °C, with a final extension of 10 min at 72 °C
OXA-F OXA-R	ACACAATACATATCAACTTCGC AGTGTGTTTAGAATGGTGATC	814	25 μ l containing: 1.5 U of <i>Taq</i> DNA polymerase, 1X buffer, 2.5 mM MgCl ₂ , 50 μ M of each dNTPs, 0.3 μ M primer and 5 μ l of DNA template ²	one cycle of 5 min at 96 °C, followed by 35 cycles of 1 min at 96 °C, 1 min at 60 °C and 2 min at 72 °C with a final extension of 10 min at 72 °C
TEM-F TEM-R	ATGAGTATTCAACATTTCCG CTGACAGTTACCAATGCTTA	868	50 μ l containing: 2.5 U <i>Taq</i> polymerase, 1X buffer, 1.5 mM MgCl ₂ , 50 mM KCl, 0.4 mM each dNTP, 0.5 mM each primer and 8 μ l of template DNA ³	one cycle of 5 min at 96 °C, followed by 30 cycles of 1 min at 96 °C, 30 s at 55 °C, 1 min at 72 °C with a final extension of 5 min at 72 °C
CTX-M-F CTX-M-R	ATGTGCAGYACCAGTAARGT TGGGTRAARTARGTSACCAGA	593	25 μ l containing: 1,5 U <i>Taq</i> DNA polymerase, 1X buffer, 1.5 mM MgCl ₂ , 200 μ M dNTPs, 0.2 μ M primer and 2 μ l DNA template ¹	one cycle of 7 min at 94 °C, followed by 35 cycles of 50 s at 94 °C, 40 s at 50 °C and 1 min at 72 °C, with a final extension

				of 5 min at 72 °C
CTX-M-15-F CTX-M-15-R	GGTTAAAAAATCACTGCGTC TTACAAACCGTCCGGTGACGA	874	50 µl containing: 1.25 U <i>Taq</i> DNA polymerase, 1X buffer, 2 mM MgCl ₂ , 200 µM dNTPs, 50 pmol primers and 2.5 µl DNA template ⁴	one cycle of 10 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 52 °C and 1 min at 72 °C, with a final extension of 10 min at 72 °C

¹ Elumalai et al., 2014

³ Safari et al., 2015

² Lim et al., 2009

⁴Weill et al., 2004

3.6. Genome Sequencing:

Genomic DNA was used as input for library preparation using the Illumina Nextera XT (Illumina, San Diego, CA, USA) library preparation kit. The subsequent clean up steps were performed using the AMPure XP PCR purification beads (Agencourt, Brea, CA). The resultant libraries were then quantified using quantitative PCR on a CFX96 (Bio-Rad, USA) utilizing the Kapa library quantification kit (Kapa Biosystems, Woburn, MA). The samples were pooled together and then sequenced on an Illumina MiSeq for paired-end 250-bp reads.

3.7. Data Analysis:

3.7.1. Genome Assembly:

The assembly of the genomes was performed *de novo* using A5 with the default parameters. This pipeline automates the procedure of data cleaning, contig assembly, error correction, quality control and scaffolding (Tritt et al., 2012).

3.7.2. RAST Genome Annotation:

The obtained *de novo* assemblies were annotated using RAST server (<http://rast.nmpdr.org>). This enabled the identification of proteins encoding rRNA, tRNA and assigning gene functions as well as predicting subsystems in the genome and assigning gene functions (Aziz et al., 2008; Overbeek et al., 2014).

3.7.3. Determination of Virulence Genes, Resistance Genes and Pathogenicity:

The ResFinder 2.1 web server (www.genomicepidemiology.org) was used to screen high throughput sequencing data and identify acquired antimicrobial resistance genes (Zankari et al., 2012). ResFinder detects the presence of whole resistance genes, but not their expression and functional integrity. Based on the ResFinder results and previously published studies, a predicted phenotype was determined. Similarly, the VirulenceFinder 1.2 web server (www.genomicepidemiology.org) was used to identify virulence genes (Joensen et al., 2014)

3.7.4. Plasmid Detection, MLST and Serotyping:

Plasmid identification was achieved using the CGE's PlasmidFinder 1.2 web service with a 95% selection threshold (Carattoli et al., 2014). MLST typing (multi-locus sequence typing) was performed by mapping the high throughput sequencing data to an online database offered at CGE website known as MLST 1.7 server (Larsen et al., 2012). *E. coli* serotypes were identified using SerotypeFinder 1.1 (Joensen et al., 2015).

3.6.5. Phage Detection:

Phage identification was performed using the publically available Phage Search Tool (PHAST) (<http://phast.wishartlab.com/index.html>) (Zhou et al., 2011). This tool enables the determination of the site of phage integration and assigns phage family.

3.6.6. Circular Visualization:

Two different tools were used to generate a circular genome. A comparative figure was generated using CGView server found in the publically available Stothard Research Group website (http://stothard.afns.ualberta.ca/cgview_server/) (Grant et al., 2008). Likewise, an informative figure was generated using DNAPlotter release 1.11 (Carver et al., 2009).

Chapter Four

RESULTS

4.1. Genome Statistics:

Paired-end libraries (Illumina) were generated from extracted DNA and fragments with size between 300-600 bp were chosen. After pooling the samples, they were sequenced on the Illumina MiSeq for paired-end reads. High-quality reads were obtained after error correction and quality trimming. Using the A5 assembly pipeline, the sequences were processed and assembled. This pipeline facilitates the procedure of error correction, data cleaning, scaffolding, contig assembly and quality control. The assembled genomes had on average 224–342 contigs, G+C % content of 50.49 % – 50.79 %, and total reads of 4,909,271 bp - 5,394,442 bp (Table 3).

Table 3. General features of the sequenced isolates.

Sample ID	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
Number of contigs	278	307	337	236	342	314	233	278	224	279
Total length (bp)	5338228	5319581	5318235	5196409	4940673	5394442	5113962	5385744	4909271	4938887
GC (%)	50.49	50.57	50.63	50.59	50.77	50.56	50.62	50.56	50.53	50.79
Number of Subsystems	617	609	604	607	604	612	615	605	607	595
Number of Coding Sequences	5135	5115	5122	5008	4771	5211	5002	5190	4743	4812
Number of RNAs	104	103	105	98	98	100	101	101	101	100

4.2. Representative and Comparative Circular Genomes:

Figure 1 represents the circular genome of isolate ECC 149. Figure 2 demonstrates the aligned regions of draft genome ECC 149 with two reference genomes: *E.coli* AA86 and *E. coli* K-12 MG1665. The rest of the individual genomes can be found in ANNEX I.

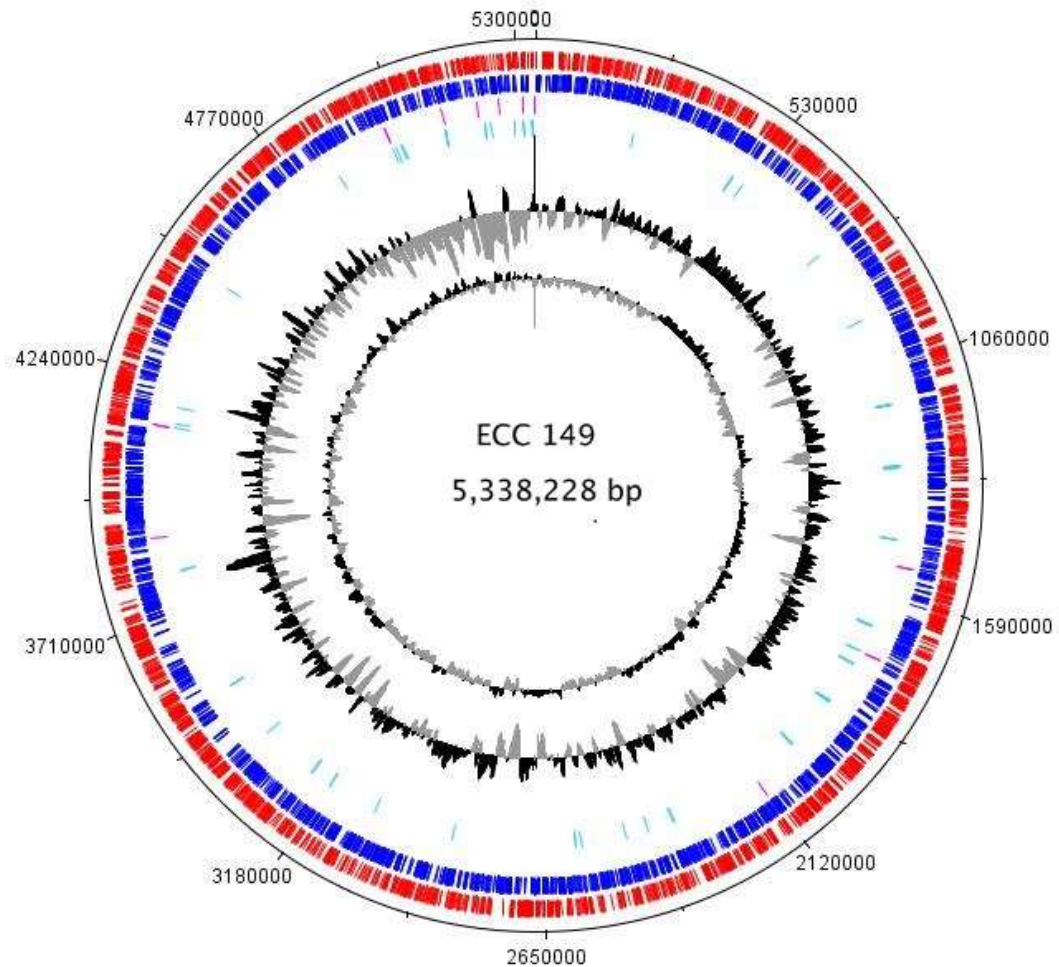


Figure 1. Genome atlas of *E. coli* ECC 149. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

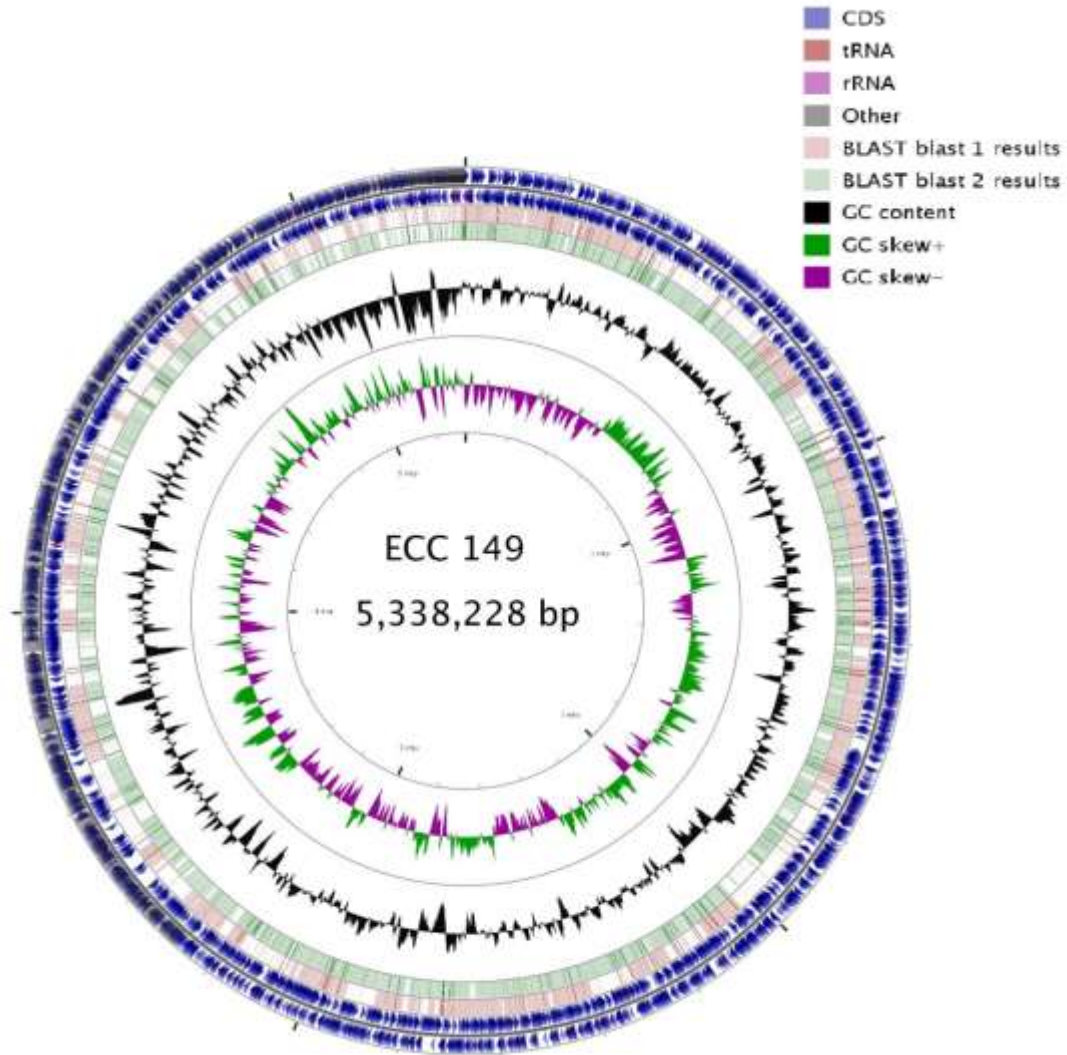


Figure 2. Comparative circular representation of the *E. coli* ECC 149 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession number NZ_AFET0100001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

4.3. Phylogenetic Grouping:

E. coli belongs to four principal phylogenetic groups: A, B1, B2 and D. Three marker genes were used for the detection of phylogenetic groups: *yjaA* gene, *chuA* gene and DNA fragment TspE4C2 (Clermont et al., 2000). A total of 27 strains were analyzed and depending on the combination of the three marker genes, they were assigned to the respective phylogenetic group (Figure 3). Accordingly, five (18.5 %) of the isolates belonged to group A, three (11.1 %) group B1, six (22.2 %) group B2 and 13 (48.2 %) group D (Table 4).

Table 4. Demography and phylogenetic grouping.

Phylogenetic Group	Name of Isolate	Age of Patient	Gender	Percentage	Site of Isolation
A	ECC 147	22	M	18.5 %	catheter
	ECC 168	83	F		urine
	ECC 186	48	F		other
	ECC 194	29	F		liver
	ECC 202	9	F		urine
B1	ECC 161	83	F	11.1 %	urine
	ECC 190	85	M		other
	ECC 174	83	F		skin
B2	ECC 167	89	F	22.2 %	other
	ECC 200	67	F		urine
	ECC 203	2	F		urine
	ECC 314	38	M		blood
	ECC 319	68	F		other
	ECC 443	87	M		other
D)	ECC 149	22	M	48.2 %	blood
	ECC 153	77	F		DTA
	ECC 157	68	M		DTA
	ECC 173	76	F		UTI
	ECC 187	47	F		urine
	ECC 188	31	F		urine
	ECC 199	54	M		liver
	ECC 205	73	M		urine
	ECC 312	68	M		blood
	ECC 322	57	M		wound
	ECC 394	31	M		abscess
	ECC 396	47	M		other
	ECC 605	8 weeks	F		other

***DTA: deep tracheal aspirate**

UTI: urinary tract infections

Other: source unidentified.

4.4. MLST and Serotyping:

The Achtman method (<http://mlst.warwick.ac.uk/mlst/dbs/Ecoli>) was used to determine the MLST for all the sequenced isolates. The method is based on sequencing seven housekeeping genes (*adh*, *fumC*, *gyrB*, *icd*, *mdh*, *purA* and *recA*). MLST typing of the isolates in question was done by mapping the NGS data to an online database using CGE's

MLST 1.8 server (Larsen et al., 2012). The results were tabulated based on the Achtman method (Table 5). The most common MLST type was ST-405 (40 %) with the rest of the STs (ST-205, ST-410, ST-448, ST-617, ST-648 and ST-1284) having the same distribution of 10 % for each. ST-405 (ECC 153, ECC 157, ECC 173 and ECC 188) belonged to O102:H6 serotype as determined by SerotypFinder 1.1 (Joensen et al., 2015). ST-205 (ECC 174) was characterized as O100:H12 serotype, while ST-410 (ECC 194), ST-448 (ECC 161), ST-617 (ECC 202) and ST-1284 (ECC 167) belonged to O8:H9, O8:H8, O89:H10 and O89:H21 serotypes, respectively. ST-648 (ECC 149) carried an H6 serotype whereas the O type gene was untypable.

Table 5. Correlation between the phylogenetic groups of the sequenced isolates and their MLST types, serotypes and clonal complexes.

Phylogenetic Groups	Sample ID	MLST Achtman method	Serotype	MLST Allele Numbering	Clonal Complex
A	ECC 167	ST-1284	O89:H21	10-4-4-8-8-13-73	10
	ECC 194	ST-410	O8:H9	6-4-12-1-20-18-7	23
	ECC 202	ST-617	O89:H10	10-11-4-8-8-13-73	10
B1	ECC 161	ST-448	O8:H8	6-6-5-16-11-8	448
	ECC 174	ST-205	O100:H12	6-23-15-16-9-8-7	205
D	ECC 149	ST-648	H6	92-4-87-96-70-58-2	648
	ECC 153	ST-405	O102:H6	35-37-29-25-4-5-73	405
	ECC 157	ST-405	O102:H6	35-37-29-25-4-5-73	405
	ECC 173	ST-405	O102:H6	35-37-29-25-4-5-73	405
	ECC 188	ST-405	O102:H6	35-37-29-25-4-5-73	405

4.5. Antibiotic Resistance:

The disk agar diffusion method revealed that all isolates undertaken in this study were carbapenem resistant (Figure 3), and more than 60% of the isolates were resistant to ciprofloxacin, tazobactam and trimethoprim/sulfamethoxazole. However, only ECC 194 (ST-410) carried a carbapenem resistance gene (*bla_{OXA-181}*), and the carriage rate of non-carbapenemase-related *bla* genes was high among the sequenced genomes (*bla_{TEM}* and *bla_{OXA}*) (Table 6). Additionally, The *bla_{CMY-42}* gene, which is a plasmid borne AmpC cephalosporinase (Feng et al., 2015), was detected in ECC 157, ECC 173 and ECC 188 (ST-405), ECC 167 (ST-1284), ECC 174 (ST-205) and ECC 202(ST-617) while *bla_{CMY-2}*, key player in broad spectrum cephalosporin resistance, was only detected in ECC 194 (ST-410). PCR assays targeting the SHV, OXA, TEM and CTX-M and CTX-M-15 encoding genes in all the isolates revealed several different patterns, with the number of genes detected varying between one to four. Results of the PCR assays of individual resistance determinants were in agreement with those detected through whole-genome sequencing (Tables 6 and 10). The carbapenem resistant isolates were additionally tested for other antimicrobial agents by the disk agar diffusion method, PCR assays and whole-genome sequencing. There was a high correlation between the detected phenotypes and genotypes (Tables 6-10 and Figure 3). Most of the isolates were multidrug-resistant due to high resistance rate to several antimicrobial classes. Among the detected genes was *aac(6')Ib-cr*, which confers resistance to aminoglycoside (tobramycin and amikacin) and fluoroquinolone (ciprofloxacin) and was found in all sequenced genomes except ECC 149 (ST-648) and ECC 173 (ST-405). However, only two isolates ECC 161 (ST-448) and ECC 194 (ST-410) were positive for *aac(3)-IId* gene and only ECC 157 and ECC 188 (ST-405)

and ECC 167 (ST-1284) had the *aac(3)-IIa* gene, both encoding for aminoglycoside resistance (Table 7). Phenicol, tetracycline lincosamide, macrolide, streptogramin B, trimethoprim, sulphonamide, rifampicin, fosfomycin and quinolone resistance were also among the detected resistant determinants (Tables 7, 8 and 9).

Table 6. Comparative analysis of β -lactams and the antibiograms of the sequenced isolates.

	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
β-lactam resistance										
<i>bla</i> _{TEM-1B}	+	-	-	+	-	-	-	-	+	-
<i>bla</i> _{CTX-M-15}	+	+	+	+	+	+	+	+	+	+
<i>bla</i> _{OXA-1}	+	+	+	-	+	+	+	+	+	+
<i>bla</i> _{OXA-10}	-	-	-	+	-	-	-	-	-	-
<i>bla</i> _{OXA-181}	-	-	-	-	-	-	-	-	+	-
<i>bla</i> _{CMY-2}	-	-	-	-	-	-	-	-	+	-
<i>bla</i> _{CMY-42}	-	-	+	-	+	+	+	+	-	+
Antimicrobial resistance										
Ciprofloxacin	R	R	R	R	R	R	R	R	R	R
Amikacin	S	S	S	R	S	S	S	S	S	S
Gentamicin	S	S	R	R	R	S	S	R	R	S
Trimethoprim-sulfamethoxazole	R	R	R	R	R	R	R	S	R	R
Tazobactam	S	I	R	R	R	R	R	R	R	R
Imipenem	R	R	R	R	R	R	R	R	R	R
Ertapenem	R	R	R	R	R	R	R	R	R	R

Table 7. Comparative analysis of fluoroquinolone, phenicol and tetracycline resistance of sequenced isolates.

	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
Fluoroquinolone and aminoglycoside resistance										
<i>aac(3)-IIa</i>	-	-	+	-	+	-	-	+	-	-
<i>aac(3)-IIId</i>	-	-	-	+	-	-	-	-	+	-
<i>aac(6')Ib-cr</i>	-	+	+	+	+	-	+	+	+	+
Phenicol resistance										
<i>catB3</i>	-	+	+	-	+	-	+	+	+	+
<i>cmlA1</i>	-	-	-	+	-	-	-	-	-	-
Tetracycline resistance										
<i>tet(A)</i>	-	-	+	-	-	-	-	+	-	-
<i>tet(B)</i>	+	+	-	+	+	+	-	-	+	+

Table 8. Comparative analysis of trimethoprim, aminoglycoside, macrolide and sulphonamide antibiotic resistance of sequenced isolates.

	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
Trimethoprim resistance										
<i>dfrA1</i>	-	-	-	-	-	-	+	-	-	-
<i>dfrA12</i>	-	-	-	-	-	-	+	-	-	-
<i>dfrA17</i>	-	+	-	+	+	+	-	-	+	+
Macrolide, Lincosamide and Streptogramin B resistance										
<i>mph(A)</i>	-	+	-	+	+	+	+	-	-	+
<i>mph(E)</i>	-	-	-	+	-	-	-	-	-	-
<i>msr(E)</i>	-	-	-	+	-	-	-	-	-	-
Sulphonamide resistance										
<i>sul1</i>	-	+	-	+	+	+	+	-	+	+
<i>sul2</i>	-	-	-	+	+	+	-	-	+	-
Aminoglycoside resistance										
<i>aadA1</i>	-	-	-	+	-	-	+	-	-	-
<i>aadA2</i>	-	-	-	-	-	-	+	-	-	-
<i>aadA5</i>	-	+	-	+	+	+	-	-	+	+
<i>strA</i>	-	-	-	+	+	+	-	-	+	-
<i>strB</i>	-	-	-	+	+	+	-	-	+	-
<i>armA</i>	-	-	-	+	-	-	-	-	-	-

Table 9. Comparative analysis of rifampicin, fosfomycin and quinolone antibiotic resistance of sequenced isolates.

	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
Rifampicin resistance										
<i>arr-2</i>	-	-	-	+	-	-	-	-	-	-
Fosfomycin resistance										
<i>fosA</i>	-	-	-	-	-	-	+	-	-	-
Quinolone resistance										
<i>qnrS1</i>	-	-	-	-	-	-	-	-	+	-

Table 10. Antimicrobial resistance gene patterns of all isolates.

