LEBANESE AMERICAN UNIVERSITY

Whole-genome sequencing based molecular characterization of multi-drug resistant isolates of *Enterobacter* spp. and *Klebsiella aerogenes* in Lebanon.

By

Georgi S. Merhi

A thesis

Submitted in partial fulfilment of the requirements for

the degree of Master of Science in Molecular Biology

School of Arts and Sciences

July 2019

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS APPROVAL FORM

Student Name: Georgi Merhi I.D. #: 201605695

Thesis Title: Whole-genome Sequencing based molecular characterization of multi-drug resistant Enterobacter spp. and Klebsiella aerogenes from Lebanon.

Program: Molecular Biology

Department: Natural Sciences

School: of 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:

Master of Science in the major of Molecular Biology.

| Thesis Advisor's Name Sima Tokajian | Signature | DATE: 31 / 07 / 2019 (Day / Month / Year) |

| Committee Member's Name GEORGES KHAZEN | Signature | DATE: 31 / 07 / 2019 (Day / Month / Year) |

| Committee Member's Name Roy Khalaf | Signature | DATE: 31 / 07 / 2019 (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 to 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 format 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: Georgi Merhi

Signature:

Date: 22/07/2019

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

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: Georgi Merhi

Signature:

Date: 22/07/2019

Acknowledgments

I would like to express my utmost gratitude to my advisor Dr. Sima Tokajian for guiding and inspiring me to dedicate myself to the love of science. I am forever grateful for her support in my times of failure, encouragement in my time of growth and belief in me when I even did not believe in myself. In the last three years, you have taught me that nothing is out of reach once you put your heart into it. Thank you!

I would also like to thank my committee members, Dr. Roy Khalaf for his lighthearted attitude and academic help. A special thank you to Dr. Georges Khazen, the person who introduced my mind to the intricacies of bioinformatics. You have shaped a part of the path I wish to walk and taught me that no dataset is too hard to tackle.

To all my colleagues in the lab especially Tamara, Sahar, Balig, Harout, Ghiwa, Yara, Elie, Leila, Jennifer, Sara, Rita, Sarah, Maria, and Asalia. You've made these past few years the best I could hope for. Thank you for all the support, laughs and fun we've shared.

Last but not least, I would like to thank my family and friends for their unconditional love and support. Thank you Miguel & Marc for always being there for me, for the constant boosts and help keeping me sane during harsh times. Thank you to my beloved Sara for always taking care of me. Thank you Zouzou for always pushing me forward. Finally, thank you Dad for giving your everything for my education and growth, I am who I am today because of what you did.

I dedicate this work to my father and brother, Sami & Miguel Merhi.

Whole genome sequencing based molecular characterization of multi-drug resistant isolates of *Enterobacter* spp. and *Klebsiella aerogenes* in Lebanon.

Georgi S. Merhi

ABSTRACT

Enterobacter cloacae complex (ECC) members and *Klebsiella aerogenes* belong to the family Enterobacteriaceae and are known as Gram-negative, rod-shaped, opportunistic pathogens. They are commonly notorious for causing healthcare-associated infections and impose a significant burden on secondary health-centers. In the last few decades, the rise of ECC and *K. aerogenes* as successful nosocomial pathogens was paralleled by the global alarming emergence of multi-drug resistant clones of these bacteria. In Lebanon only few sporadic, and somewhat lacking, reports exist on multi-drug resistant ECC and *K. aerogenes* isolates. In our study whole-genome sequencing (WGS) was used for the molecular characterization of 15 *E. cloacae* complex and *K. aerogenes* recovered from clinical settings. The isolates were initially collected and characterized through *hsp60* genotyping and 16S rRNA gene sequencing. Pulse field gel electrophoresis (PFGE) and multi-locus sequence typing (MLST) were used to investigate the clonal relatedness and antimicrobial susceptibility was tested and confirmed through *in silico* analysis. Results were further confirmed through complimentary PCR assays and plasmid-based replicon typing (PBRT). Phylogenetic relatedness was assessed through whole-genome average nucleotide identity (wgANI) and core-genome single nucleotide polymorphism (cgSNP) based phylogenetic analysis. Eight isolates were identified as *Enterobacter hormaechei* (clusters III, VI, VIII) representing 88% (n=8) of the total studied *Enterobacter* isolates. 53% (n=8) of isolates exhibited resistance to one or more of the tested penicillin and cephalosporins, while only 20% (n=3) showed complete resistance to carbapenems. Novel STs were determined for ECC and *K. aerogenes* isolates and one ECC isolate (ST114) matched an epidemic clone. IncFII plasmids were detected in 47% (n=7) of the isolates, which conforms with the most common plasmid families in both ECC and *K. aerogenes*. One *bla*$_{CTX-M-15}$ positive ECC isolate had the ESBL gene integrated in the chromosome through an IS*Ecp1*-*bla*$_{CTX-M-15}$-*orf477*Δ transposition unit, while a second extremely resistant ECC isolate coharbored *bla*$_{NDM-1}$ and *bla*$_{CTX-M-15}$. Carbapenem resistance in *K. aerogenes* isolates was due to the additive effect of osmoporin inactivation (*omp36*) and constitutive chromosomal *ampC* expression. The results of this study showed the polyphyletic nature of the ECC and its dissemination in Lebanon and revealed its resistance mechanisms. The

generated data could be used for other comparative genomic studies to better understand the molecular epidemiology, the dynamics of dissemination and the genetic basis of multidrug resistance within nosocomial pathogens in Lebanon.

Keywords: *Enterobacter cloacae* complex, *Klebsiella aerogenes*, MDR, NDM-1, SNP, Phylogeny, Lebanon.