Resistance Pattern	Isolate ID
SHV, OXA, CTX-M, CTX-M-15	ECC 147
TEM, OXA, CTX-M, CTX-M-15	ECC 149, ECC 161, ECC 186, ECC 187, ECC 194
SHV, CTX-M, CTX-M-15	ECC 319
TEM, CTX-M, CTX-M-15	ECC 203
OXA, CTX-M, CTX-M-15	ECC 153, ECC 157, ECC 167, ECC 173, ECC 174, ECC 188, ECC 190, ECC 202
TEM, OXA, CTX-M	ECC 314, ECC 605
CTX-M, CTX-M-15	ECC 200
TEM, CTX-M	ECC 199, ECC 312, ECC 322, ECC 394, ECC 396
TEM	ECC 205

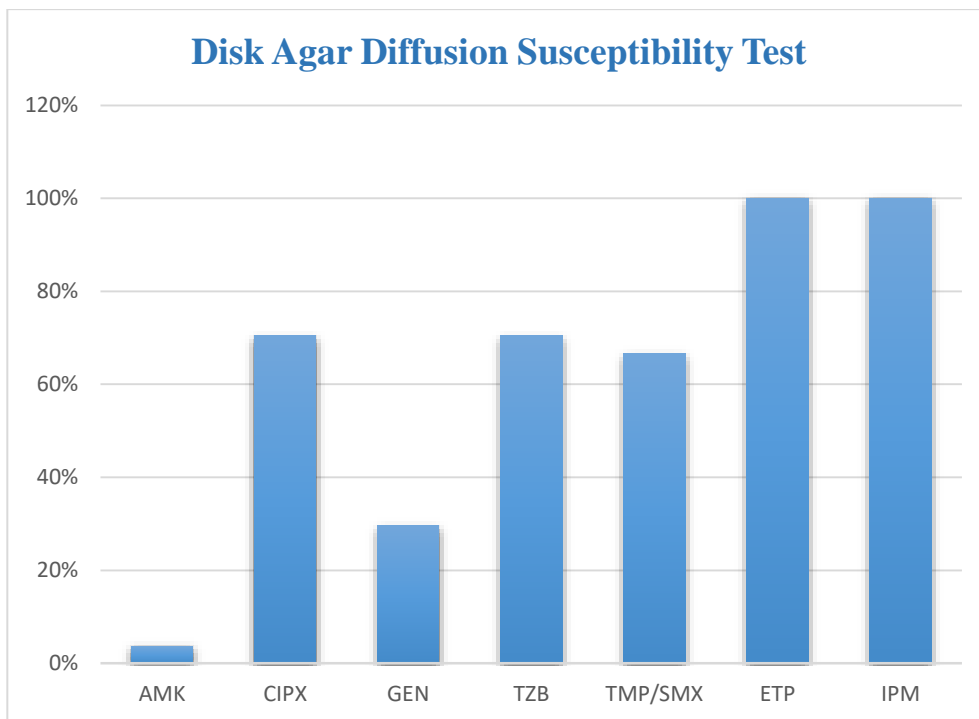


Figure 3. Bar diagram showing the percentage resistance of all 27 *E. coli* isolates to the tested antibiotics by the disk agar diffusion method. AMK- Amikacin, CIPX- Ciprofloxacin, GEN- Gentamicin, TZB- Tazobactam (β -lactamase inhibitor), TMP/SMX- Trimethoprim/Sulfamethoxazole, ETP- Ertapenem and IPM- Imipenem.

4.6. Virulence Factors:

Virulence factors encoding genes were detected using the CGE VirulenceFinder 1.2 tool (Table 11). *gad* was detected in all the sequenced isolates, which is a glutamate decarboxylase that aids *E. coli* to survive under highly acidic conditions (Capitani et al., 2003). Adhesins were also detected in all the sequenced genomes. As for toxins, *senB* was identified in ST-405 isolates (ECC153, ECC 157, ECC 173 and ECC 188) and the Enterotoxigenic *E. coli* heat-stable enterotoxin 1 (EAST1) encoded by *astA* (Silva et al., 2014) was found in sequence types ST-205 (ECC 174) and ST-1284 (ECC 167). *iss* encodes an outer membrane lipoprotein which enhances serum resistance (Johnson et al.,

2008) and it was detected in sequence types ST-648 (ECC 149), ST-1284 (ECC 167) and ST-617 (ECC 202). Finally, *kpsM* capsular gene was detected in sequence type ST-648 (ECC 149), and siderophores *iutA* and *iucD* were identified in sequence types ST-1284 (ECC 167) and ST-617 (ECC 202).

Table 11. Genes encoding the most important virulence factors on the *E. coli* genomes.

MLST	Isolate ID	Serotypes	Phylogenetic group	Virulence factor genes
ST205	ECC 174	O100:H12	B1	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹ , <i>lpfa</i> ²), toxins (<i>astA</i> ³), serum resistance-associated (<i>traT</i> ⁴), pathogenicity island marker (<i>malX</i> ⁵) and increased acid resistance (<i>gad</i> ⁶)
ST405	ECC 153 ECC 157 ECC 173 ECC 188	O102:H6	D	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> , <i>air</i> ¹), toxins (<i>senB</i> ³), pathogenicity island marker (<i>malX</i> ⁵), increased acid resistance (<i>gad</i> ⁶) and expression of invasion genes (<i>eilA</i> ⁷)
ST410	ECC 194	O8:H9	A	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹ , <i>lpfa</i> ²), pathogenicity island marker (<i>malX</i> ⁵) and increased acid resistance (<i>gad</i> ⁶)
ST448	ECC 161	O8:H8	B1	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹ , <i>lpfa</i> ²), pathogenicity island marker (<i>malX</i> ⁵) and increased acid resistance (<i>gad</i> ⁶).
ST617	ECC 202	O89:H10	A	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹), siderophores (<i>iucD</i> , <i>iutA</i> ⁸), serum resistance-associated (<i>traT</i> ⁴), pathogenicity island marker (<i>malX</i> ⁵), increased serum resistance (<i>iss</i> ⁹) and increased acid resistance (<i>gad</i> ⁶)

ST648	ECC 149	H6	D	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹ , <i>lpfa</i> ² , <i>air</i> ¹), serum resistance-associated (<i>traT</i> ⁴), pathogenicity island marker (<i>malX</i> ⁵), increased serum resistance (<i>iss</i> ⁹), increased acid resistance (<i>gad</i> ⁶), expression of invasion genes (<i>eilA</i> ⁷) and capsular gene (<i>kpsM II</i> ¹⁰)
ST1284	ECC 167	O89:H21	A	Adhesins (<i>fimH</i> , <i>fimB</i> , <i>fimA</i> ¹ , <i>lpfa</i> ²), toxins (<i>astA</i> ³), siderophores (<i>iucD</i> , <i>iutA</i> ⁸), serum resistance-associated (<i>traT</i> ⁴), pathogenicity island marker (<i>malX</i> ⁵), increased serum resistance (<i>iss</i> ⁹), increased acid resistance (<i>gad</i> ⁶) and capsular gene (<i>capU</i> ¹¹)

¹Harrington et al., 2009

² Baranzoni et al., 2016

³ Silva et al., 2014

⁴Nolan et al., 2003

⁵Östblom et al., 2011

⁶ Capitani et al., 2003

⁷Hüttener et al., 2014

⁸ Gao et al., 2012

⁹Johnson et al., 2008

¹⁰Johnson et al., 2004

¹¹Afset et al., 2006

4.7. Plasmids:

All sequenced genomes matched with different plasmid incompatibility groups including IncA/C2, IncFIA, IncFIB, IncFII, IncI1, IncQ1, IncX3 and IncX4. Col(BS512) was detected in ECC 167 and ECC 194 with ECC 194 additionally carrying ColKP3, while Col156 was found in ECC 157 and ECC 173 (Table 12).

Table 12. Comparative analysis of plasmids in sequenced isolates.

Plasmids	ECC 149	ECC 153	ECC 157	ECC 161	ECC 167	ECC 173	ECC 174	ECC 188	ECC 194	ECC 202
IncFIA		+	+	+	+	+	+	+	+	+
IncFIB	+	+	+	+	+	+	+	+	+	+
IncFII	+	+	+	+	+	+	+	+		+
IncI1			+		+	+	+	+		+
IncX3									+	
IncX4							+			
IncQ1				+					+	
IncA/C2				+						
Col156			+			+				
Col(BS512)					+				+	
ColKP3									+	
Col(MG828)										+
ColpVC		+								
p0111	+	+				+		+		

4.8. Phages:

Several multi-genic regions were identified as being associated with known bacteriophages. *stx2* converting I phage was detected in ECC 149 and ECC 194, and *stx2* converting 1717 phage in ECC 153, ECC 157, ECC 161 and ECC 167. ECC 149 also carried the *Burkholderia cenocepacia* BcepMu phage. The lysogenic phages P1 (ECC 153, ECC 157, ECC 167, ECC 173, ECC 188 and ECC 202) and P4 (ECC 161, ECC 167 and ECC 202) were also identified. Finally, the temperate phage P2 was detected in ECC 153 and ECC 157 and some other phages including: *mEp460*, *fiAA91*, *ss*, *P88*, *Fels 2*, *HK629*, *HK630*, *SfII*, *SFIV*, *SfV*, *Sf6* and *SSU5* were seen among the sequenced genomes.

4.9. Insertion Sequences:

Figure 4 represents all insertion sequence families of ECC 149. The rest of the genomes can be found in ANNEX I.

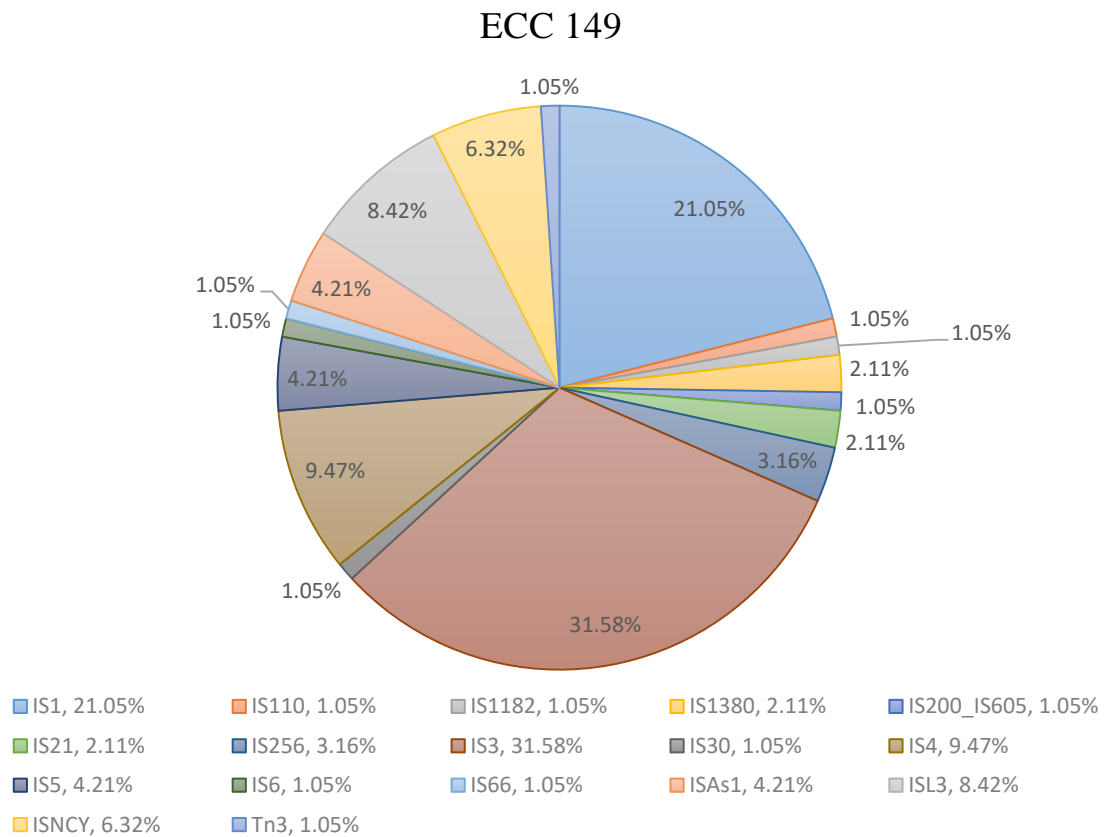


Figure 4. A pie chart demonstrating the different percentages of major insertion sequence families for ECC 149.

4.10. Phylogenetic Analysis:

A concatenated maximum likelihood tree was generated to establish the phylogenetic relatedness of the sequenced isolates. ST131 and non-ST131 belonging to B2

phylogenetic group were used as reference genomes. Phylogenetic analysis showed the grouping of ECC 153, ECC 157, ECC 173 and ECC 188 into the same clade with reference strain UMN026, while ECC 161 and ECC 174 grouped with SE11 (Figure 5). ECC 167 and ECC 202 were clustered together and in close association with MG1655, whereas ECC 194 was clustered with APECO78.

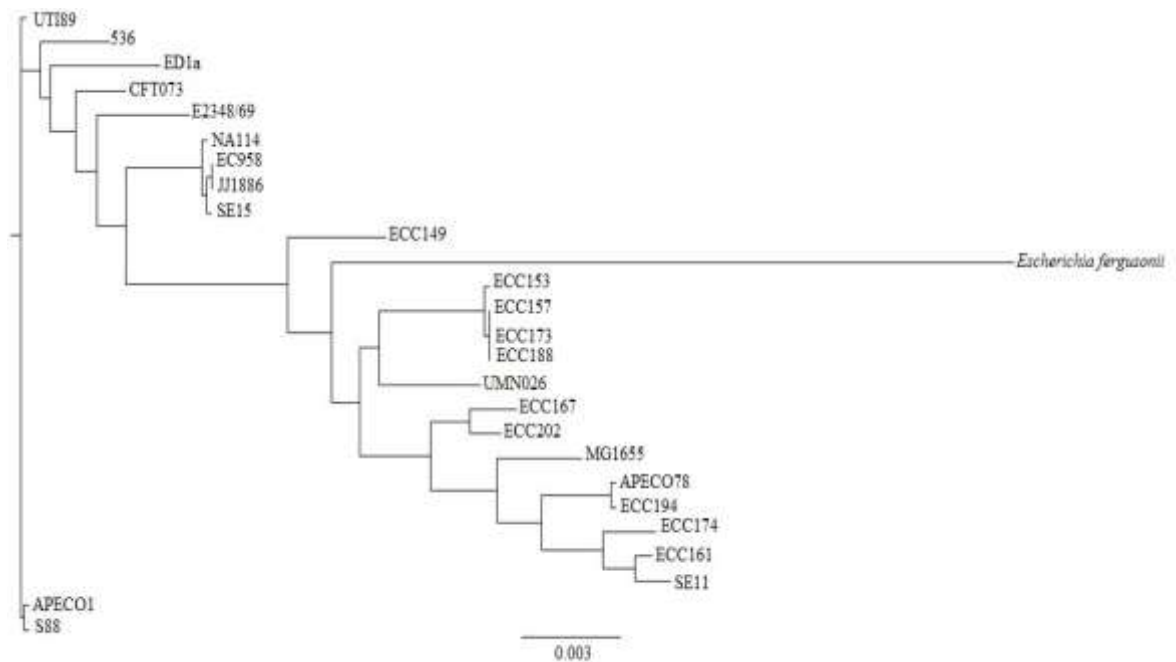


Figure 5. Maximum likelihood tree generated using PhyloSift. NCBI was used to download the genomes. PhyloSift concatenates the alignments of 37 elite markers. The maximum likelihood tree was then inferred using FigTree. *E. coli* belonging to B2 phylogenetic group and non-ST131 genomes used: UT189, CFT073, ED1a, E2348/69, S88, APEC O1 and 536. ST131 reference genomes: NA114, SE15 and EC958

Chapter Five

DISCUSSION

One of the most common Gram-negative bacterial pathogens is ExPEC, which affects all age groups and causes a broad range of clinical diseases (Poolman et al., 2015). Distinct from intestinal pathogens and commensals, ExPEC causes infections of the bloodstream, respiratory tract, urinary tract, peritoneum and cerebrospinal fluid. ExPEC associated infections can be both community and hospital acquired (Bajaj et al., 2016), and the increased occurrence of ESBL- and carbapenemase producing ExPEC isolates are a serious threat to the public health (Abraham et al., 2014). Multidrug resistant *E. coli* pose a significant challenge to manage and control the spread of infections (Poolman et al., 2016). ExPEC isolates have acquired resistance towards nearly all antibiotic classes (Nicolas-Chanoine et al., 2014) with the pandemic potential of these strains particularly in Lebanon being of great concern. In this work we examined the genomic attributes of 27 carbapenem resistant ExPEC isolates.

5.1. Antibiotic resistance:

An ideal method for investigating and tracking antibiotic resistant bacteria is via MLST typing, which also demonstrates the genetic variety between different multidrug resistant ExPEC lineages (Hazen et al., 2014). The most notorious ExPEC pathogen is the virulent clonal group B2-O25:H4-ST-131, but it should be acknowledged that beside B2-ST-131, there are group D strains that constitute around two thirds of all ExPEC infections and they include ST-405 and ST-648 (Ewers et al., 2014). Among our isolates, the most

common ST was ST-405, which was also detected in Europe, Asia and in Saudi Arabia (Algoribi et al., 2015). ExPEC's most prevalent ESBL is CTX-M-15 (Poolman et al., 2016) and clonal outbreaks of *bla*_{CTX-M-15} carrying *Enterobacteriaceae* were reported in Italy, Portugal, Norway, Austria, France, Spain, Tunisia, Canada, South Korea and the United Kingdom (Coque et al., 2008). In this study, 63 % of all *E. coli* isolates (Table 10), harbored the *bla*_{CTX-M-15} gene, which reemphasizes the countrywide spread of CTX-M-15 (Baroud et al., 2011). According to Coque et al. the worldwide dissemination of *bla*_{CTX-M-15} gene is attributed to the two epidemic strains of *E. coli* belonging to phylogenetic groups B2 (ST-131) and D (ST-405), and the gene is carried on an IncF plasmid, which facilitates the transfer of resistance among isolates (Coque et al., 2008). All the ST-405 types detected in this study carried the *bla*_{CTX-M-15} gene (Table 6) and the IncF plasmid (Table 12). Further, 75 % of this study's ST-405 isolates harbored the CMY-42 gene in addition to OXA-1. CMY-42 is an AmpC enzyme that differs from CMY-2 by a substitution in a single amino acid (Ser instead of Val at Ambler's position 211) (Feng et al., 2015). ST-405 *E. coli* clone harboring CTX-M-15, CMY-42 and OXA-1 was reported recently in India (Alm et al., 2015). Moreover, isolate ECC 149 (ST-648) was also a CTX-M producer (Table 6) and resistant to trimethoprim-sulfamethoxazole. Multidrug resistance was previously reported to be common among ST-648 isolates particularly towards non-lactams such as levofloxacin, trimethoprim-sulfamethoxazole, minocycline and gentamicin (Sherchan et al., 2015). ST-648 is widespread being detected in several geographical regions including Asia, Africa, North and South America and Europe (Sherchan et al., 2015).

Other multidrug resistant ST types were also detected in this study namely ST-410 (CC23), ST-617 (CC10) and ST-1284 (CC10) (Table 5). Peirano et al. (2011) showed that ST-410 isolated from Rio de Janeiro was positive for CTX-M-15, *aac(6')-Ib-cr*, OXA-1 and TEM-1 β -lactamases (Peirano et al., 2011), all of which were also detected in ECC 194 (ST-410) (Tables 6 and 7). Although all the isolates in this study were carbapenem resistant, only ECC 194 (ST-410) harbored at the molecular level a carbapenem resistant determinant, *bla_{OXA-181}* gene. *bla_{OXA-181}* gene was not previously detected in Lebanon, and differs from OXA-48 by four amino acid substitutions (Potron et al., 2011), but had a similar β -lactamase hydrolysis spectrum, which included penicillins and carbapenems (Liu et al., 2015). The *bla_{OXA-181}* originated from an environmental bacterium, *Shewanella xiamenensis*. It was first detected in *Klebsiella pneumoniae* and *Enterobacter cloacae* isolated from India, and was later found in several species of *Enterobacteriaceae* in other countries including: France, Norway, Oman, Canada, Romania, United Kingdom, New Zealand, Singapore, Sri Lanka, South Africa, Bangladesh and Nepal (Liu et al., 2015). *bla_{OXA-181}* was primarily detected on an IncT plasmid (84 kb) or ColE2 type plasmid (7.6 kb), with a more recent study reporting its detection on an IncX3 plasmid (Liu et al., 2015). ECC 194 (ST-410) was IncX3 positive (Table 12), and the detection of *bla_{OXA-181}* in this isolate and on a self-transmissible IncX3 plasmid is alarming due to the potential of rapid dissemination (Liu et al., 2015). In China, strain WCHEC14828 was also shown to carry the *bla_{OXA-181}* gene on the IncX3 plasmid along with *bla_{OXA-1}* (non-ESBL oxacillinase gene), *bla_{CTX-M-15}* (ESBL gene), *bla_{CMY-2}* (plasmid-borne AmpC gene), *bla_{TEM-1b}* (non-ESBL β -lactamase gene), *aac(6')-Ib-cr* (codes for aminoglycoside acetyltransferase which results in low level activity against fluoroquinolones), *qnrS1* (low level of resistance to fluoroquinolones), *bla_{ampC}* (chromosomally carried AmpC gene) and *tetA* (tetracycline

resistance gene) (Liu et al., 2015). ECC 194 (ST-410) had the same set of resistant determinants except for *bla*_{ampC} and *tetA*.

Mshana et al. revealed year 2016 that ST-617 *E. coli* isolates harbored *bla*_{CTX-M-15}, *bla*_{OXA-1}, *bla*_{TEM-1B}, *aac(6')Ib-cr*, *aac(3)-IIa*, *aadA5*, *strA* and *strB* as well as IncFIA, IncFIB and IncFII, which was similar to ECC 202 (ST-617) with the exception of *bla*_{TEM-1B}, *aac(3)-IIa*, *strA* and *strB*. ST-617 (CC10) was also reported in Belgium, with high prevalence in Nigeria (Brolund et al., 2014). Usually, CC10 and 23 are linked to multidrug resistance and in particular ciprofloxacin resistance among sepsis and UTI clones (Giufre et al., 2012), which was consistent with the results of this study with all the isolates being ciprofloxacin resistant (Table 6).

CC10 and CC23 were both frequently associated with clinical ESBL producing *E. coli* (Oteo et al. 2009). ST-1284 (CC10) *E. coli* isolates were CTX-M-15 producers (Novais et al., 2012), which was similar to ECC 167 (ST-1284) (Table 11). *E. coli* ST-1284 was previously isolated from raw milk in Algeria (Yaici et al., 2016) and commercial swine in Brazil (Silva et al., 2016). However, to the best of our knowledge, this is the first isolation of ST-1284 from humans. Moreover, *E. coli* strain (WCHEC13-8) isolated in China, was closely related to ST-1284. WCHEC13-8 belonged to ST-3835 with an allelic profile of: 10-4-4-411-8-13-73, differing from ST-1284 by a single allele (10-4-4-8-8-13-73) (Feng et al., 2015). It is noteworthy, that ST-3835 had some common resistance determinants as that of ECC 167 (ST-1284), in addition to *bla*_{NDM-1}, *bla*_{SHV-12} and *bla*_{ampC}. Finally, ST-3835 was reported in several other countries including: Germany (dog), Spain (human) and Korea (human) (Feng et al., 2015).

ST-205 reported by Aizawa et al. were identified as CMY-2 producing *E. coli* isolates (Aizawa et al., 2014), whereas ECC 174 (ST-205) was a CMY-42 producer rather than CMY-2 (Table 6) (Mshana et al., 2016). On the other hand in Spain, ST-448 carbapenem resistant *E. coli* was detected and it showed resistance towards all antibiotics including: imipenem, meropenem, ertapenem, gentamicin, amikacin, ciprofloxacin and trimethoprim-sulfamethoxazole as well as towards the β -lactamase inhibitor tazobactam (Porres-Osante et al., 2014), which was similar to ECC 161 (ST-448) resistance pattern (Table 5).