Table of contents

Chapter **Page**

I- Introduction……………………………………………………………..1

1.1. Overview of *Enterobacter* spp. ……………………………………1
1.1.1. Background………………………………………………………1
1.1.2. *E. cloacae* complex and its major clinically significant members....2
1.1.2.1. *E. cloacae*……………………………………………………..2
1.1.2.2. *E. hormaechei*…………………………………………………3
1.1.3. Taxonomy, whole-genome sequencing (WGS) & current phylogenetic clusters……………………………………………………………3
1.1.3.1. The Hoffmann clusters……………………………………………..3
1.1.3.2. Novel phylogenetic clusters………………………………………..5
1.1.4. *Enterobacter aerogenes,* reclassified………………………………7
1.1.4.1. *E. aerogenes* overview……………………………………………7
1.1.4.2. Reclassification into *Klebsiella aerogenes*………………………..7
1.2. Epidemiology and Typing…………………………………………..8
1.3. Virulence factors in *K. aerogenes*…………………………………9
1.4. Antibiotic resistance………………………………………………..10
1.4.1. Intrinsic resistance…………………………………………………10
1.4.2. Acquired resistance…………………………………………………11
1.4.2.1. ESBLs in ECC and *K. aerogenes*…………………………………12
1.4.2.2. Carbapenem resistance……………………………………………12
1.4.2.2.1. Carabapenemase plasmid families in ECC……………………..13
1.4.2.2.2. Porin Mutations in *K. aerogenes*………………………………..13
1.5. Incidence in Lebanon and Aim……………………………………14

II- Materials and Methods………………………………………………16

2.1. Bacterial Isolates……………………………………………………16
2.2. Antimicrobial susceptibility testing………………………………17
2.3. DNA Extraction……………………………………………………..18
2.4. PCR amplification and sequencing of the *hsp60* gene…………...18
2.5. Pulse-Field Gel Electrophoresis (PFGE)…………………………18
2.6. Multi-Locus sequence typing (MLST)…………………………..19
2.7. PCR based replicon typing (PBRT)……………………………...20
2.8. Whole-Genome Sequencing……………………………………..20
2.8.1. Library Preparation………………………………………………20
2.8.2. Quality control and assembly……………………………………20
2.8.3. Annotation…………………………………………………………21
2.8.4. *In silico* resistance, virulence, MLST profiling and mobile genetic elements (MGEs)……………………………………………………21

2.8.5. PacBio sequencing and novel plasmid sequence………………22
2.8.6. Pan genome and Recombination Analysis……………………..22
2.8.7. Core genome (cg) SNP phylogenetic analysis…………………22

III- Results………………………………………………………………24

3.1. *In silico* primary identification…………………………………..24
3.2. Susceptibility profiling…………………………………………..25
3.3. *hsp60* identification and clustering………………………………27
3.4. wgANI identification…………………………………………….28
3.5. Pan genome and recombination analysis…………………………29
3.6. Core genome (cg) SNP phylogenetic analysis……………………35
3.7. Virulence Factors…………………………………………………37
3.8. PFGE……………………………………………………………..38
3.9. ST characterization and molecular resistance profiling…………41
3.10. Plasmid Typing…………………………………………………...44
3.11. Novel bla_{NDM-1} plasmid…………………………………………45

IV- Discussion…………………………………………………………48

References……………………………………………………………..…55-66

LIST OF TABLES

Table **Page**

1. **Introduction:** Novel phylogenetic and Hoffmann clusters for the *E. cloacae* complex…………………………………………………………......6
2. **Material and Methods:** Primary isolate information including the corresponding patient's sex and age, isolation source and whole-genome NCBI accession numbers..16
3. **Material and Methods:** *hsp60* and MLST amplification and sequencing primers with annealing positions in the target gene………………………………….……………………..19
4. **Results:** *16S rDNA* Sequencing identification results versus the in silico KmerFinder tool output……………………………………….…24
5. **Results:** Kirby-Bauer disk diffusion assay for the 18 antibiotics used in this study. The antimicrobial agents used belonged to seven classes of antimicrobial agents. S: Sensitive; R: Resistant; I: Intermediate Resistance ……………………………………………………………………26
6. **Results:** Kleborate analysis output. Virulence and resistance systems detected along the yersiniabactin (*ybt*) and colibactin (*cbl*) loci sequence type. A value of 0 indicates no hits, 1 = one identified system, 2 = two identified systems……………………………………………………………...38
7. **Results:** MLST allelic variations for all housekeeping genes loci for all isolates (N=15). Allelic combinations with no match were termed as novel and submitted to the database curators………………………………...42

8. **Results:** Plasmid based replicon typing and in silico based plasmid incompatibility profiling of study isolates (n=15). IncF isoforms were identified in 46% (n=7) of isolates……………………………………..44

LIST OF FIGURES

Figure **Page**

1. **Results:** A rooted UPGMA tree based on the hsp60 sequence analysis of all ECC (n=9) isolates against references matching (n=10) each Hoffmann cluster (Acession # AJ417108 to AJ417143). The tree was generated using clustal Omega multiple sequence alignment tool on EMBL-EBI……………26
2. **Results:** All vs All Average Nucleotide Identity (ANI) percentage relatedness matrix of ECC isolates and 6 reference genomes representing the five subspecies of *E. hormaechei* and *E. cloacae subsp. cloacae*. Red blocks represent isolate relatedness higher than 97% and white blocks indicate relatedness values lower than 96%..29
3. **Results:** Representation of pan genome analysis denoting gene absence and presence (right) with its associated accessory binary genes tree (left) of ECC isolates (ENM1 → ENM9). Adjacent blue blocks on the left represent core genes shared between isolates whilst scattered block on the right represent cloud and shell accessory genes………………………………………31
4. **Results:** Representation of pan genome analysis denoting gene absence and presence (right) with its associated accessory binary genes tree (left) of K. aerogenes isolates (KAM1 → KAM6). Adjacent blue blocks on the left represent core genes shared between isolates whilst scattered block on the right represent cloud and shell accessory genes…………………………….32
5. **Results:** Recombination blocks within the core genome alignment of ECC (ENM1→ENM9) isolates with relative metadata heatmap. Red blocks denote shared recombination sites and blue block signal independent recombination

sites. Major recombination hotspots are marked with dark blue circles (n=2) emphasizing presence/absence of shared recombination……………..33

6. **Results:** Recombination blocks within the core genome alignment of *K. aerogenes* (KAM1→KAM6) isolates with relative metadata heatmap. Red blocks denote shared recombination sites and blue block signal independent recombination sites. Major recombination hotspots are marked with dark blue circles (n=3) emphasizing presence/absence of shared recombination……………………………………………………………34

7. **Results:** RAxML generated clean core maximum-likelihood tree with 1000 bootstrap support elucidating the phylogenetic relationship between ECC isolates (ENM1→ENM9) with one *E. hormaechei* subsp. *oharae* reference genome (accession # LRJW01000000). MGEs and recombination events were removed prior to generating the tree………………………………….35

8. **Results:** RAxML generated clean core maximum-likelihood tree with 1000 bootstrap support elucidating the phylogenetic relationship between ECC isolates (ENM1→ENM9) with one *E. hormaechei* subsp. *oharae* reference genome (accession # LRJW01000000). MGEs and recombination events were removed prior to generating the tree…………………………………..36

9. **Results:** *wzi* loci alignment for all K. aerogenes isolates (n=6) with the locus allelic type shown. Purple color indicates genes that are present while beige indicates gene absence and deletions………………………………….37

10. **Results:** Pulse-field gel electrophoresis elucidating the clonality of isolates based on their banding pattern obtained through *Xba*I restriction digestion. Dendograms were generated using BioNumerics v7.6.1. A) PFGE-based dendogram for ECC isolates (n=9) showcasing 7 different pulsotypes. B)

PFGE-based dendogram for *K. aerogenes* isolates (n=6) demonstrating 3 different puslotypes…………………………………………………...39

11. **Results:** *bla*$_{CTX\text{-}m\text{-}15}$ genetic environment comparison between ENM1 and ENM2. Both chromosomal and plasmid encoded copies of *bla*$_{CTX\text{-}m\text{-}15}$ are carried by an IS*Ecp1*-*bla*$_{CTX\text{-}M\text{-}15}$-*orf477*Δ transposition unit………..42
12. **Results:** Alignment of outer membrane protein (omp36) encoding genes and their genetic environment. KAM1 → KAM4 show a functional *omp36* gene sequence (red arrow) while KAM5 and KAM6 carry a truncated form of the gene (truncated red arrows)…………………………………………42
13. **Results:** Molecular characterization and attributes of all study isolates (n=15). Correct nomenclature and relative subspecies. 14 sequence types (including novel STs) and their molecular characteristics such as plasmid carriage and different categories of resistance determinants………………………..44
14. **Results:** Overview of plasmid pNDM1_ENM2. Complete sequence of the 104318 bp plasmid, fully annotated. *tra* genes are colored in blue, *parA* and *parB* genes colored in teal and IncFII *repA* gene colored in navy blue. *bla*$_{NDM\text{-}1}$ is flanked a bleomycin resistance gene (*mbl*) downstream and an IS*Aba125* upstream……………………………………………………….45
15. **Results:** Genetic environment of *bla*$_{NDM\text{-}1}$ allocated on the pDNM1_ENM2 plasmid. IS*Ab125* (cyan) is found upstream of *bla*$_{NDM\text{-}1}$ (red), while a bleomycin resistance gene (orange) and a Tn*5403* (navy blue) are found downstream………………………………………………………….46

LIST OF ABBREVIATIONS

LRTI	Lower Respiratory Tract Infection
UTI	Urinary Tract Infection
AMR	Antimicrobial resistance
ECC	*Enterobacter cloacae* complex
HAI	Healthcare Associated Infection
WGS	Whole Genome Sequencing
DNA	Deoxyribonucleic Acid
ANI	Average Nucleotide Identity
SNP	Single Nucleotide Polymorphism
HGT	Horizontal Gene Transfer
ORF	Open Reading Frame
ODC	Ornithine Decarboxylase
PFGE	Pulse-Field Gel Electrophoresis
MLSA	Multi-Locus Sequence Analysis
ST	Sequence Type
ESBL	Extended Spectrum β-lactamases
CREC	Carbapenem Resistant *E. cloacae*
VFs	Virulence factors

CDC	Center for Disease Control and Prevention
CRE	Carbapenem Resistant *Enterobacteriaceae*
CP-CRE	Carbapenemase Producing-Carbapenem Resistant *Enterobacteriaceae*
MGE	Mobile Genetic Elements
Kbp	Kilo-Base Pairs
Inc	Incompatibility
PBRT	PCR-based Replicon Typing
IS	Insertion Sequence
KPC	*Klebsiella pneumoniae* Carbapenemase
MBL	Metallo-β-Lactamase
NDM	New-Delhi Metallo-β-Lactamase
IMP	Imipenemase
VIM	Verona integron-encoded Metallo-β-Lactamase
CR-KE	Carbapenem Resistant-*Klebsiella aerogenes*
Omp	Outer Membrane Pori
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time of Flight
CLSI	Clinical & Laboratory Standard Institute
gDNA	Genomic DNA
XDR	Extremely Drug Resistant

Chapter One

Introduction

1.1. Overview of *Enterobacter sp.*

1.1.1. Background

Little did researchers know that the discovery of microorganisms, almost 400 years ago, would introduce humanity to a vast new dimension dubbed the science of microbes or microbiology. Fast forward to our current day and age, our interactions with the microscopic entities known as bacteria has grown exponentially, yet can only be summarized with one question: are they friend or foe (Gest, 2004)?
Bacteria belonging to the *Enterobacter* genus, members of the family *Enterobacteriaceae*, are a prime example of the mentioned duality. They are generally characterized as Gram-negative, rod-shaped, motile, non-spore forming and engage in saprophytic activity in the environment. These bacteria are usually found in soil and sewage but most importantly, they are enteric commensals in the human gastrointestinal tract (Sanders & Sanders, 1997). Currently, the *Enterobacter* genus encompasses about 19 species which include: *Enterobacter cloacae, E. asburiae, E. cancerogenus, E. chengduensis, E. chuandaensis, E. hormaechei, E. kobei, E. ludwigii, E. roggenkampii, E. sichuanensis, E. mori, E. huaxiensis, E. nickellidurans, E. oligotrophica, E. siamensis, E. soli, E. tabaci* and *E. timonensis* (Federhen, 2012). Amongst the aforementioned species, many are considered as plant pathogens and are rarely linked to clinical settings and nosocomial scenarios. However, in the last few

decades, a few *Enterobacter* species surfaced as potent opportunistic pathogens causing the like of lower respiratory and urinary tract infections (LRTI & UTI) while also demonstrating an exponential increase in prevalence within a limited span of time (Hoffmann & Roggenkamp, 2003). As a result, the *Enterobacter* genus has been recently included into the ESKAPE group of pathogens and labeled as posing critical threat through worsening the spread of global antimicrobial resistance (AMR) (WHO, 2017). Moreover, these clinically relevant species were represented mainly by *E. cloacae* and named as the *Enterobacter cloacae* complex or ECC (Mezzatesta, Gona, & Stefani, 2012).