Lebanon, similar to many countries worldwide, is facing a serious threat with an increase in carbapenem resistant *Enterobacteriaceae* (El-Herte et al., 2012). Carbapenem resistance in *Enterobacteriaceae* can be linked to at least one of the following: acquisition of carbapenem resistance genes, alteration of penicillin binding proteins due to mutations and/or reduction of outer membrane permeability as a cause of porin loss in addition to overexpression of an ESBL or an AmpC enzyme (Adler et al., 2013). The fact that all isolates in this study were carbapenem resistant and yet only one carried a carbapenem resistant determinant could be possibly attributed to several other resistance determinants. Generally carbapenems are not sensitive to the CTX-M enzymes, but overexpression of CTX-M-15 along with decreased membrane permeability, could be linked to an increase in carbapenem resistance (D'Andrea et al., 2013). Similarly, Adler et al. showed that the overexpression of OXA-1 (carried by all sequenced isolates except ECC 161) and TEM-1 (ECC 161 and ECC 194) led to higher catalytic activity towards carbapenems when combined with the loss of OmpF and OmpC porins (Adler et al., 2013). All sequenced isolates in this study didn't have the OmpF and OmpC porins. Whereas some class D β -

lactamases confer high resistance to carbapenems, others like OXA-2 and OXA-10 (ECC 161) β -lactamases were not considered as major elements contributing to carbapenem resistance. However, Antunes et al. demonstrated in year 2014 that in *E. coli*, OXA-48 and OXA-58 carbapenemases showed the same level of resistance to carbapenems as that seen with OXA-2 and OXA-10. Additionally, plasmid mediated AmpC β -lactamase in *E. coli* encodes resistance to third generation cephalosporins, and when combined with loss of porins the net outcome was carbapenem resistance with the most common two enzymes involved being *bla*_{CMY-2} (ECC 194) and *bla*_{CMY-42} (ECC 157, ECC 167, ECC 173, ECC 174, ECC 188 and ECC 202) (Rocha et al., 2015). Thus, carbapenem resistance with the lack of detectable genes could be attributed to the other set of ESBLs and β -lactamases (OXA-1, OXA-10, TEM-1, CTX-M-15, CMY-2 and CMY-42) detected in the sequenced genomes. This potential resistance reservoir, which caused reduced susceptibility to carbapenems, is a major concern knowing that carbapenems are the antibiotics of choice in treating ESBL- and β -lactamase producing *E. coli* (Adler et al., 2013).

5.2. Virulence:

Our study revealed that among the 27 isolates, the distribution of phylogenetic groups was as follows: five (18.5 %) belonged to group A, three (11.1 %) belonged to group B1, six (22.2 %) belonged to group B2 and 13 (48.2 %) belonged to group D (Table 4). Although it seems that ExPEC strains derive from diverse origins, phylogenetic studies established that these bacteria fall into four main phylogenetic groups (A, B1, B2 and D) with the most virulent strains belonging to B2 and D (Bashir et al., 2012).

Typically, ExPEC strains carry specialized virulence factors like adhesins, polysaccharide coats, invasins, toxins and iron acquisition systems (Sannes et al., 2004) that enable

ExPEC isolates to colonize, injure host tissues and avoid the host's immune response (Blanco et al., 2013). Various studies have suggested that phylogenetic groups A and B1 carry fewer virulence determinants, but are more resistant, whereas phylogenetic groups B2 and D possess and express more pathogenicity islands and virulence factors (Chakraborty et al., 2015).

Adhesins, are an important group of virulence factors expressed by ExPECs to avoid removal and to enable tissue tropism. ExPEC's most important adhesins are the fimbriae, which are organelles belonging to the chaperone-usher subclass. This class includes type 1, S, F1C and P fimbriae (Klemm et al., 2010). Type 1 fimbriae are extremely conserved and common among commensals as well as UPEC isolates. Particularly, this virulence factor is crucial in the establishment of a UTI (Wiles et al., 2008). In this study, isolates of phylogenetic groups A, B1, B2 and D had equal distribution of *fimA*, *fimB* and *fimH*. All were classified as type 1 fimbriae, which mediates mannose-specific binding to receptors on host cells. *fimA*, *fimB*, *fimC* and *fimD* genes play a role in fimbriae synthesis, while *fimB* and *fimE* are regulatory genes that control type 1 fimbriae phase variation. Additionally, *fimF*, *fimG*, and *fimH* were needed in longitudinal regulation as well as in adhesion (Antão et al., 2009). Although *fim* is found in all *E. coli* (pathogens and commensals), its expression and function varies significantly between commensal strains and pathogens as well as between clinical isolates from diverse syndromes. UTI associated clinical isolates express a form of *fimH* that enables it to attach to monomannose and trimannose receptors, which is an adaptive variation that aids it to colonize the bladder whereas commensal strains express a variant of *fimH* that binds trimannose receptors resulting in gut colonization (Johnson et al., 2002). *lpfa* is another adhesin detected in 66.7

% of the isolates in this study belonging to phylogroup A, all of phylogroup B1 and 20 % of phylogroup D. *lpfa* plays a role in host cell adhesion (Baranzoni et al., 2016), and is the second most common virulence factor among phylogenetic groups A and B1 (Campos et al., 2008). On the other hand, *air*, an adhesin involved in aggregation, colonization and invasion, was found in all the isolates belonging to phylogroup D (Harrington et al., 2009), which were additionally positive for *eilA* gene. EilA, a homolog of HilA, plays a role in the expression of invasion genes (Hüttener et al., 2014).

Proteases and bacterial toxins are important virulence factors. *senB*, a plasmid encoded enterotoxin, and *astA* (Enteroaggregative *E. coli* heat stable enterotoxin 1) were the only detected toxins in the sequenced genomes. It is noteworthy that both *senB* and *astA* genes are frequently correlated with diarrhoeagenic pathovars rather than ExPEC strains (Sousa, 2003), but it was established that generally, some of the toxin encoding genes have the potential to transfer from diarrhoeagenic *E. coli* to ExPEC strains and particularly to UPEC (Mirzarazi et al., 2015). *senB* was found in 80 % of the isolates in this study belonging to phylogenetic group D, whereas, *astA* was found in 50 % of B1 and 33.3 % of A. In harmony with our results, Mirzarazi et al. (2015) showed that *senB* was most prevalent in phylogroup D and *astA* in B2 followed by B1 and A.

Moreover, serum resistance associated protein (*iss*) was detected in ECC 149, ECC 167 and ECC 202, the aerobactin siderophore receptor ferric aerobactin uptake (*iutA*) in ECC 167 and ECC 202 and hydroxamate siderophore aerobactin (*iucD*) in ECC 167 and ECC 202. *iss* is more frequently associated with isolates of phylogenetic group D than A (Hiki et al., 2014), but in this study it was detected in 20 % of isolates in phylogenetic group D and 66.7 % of isolates in phylogenetic group A. *iutA* and *iucD* were detected in 66.7 % of

isolates belonging to phylogenetic group A, with both siderophores being associated with phylogenetic group B2 followed by A (Staji et al., 2016). Iron is vital for bacterial growth and in low iron environments such as in the urinary bladder, the acquisition of siderophore systems will be a great advantage during colonization (Watts et al., 2012). Aerobactin is a hydroxamate siderophore synthesized by the majority of pathogenic *E. coli* and is encoded by *iucABCD* genes. Once aerobactin is synthesized, it is taken up by the receptor protein encoded by *iutA* (Gao et al., 2012). Iron starvation is the instantaneous trigger for the activation of both of these components (Ling et al., 2013). Usually, these genes are only found in virulent strains and play an important role in pathogenesis particularly in the case of UPEC (Gao et al., 2012). As for the rest of the virulence factors, *malX* (pathogenicity island marker) and *gad* (glutamate decarboxylase) were detected in all isolates of phylogenetic groups A, B1 and D. *traT* (serum resistance-associated) (Riley, 2014) is most commonly associated with phylogroups A (Derakhshandeh et al., 2015), which was the case in our isolates as it was found in 66.7 % of phylogroup A, 50 % of phylogroup B1 and 20 % of phylogroup D. Finally, *kpsM II* (group 2 capsule) and *capU* (capsular gene) were identified by next generation sequencing in phylogenetic groups D (20 %) and B1 (33.3 %) respectively. *kpsM II* gene is most frequently linked to phylogenetic groups B2 and D (Petersen et al., 2009) and it aids in the protection of ExPEC against complement-mediated killing and phagocytosis (Johnson et al., 2004). Understanding how ExPEC isolates utilize their virulence factors in colonizing and persisting within a host will enable the customization of treatments based on the set of virulence genes present in the infecting strain (Wiles et al., 2008).

5.3. Plasmids:

Sequences correlating to IncA/C2, IncFIA, IncFIB, IncFII, IncI1, IncQ1, IncX3, IncX4, Col(BS512) and Col156 plasmids were identified, with incompatibility groups IncFIA and IncFIB being present in all isolates (Table 12). The widespread emergence of the *bla*_{CTX-M-15} gene, which is frequently interrelated with TEM-1, OXA-1 and *aac(6')-Ib-cr*, was mainly linked to IncF replicons (Carattoli, 2009). As previously established, among the isolates carrying IncF plasmids, all were positive for CTX-M-15, 30 % for TEM-1, 90 % for OXA-1 and 80 % for *aac(6')-Ib-cr* (Tables 6 and 7). CTX-M-15 carrying IncF plasmids are not exclusively for ST-131, but have also been detected in several other STs including ST-405, ST-410, ST-617 and ST-1284 (Carattoli, 2009 and Silva et al., 2016), all of which were in harmony with this study's results (Table 6 and 11). Furthermore, ECC 167 and ECC 194 carried the pBS512_2 plasmid, which is linked to type III secretion system (Villa et al., 2010). Type III secretion systems are used by pathogenic Gram-negative bacteria to subvert the signaling pathways of host cells and inject virulence proteins into their cytoplasm (Zhou et al., 2014). Moreover, isolates ECC 153, ECC161, ECC 173 and ECC 188 harbored the pRSB107 plasmid with an allelic profile of F2:A1:B1 which was associated with aerobactin (Woodford et al., 2009). IncX3 and IncX4 were found in ECC 194 and ECC 174, respectively. Resistance genes including ESBLs (particularly CTX-M), carbapenemases and plasmid encoded quinolone resistance (*qnrS1*) were previously reported to be localized on IncX plasmids (Lo et al., 2014). *qnrS1* is mainly associated with IncX3 plasmids (Dobiasova and Dolejska, 2016), which was the case for ECC 194. It is noteworthy that ECC 194 was the only isolate carrying a carbapenemase gene *bla*_{OXA-181}, and this is in harmony with the fact that it was the only

isolate harboring the ColKP3 plasmid. ColKP3 plasmid was initially isolated from *Klebsiella pneumoniae* and was shown to harbor the *bla*_{OXA-181} gene (Potron et al., 2011). Further, only isolates ECC 157, ECC 167, ECC 173, ECC 174, ECC 188 and ECC 202 that harbored the *bla*_{CMY-42} gene (encodes AmpC enzyme) were found to also be positive for the IncII plasmid which was in concordance with Feng et al. (2015).

5.4. Phages and Insertion sequences:

All ST-410 and ST-648 isolates in this study possessed *stx2* converting I phage. Furthermore, all ST-448 and ST-1284 as well as half ST-405 isolates carried the *stx2* converting 1717 phage. It is known that ExPEC strains carry the Shiga toxin gene (Wester et al., 2013). STEC (Shiga toxin producing *E. coli*) or closely related pathogens produce the Shiga toxin (*stx1*, *stx2* and *stx2* variants), which is encoded by temperate double stranded DNA lambdoid phages and which causes drastic human diseases including hemolytic uremic syndrome (Muniesa et al., 2004).

Resistance and virulence genes are fitness traits that are usually linked with transposable elements (Ho et al., 2016). The most common insertion sequence families detected among all our isolates were both IS1 and IS3 (ANNEX I). IS1 elements are correlated with heat stable toxins (Mahillon et al., 1998) and they are usually found flanking siderophore systems such as *iuc/iutABCD* (Ho et al., 2016).

5.5. Phylogenetic Analysis:

Phylogenetic analysis revealed that the isolates were mainly grouped with the reference strains based on their respective phylogenetic groups. Phylogenetic group D *E. coli* ST-405 isolates (ECC 153, ECC 157, ECC 173 and ECC 188) clustered together and with the

reference ExPEC strain UMN026 (ST-69) (Doumith et al., 2015) of phylogenetic group D, but separately from phylogenetic group D ECC 149 (ST-648). The clade containing group B2 ST-131 strains (SE15, EC958, NA114 and JJ1886) was also located separately from phylogenetic group D isolates (Forde et al., 2014). Moreover, ECC 174 and ECC 161 were grouped together with SE11, which was supported by the fact that they all belonged to phylogenetic group B1 (Oshima et al., 2008). On the other hand, ECC 194 and APECO78 clustered together although they belonged to different phylogenetic groups A and B1, respectively. However, this could be attributed to both having the same ST type (ST-410 and CC23) (Wyrsh et al., 2015). ECC 194 was also phylogenetically close to reference strain MG1655 and to isolates ECC 167 and ECC 202 with all being of phylogenetic group A (Forde et al., 2014).

5.6. Future Work:

The pandemic potential of ExPEC strains especially in Lebanon is of great concern and poses a serious threat to the public health. A significant threat is the dispersal of transmissible plasmids carrying different sets of resistance and virulence determinants, thus the detection and identification of such replicons is critical. Moreover, since carbapenem resistant ExPEC strains are now widespread and being linked to community and hospital acquired infections, further investigations are needed to better understand their dispersal, molecular epidemiology and pathogenicity. Additionally, the role of recombination and mobile genetic elements in the diversification and spread of carbapenem resistant ExPEC clonal lineages should be better studied. Finally, in Lebanon efforts should be focused at implementing active antibiotic surveillance and infection control programs to limit and avoid outbreaks of carbapenem resistant *E. coli*.

Chapter Six

CONCLUSION

- This study is the first detection of *bla*_{OXA-181} encoding *E. coli* from Lebanon in addition to it being the first comprehensive genome-wide comparative analysis of carbapenem resistant ExPEC isolates.
- The distribution of phylogenetic groups was as follows: five (18.5 %) belonged to group A, three (11.1 %) belonged to group B1, six (22.2 %) belonged to group B2 and 13 (48.2 %) belonged to group D.
- The average sequence size of the ten isolates was 5.2 Mb with an average of 224–342 contigs and G+C % content of 50.49 % – 50.79 %.
- The isolates were derived from 7 different lineages: ST-405 (ECC 153, ECC 157, ECC 173 and ECC 188), ST-205 (ECC 174), ST-410 (ECC 194), ST-448 (ECC 161), ST-617 (ECC 202), ST-1284 (ECC 167) and ST-648 (ECC 149).
- Out of the 27 isolates, 59.6 %, 51.9 %, 7.4 % , 88.9 % and 63 % carried the *bla*_{OXA}, *bla*_{TEM}, *bla*_{SHV}, *bla*_{CTX-M} and *bla*_{CTX-M-15} genes respectively.
- Analysis revealed that the sequenced isolates harbored different β-lactamase genes including *bla*_{OXA-1}, *bla*_{OXA-10}, *bla*_{OXA-181}, *bla*_{TEM-1b}, *bla*_{CTX-M-15}, *bla*_{CMY-2}, *bla*_{CMY-42} in addition to *aac(6')Ib-cr* gene which confers tetracycline and aminoglycoside resistance among others.
- Although all isolates were carbapenem resistant according to disk agar diffusion, only ECC194 (ST-410) was positive for carbapenem resistance gene *bla*_{OXA-181}.

This can be possibly attributed to loss of OmpF and OmpC porins in all sequenced isolates in combination with increased production or expression of certain ESBLs.

- Adhesins *fimA*, *fimB* and *fimH* were also identified in all isolates.
- Other virulence factors detected include: *air*, *lpfa*, *senB*, *astA*, *iss*, *malX*, *capU*, *kpsM*, *iutA*, *iucD* and *traT*.
- All isolates had results that matched with different incompatibility groups including IncA/C2, IncFIA, IncFIB, IncFII, IncI1, IncQ1, IncX3 and IncX4.
- Col(BS512) was detected in ECC 167 and ECC 194 with ECC 194 additionally carrying ColKP3, while Col156 was found in isolates ECC 157 and ECC 173.
- *stx2* converting I phage was seen in isolates ECC 149 and ECC 194. Additionally, *stx2* converting 1717 phage was identified in isolates: ECC 153, ECC 157, ECC 161 and ECC 167.
- The most common insertion sequence families detected among all our isolates were both IS1 and IS3 (ANNEX I).
- Phylogenetic analysis revealed that the isolates were mainly grouped with the reference strains based on their respective phylogenetic groups.
- The pandemic potential of ExPEC strains especially in Lebanon is of great concern and poses a serious threat to the public health. A significant threat is the dispersal of transmissible plasmids carrying different sets of resistance and virulence determinants, thus the detection and identification of such replicons is critical.
- Finally, in Lebanon efforts should be focused at implementing active antibiotic surveillance and infection control programs to limit and avoid outbreaks of carbapenem resistant *E. coli*.

BIBLIOGRAPHY

- Abraham, S., Wong, HS., Turnidge, J., Johnson, JR., Trott, DJ. (2014). Carbapenemase-producing bacteria in companion animals: a public health concern on the horizon. *The Journal of Antimicrobial Chemotherapy*, 69(5), 1155-7. DOI: 10.1093/jac/dkt518.
- Adler, M., Anjum, M., Andersson, DI., Sandergreen L. (2013). Influence of acquired β -lactamases on the evolution of spontaneous carbapenem resistance in *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*, 68(1), 51-59. DOI: 10.1093/jac/dks368.
- Afset, JE., Bruant, G., Brousseau, R., Harel, J., Anderssen, E., Bevanger, L., Bergh, K. (2006). Identification of Virulence Genes Linked with Diarrhea Due to Atypical Enteropathogenic *Escherichia coli* by DNA Microarray Analysis and PCR. *Journal of Clinical Microbiology*, 44(10), 3703–3711. DOI: 10.1128/JCM.00429-06.
- Aizawa, J., Neuwirt, N., Barbato, L., Neves, PR., Leigue, L., Padilha, J., Pestana de Castro, AF., Gregory, L., Lincopan, N. (2014). Identification of fluoroquinolone-resistant extended-spectrum β -lactamase (CTX-M-8)-producing *Escherichia coli* ST224, ST2179 and ST2308 in buffalo (*Bubalus bubalis*). *Journal of Antimicrobial Chemotherapy*, 69(10):2866-9. DOI: 10.1093/jac/dku218.
- Alghoribi, MF., Gibreel, TM., Farnham, G., Al Johani, SM., Balkhy, HH., Upton, M. (2015). Antibiotic-resistant ST38, ST131 and ST405 strains are the leading uropathogenic *Escherichia coli* clones in Riyadh, Saudi Arabia. *Journal of Antimicrobial Chemotherapy*, 70(10), 2757-62. DOI: 10.1093/jac/dkv188.
- Alm, RA., Johnstone, MR., Lahiri, SD. (2015). Characterization of *Escherichia coli* NDM isolates with decreased susceptibility to aztreonam/avibactam: role of a novel insertion in PBP3. *The Journal of Antimicrobial Chemotherapy*, 70(5), 1420-8. DOI: 10.1093/jac/dku568.

- Antão, EM., Wieler, LH., Ewers, C. (2009). Adhesive threads of extraintestinal pathogenic *Escherichia coli*. *Gut Pathogens*, 10(1), 22. DOI: 10.1186/1757-4749-1-22.
- Antunes, NT., Lamoureaux, TL., Toth, M., Stewart, NK., Frase, H., Vakulenko, SB. (2014). Class D β -Lactamases: Are They All Carbapenemases? *Antimicrobial Agents and Chemotherapy*, 58(4), 2119-25. DOI: 10.1128/AAC.02522-13.
- Alteri, CJ., Smith, SN., Mobley, HLT. (2009), Fitness of *Escherichia coli* during Urinary Tract Infection Requires Gluconeogenesis and the TCA Cycle. *PLoS Pathogen*, 5(5). DOI: 10.1371/journal.ppat.1000448.
- Araj, GF., Ibrahim, GY. (2008). Tigecycline in vitro activity against commonly encountered multidrug-resistant Gram-negative pathogens in a Middle Eastern country. *Diagnostic Microbiology and Infectious Disease*, 62(4), 411–415.
- Araj, GF., Uwaydah, MM., Alami, SY. (1994). Antimicrobial susceptibility patterns of bacterial isolates at the American University Medical Center in Lebanon. *Diagnostic Microbiology and Infectious Disease*, 20(3), 151-8.
- Asadulghani, M., Ogura, Y., Ooka, T., Itoh, T., Sawaguchi, A., Iguchi, A., Nakayam, K., Hayashi, T. (2009). The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathogen*, 5(5):e1000408. DOI: 10.1371/journal.ppat.1000408.
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Zagnitko, O. (2008). The RAST server: Rapid annotations using subsystems technology. *BioMed Central Genomics*, 9, 75. DOI: 10.1186/1471-2164-9-75.
- Bajaj, P., Singh, N. S., & Viridi, J. S. (2016). *Escherichia coli* β -Lactamases: What Really Matters. *Frontiers in Microbiology*, 7, 417. <http://doi.org/10.3389/fmicb.2016.00417>
- Baranzoni, G. M., Fratamico, P. M., Gangiredla, J., Patel, I., Bagi, L. K., Delannoy, S., ... Pepe, T. (2016). Characterization of Shiga Toxin Subtypes and Virulence Genes

in Porcine Shiga Toxin-Producing *Escherichia coli*. *Frontiers in Microbiology*, 7, 574. <http://doi.org/10.3389/fmicb.2016.00574>

Baroud, M., Araj, GF., Matar, GM. (2011). Spread of CTX-M-15 Extended Spectrum β -lactamases encoding genes among *Enterobacteriaceae* in the Middle Eastern Region. *The International Arabic Journal of Antimicrobial Agents*, 1(4). DOI: 10:3823/703.

Baroud M., Dandache, I., Araj, GF., Wakim, R., Kanj, S., Kanafani, Z., Khairallah, M., Sabra, A., Shehab, M., Dbaibo, G., Matar, GM. (2012). 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. *International Journal of Antimicrobial Agents*, 41(1), 75 – 79.

Bashir, S., Haque, A., Sarwar, Y., Ali, A., Anwar, MI. (2012). Virulence profile of different phylogenetic groups of locally isolated community acquired uropathogenic *E. coli* from Faisalabad region of Pakistan. *Annals of Clinical Microbiology and Antimicrobials*, 11, 23. DOI: 10.1186/1476-0711-11-23.

Ben-Ami, R., Rodríguez-Baño, J., Arslan, H., Pitout, JD., Quentin, C., Calbo, ES., Azap, OK., Arpin, C., Pascual, A., Livermore, DM., Garau, J., Carmeli, Y. (2009). A multinational survey of risk factors for infection with extended-spectrum beta-lactamase-producing *Enterobacteriaceae* in nonhospitalized patients. *Clinical Infectious Diseases*, 49(5), 682-90. DOI: 10.1086/604713.

Beyrouthy, R., Robin, F., Dabboussi, F., Mallat, H., Hamzé, M., Bonnet, R. (2014). Carbapenemase and virulence factors of *Enterobacteriaceae* in North Lebanon between 2008 and 2012: evolution via endemic spread of OXA-48. *Journal of Antimicrobial Chemotherapy*, 69(10), 2699-705. DOI: 10.1093/jac/dku181.

Blanco, J., Mora, A., Mamani, R., López, C., Blanco, M., Dahbi, G., Herrera, A., Marzoa, J., Fernández, V., de la Cruz, F., Martínez-Martínez, L., Alonso, MP., Nicolas-Chanoine, MH., Johnson, JR., Johnston, B., López-Cerero, L., Pascual, A., Rodríguez-Baño, J., Spanish Group for Nosocomial Infections (GEIH). (2013). Four Main Virotypes among Extended-Spectrum- β -Lactamase-Producing Isolates of *Escherichia coli* O25b:H4-B2-ST131: Bacterial, Epidemiological, and Clinical Characteristics. *Journal of Clinical Microbiology*, 51(10), 3358-67. DOI: 10.1128/JCM.01555-13.