1.1.2. *E. cloacae* complex and its major clinically significant members.

The ECC is a polyphyletic complex mirrored by ten species from the *Enterobacter* genus (Mezzatesta, Gona, & Stefani, 2012). However, not all species can be the direct cause of healthcare associated infections (HAIs), such as the recently classified *E. mori*, which shared high relatedness with other members of the ECC yet was found to be a mulberry tree pathogen (Zhu et al., 2011). Nevertheless, the two most prominent nosocomial pathogens from the ECC remain, to this date, *E. cloacae and E. hormaechei* and subsequent subspecies (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019).

1.1.2.1. *E. cloacae*

E. cloacae was first identified in 1890 under the name *Bacillus cloacae* and has, since then, undergone multiple taxonomic reassignments (Mezzatesta, Gona, & Stefani, 2012). In 1960, Hormaeche and Edwards defined the genus *Enterobacter* hence reclassifying the bacterium as it is currently named (Hormaeche & Edwards, 1960). A major advantage for the

E. cloacae species is its large genetic heterogeneity which, in turn, confers the ability to survive and adapt in multiple niches. This enables it to colonize and contaminate medical equipment in hospitals with relative ease, thus, making *E. cloacae* a successful common and nosocomial pathogen with a significant looming threat (Dalben et al., 2008; Mezzatesta, Gona, & Stefani, 2012; Musil, Jensen, Schilling, Ashdown, & Kent, 2010).

1.1.2.2. *E. hormaechei*

From the early 1970's until the late 1980's, 22 strains from this species (unclassified at that time) were isolated from multiple outbreak scenarios and bacteremia cases notably in Brazil and USA (Campos, Lobianco, Seki, Santos, & D Asensi, 2007; Wenger et al., 1997). These strains presented biochemical profiles that were unfamiliar, yet, highly similar to isolates belonging the *Enterobacter* genus. This led to its reclassification from "enteric group 75" to *E. hormaechei* in 1989 (Mezzatesta, Gona, & Stefani, 2012; O'Hara et al., 1989). From that point onward, 5 subspecies were identified and classified as follows: *E. hormaechei* subsp. *oharae*, *E. hormaechei* subsp. *steigerwaltii, E. hormaechei* subsp *hormaechei, E. hormaechei* subsp *hoffmannii* and *E. hormaechei* subsp *xiangfangensis* (Hoffmann et al., 2005; Sutton, Brinkac, Clarke, & Fouts, 2018). However, the prevalence of *E. hormaechei* remains significantly underestimated due to the lack of sensitivity in routine surveillance studies and reliance on elementary phenotypic identification techniques in clinical settings (Morand et al., 2009).

1.1.3. Taxonomy, whole-genome sequencing (WGS) & current phylogenetic clusters.

1.1.3.1. The Hoffman clusters

Primitive attempts at differentiating and classifying members of ECC were solely reliant on the biochemical properties and antimicrobial resistance patterns of each distinct bacterial species. However, these methods proved ineffective in resolving a clear and curated scheme for clinical characterization of this complex (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019). Moreover, utilizing the traditional 16S ribosomal DNA sequence analysis technique yielded poor results in resolving this diverse cluster (Mezzatesta, Gona, & Stefani, 2012).

By sequencing a fragment of the *hsp60* heat-shock gene, Hoffmann and Roggenkamp were able to propose a genetic structure for the ECC that encompasses 13 genetic clusters (Hoffmann & Roggenkamp, 2003). The distribution of the clusters is as follows: Cluster I as *E. asburiae*, cluster II as *E. kobei*, cluster III as *E. hormaechei* subsp *hoffmannii*, cluster IV as *E. roggenkampii*, cluster V as *E. ludwigii*, cluster VI as *E. hormaechei* subsp *oharae* or subsp *xiangfangensis*, cluster VII as *E. hormaechei* subsp *hormaechei*, cluster VIII as *E. hormaechei* subsp *sterigerwaltii*, cluster IX as *E. bugandensis*, cluster XI as *E. cloacae* subsp *cloacae* and cluster XII as *E. cloacae* subsp *dissolvens* (Hoffmann & Roggenkamp, 2003; Hoffmann et al. 2005a,b,c; Sutton, Brinkac, Clarke, & Fouts, 2018). Initially, cluster X was classified as native to *E. nimipressuralis*. However, following taxonomic re-evaluation the bacterium was reclassified as *Lelliotia nimipressuralis* (Brady, Cleenwerck, Venter, Coutinho, & De Vos, 2013).

1.1.3.2. Novel phylogenetic clusters

The use of a single locus for species identification, within the ECC, proved somewhat limiting in terms of sensitivity. Therefore, this warranted an in-depth and comprehensive approach for precise species identification and novel insights into phylogeny (Chavda et al., 2016). WGS has revolutionized the way we generate and interpret data, especially in the field of microbiology. With the rapid advancements in novel sequencing technologies, it is now possible to sequence a single bacterial genome at the modest cost of 50 US dollars (Land et al., 2015). These factors have massively increased data output and established an international consensus for the need of WGS in microbial research and routine pathogen diagnostics (Kluytmans-van den Bergh et al., 2016; Köser et al., 2012). Incorporating WGS data analysis in routine clinical microbiological procedures has significantly enhanced accurate species identification, antimicrobial resistance testing and epidemiological typing (Köser et al., 2012).