- Bradford, P. A. (2001). Extended-Spectrum β -Lactamases in the 21st Century: Characterization, Epidemiology, and Detection of This Important Resistance Threat. *Clinical Microbiology Reviews*, 14(4), 933–951. <http://doi.org/10.1128/CMR.14.4.933-951.2001>.
- Brolund, A., Edquist, P.J., Mäkitalo, B., Olsson-Liljequist, B., Söderblom, T., Wisell, K.T., Giske, C.G. (2014). Epidemiology of extended-spectrum β -lactamase-producing *Escherichia coli* in Sweden 2007-2011. *Clinical Microbiology and Infection*, 20(6), O344-52. DOI: 10.1111/1469-0691.12413.
- Bush, K., Jacoby, G.A., Medeiros, A.A. (1995). A functional classification scheme for beta-lactamases and its correlation with molecular structure. *Antimicrobial Agents and Chemotherapy*, 39, 1211–1233.
- Campos, T.A., Carvalhães Lago, J., Nakazato, G., Guedes Stehling, E., Brocchi, M., Castro, A.F.P., Silveira, W.D. (2008). Occurrence of virulence-related sequences and phylogenetic analysis of commensal and pathogenic avian *Escherichia coli* strains (APEC). *Pesquisa Veterinária Brasileira*, 28(10), <http://dx.doi.org/10.1590/S0100-736X2008001000015>.
- Capitani, G., De Biase, D., Aurizi, C., Gut, H., Bossa, F., & Grütter, M. G. (2003). Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *The EMBO Journal*, 22(16), 4027–4037. <http://doi.org/10.1093/emboj/cdg403>
- Carattoli, A. (2009). Resistance Plasmid Families in *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*, 53(6), 2227–2238. DOI: 10.1128/AAC.01707-08.
- Carattoli, A., Zankari, E., Garcia-Fernandez, A., Volby Larsen, M., Lund, Villa, L., Hasman, H. (2014). Plasmid Finder and pMLST: In silico detection and typing of plasmids. *Antimicrobial Agents and Chemotherapy*, 58(7), 3895-3903. DOI: 10.1128/AAC.02412-14.
- Carrer A., Poirel, L., Yilmaz, M., Akan, O.A., Feriha, C., Cuzon, G., Matar, G., Honderlick, P., Nordmann, P. (2010). Spread of OXA-48-encoding plasmid in Turkey and beyond. *Antimicrobial Agents Chemotherapy*, 54, 1369–1373.

- Carver, T., Thomson, N., Bleasby, A., Berriman, M., Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*, 25(1), 119–120. <http://doi.org/10.1093/bioinformatics/btn578>.
- Chakraborty, A., Saralaya, V., Adhikari, P., Shenoy, S., Baliga, S., & Hegde, A. (2015). Characterization of *Escherichia coli* Phylogenetic Groups Associated with Extraintestinal Infections in South Indian Population. *Annals of Medical and Health Sciences Research*, 5(4), 241–246. <http://doi.org/10.4103/2141-9248.160192>.
- Clermont, O., Bonacorsi, S., & Bingen, E. (2000). Rapid and Simple Determination of the *Escherichia coli* Phylogenetic Group. *Applied and Environmental Microbiology*, 66(10), 4555–4558.
- Coque, T. M., Novais, Â., Carattoli, A., Poirel, L., Pitout, J., Peixe, L., ... Nordmann, P. (2008). Dissemination of Clonally Related *Escherichia coli* Strains Expressing Extended-Spectrum β -Lactamase CTX-M-15. *Emerging Infectious Diseases*, 14(2), 195–200. <http://doi.org/10.3201/eid1402.070350>
- Croxen, MA., Finlay, BB. (2010). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews Microbiology*, 8(1), 26-38. DOI: 10.1038/nrmicro2265.
- D'Andrea, MM., Arena, F., Pallecchi, L., Rossolini, GM. (2013). CTX-M-type β -lactamases: A successful story of antibiotic resistance. *International Journal of Medical Microbiology*, 303(6-7), 305-317. DOI: 10.1016/j.ijmm.2013.02.008
- Daoud, Z., Hakime, N. (2003). Prevalence and susceptibility patterns of extended-spectrum β -lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae* in a general university hospital in Beirut, Lebanon. *Revista Española de Quimioterapia*, 16, 233–8.
- Daoud, Z., Moubareck, C., Hakime, N., Doucet-Populaire, F. (2006). Extended spectrum β -lactamase producing *Enterobacteriaceae* in Lebanese ICU patients: epidemiology and patterns of resistance. *The Journal of General and Applied Microbiology*, 52, 169–78.

- Daoud Z, Afif C. (2011). *Escherichia coli* isolated from urinary tract infections of Lebanese patients between 2000 and 2009: epidemiology and profiles of resistance. *Chemotherapy Research and Practice*, 2011, 1–6.
- Daoud, Z., Salem Sokhn, E., Masri, K., Cheaito, K., Haidar-Ahmad, N., Matar, G. M., & Doron, S. (2015). *Escherichia coli* Isolated from Urinary Tract Infections of Lebanese Patients between 2005 and 2012: Epidemiology and Profiles of Resistance. *Frontiers in Medicine*, 2, 26.
- Derakhshandeh, A., Firouzi, R., Motamedifar, M., Motamedi Borojeni, A., Bahadori, M., Arabshahi, S., Novinrooz, A., Heidari, S. (2015). Distribution of virulence genes and multiple drug-resistant patterns amongst different phylogenetic groups of uropathogenic *Escherichia coli* isolated from patients with urinary tract infection. *Letters in Applied Microbiology*, 60(2), 148-54. DOI: 10.1111/lam.12349.
- Dhillon, R. H.-P., & Clark, J. (2012). ESBLs: A Clear and Present Danger? *Critical Care Research and Practice*, 2012, 625170. <http://doi.org/10.1155/2012/625170>.
- Dobiasova, H., Dolejska, M. (2016). Prevalence and diversity of IncX plasmids carrying fluoroquinolone and β -lactam resistance genes in *Escherichia coli* originating from diverse sources and geographical areas. *The Journal of Antimicrobial Chemotherapy*, 71(8), 2118-24. DOI: 10.1093/jac/dkw144.
- Domínguez, E., Zarazaga, M., Sáenz, Y., Briñas, L., Torres, C. (2002). Mechanisms of antibiotic resistance in *Escherichia coli* isolates obtained from healthy children in Spain. *Microbial Drug Resistance*, 8(4), 321-7.
- Doumith, M., Day, M., Ciesielczuk, H., Hope, R., Underwood, A., Reynolds, R., ... Woodford, N. (2015). Rapid Identification of Major *Escherichia coli* Sequence Types Causing Urinary Tract and Bloodstream Infections. *Journal of Clinical Microbiology*, 53(1), 160–166. <http://doi.org/10.1128/JCM.02562-14>
- Doumith M, Ellington MJ, Livermore DM, Woodford N. (2009). Molecular mechanisms disrupting porin expression in ertapenem-resistant *Klebsiella* and *Enterobacter* spp. clinical isolates from the UK. *Journal of Antimicrobial Chemotherapy*, 63(4):659-67. DOI: 10.1093/jac/dkp029.

- El-Herte, RI., Araj, GF., Matar, GM., Baroud, M., Kanafani, ZA., Kanj, SS. (2012). Detection of carbapenem-resistant *Escherichia coli* and *Klebsiella pneumoniae* producing NDM-1 in Lebanon. *The Journal of Infection in Developing Countries*, 6(5), 457-61.
- Elumalai, S., Muthu, G., Selvam, RE., Ramesh, S. (2014). Detection of TEM-, SHV- and CTX-M-type β -lactamase production among clinical isolates of *Salmonella* species. *Journal of Medical Microbiology*, 63, 962-7. DOI: 10.1099/jmm.0.068486-0.
- Eto, DS., Jones, TA., Sundsbak, JL., Mulvey, MA. (2007). Integrin-mediated host cell invasion by type 1-piliated uropathogenic *Escherichia coli*. *PLoS Pathogen*, 3(7), e100.
- Evans, B. A., & Amyes, S. G. B. (2014). OXA β -Lactamases. *Clinical Microbiology Reviews*, 27(2), 241–263. <http://doi.org/10.1128/CMR.00117-13>.
- Ewers C., Bethe A., Stamm I., Grobbel M., Kopp P. A., Guerra B., Stubbe, M., Doi, Y., Zong, Z., Kola, A., Schaufler, K., Semmler, T, Fruth, A., Wieler, LH., Guenther, S. (2014). CTX-M-15-D-ST648 *Escherichia coli* from companion animals and horses: another pandemic clone combining multiresistance and extraintestinal virulence? *Journal of Antimicrobial Chemotherapy*. 69, 1224–1230. DOI: 10.1093/jac/dkt516
- Feng, Y., Yang, P., Xie, Y., Wang, X., McNally, A., Zong, Z. (2015). *Escherichia coli* of sequence type 3835 carrying *bla*_{NDM-1}, *bla*_{CTX-M-15}, *bla*_{CMY-42} and *bla*_{SHV-12}. *Scientific Reports*, 5:12275. DOI: 10.1038/srep12275.
- Flores-Mireles, AL., Walker, JN., Caparon, M., Hultgren, SJ. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nature Reviews Microbiology*, 13(5), 269-84. DOI: 10.1038/nrmicro3432.
- Forde, BM., Ben Zakour, NL., Stanton-Cook, M., Phan, MD., Totsika, M., Peters, KM., Chan, KG., Schembri, MA., Upton, M., Beatson, SA. (2014). The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant E. coli O25b:H4-ST131 clone. *PLoS One*, 9(8), e104400. DOI: 10.1371/journal.pone.0104400.

- Gao, Q., Wang, X., Xu, H., Xu, Y., Ling, J., Zhang, D., Gao, S., Liu, X. (2012). Roles of iron acquisition systems in virulence of extraintestinal pathogenic *Escherichia coli*: salmochelin and aerobactin contribute more to virulence than heme in a chicken infection model. *BMC Microbiology*, 12, 143. DOI: 10.1186/1471-2180-12-143
- Grad, Y. H., Lipsitch, M., Feldgarden, M., Arachchi, H. M., Cerqueira, G. C., FitzGerald, M., ... Hanage, W. P. (2012). Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proceedings of the National Academy of Sciences of the United States of America*, 109(8), 3065–3070.
- Grant JR., Stothard P. (2008). The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Research*. DOI: 10.1093/nar/gkn179.
- Gröbner S, Linke D, Schütz W, Fladerer C, Madlung J, Autenrieth IB, Witte W, Pfeifer Y. (2009). Emergence of carbapenem-non-susceptible extended-spectrum beta-lactamase-producing *Klebsiella pneumoniae* isolates at the university hospital of Tübingen, Germany. *Journal of Medical Microbiology*. DOI: 10.1099/jmm.0.005850-0.
- Giufrè, M., Graziani, C., Accogli, M., Luzzi, I., Busani, L., Cerquetti, M. (2012). *Escherichia coli* of human and avian origin: detection of clonal groups associated with fluoroquinolone and multidrug resistance in Italy. *Journal of Antimicrobial Chemotherapy*, 67(4), 860-7. DOI: 10.1093/jac/dkr565.
- Hannan, TJ., Mysorekar, IU., Hung, CS., Isaacson-Schmid, ML., Hultgren, SJ. (2010). Early severe inflammatory responses to uropathogenic *E. coli* predispose to chronic and recurrent urinary tract infection. *PLoS Pathogen*, 6(8), e1001042. DOI: 10.1371/journal.ppat.1001042.
- Harrington, S. M., Sheikh, J., Henderson, I. R., Ruiz-Perez, F., Cohen, P. S., Nataro, J. P. (2009). The Pic Protease of Enteroaggregative *Escherichia coli* Promotes Intestinal Colonization and Growth in the Presence of Mucin. *Infection and Immunity*, 77(6), 2465–2473. <http://doi.org/10.1128/IAI.01494-08>
- Hasman H, Saputra D, Sicheritz-Ponten T, Lund O, Svendsen CA, Frimodt-Møller N, Aarestrup FM. (2014). Rapid Whole-Genome Sequencing for Detection and Characterization of Microorganisms Directly from Clinical Samples. *Journal of Clinical Microbiology*, 52(1): 139-146.

- Hazen, TH., Zhao, L., Boutin, MA., Stancil, A., Robinson, G., Harris, AD., Rasko, DA., Johnson, JK. (2014). Comparative genomics of an IncA/C multidrug resistance plasmid from *Escherichia coli* and *Klebsiella* isolates from intensive care unit patients and the utility of whole-genome sequencing in health care settings. *Antimicrobial Agents and Chemotherapy*, 58(8), 4814-25. DOI: 10.1128/AAC.02573-14.
- Hiki, M., Usui, M., Akiyama, T., Kawanishi, M., Tsuyuki, M., Imamura, S., ... Asai, T. (2014). Phylogenetic grouping, epidemiological typing, analysis of virulence genes, and antimicrobial susceptibility of *Escherichia coli* isolated from healthy broilers in Japan. *Irish Veterinary Journal*, 67(1), 14. <http://doi.org/10.1186/2046-0481-67-14>
- Ho, WS., Yap, KP., Yeo, CC., Rajasekaram, G., Thong, KL. (2016). The Complete Sequence and Comparative Analysis of a Multidrug-Resistance and Virulence Multireplicon IncFII Plasmid pEC302/04 from an Extraintestinal Pathogenic *Escherichia coli* EC302/04 Indicate Extensive Diversity of IncFII Plasmids. *Frontiers in Microbiology*, 6, 1547. DOI: 10.3389/fmicb.2015.01547
- Hüttener, M., Paytubi, S., Juárez, A. (2014). HilA-like regulators in *Escherichia coli* pathotypes: the YgeH protein from the enteroaggregative strain 042. *BMC Microbiology*, 14, 268.
- Joensen, KG., Scheutz, F., Lund, O., Hasman, H., Kaas, RS., Nielsen, EM., Aarestrup, FM. (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology*, 52(5), 1501-1510.
- Joensen, KG., Tetzschner, AM., Iguchi, A., Aarestrup, FM., Scheutz, F. (2015). Rapid and easy in silico serotyping of *Escherichia coli* using whole genome sequencing (WGS) data. *Journal of Clinical Microbiology*, 53, 8, 2410-26. DOI: 10.1128/JCM.00008-15.
- Johnson, JR., Russo, TA. (2002). Extraintestinal pathogenic *Escherichia coli*: "The other bad *E. coli*". *The Journal of Laboratory and Clinical Medicine*, 139(3), 155-62.

- Johnson, J. R., & 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. *Journal of Clinical Microbiology*, 42(4), 1773–1776. <http://doi.org/10.1128/JCM.42.4.1773-1776.2004>
- Johnson, T. J., Wannemuehler, Y. M., & Nolan, L. K. (2008). Evolution of the *iss* Gene in *Escherichia coli*. *Applied and Environmental Microbiology*, 74(8), 2360–2369. <http://doi.org/10.1128/AEM.02634-07>.
- Justice, SS., Hung, C., Theriot, JA., Fletcher, DA., Anderson, GG., Footer, MJ., Hultgren, SJ. (2004). Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(5), 1333-8.
- Kaper, JB., Nataro, JP., Mobley, HL. (2004). Pathogenic *Escherichia coli*. *Nature Reviews Microbiology*, 2(2), 123-40.
- Klemm, E., Dougan, G. (2016). Advances in Understanding Bacterial Pathogenesis Gained from Whole-Genome Sequencing and Phylogenetics. *Cell & Host Microbe*, 19(5):599-610. DOI: 10.1016/j.chom.2016.04.015.
- Klemm, P., Hancock, V., Schembri, MA. (2010). Fimbrial adhesins from extraintestinal *Escherichia coli*. *Environmental Microbiology Reports*, 2(5), 628-40. DOI: 10.1111/j.1758-2229.2010.00166.x.
- Knothe, H., Shah, P., Krcmery, V., Antal, M., Mitsuhashi, S. (1983). Transferable resistance to cefotaxime, cefoxitin, cefamandole and cefuroxime in clinical isolates of *Klebsiella pneumoniae* and *Serratia marcescens*. *Infection*, 11(6), 315-7.
- Köhler, CD., Dobrindt, U. (2011). What Defines Extraintestinal Pathogenic *Escherichia coli*? *International Journal of Medical Microbiology*, 301(8), 642-7. DOI: 10.1016/j.ijmm.2011.09.006.
- Köser, CU., Ellington, MJ., Cartwright, EJ., Gillespie, SH., Brown, NM., Farrington, M., Holden, MT., Dougan, G., Bentley, SD., Parkhill, J., Peacock, SJ. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health

microbiology. *PloS Pathogen*, 8(8):e1002824. DOI:
10.1371/journal.ppat.1002824.

- Lane, MC., Alteri, CJ., Smith, SN., Mobley, HL. (2007). Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract. *Proceedings of the National Academy of Sciences of the United States of America*, 104(42), 16669-74.
- Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., Lund, O. (2012). Multilocus sequence typing of total genome sequenced bacteria. *Journal of Clinical Microbiology*, 50(4), 1355-1361.
- Lartigue MF, Poirel L, Poyart C, Réglie-Poupet H, Nordmann P. (2007). Ertapenem resistance of *Escherichia coli*. *Emerging Infectious Diseases*, 13(2), 315-7.
- Lim, KT., Yasin, R., Yeo, CC., Puthuchery, S., Thong, KL. (2009). Characterization of Multidrug Resistant ESBL-Producing *Escherichia coli* Isolates from Hospitals in Malaysia. *Journal of Biomedicine and Biotechnology*, 2009, 165637.
- Ling, J., Pan, H., Gao, Q., Xiong, L., Zhou, Y., Zhang, D., Gao, S., Liu, X. (2013). Aerobactin synthesis genes *iucA* and *iucC* contribute to the pathogenicity of avian pathogenic *Escherichia coli* O2 strain E058. *PLoS One*, 8(2), e57795. DOI: 10.1371/journal.pone.0057794.
- Liu, Y., Feng, Y., Wu, W., Xie, Y., Wang, X., Zhang, X., Xinchuan, C., Zong, Z. (2015). First Report of OXA-181-Producing *Escherichia coli* in China and Characterization of the Isolate Using Whole-Genome Sequencing. *Antimicrobial Agents and Chemotherapy*, 59(8), 5022–5025. <http://doi.org/10.1128/AAC.00442-15>
- Livermore DM, Woodford N. (2006). The beta-lactamase threat in *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter*. *Trends in Microbiology*, 14(9), 413-20.
- Lo, WU., Chow, KH., Law, PY., Ng, KY., Cheung, YY., Lai, EL., Ho, PL. (2014). Highly conjugative IncX4 plasmids carrying blaCTX-M in *Escherichia coli* from humans and food animals. *Journal of Medical Microbiology*, 63(6), 835-40. DOI: 10.1099/jmm.0.074021-0.

- Mahillon, J., & Chandler, M. (1998). Insertion Sequences. *Microbiology and Molecular Biology Reviews*, 62(3), 725–774.
- Matar GM, Cuzon G, Araj GF, Naas T, Corkill J, Kattar MM, Nordmann P. (2008). Oxacillinase-mediated resistance to carbapenems in *Klebsiella pneumoniae* from Lebanon. *Clinical Microbiology and Infection*. DOI: 10.1111/j.1469-0691.2008.02059.x.
- Matar, GM., Dandache, I., Carrër, A., Khairallah, MT., Nordmann, P., Sabra, A., Araj, GF. (2010). Spread of OXA-48-mediated resistance to carbapenems in Lebanese *Klebsiella pneumoniae* and *Escherichia coli* that produce extended spectrum beta-lactamase. *Annals of Tropical Medicine and Parasitology*, 104(3), 271-4. DOI: 10.1179/136485910X12647085215651.
- Mirzarazi, M., Rezafofighi, S. E., Pourmahdi, M., & Mohajeri, M. R. (2015). Occurrence of genes encoding enterotoxins in uropathogenic *Escherichia coli* isolates. *Brazilian Journal of Microbiology*, 46(1), 155–159. <http://doi.org/10.1590/S1517-838246120130860>
- McAdam, P. R., Richardson, E. J., & Fitzgerald, J. R. (2014). High-throughput sequencing for the study of bacterial pathogen biology. *Current Opinion in Microbiology*, 19(100), 106–113. <http://doi.org/10.1016/j.mib.2014.06.002>
- Moubareck C, Daou Z, Hakimé NI, Hamzé M, Mangeney N, Matta H, Mokhbat JE, Rohban R, Sarkis DK, Doucet-Populaire F. (2005). Countrywide spread of community- and hospital-acquired extended-spectrum beta-lactamase (CTX-M-15)-producing *Enterobacteriaceae* in Lebanon. *Journal of Clinical Microbiology*, 43(7), 3309-13.
- Mshana, S. E., Falgenhauer, L., Mirambo, M. M., Mushi, M. F., Moremi, N., Julius, R., ... Chakraborty, T. (2016). Predictors of *bla*_{CTX-M-15} in varieties of *Escherichia coli* genotypes from humans in community settings in Mwanza, Tanzania. *BMC Infectious Diseases*, 16, 187. <http://doi.org/10.1186/s12879-016-1527-x>
- Mulvey, M. A., Schilling, J. D., & Hultgren, S. J. (2001). Establishment of a Persistent *Escherichia coli* Reservoir during the Acute Phase of a Bladder Infection. *Infection and Immunity*, 69(7), 4572–4579. <http://doi.org/10.1128/IAI.69.7.4572-4579.2001>.

- Muniesa, M., Blanco, JE., De Simón, M., Serra-Moreno, R., Blanch, AR., Jofre, J. (2004). Diversity of *stx2* converting bacteriophages induced from Shiga-toxin-producing *Escherichia coli* strains isolated from cattle. *Microbiology*, 150(9), 2959-71.
- Munk Vejborg, R., Hancock, V., Petersen, AM., Krogfelt, KA., Klemm, P. (2011). Comparative genomics of *Escherichia coli* isolated from patients with inflammatory bowel disease. *BMC Genomics*, 12, 316. DOI: 10.1186/1471-2164-12-316
- Mysorekar, IU., Isaacson-Schmid, M., Walker, JN., Mills, JC., Hultgren, SJ. (2009). Bone morphogenetic protein 4 signaling regulates epithelial renewal in the urinary tract in response to uropathogenic infection. *Cell Host & Microbe*, 5(5), 463-75. DOI: 10.1016/j.chom.2009.04.005.
- Nicolas-Chanoine, MH., Bertrand, X., and Madec, JY. (2014). *Escherichia coli* ST131, an intriguing clonal group. *Clinical Microbiology Review*, 27(3), 543-74. doi.org/10.1128/CMR.00125-13.
- Nolan LK., Horne, SM., Giddings, CW., Foley, SL., Johnson, TJ., Lynne, AM., Skyberg, J. (2003). Resistance to serum complement, iss, and virulence of avian *Escherichia coli*. *Veterinary Research Communications*, 27(2), 101-10.
- Nordmann, P., Dortet, L., Poirel, L. (2012). Carbapenem resistance in Enterobacteriaceae: here is the storm! *Trends in Molecular Medicine*, 18(5), 263-72. DOI: 10.1016/j.molmed.2012.03.003.
- Novais, Â., Viana, D., Baquero, F., Martínez-Botas, J., Cantón, R., & Coque, T. M. (2012). Contribution of IncFII and Broad-Host IncA/C and IncN Plasmids to the Local Expansion and Diversification of Phylogroup B2 *Escherichia coli* ST131 Clones Carrying *bla*_{CTX-M-15} and *qnrS1* Genes. *Antimicrobial Agents and Chemotherapy*, 56(5), 2763–2766. <http://doi.org/10.1128/AAC.06001-11>
- Oteo, J., Diestra, K., Juan, C., Bautistaa, V., Novais, A., Pérez-Vázquez, M., Moyác, B., Miró, E., Coque, T., Oliver, A., Cantón, R., Navarr, F., Campos, J. (2009). Extended-spectrum β -lactamase producing *Escherichia coli* in Spain belong to a large variety of multilocus sequence typing types, including ST10 complex/A, ST23 complex/A and ST131/B2. *International Journal of Antimicrobial Agents*, 34, 173–6.