The genomes of 97 *Enterobacter* isolates were sequenced and measured against available genomes in the genbank database using average nucleotide identity (ANI) comparison and single nucleotide polymorphism (SNP) phylogenetic analysis (Chavda et al., 2016). Following which, the ECC was divided into 18 phylogenetic clusters (A to R) with additional amendments being added to the scheme in subsequent and more recent studies as presented below in table 1 (Chavda et al., 2016; Sutton, Brinkac, Clarke, & Fouts, 2018).

Table 1 Novel phylogenetic and Hoffmann clusters for the *E. cloacae* complex.

Species and Subspecies Names	Phylogenetic Groups	Hoffmann Clusters
E. hormaechei ssp. *xiangfangensis*	A	VI
E. hormaechei ssp. *sterigerwaltii*	B	VIII
E. hormaechei ssp. *oharae*	C	VI
E. hormaechei ssp. *hoffmannii*	D	III
E. hormaechei ssp. *hormaechei*	E	VII
E. mori	F	N/A
E. cloacae ssp. *cloacae*	G	XI
E. cloacae ssp. *dissolvens*	H	XII
E. ludwigii	I	V
E. asburiae	J	I
E. cloacae complex clade K	K	N/A
E. cloacae complex clade L	L	N/A
E. roggenkampii	M	IV
E. cloacae complex clade N	N	N/A
E. cloacae complex clade O	O	N/A
E. cloacae complex clade P	P	N/A
E. kobei	Q	II
E. bugandensis	R	IX

1.1.4. *Enterobacter aerogenes,* reclassified.

1.1.4.1. *E. aerogenes* overview

E. aerogenes is, also, a motile, non-spore forming, rod shaped, Gram-negative bacterium. *E. aerogenes* takes on clinical significance as an opportunistic pathogen that it is often the cause of HAIs. It is re-emerging as a hazardous "killer bug" and is responsible for urinary tract infections, wound/skin infections and central nervous system infections around the globe (Davin-Regli & Pagès, 2015; Sanders & Sanders, 1997). The potency of *E. aerogenes* as a human pathogen is underlined by its ability to efficiently colonize multiple reservoirs in hospitals and inflate risk factors in such environments (Malek et al., 2019).

1.1.4.2. Reclassification into *Klebsiella aerogenes*

In the light of WGS data analysis, *E. aerogenes* was recently classified as *K. aerogenes* (Tindall, Sutton, & Garrity, 2017). The premature classification of *K. aerogenes* was primarily based on phenotypic discrepancies between the *Enterobacter* and *Klebsiella* genera. However, genomic insights into multiple *K. aerogenes* isolates demonstrated that these differences are caused by horizontal gene transfer (HGT) events (Diene et al., 2013). *K. aerogenes* has acquired a 48 functional open reading frames (ORF) operon, native to the *Serratia* genus, which codes for a flagellar assembly system. Similarly, ornithine decarboxylase (ODC) and urease activity were acquired through genetic material exchange (Diene et al., 2013). Chavda et al, examined the phylogenetic relationship between all members of the family *Enterobacteriaceae* through twenty-six universal marker gene analysis, in which *K. aerogenes* genomes clustered

with isolates form the *Klebsiella* genus (Chavda et al., 2016). The combined results of these studies clearly warranted the taxonomical reassignment of the bacterium as *K. aerogenes* (Tindall, Sutton, & Garrity, 2017).

1.2. Epidemiology and Typing

Molecular epidemiological typing of isolates, in outbreak scenarios, is a key step for understanding the pattern of dissemination of threatening clones locally and globally (Haertl & Bandlow, 1993). Outbreaks reports due to ECC clones are steadily increasing and are cause for major concern. Pulse-Field Gel Electrophoresis (PFGE) has long been the gold standard for fingerprinting ECC pathogens. However, PFGE's lack of reproducibility prevented microbial epidemiologists from curating a global scheme for typing ECC epidemic clones (Kluytmans-van den Bergh et al., 2016). In 2008, multi-locus sequence analysis (MLSA) revealed two timely ancestral clades and identified a 1.04 recombination ratio which accounts for the distance within the ECC (Paauw et al., 2008). A multi-locus sequence typing scheme (MLST) emerged by adding *dnaA* to the 6 housekeeping genes used in the MLSA (Miyoshi-Akiyama, Hayakawa, Ohmagari, Shimojima, & Kirikae, 2013). MLST provided a common ground where endemic and epidemic ECC clones can be categorized and assessed by their geographical distribution. Currently, four major sequence types (ST) are associated with extended spectrum β-lactamases (ESBLs) secretion (ST66, ST78, ST108 and ST114). Whereas, ST90, ST93, ST105 and ST114 were detected as clones exhibiting carbapenem resistance also known as carbapenem resistance *E. cloacae* (CREC) (Izdebski et al., 2015; Peirano et al., 2018).

Recently, a MLST scheme was developed for *K. aerogenes* ("Klebsiella aerogenes MLST", 2018). However, data related to the global distribution of threatening *K. aerogenes* clones is still scarce and lacking (Malek et al., 2019).

1.3. Virulence factors in *K. aerogenes*

When phylogenetically screened against other members of the family *Enterobacteriaceae*, *K. aerogenes* isolates clustered closely to *Klebsiella pneumoniae* (Chavda et al., 2016). Consequently, further analysis in the virulome of *K. aerogenes* isolates showed that this bacterium demonstrates similar virulence factors (VFs) loci with its closest phylogenetic match (Malek et al., 2019).

Pathogenic bacteria are in need of iron while partaking in cellular growth and replication within a host. Human hosts, on the other hand, possess a myriad of iron sequestering proteins which serve as a defense mechanism against invading pathogens (Podschun & Ullmann, 1998; Shon, Bajwa, & Russo, 2013). Therefore, iron acquisition systems or siderophores are vital for securing iron ions within the host (Podschun & Ullmann, 1998). Hence, *K. aerogenes* relies on siderophores for its proliferation and establishing an infection within its targeted host (Malek et al., 2019).

Yersiniabactin, a siderophore encoded by the *ybt* locus, has been recently detected in *K. aerogenes* isolates. The *ybt* is generally mobilized through an integrative conjugative element (ICE) designated ICE*Kp* (Holt et al., 2015; Lam et al., 2018; Malek et al., 2019). The *clb* locus, which encodes the secretion of colibactin, is another VF that is usually co-transferred within

the ICE*Kp*. Colibactin is a genotoxic polyketide that induces DNA damage in host cells during bacterial infections (Lam et al., 2018).