- Oshima, K., Toh, H., Ogura, Y., Sasamoto, H., Morita, H., Park, S.-H., ... Hattori, M. (2008). Complete Genome Sequence and Comparative Analysis of the Wild-type Commensal *Escherichia coli* Strain SE11 Isolated from a Healthy Adult. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, 15(6), 375–386. <http://doi.org/10.1093/dnares/dsn026>
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, 42, 206-214.
- Paterson, DL., Bonomo RA. (2005). Extended-spectrum beta-lactamases: a clinical update. *Clinical Microbiology Reviews*, 18(4), 657-86.
- Peirano, G., Bradford, PA., Kazmierczak, KM., Badal, RE., Hackel, M., Hoban, DJ., Pitout, JD. (2014). Global incidence of carbapenemase-producing *Escherichia coli* ST131. *Emerging Infectious Diseases*, 20(11), 1928-31. DOI: 10.3201/eid2011.141388.
- Peirano, G., Asensi, MD., Pitondo-Silva, A., Pitout, JD. (2011). Molecular characteristics of extended-spectrum beta-lactamase-producing *Escherichia coli* from Rio de Janeiro, Brazil. *Clinical Microbiology and Infection*, 17, 1039–43.
- Petersen, AM., Nielsen, EM., Litrup, E., Brynskov, J., Mirsepasi, H., Krogh, KA. (2009). A phylogenetic group of *Escherichia coli* associated with active left-sided inflammatory bowel disease. *BMC Microbiology*, 9, 171. DOI: 10.1186/1471-2180-9-171.
- Poirel, L., Naas, T., Nordmann, P. (2010). Diversity, epidemiology, and genetics of class D beta-lactamases. *Antimicrobial Agents and Chemotherapy*, 54(1), 24-38. doi: 10.1128/AAC.01512-08.
- Poolman, JT., Wacker, M. (2015). Extraintestinal Pathogenic *Escherichia coli*, a Common Human Pathogen: Challenges for Vaccine Development and Progress in the Field. *The Journal of Infectious Diseases*, 213(1):6-13. DOI: 10.1093/infdis/jiv429.
- Poolman, J. T., & Wacker, M. (2016). Extraintestinal Pathogenic *Escherichia coli*, a Common Human Pathogen: Challenges for Vaccine Development and Progress in the Field. *The Journal of Infectious Diseases*, 213(1), 6–13. <http://doi.org/10.1093/infdis/jiv429>

- Porres-Osante, N., Azcona-Gutiérrez, JM., Rojo-Bezares, B., Undabeitia, E., Torres, C., Sáenz, Y. (2014). Emergence of a multiresistant KPC-3 and VIM-1 carbapenemase-producing *Escherichia coli* strain in Spain. *Journal of Antimicrobial Chemotherapy*, 69(7), 1792-5. DOI: 10.1093/jac/dku055.
- Potron, A., Poirel, L., Nordmann, P. (2011). Origin of OXA-181, an emerging carbapenem-hydrolyzing oxacillinase, as a chromosomal gene in *Shewanella xiamenensis*. *Antimicrobial Agents and Chemotherapy*, 55(9), 4405-7. DOI: 10.1128/AAC.00681-11.
- Potron, A., Nordmann, P., Lafeuille, E., Al Maskari, Z., Al Rashdi, F., & Poirel, L. (2011). Characterization of OXA-181, a Carbapenem-Hydrolyzing Class D β -Lactamase from *Klebsiella pneumoniae*. *Antimicrobial Agents and Chemotherapy*, 55(10), 4896–4899. <http://doi.org/10.1128/AAC.00481-11>
- Queenan AM., Bush K. (2007). Carbapenemases: the versatile beta-lactamases. *Clinical Microbiology Reviews*, 3, 440-58.
- Rasmussen, HS., McCann, PP. (1997). Matrix metalloproteinase inhibition as a novel anticancer strategy: a review with special focus on batimastat and marimastat. *Pharmacology & Therapeutics*, 75(1), 69-75.
- Riley, L.W. (2014). Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clinical Microbiology and Infection*, 20(5), 380-90. DOI: 10.1111/1469-0691.12646
- Rocha, DA., Campos, JC., Passadore, LF., Sampaio, SC., Nicodemo, AC., Sampaio, JL. (2015). Frequency of Plasmid-Mediated AmpC β -Lactamases in *Escherichia coli* Isolates from Urine Samples in São Paulo, Brazil. *Microbial Drug Resistance*, 22(4), 321-7. DOI: 10.1089/mdr.2015.0190.
- Rupp, ME., Fey, PD. (2003). Extended spectrum beta-lactamase (ESBL)-producing *Enterobacteriaceae*: considerations for diagnosis, prevention and drug treatment. *Drugs*, 63(4), 353-65.

- Safari, M., Mozaffari Nejad, A. S., Bahador, A., Jafari, R., & Alikhani, M. Y. (2015). Prevalence of ESBL and MBL encoding genes in *Acinetobacter baumannii* strains isolated from patients of intensive care units (ICU). *Saudi Journal of Biological Sciences*, 22(4), 424–429.
- Saldaña, Z., Erdem, AL., Schüller, S., Okeke, IN., Lucas, M., Sivananthan, A., Phillips, AD., Kaper, JB., Puente, JL., Girón, JA. (2009). The *Escherichia coli* common pilus and the bundle-forming pilus act in concert during the formation of localized adherence by enteropathogenic *E. coli*. *Journal of Bacteriology*, 191(11), 3451–61. DOI: 10.1128/JB.01539-08.
- Samaha-Kfoury, J. N., & Araj, G. F. (2003). Recent developments in β lactamases and extended spectrum β lactamases. *British Medical Journal*, 327(7425), 1209–1213.
- Sannes, MR., Kuskowski, MA., Owens, K., Gajewski, A., Johnson, JR. (2004). Virulence factor profiles and phylogenetic background of *Escherichia coli* isolates from veterans with bacteremia and uninfected control subjects. *The Journal of Infectious Diseases*, 190(12), 2121–8.
- Shaikh, S., Fatima, J., Shakil, S., Rizvi, S. M. D., & Kamal, M. A. (2015). Antibiotic resistance and extended spectrum beta-lactamases: Types, epidemiology and treatment. *Saudi Journal of Biological Sciences*, 22(1), 90–101.
- Shames, S. R., & Finlay, B. B. (2010). Breaking the Stereotype: Virulence Factor–Mediated Protection of Host Cells in Bacterial Pathogenesis. *PLoS Pathogens*, 6(9), e1001057. <http://doi.org/10.1371/journal.ppat.1001057>
- Sherchan, JB., Hayakawa, K., Miyoshi-Akiyama, T., Ohmagari, N., Kirikae, T., Nagamatsu, M., Tojo, M., Ohara, H., Sherchand, JB., Tandukar, S. 2015. Clinical epidemiology and molecular analysis of extended-spectrum—lactamase producing *Escherichia coli* in Nepal: characteristics of sequence types 131 and 648. *Antimicrobial Agents and Chemotherapy* 59, 3424–3432. DOI:10.1128/AAC.00270-15.
- Silva, KC., Moreno, M., Cabrera, C., Spira, B., Cerdeira, L., Lincopan, N., Moreno, AM. (2016). First Characterization of CTX-M-15-Producing *Escherichia coli* Strains Belonging to Sequence Type (ST) 410, ST224, and ST1284 from Commercial Swine in South America. *Antimicrobial Agents and Chemotherapy*, 60(4), 2505–8. DOI: 10.1128/AAC.02788-15.

- Silva, L. E., Souza, T. B., Silva, N. P., & Scaletsky, I. C. (2014). Detection and genetic analysis of the enteroaggregative *Escherichia coli* heat-stable enterotoxin (EAST1) gene in clinical isolates of enteropathogenic *Escherichia coli* (EPEC) strains. *BMC Microbiology*, *14*, 135. <http://doi.org/10.1186/1471-2180-14-135>.
- Staji, H., Khoshgoftar, J., Vayeghan, AJ., Bejestani, MRS. (2016). Phylogenetic Grouping and Assessment of Virulence Genotypes, With Antibiotic Resistance Patterns, of *Escherichia Coli* Strains Implicated in Female Urinary Tract Infections. *International Journal of Enteric Pathogens*, *4*(1). DOI: 10.17795/ijep31609.
- Stedt, J., Bonnedahl, J., Hernandez, J., McMahon, B. J., Hasan, B., Olsen, B., Drobni, M., Waldenström, J. (2014). Antibiotic resistance patterns in *Escherichia coli* from gulls in nine European countries. *Infection Ecology & Epidemiology*, *4*, 10. <http://doi.org/10.3402/iee.v4.21565>
- Touchon, M., Hoede, C., Tenailon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, ME., Frapy, E., Garry, L., Ghigo, JM., Gilles, AM., Johnson, J., Le Bouguéneq, C., Lescat, M., Mangenot, S., Martinez-Jéhanne, V., Matic, I., Nassif, X., Oztas, S., Petit, MA., Pichon, C., Rouy, Z., Ruf, CS., Schneider, D., Tourret, J., Vacherie, B., Vallenet, D., Médigue, C., Rocha, EP., Denamur, E. (2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genetics*, *5*(1), e1000344. DOI: 10.1371/journal.pgen.1000344.
- Tritt, A., Eisen, J., Facciotti, M., Darling, A. (2012). An Integrated Pipeline for *de Novo* Assembly of Microbial Genomes. *PLOS One*, *7*(9), e42304.
- Turner, JC. (2005). Explaining the nature of power: a three-process theory. *European Journal of Social Psychology*, *35*(1), 1–22.
- Villa, L., García-Fernández, A., Fortini, D., Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *The Journal of Antimicrobial Chemotherapy*, *65*(12), 2518-29. DOI: 10.1093/jac/dkq347.
- Watts, RE., Totsika, M., Challinor, VL., Mabbett, AN., Ulett, GC., De Voss, JJ., Schembri, MA. (2012). Contribution of siderophore systems to growth and urinary tract colonization of asymptomatic bacteriuria *Escherichia coli*. *Infection and Immunity*, *80*(1), 333-44. DOI: 10.1128/IAI.05594-11.

- Wayne, PA. (2009). Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing. Nineteenth informational supplement M100-S19.
- Weill, FX., Perrier-Gros-Claude, JD., Demartin, M., Coignard, S., Grimont, PA. (2004). Characterization of extended-spectrum-beta-lactamase (CTX-M-15)-producing strains of *Salmonella enterica* isolated in France and Senegal. *FEMS Microbiology Letters*, 238(2), 353-8
- Wester, A. L., Brandal, L. T., & Dahle, U. R. (2013). Extraintestinal Pathogenic *Escherichia coli* Carrying the Shiga Toxin Gene *stx*₂. *Journal of Clinical Microbiology*, 51(12), 4279–4280. <http://doi.org/10.1128/JCM.01349-13>
- Wiles, TJ., Kulesus, RR., Mulvey, MA. (2008). Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Experimental and Molecular Pathology*, 85(1), 11-9. DOI: 10.1016/j.yexmp.2008.03.007.
- Woodford, N., Carattoli, A., Karisik, E., Underwood, A., Ellington, M. J., & Livermore, D. M. (2009). Complete Nucleotide Sequences of Plasmids pEK204, pEK499, and pEK516, Encoding CTX-M Enzymes in Three Major *Escherichia coli* Lineages from the United Kingdom, All Belonging to the International O25:H4-ST131 Clone. *Antimicrobial Agents and Chemotherapy*, 53(10), 4472–4482. <http://doi.org/10.1128/AAC.00688-09>.
- Wyrsh, E., Roy Chowdhury, P., Abraham, S., Santos, J., Darling, AE., Charles, IG., Chapman, TA., Djordjevic, SP. (2015). Comparative genomic analysis of a multiple antimicrobial resistant enterotoxigenic *E. coli* O157 lineage from Australian pigs. *BMC Genomics*, 15, 165. DOI: 10.1186/s12864-015-1382-y.
- Yaici, L., Haenni, M., Saras, E., Boudehouche, W., Touati, A., Madec, JY. (2016). *bla*_{NDM-5}-carrying IncX3 plasmid in *Escherichia coli* ST1284 isolated from raw milk collected in a dairy farm in Algeria. *Journal of Antimicrobial Chemotherapy*, 71(9), 2671-2. DOI: 10.1093/jac/dkw160.
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O. Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *Antimicrobial Agents and Chemotherapy*, 67(11), 2640-2644.

Zhou, Y., Liang, Y., Lynch, H. K., Dennis, J. J., Wishart, D. S. (2011). PHAST: A fast phage search tool. *Nucleic Acids Research*, 39(2), 347-352.

Zhu L, Yan Z, Zhang Z, Zhou Q, Zhou J, et al. (2013) Complete Genome Analysis of Three *Acinetobacter baumannii* Clinical Isolates in China for Insight into the Diversification of Drug Resistance Elements. *PLOS One*, 8(6), e66584. DOI: 10.1371/journal.pone.0066584.

Östblom, A., Adlerberth, I., Wold, A. E., & Nowrouzian, F. L. (2011). Pathogenicity Island Markers, Virulence Determinants *malX* and *usp*, and the Capacity of *Escherichia coli* To Persist in Infants' Commensal Microbiotas. *Applied and Environmental Microbiology*, 77(7), 2303–2308. <http://doi.org/10.1128/AEM.02405-10>

ANNEX I

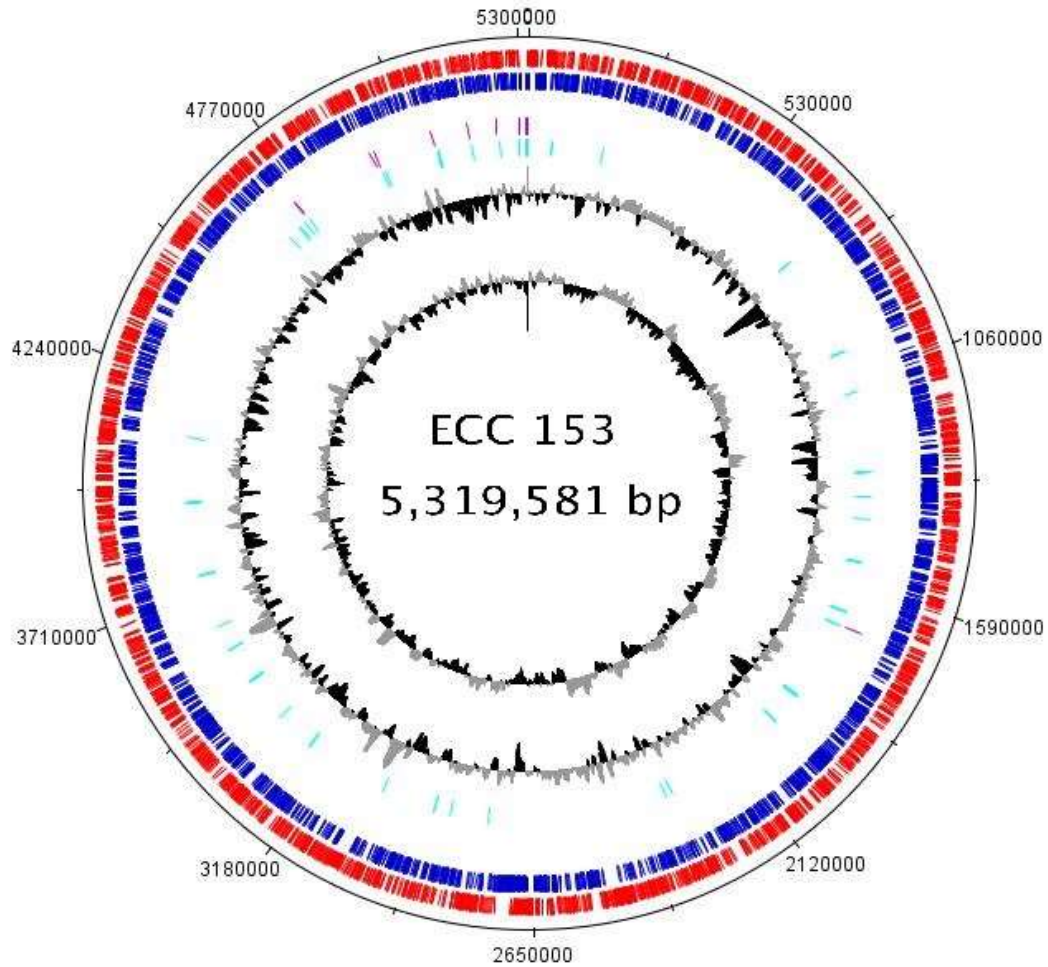


Figure 1. Genome atlas of *E. coli* ECC 153. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli* coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

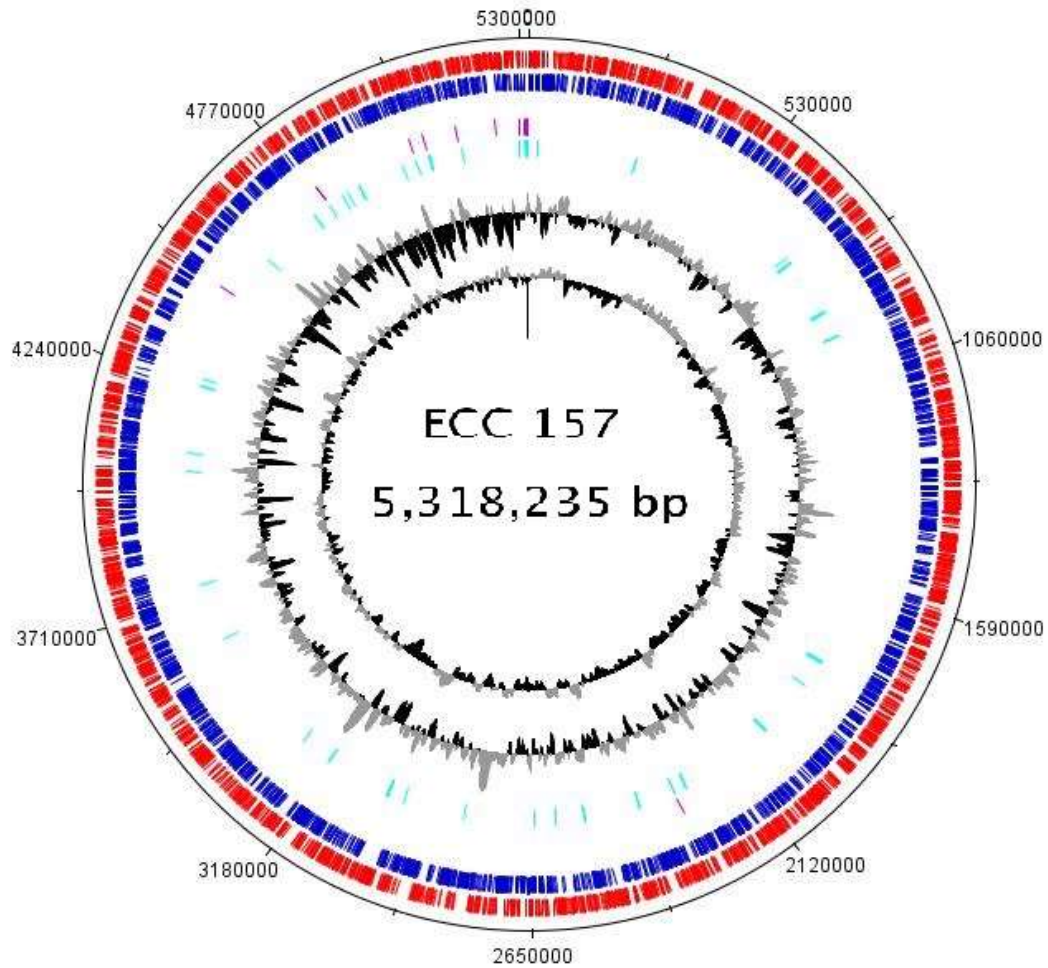


Figure 2. Genome atlas of *E. coli* ECC 157. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

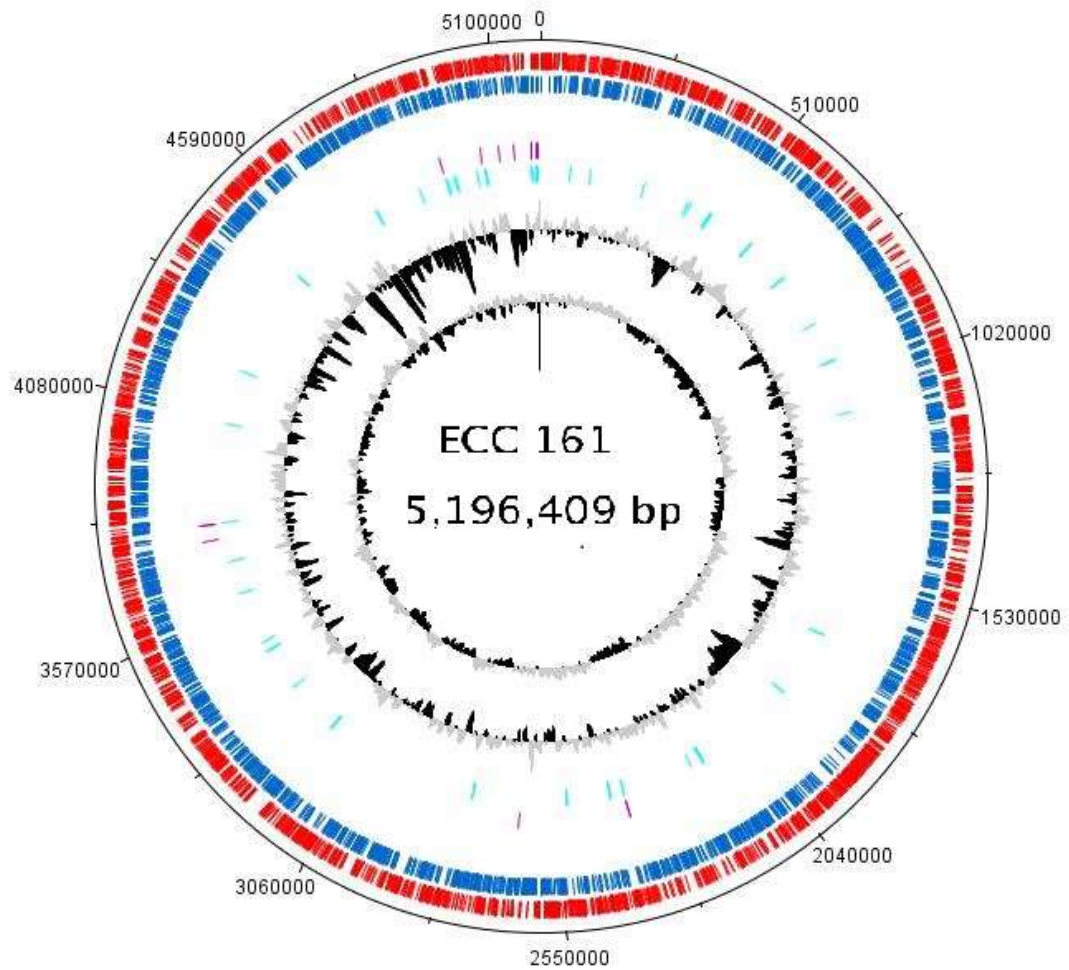


Figure 3. Genome atlas of *E. coli* ECC 161. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

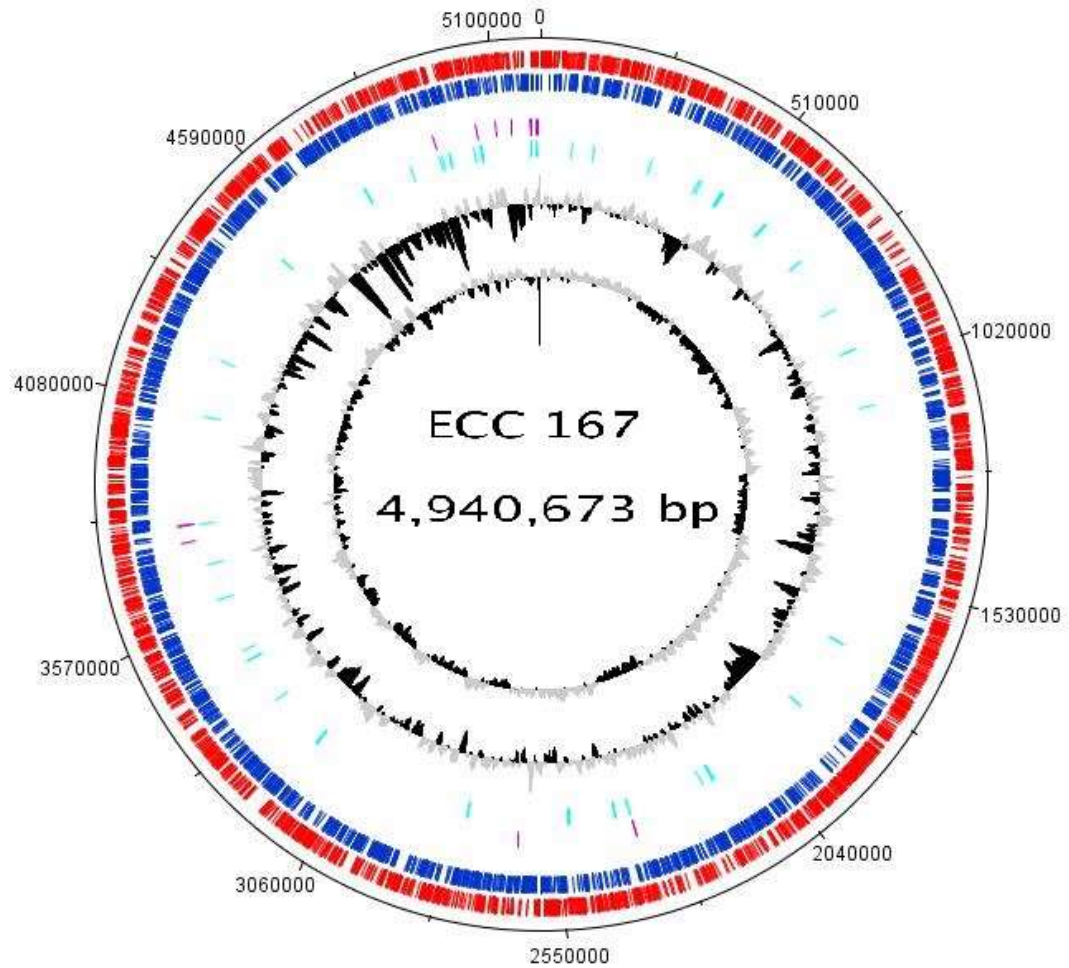


Figure 4. Genome atlas of *E. coli* ECC 167. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

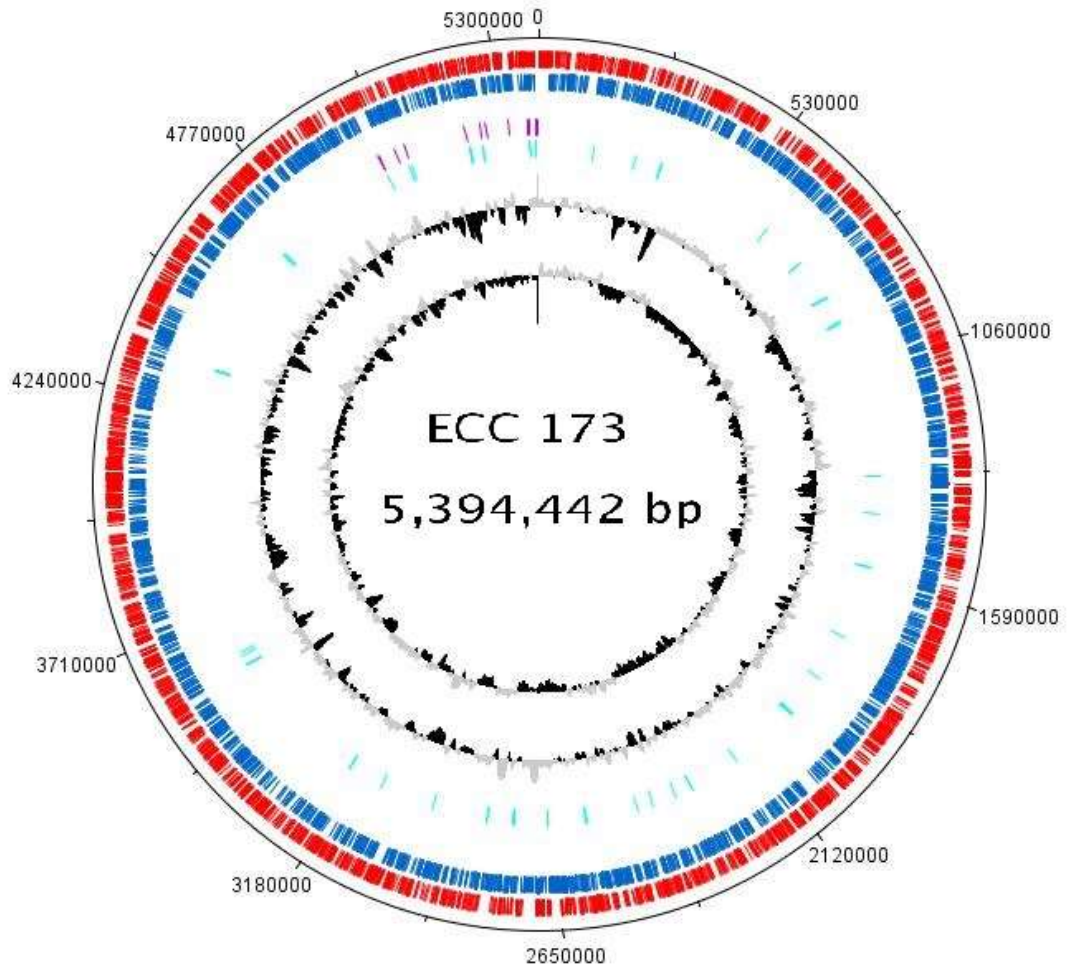


Figure 5. Genome atlas of *E. coli* ECC 173. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

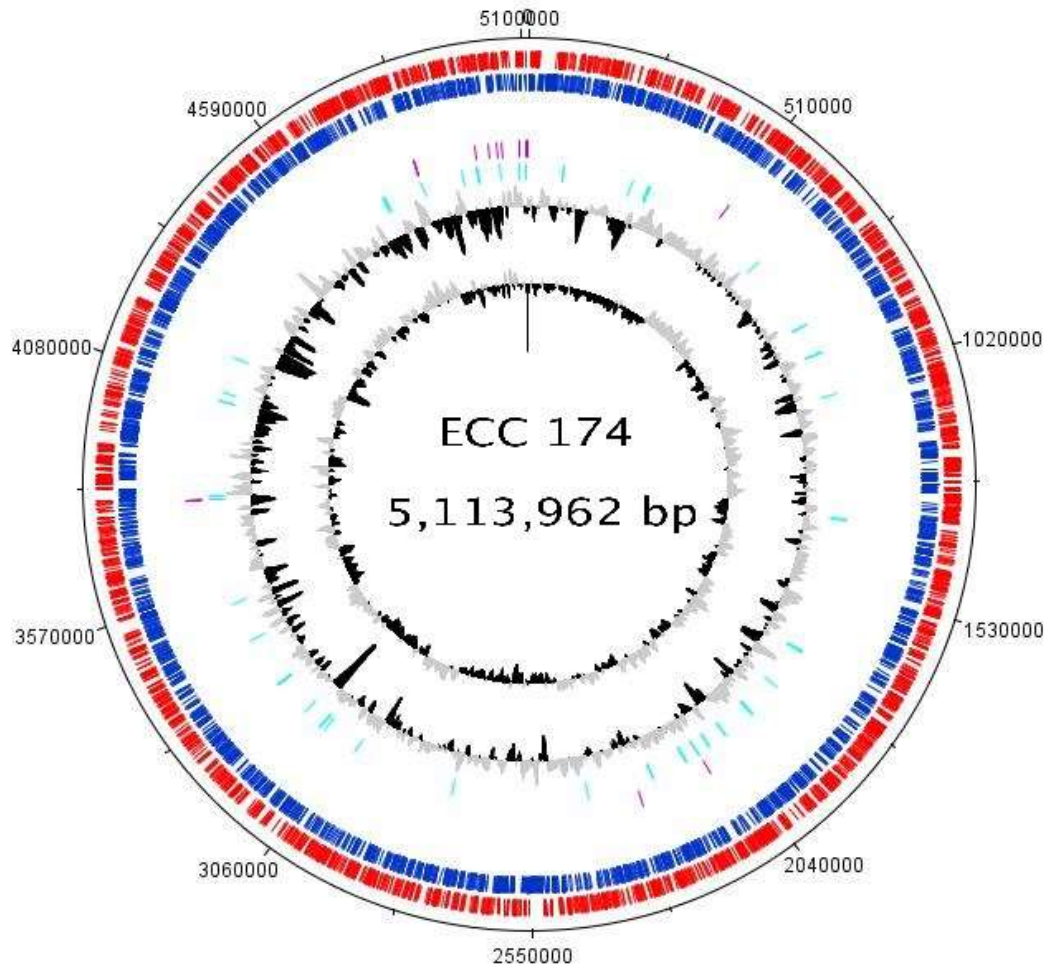


Figure 6. Genome atlas of *E. coli* ECC 174. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

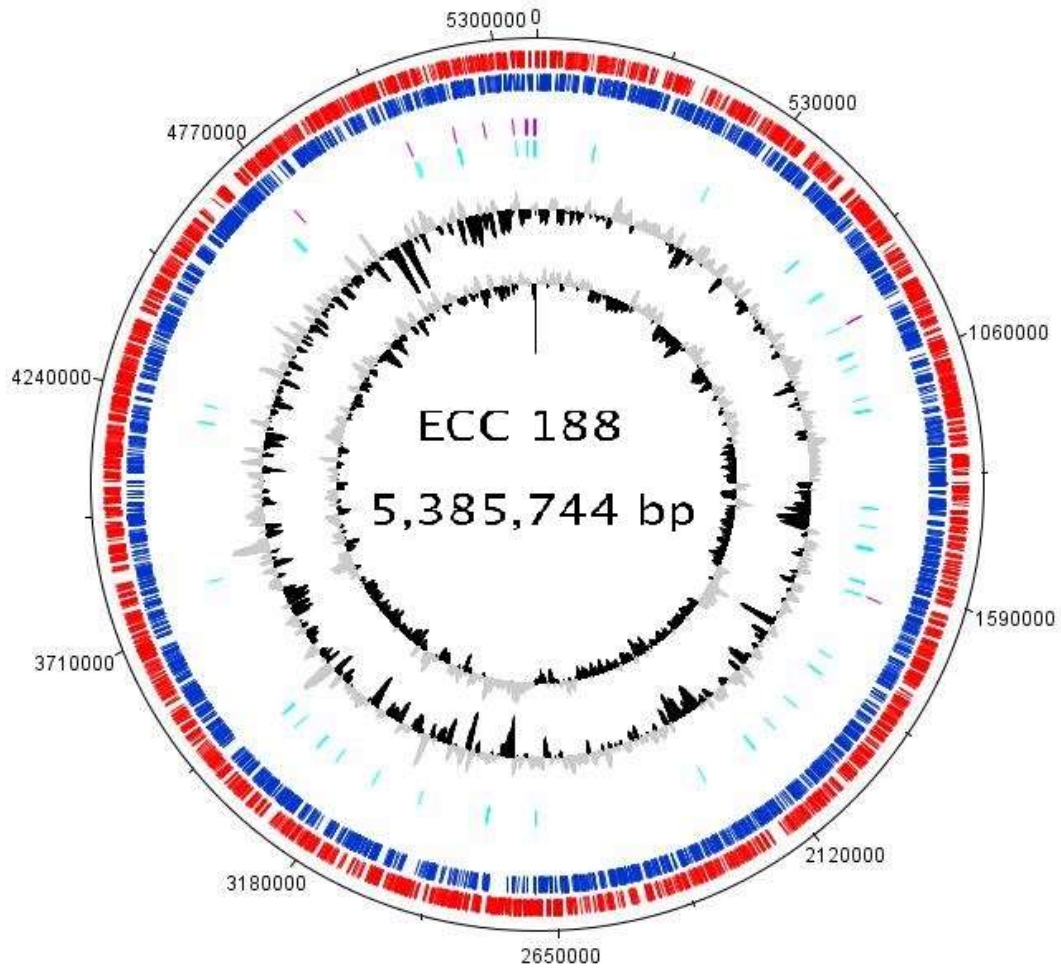


Figure 7. Genome atlas of *E. coli* ECC 188. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

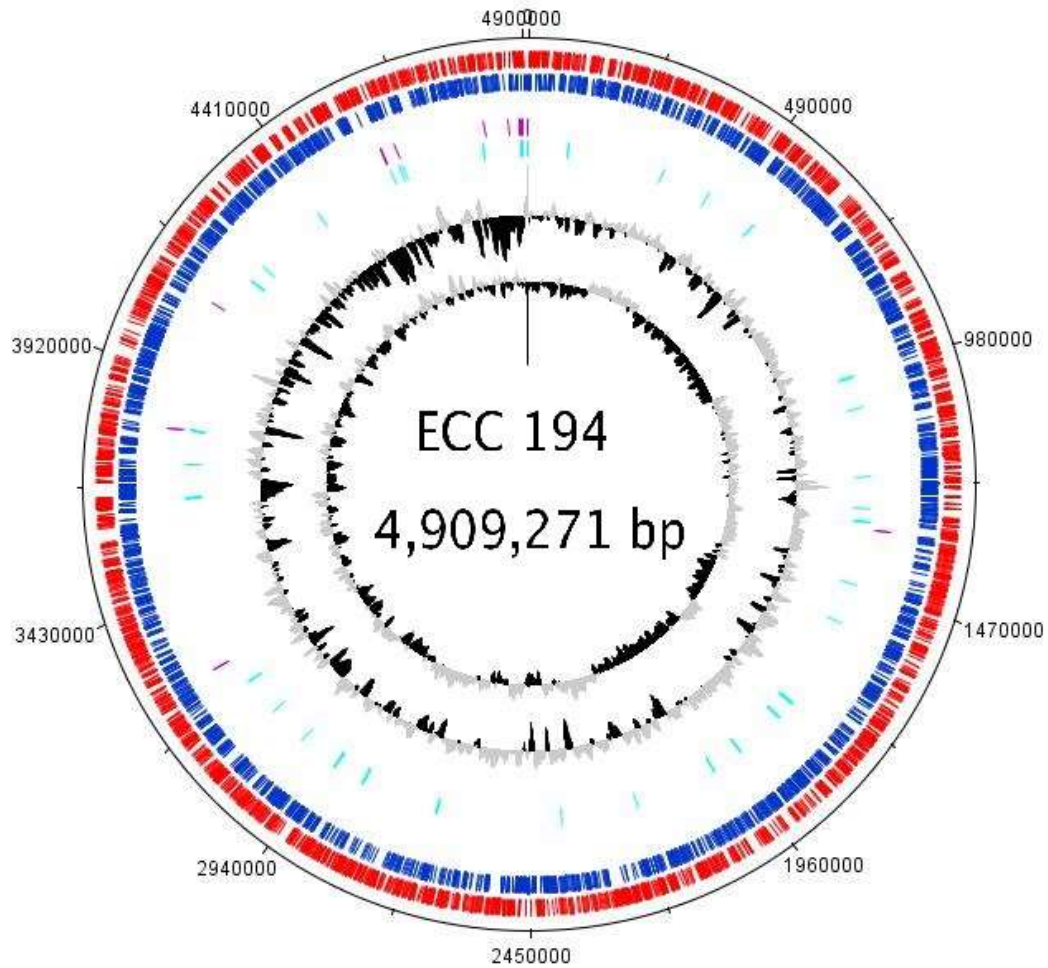


Figure 8. Genome atlas of *E. coli* ECC 194. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

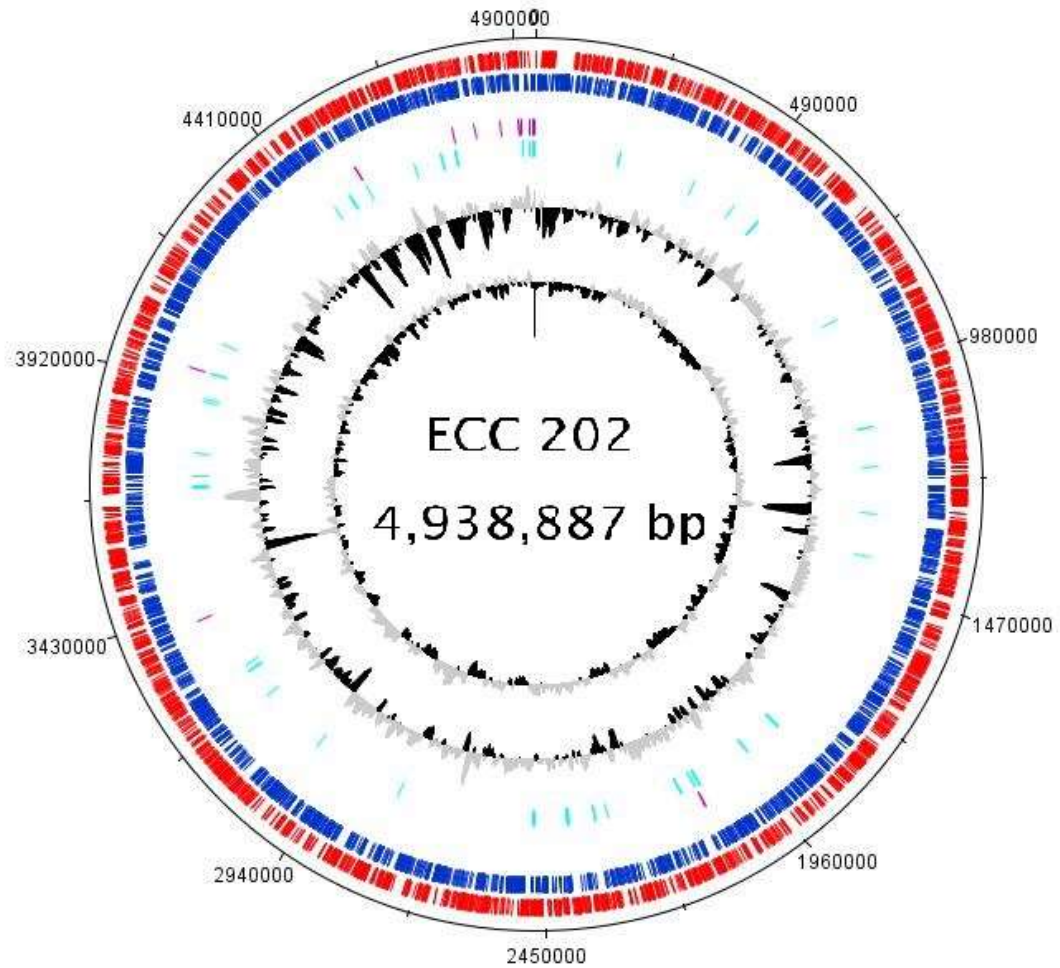


Figure 9. Genome atlas of *E. coli* ECC 202. This graphical representation of the genome was generated using DNAPLOTTER. From outside to inside: *E. coli*, coding DNA sequence on the forward strand (red); *E. coli* coding DNA sequence on the reverse strand (blue); ribosomal RNA genes (purple; in this case none present); tRNA genes (light blue); GC plot (black above mean and grey below mean); GC skew (black above mean and grey below mean).

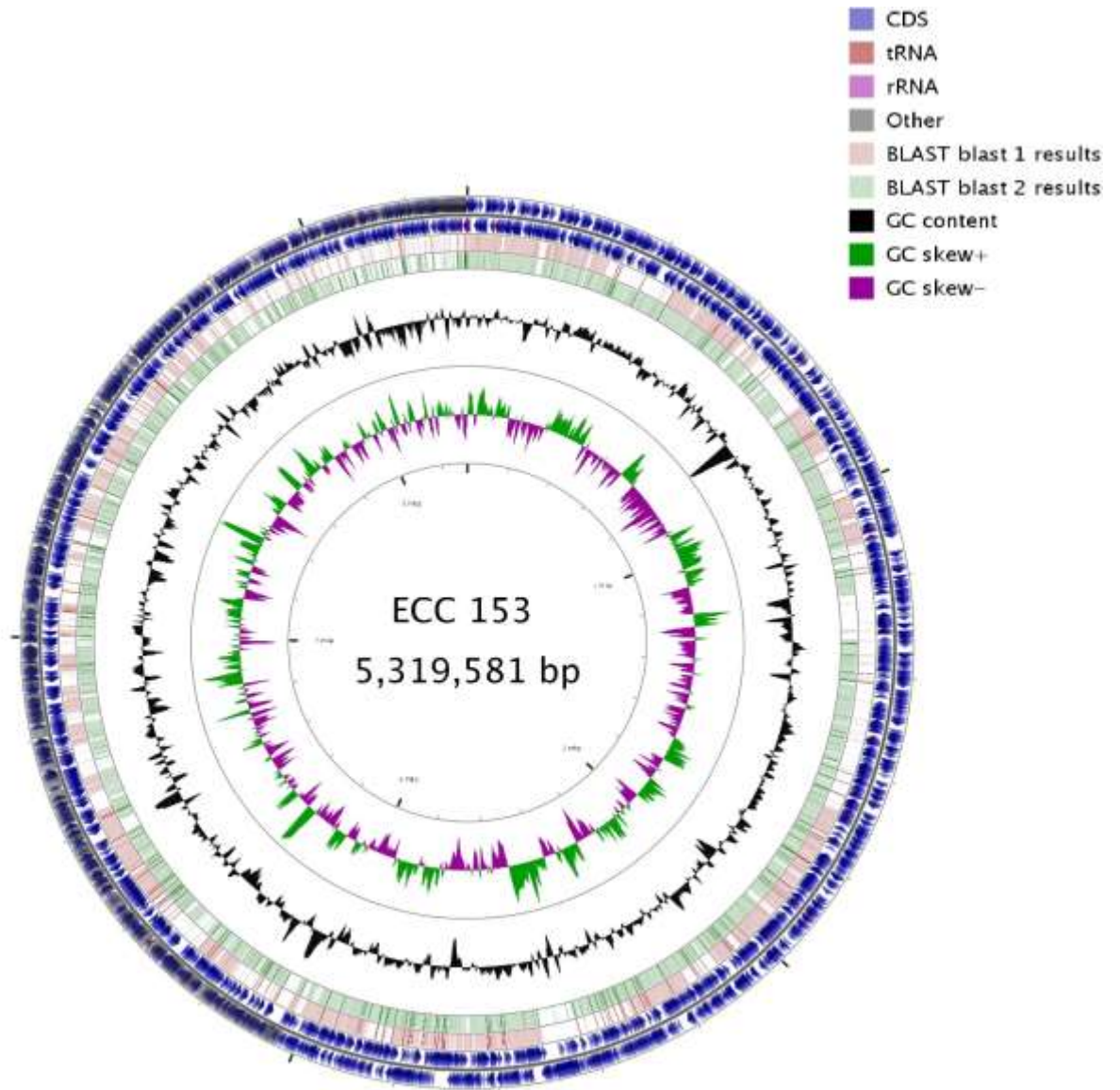


Figure 10. Comparative circular representation of the *E. coli* ECC 153 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession number NZ_AFET01000001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits).