1.4. Antibiotic resistance

The center for disease prevention and control (CDC) has classified antimicrobial resistance as reaching a critical juncture and posing alarming threat to public health sectors (CDC, 2019). Multiple important industries stand to be affected and by 2050 researchers estimate that, if no measures are taken, 10 million deaths per year will be caused by pan resistant bacteria (Tangcharoensathien et al., 2017).

One of the key protagonists in AMR are bacteria belonging to the family *Enterobacteriaceae*. They can be classified as: ESBL-producing *Enterobacteriaceae* and carbapenem resistant *Enterobacteriaceae* (CRE). The latter can be subdivided into: carbapenemase producing (CP)-CRE and non CP-CRE (CDC, 2019).

1.4.1. Intrinsic resistance

ECC members and *K. aerogenes* display natural intrinsic resistance to penicillins, first and second generation cephalosporins. This native resistance is owed to a chromosomally encoded Ambler class C cephalosporinase gene cluster assembled from *ampR-ampD-ampC* genes. Treatment with third generation cephalosporins will induce positive selective pressure on the complex and introduce non-synonymous mutations in the *ampR* and *ampD* genes. This will result in the depression of the *ampC* gene and, consequently, the hyperproduction of the AmpC cephalosporinase (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019; Davin-Regli & Pagès, 2015; Mezzatesta, Gona, & Stefani, 2012). *ampC* depressed ECC and *K. aerogenes* strains

account for more than half of clinically detected cases. However, most of these strains often are found to harbor additional resistance determinants acquired through horizontal gene transfer (HGT) (Davin-Regli & Pagès, 2015).

1.4.2. Acquired resistance

Acquiring antibiotic resistance through HGT is a phenomenon primarily mediated by mobile genetic elements (MGEs). Conjugative plasmids and conjugative resistance transposons are the two key structures that are capable of self-intercellular transfer (Bennett, 2008). Bacterial plasmids are defined as extracellular "mini" chromosomes which carry accessory genes that enhance fitness, adaption and resistance of its host (Rozwandowicz et al., 2018). Their size varies from 30 kbp to 300 kbp and can be carriers of resistance determinants that can halt the action of all currently available antibiotics (Bennett, 2008). Plasmid typing is based on the concept of incompatibility (Inc) groups and replication control. Two plasmids with similar Inc groups cannot coexist in the same bacterial cell. Based on this rationale, a PCR based replicon typing (PBRT) scheme was developed to detect different plasmid families within *Enterobacteriaceae* swiftly. The scheme encompasses the following plasmid families: HI2, HI1, I1-γ, X, L/M, N, FIA, FIB, FIC, FII, FIIK, W, Y, P, A/C, T, K, B/O (Carattoli, 2009; Carattoli et al., 2005).

Transposons are dynamic gene structures capable of engaging in intercellular and intracellular DNA exchange. They're classified as transposable elements, similarly to insertion sequences (IS) and integrons. Transposons manifest diverse genetic conformations and are extremely prone to recombination with each transposition event (Bennett, 2004).

1.4.2.1. ESBLs in ECC and *K. aerogenes*

Genes encoding for ESBLs are commonly plasmid borne. These enzymes confer resistance to all third generation cephalosporins and monobactams (aztreonam) (Davin-Regli & Pagès, 2015). Strains expressing ESBL genes along with overproduction of AmpC cephalosporinase show pan-resistance to all β-lactams in the exception of carbapenems (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019).

The most common and prevalent ESBL encoding gene within the ECC is $bla_{CTX\text{-}M\text{-}15}$, which belongs to the CTX-M-1 family. Other significant genes are $bla_{SHV\text{-}1}$ and bla_{TEM1} (Cantón, González-Alba, & Galán, 2012). As for *K. aerogenes*, it is commonly associated with the TEM variant $bla_{TEM\text{-}24}$ (Davin-Regli & Pagès, 2015).

1.4.2.2. Carbapenem resistance

Carbapenems were thought about as last resort antibiotics in the treatment of ESBL producing *Enterobacteriaceae* isolates. However, exaggerated misuse of these drugs led to positive selective pressure eliciting resistance to carbapenems (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019).

ECC members were amongst the first *Enterobacteriaceae* to harbor carbapenem resistance determinants. Current epidemiological studies show that *Enterobacter* isolates are typically the third or second most common cause of carbapenem resistance in HAIs (Peirano et al., 2018).

Plasmid borne carbapenem resistance genes can be categorized in three groups based on the Ambler classification. Class A serine carbapenemases including the *K. pneumoniae* carbapenemase (bla_{KPC}). Class B metallo-β-

lactamases (MBLs) including the predominant *bla*$_{NDM}$, *bla*$_{IMP}$ and *bla*$_{VIM}$. Finally, class D oxacillinase with the predominant *bla*$_{OXA-48}$ variant (Nordmann, Naas, & Poirel, 2011).

1.4.2.2.1. Carabapenemase plasmid families in ECC

Worldwide, members of the ECC have been found to harbor the *bla*$_{KPC}$ and *bla*$_{NDM}$ variants predominately. The former is endemic to the United States, whereas the latter is mainly identified in the Indian sub-continent and South-Eastern Asia (Nordmann, Dortet, & Poirel, 2012). In the Middle East, *bla*$_{OXA-48}$ is the most prevalent especially with its association with *bla*$_{CTX-M-15}$ on the same plasmids, within the same gene cassette (Mezzatesta, Gona, & Stefani, 2012; Nordmann, Naas, & Poirel, 2011). NDM encoding genes are mostly transported through IncFIA/B, IncFII, IncN and IncX plasmids (Khan, Maryam, & Zarrilli, 2017). As for *bla*$_{OXA-48}$, it is commonly linked to IncL or IncM endemic plasmid in Europe and the Middle East (Evans & Amyes, 2014).