These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

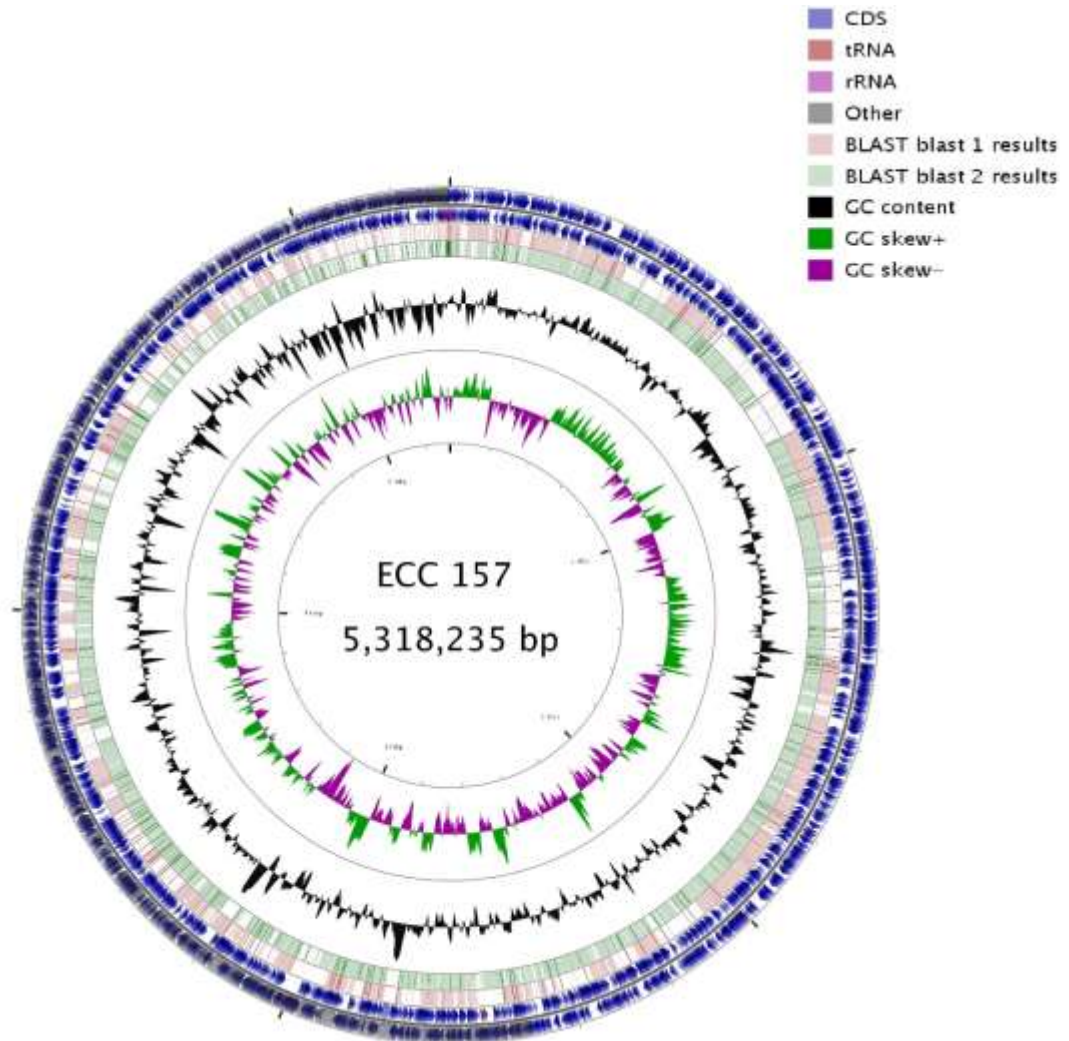


Figure 11. Comparative circular representation of the *E. coli* ECC 157 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession number NZ_AFET01000001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green

regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

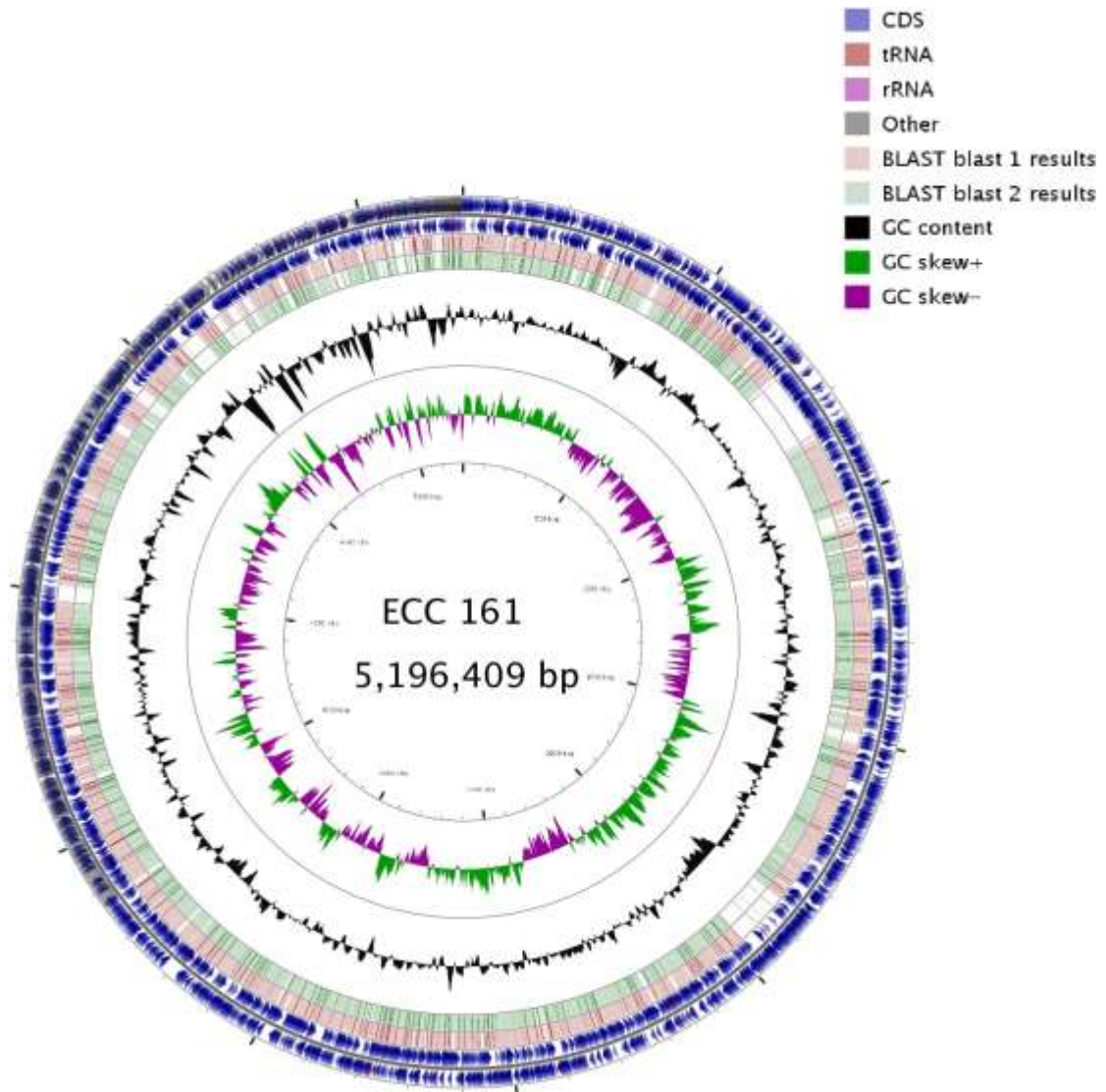


Figure 12. Comparative circular representation of the *E. coli* ECC 161 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* PCN033 (accession number AFAT000000000) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* PCN033 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions

indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

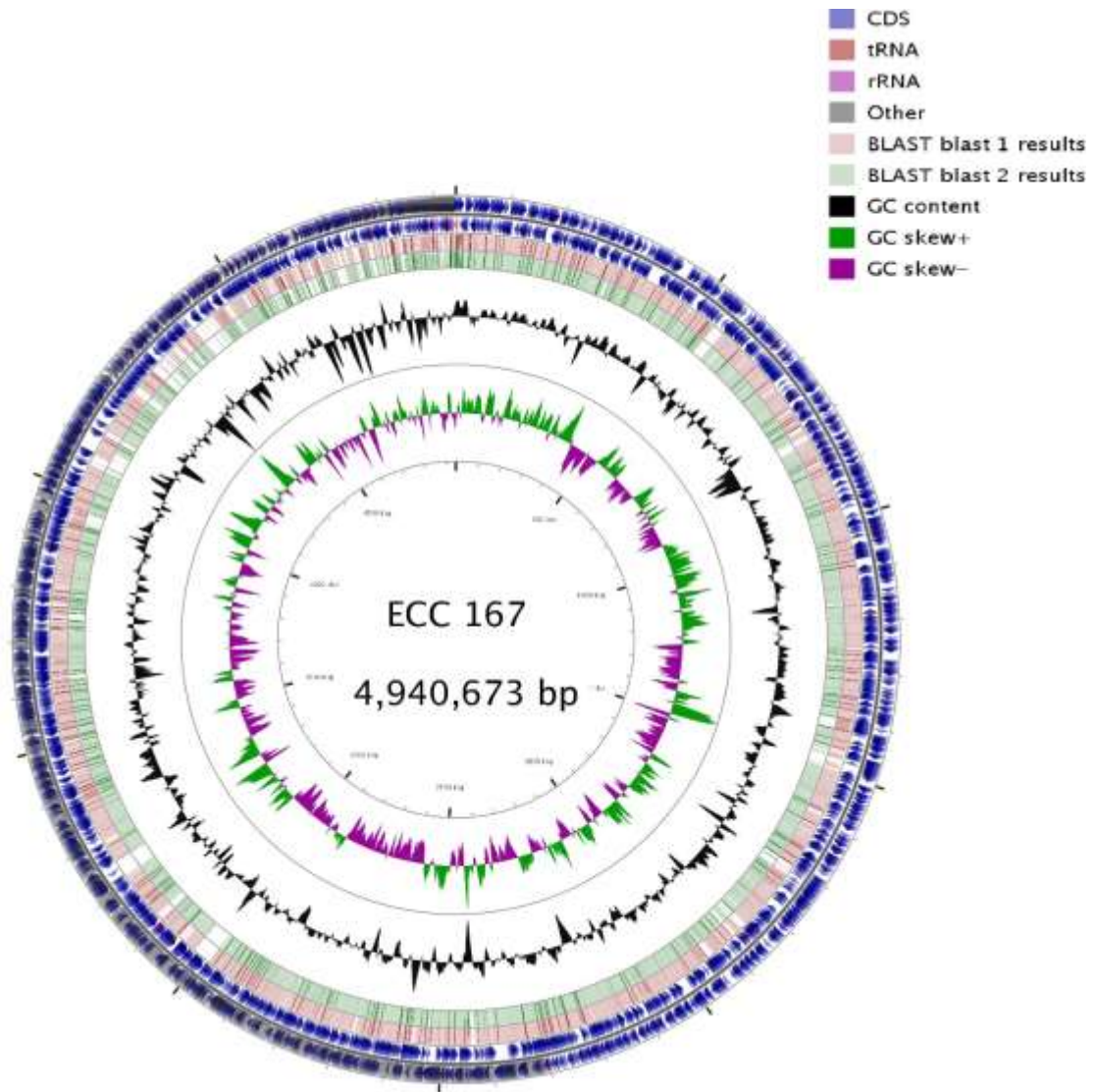


Figure 13. Comparative circular representation of the *E. coli* ECC 167 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* PCN033 (accession number AFAT000000000) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* PCN033 track, fourth

track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

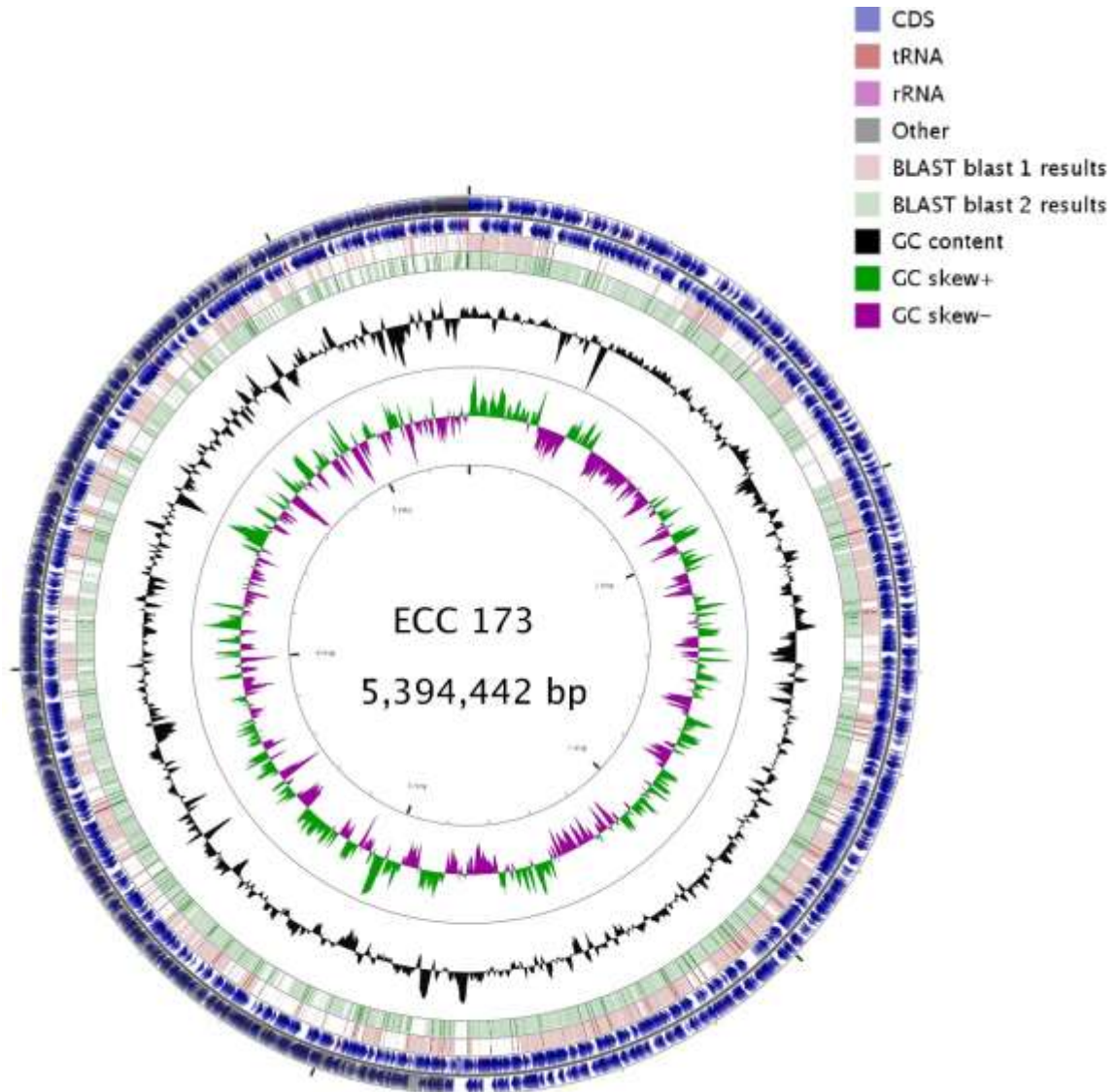


Figure 14. Comparative circular representation of the *E. coli* ECC 173 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession number NZ_AFET01000001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical

representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

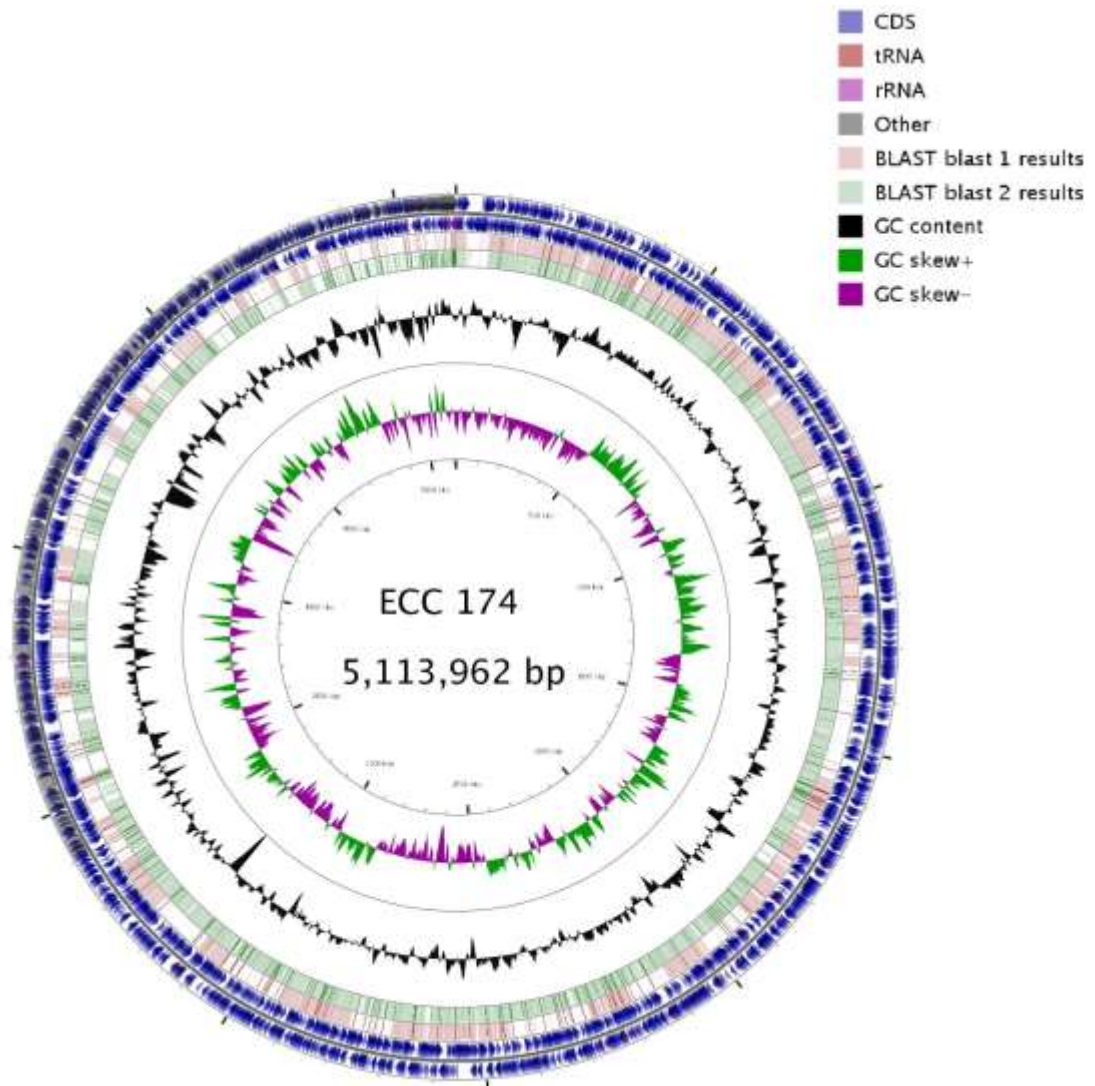


Figure 15. Comparative circular representation of the *E. coli* ECC 174 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession

number NZ_AFET01000001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

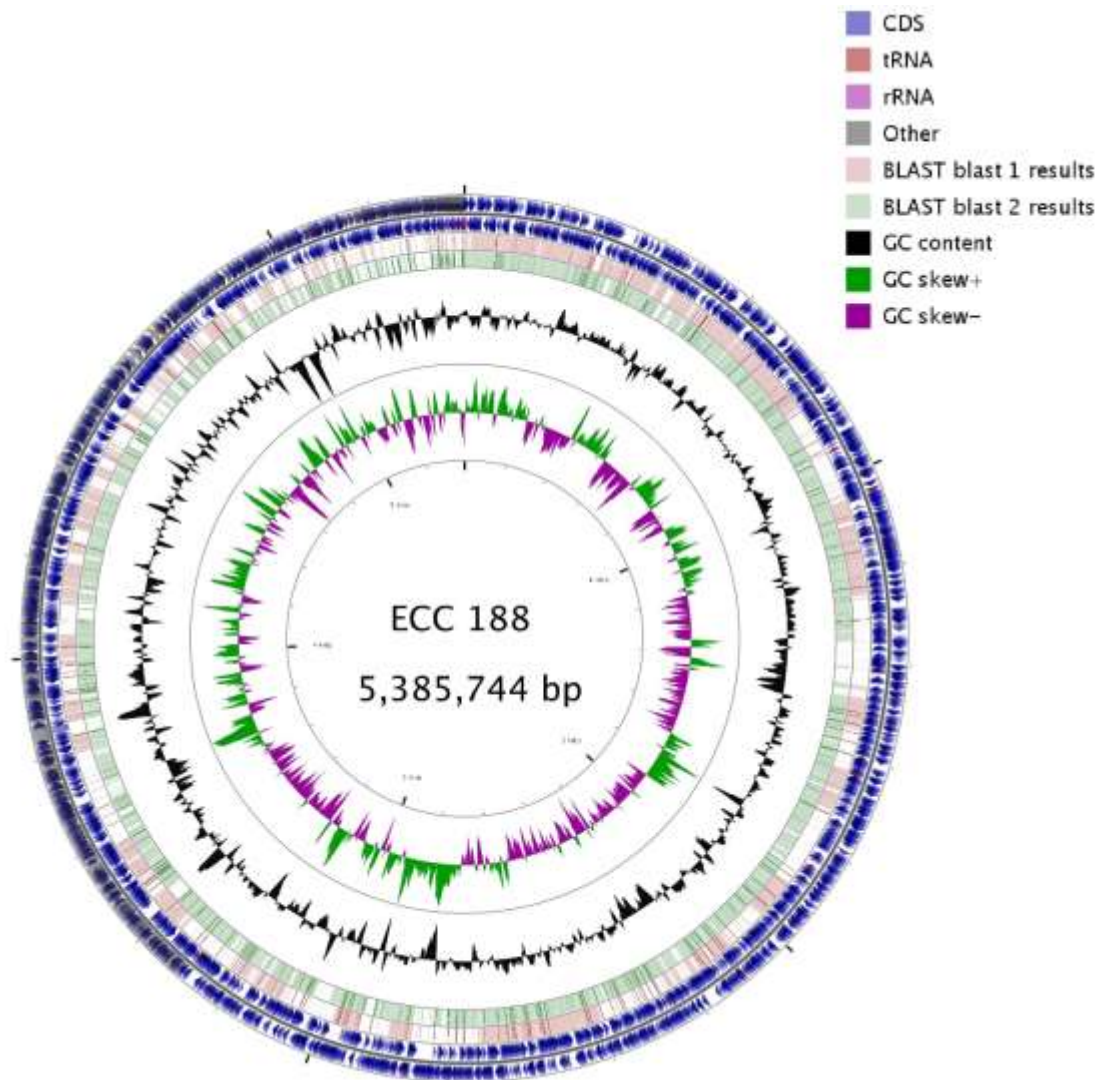


Figure 16. Comparative circular representation of the *E. coli* ECC 188 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* AA86 (accession number NZ_AFET01000001.1) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* AA86 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

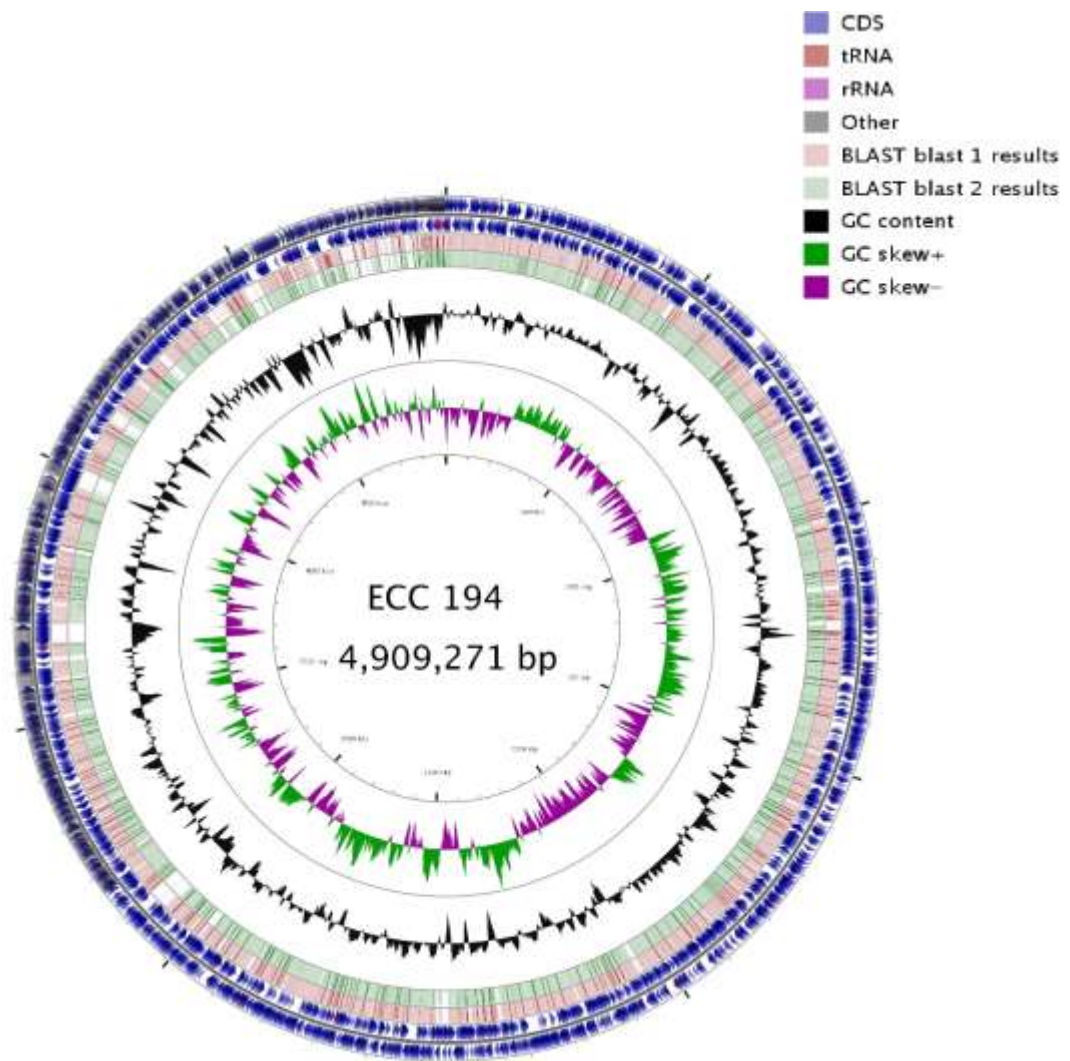


Figure 17. Comparative circular representation of the *E. coli* ECC 194 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* PCN033 (accession number AFAT000000000) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* PCN033 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

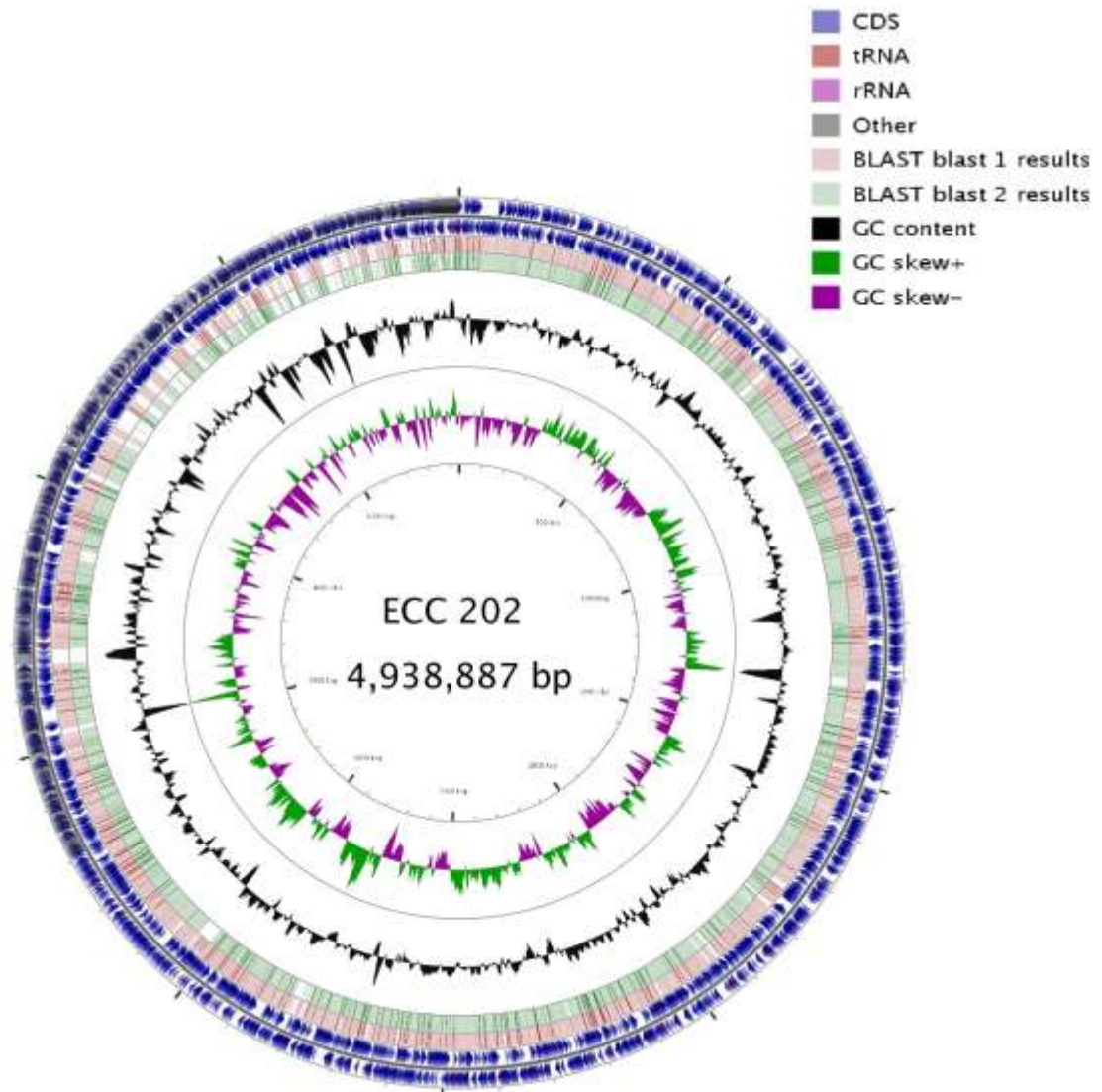


Figure 18. Comparative circular representation of the *E. coli* ECC 202 isolate genome BLASTed against the genomes of two reference strains. Blast 1: *E. coli* PCN033 (accession number AFAT000000000) and Blast 2: *E. coli* K-12 MG1665. This graphical representation of the genome was generated using CGview server. Circular tracks show (from outside inwards): *E. coli* coding DNA sequence on the forward strand, *E. coli* coding DNA sequence on reverse strand (rRNA and tRNA are colored according to the legend inside the first two tracks), third track is the BLAST 1 *E. coli* PCN033 track, fourth track is the BLAST 2 *E. coli* K-12 MG1665 track. Both third and fourth tracks represent the positions covered by the BLASTN alignment. Inside these tracks, white regions indicate parts of the input sequence that did not yield a blast hit, light pink/light green represents parts of the input sequence that yield one blast hit, and darker pink/green regions indicate parts of the input sequence that yield several blast hits (overlapping hits). These often include rRNA or tRNA genes or repetitive sequences which represents the positions covered by the BLASTN alignment. Then the GC content is shown in black and

finally the positive and negative GC skew are colored green and purple, respectively. Image created using CGview Server V 1.0 (Grant & Stothard, 2008).

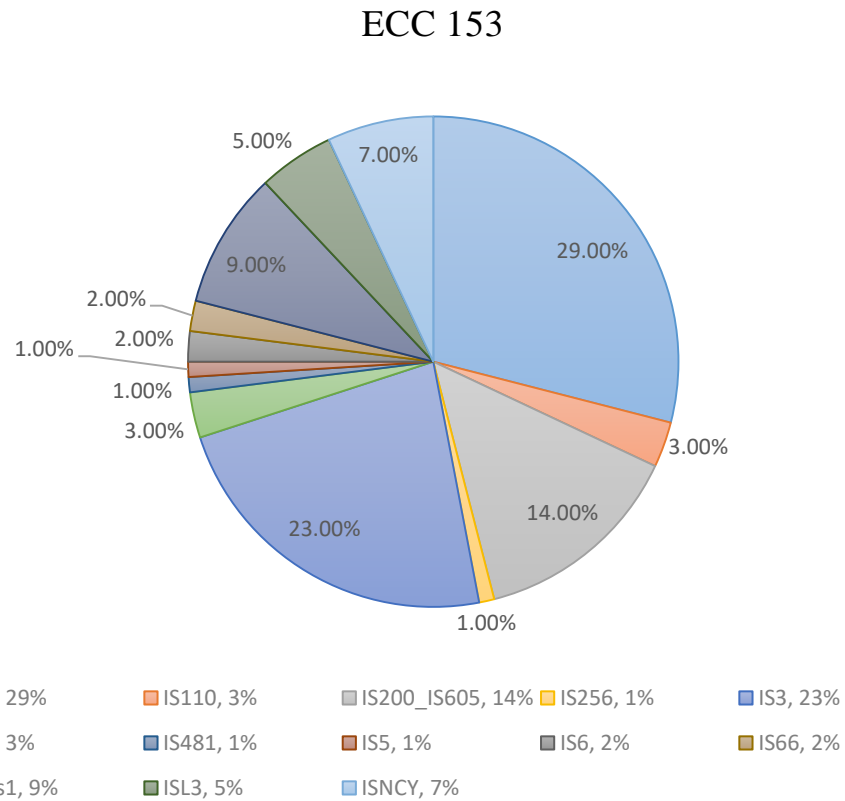


Figure 19. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 153.

ECC 157

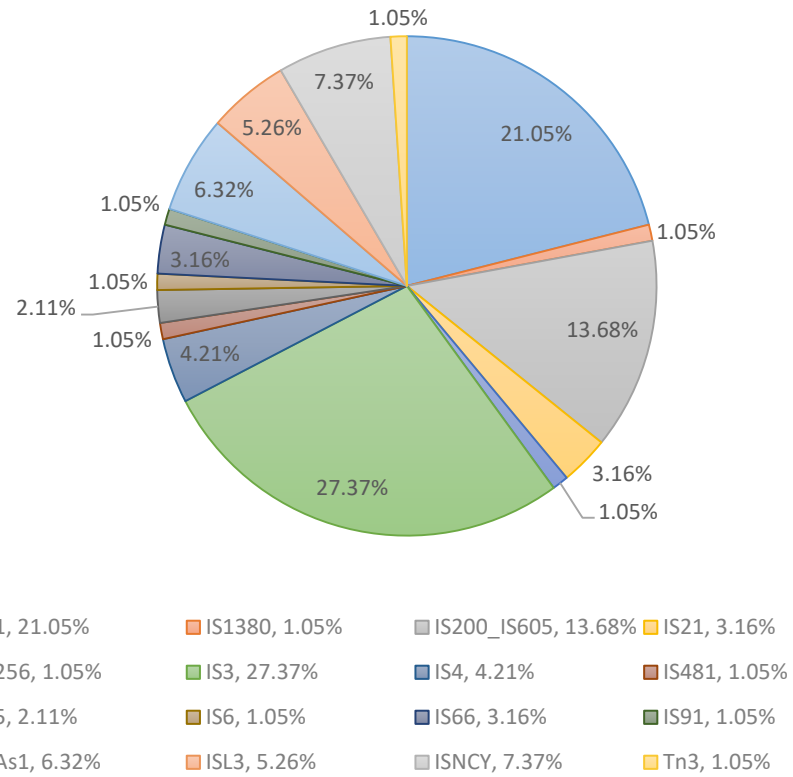


Figure 20. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 157.

ECC 161

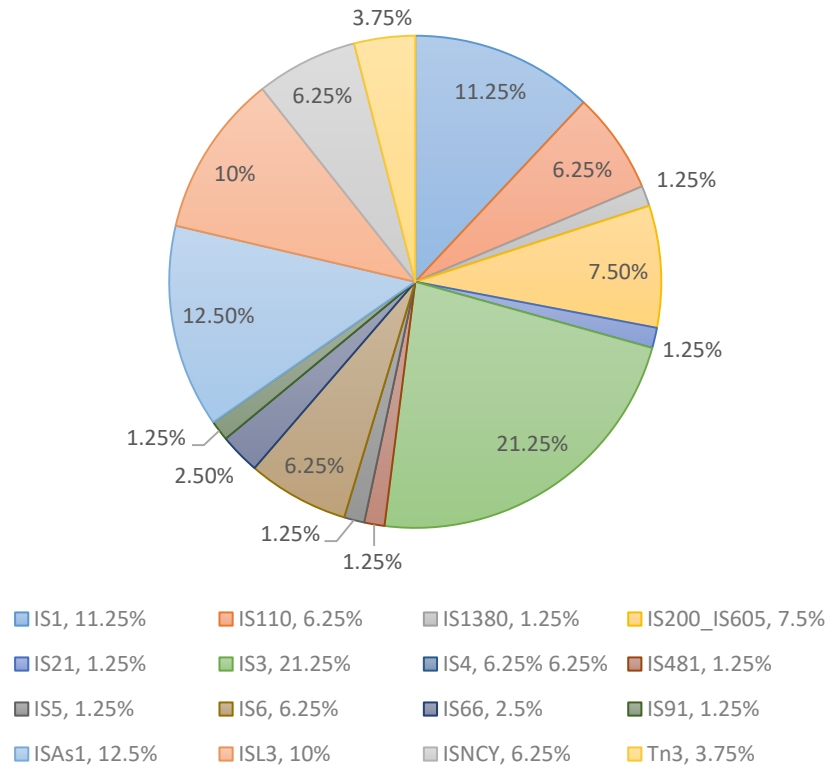


Figure 21. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 161.

ECC 167

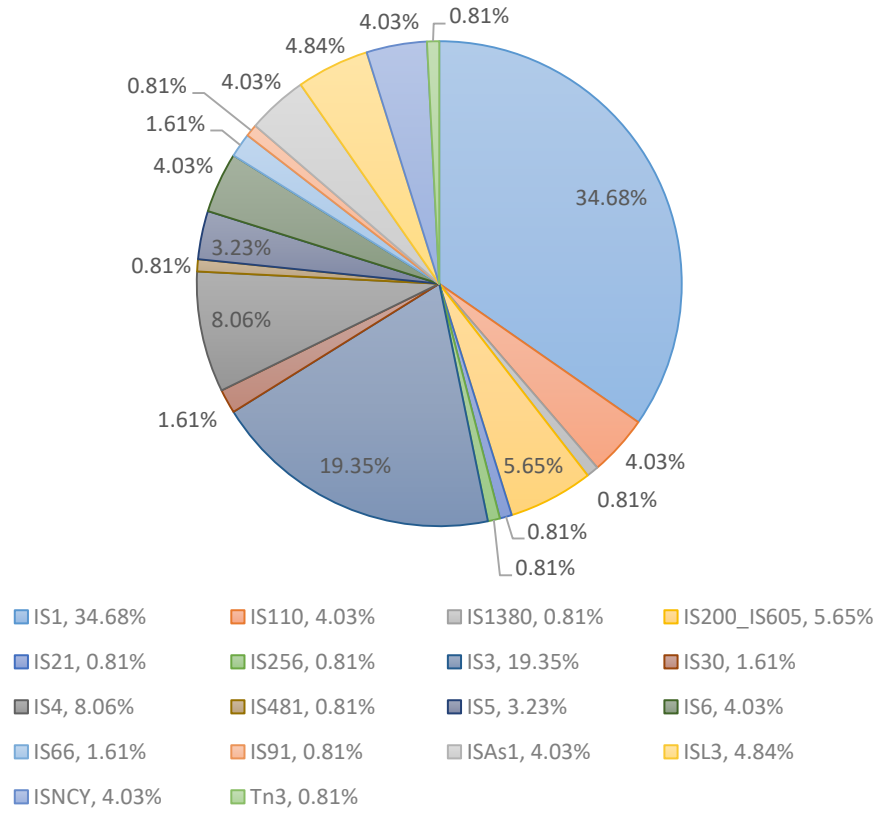


Figure 22. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 167.

ECC 173

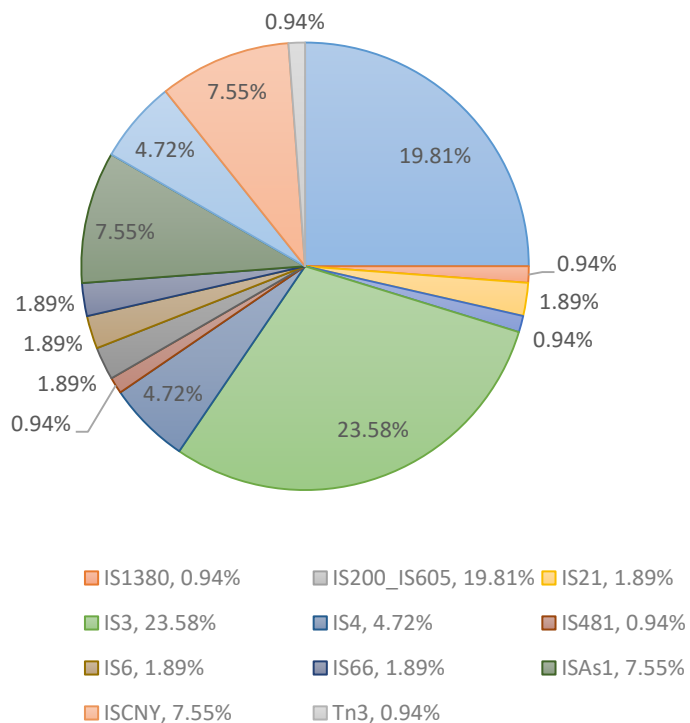


Figure 23. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 173.

ECC 174

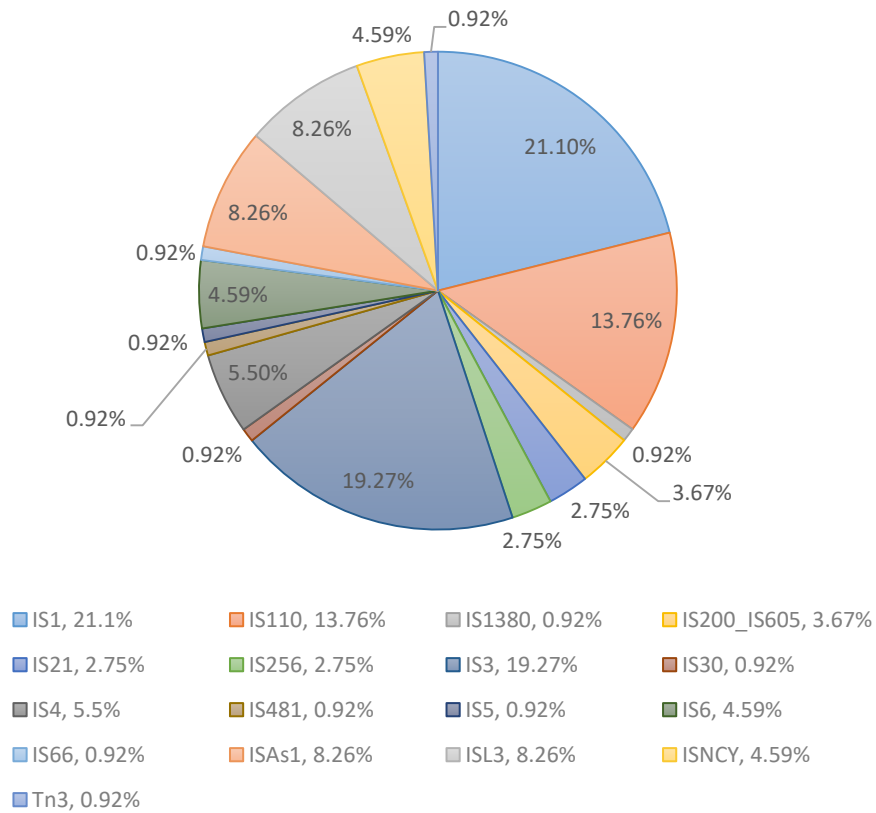


Figure 24. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 174.

ECC 188

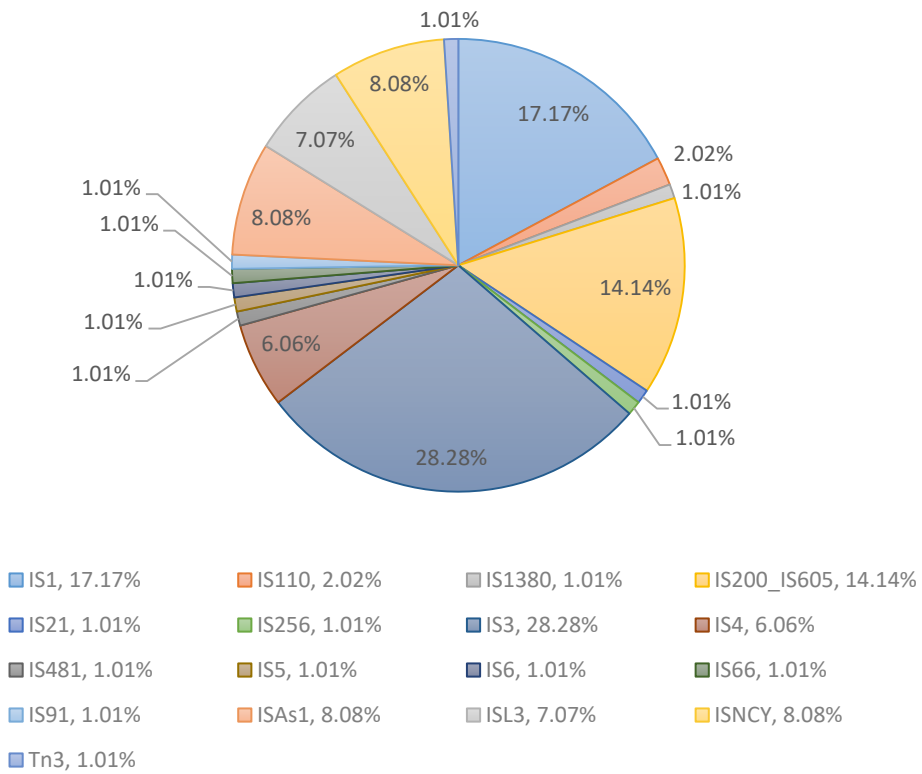


Figure 25. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 188.

ECC 194

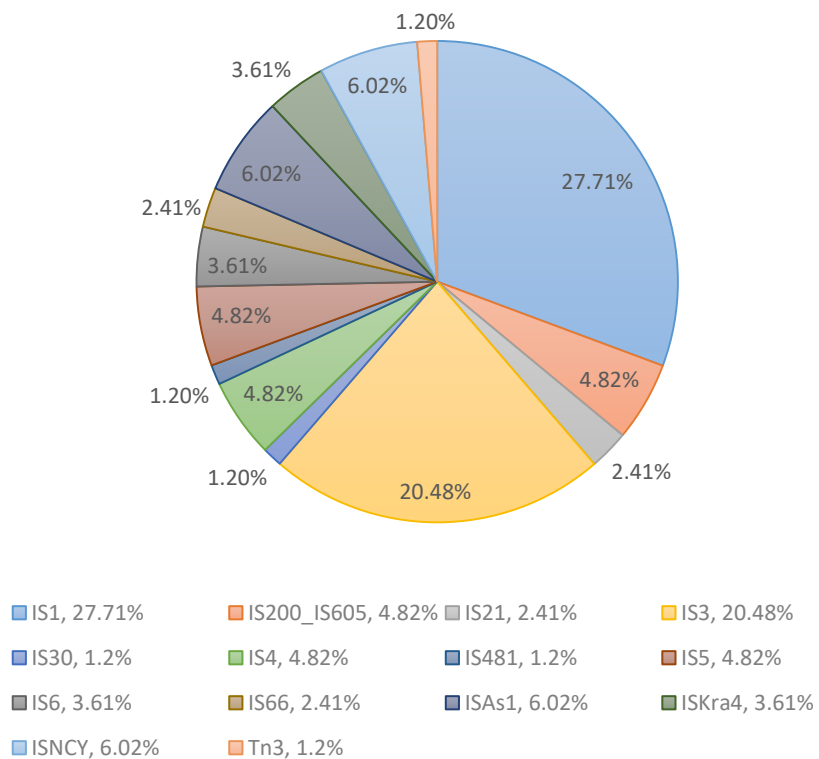


Figure 26. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 194.

ECC 202

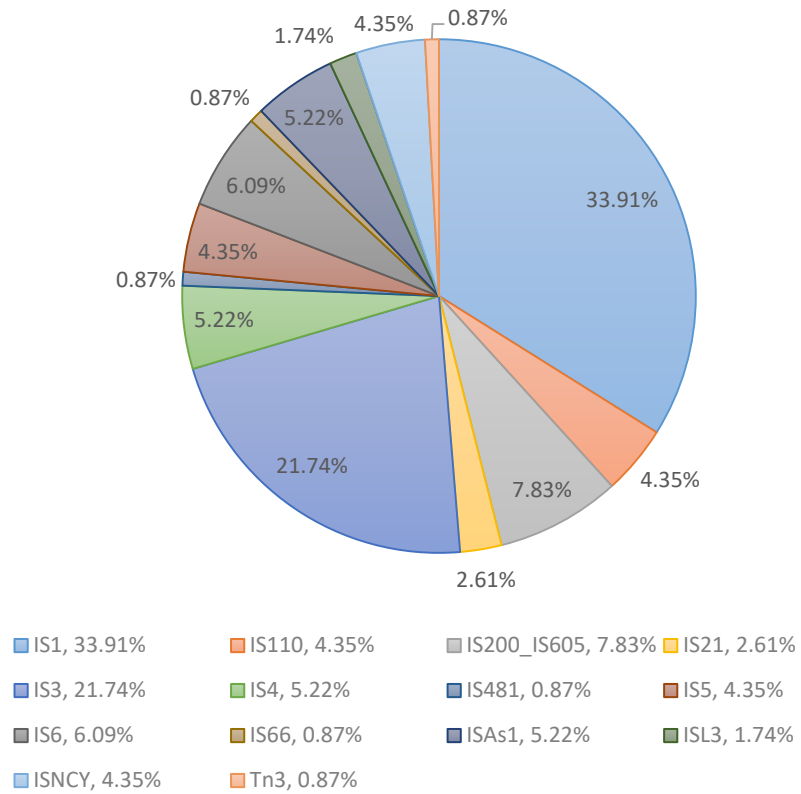


Figure 27. A pie chart demonstrating the different percentages of major insertion sequence families for isolate ECC 202.