1.4.2.2.2. Porin Mutations in *K. aerogenes*

Carbapenem resistant *K. aerogenes* (CR-KE) clinical cases due to plasmid borne carbapenemases are well documented in US, Europe and Asia (Davin-Regli & Pagès, 2015). However, CR-KE's main mechanism of carbapenem resistance, principally relies on combining the constitutive overexpression of chromosomal *ampC* and altering membrane permeability through porin regulation (Malek et al., 2019).

Porin regulation involves two membrane structures, Omp35 and Omp36. Treatment with carbapenems initially induces mutations, and consequent loss, in Omp35 which will confer partial resistance to

carbapenems (resistance to ertapenem). Further positive selective pressure (continued treatment) will induce non-synonymous mutation in Omp36 bringing about resistance to all carbapenems (Lavigne et al., 2013). It is worthy to note that this phenotypic change is reversible once the course of treatment is stopped (Bornet, Davin-Regli, Bosi, Pages, & Bollet, 2000).

1.5. Incidence in Lebanon and Aim

Despite the importance of *Enterobacter* sp and *K. aerogenes* as global pathogens, to our knowledge, there is a lack in studies that tackle the presence of these bacteria in Lebanon. A significant amount of local studies that tackled ESBL and carbapenemase production only utilizes phenotypic screening and PCR assays (Daoud & Afif, 2011; Matar et al., 2008). Consequently, and due to its polyphyletic nature, *Enterobacter* sp. presence and carriage are heavily underreported locally (Daoud et al., 2017). Recently, few studies characterized *Escherichia coli* and *K. pneumoniae* isolates from Lebanon through the use of WGS and bioinformatics analysis, however information related to *Enterobacter* sp. and *K. aerogenes* is still limited (Arabaghian et al., 2019; Dagher et al., 2018). A recent report about antimicrobial susceptibility data from various Lebanese regions and hospitals only showcased statistics for ESBL and carbapenemase carriage in *Acinetobacter* spp.*, Pseudomonas aeruginosa, Salmonella* spp., *Klebsiella* spp., and *E. coli* (Moghnieh et al., 2019). Therefore, there is a major local need to monitor and limit outbreak scenarios that could be caused by *Enterobacter* sp and *K* aerogenes. We aimed in this study to setup an accurate workflow for the correct identification and molecular characterization of these bacteria in Lebanon

and the region through the use of WGS. To our knowledge this is the first study in Lebanon that will rely on in-depth genome scavenging to investigate the molecular resistome, encompassing acquisition of horizontally transferred elements, that underly the observed phenotypic resistance of these isolates. And also perform comprehensive comparative genome analyses to elucidate the clonality, phylogenetic relationships and trace the molecular evolutionary path of the different isolates that are circulating in Lebanon.

Chapter Two

Materials and Methods

2.1. Bacterial Isolates

A total of nine ECC and six *K. aerogenes* isolates collected from 2015 to 2016 were provided by the American university of Beirut medical center (AUBMC) and were labeled as ENM1-ENM9 and KAM1-KAM6, respectively. Isolates were collected from different infection sites such as blood, abdominal fluid and skin. 86% (n=13) were Males and 14% (n=2) were Females (Table 1), ranging from 2 months to 83 years, with the mean patient's age being 47 ± 29 years old.

Table 2 Primary isolate information including the corresponding patient's sex and age, isolation source and whole-genome NCBI accession numbers

Isolates	**Isolation Source**	**Gender**	**Age**	**Accession Number**
ENM-1	Blood	Male	21	PYGW00000000
ENM-2	skin	Male	82	PYGX00000000
ENM-3	Abdominal fluid	Male	19	PYGY00000000
ENM-4	Abdominal fluid	Male	19	PYGZ00000000
ENM-5	Abdominal fluid	Female	62	PYHA00000000
ENM-6	Skin	Male	2 months	PYHB00000000
ENM-7	Urine	Male	66	PYHC00000000

ENM-8	Skin	Male	1	PYHD00000000
ENM-9	Blood	Male	59	PYHE00000000
KAM-1	Sputum	Male	83	PYHF00000000
KAM-2	Abdominal fluid	Female	28	PYHG00000000
KAM-3	Sputum	Male	71	PYHI00000000
KAM-4	Sputum	Male	83	PYHA00000000
KAM-5	Urine	Male	64	PYHJ00000000
KAM-6	Urine	Male	52	PYHK00000000

Preliminary identification was performed through Matrix-Assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) system (Bruker Daltonik, GmbH, Bremen, Germany) following the manufacturer's instructions.

2.2. Antimicrobial susceptibility testing

Antimicrobial phenotypic testing was performed through the Kirby-Bauer disk diffusion technique on Mueller-Hinton agar. Isolates were tested against 20 clinically relevant antimicrobials. These included piperacillin, ticarcillin, amoxicillin/clavulanic acid, piperacillin/tazobactam, cefuroxime, ceftriaxone. Cefotaxime, ceftazidime, cefepime, imipenem, meropenem, ertapenem, aztreonam, amikacin, gentamicin, tobramycin, tetracycline, tigecyline, ciprofloxacin and trimethoprim/sulfamethoxazole. Minimum-inhibitory concentrations were assessed with the Vitek2 system (BioMerieux). All the data was interpreted according to the clinical & laboratory standard institute (CLSI) recommendations (CLSI, 2018).

2.3. DNA Extraction

DNA was extracted from fresh bacterial cultures grown overnight in tryptic soy broth (TSB) using the Nucleospin® Tissue Kit (Macherey-Nagel, Germany) following the manufacturer's instructions. Genomic DNA was stored at -20° C for further downstream applications.

2.4. PCR amplification and sequencing of the *hsp60* gene

Hsp60 gene amplification was used to assess the Hoffmann cluster of the ECC members. PCR amplification and sequencing were performed with the *hsp60*-F and *hsp60*-R primers as described by Hoffman et al (Hoffmann & Roggenkamp, 2003) (Table 3). *hsp60* sequences were aligned using Clustal-Omega as part of the EMBL-EBI toolset (https://www.ebi.ac.uk/Tools/msa/clustalo/) generating a multiple sequence alignment output file (Madeira et al., 2019). The Hoffman cluster tree was generated using the interactive tree of life (iTOL) tool (Letunic & Bork, 2019).

2.5. Pulse-field Gel electrophoresis (PFGE)

PFGE carried out with the *Xba*I restriction enzyme (ThermoScientific, Waltham, MA, USA) for both ECC and *K. aerogenes* isolates using separate runs. *Salmonella enterica* subsp. *enterica serovar branenderup* was used as a reference and electrophoresis was performed with the CHEF DR-III system (Bio-Rad Laboratories,Bio-Rad Laboratories Inc., Hercules, CA, USA). Run time for both the ECC and *K. aerogenes* was 18h with switch times of 6.76s to 35.38s according to the pulseNet standard protocol. Gels were stained with Ethidium Bromide and de-stained before visualization.

Bacterial fingerprint patterns were analyzed with the BioNumerics software version 7.6.1 (Applied Maths, Belgium) and the resulting pulsotypes were characterized as different based on the threshold of three or more divergent bands (Tenover et al., 1995). Fingerprint clusters were inferred with the BioNumerics software based on the Dice correlation with optimization and tolerance values set for both as 1.5%.

2.6. Multi-Locus sequence typing (MLST)

MLST was performed based on the scheme developed by Miyoshi-Akiyama et al. using seven housekeeping genes (*dnaA, fusA, gyrB, leuS, pyrG, rplB* and *rpoB*) with the forward and reverse amplification/sequencing primers as shown in Table 3 (Miyoshi-Akiyama, Hayakawa, Ohmagari, Shimojima, & Kirikae, 2013). Sequence types were determined through the pubmlst databases for ECC and *K. aerogenes* (https://pubmlst.org/ecloacae/; https://pubmlst.org/kaerogenes/).

Table 3 *hsp60* and MLST amplification and sequencing primers with annealing positions in the target gene.

Amplification primers		
Name	Sequence (5'->3')	Position in the target gene
dnaA-f2	AYAACCCGCTGTTCCTBTATGGCGGCAC	500-527
dnaA-r	KGCCAGCGCCATCGCCATCTGACGCGG	1222-1248
fusA-f2	TCGCGTTCGTTAACAAAATGGACCGTAT	413-440
fusA-r2	TCGCCAGACGGCCCAGAGCCAGACCCAT	1291-1318
gyrB-f	TCGACGAAGCGCTCGCGGGTCACTGTAA	143-170
gyrB-r	GCAGAACCGCCCGCGGAGTCCCCTTCCA	1268-1295
leuS-f2	GATCARCTSCCGGTKATCCTGCCGGAAG	1342-1369
leuS-r	ATAGCCGCAATTGCGGTATTGAAGGTCT	2159-2186
pyrG-f	AYCCBGAYGTBATTGCRCAYMAGGCGAT	56-83
pyrG-r	GCRCGRATYTCVCCCTSHTCGTCCCAGC	563-590
rplB-f	GTAAACCGACATCTCCGGGTCGTCGCCA	17-44
rplB-r	ACCTTTGGTCTGAACGCCCCACGGAGTT	735-762
rpoB-f	AAAAACGTATTCGTAAGGATTTGGTAA	252-280

rpoB-r2	CCAGCAGATCCAGGCTCAGCTCCATGTT	973-1000
Sequencing primers		
gyrB-r3-seq	GCAGAACCGCCCGCGGAGTCCCCTTCC	1269-1295
gyrB-f3-seq	AAAACCGGTACYATGGTGCGTTTCTGG	484-510
fusA-r2-seq	ATCTCTTCACGYTTGTTAGCGTGCATCT	1094-1121
hsp60 primers		
Hsp60-F	GGT AGA AGA AGG CGT GGT TGC	
Hsp60-R	ATG CAT TCG GTG GTG ATC ATC AG	

2.7. PCR based replicon typing (PBRT)

Plasmidic profiles were inferred using the DIATHEVA PBRT kit (Diatheva, Fano, Italy). All isolates were subjected to eight multiplex PCR assays accounting for the 25 plasmid replicons identified within the *Enterobacteriaceae* family (Carattoli et al., 2005). PCR assays were performed according to the manufacturer's instructions and amplicons were visualized on a 2.5% agarose gel stained with ethidium bromide.

2.8. Whole-Genome Sequencing

2.8.1. Library Preparation

Library preparation was performed using the Illumina Nextera XT DNA library preparation kit (Illumina) using 1 µg of extracted genomic DNA (gDNA). The generated gDNA library was quantified with the Qubit 2.0 fluorometer and subsequently sequenced on an Illumina MiSeq using the paired-end protocol with 250 bp read length.

2.8.2. Quality control and assembly

Raw reads were quality controlled with fastQC (Andrews, 2010). Adapter sequences and low-quality bases (Phred score < 30) were trimmed with Trimmomatic v0.38 (Bolger, Lohse, & Usadel, 2014). Contigs were

assembled *de novo* using SPAdes v3.13.1 and applying read error correction. The option *–plasmid* was additionally used to filter out plasmid sequences from raw reads (Bankevich et al., 2012).

2.8.3. Annotation

Annotation for the fasta files was carried out through Prokka and the RASTtk pipeline (Brettin et al., 2015; Seemann, 2014). Annotation through prokka was generated with the following flags --*force -- compliant – locustag* and *–addgenes*. Annotation through RASTtk was initiated with the default pipeline parameters with the added mobile genetic element annotation option.

2.8.4. *In silico* resistance, virulence, MLST profiling and mobile genetic elements (MGEs).

Primary *in silico* species identification was performed based on the k-mer genome parsing tool KmerFinder 3.1 (Hasman et al., 2014).

For the ECC, *in silico* resistance and plasmidic profiles were assessed with the ResFinder 3.0 and PlasmidFinder 2.0 databases (Carattoli et al., 2014; Zankari et al., 2012). *In silico* MLST types were confirmed with the blast option available on pubmlst. Insertion sequences were identified through the ISfinder tool (https://www-is.biotoul.fr/index.php) (Siguier, Perochon, Lestrade, Mahillon, & Chandler, 2006).

Plasmid presence was also investigated with the plasmid constellation network (PLACNETw) tool (Vielva, de Toro, Lanza, & de la Cruz, 2017).

Resistance and virulence in *K. aerogenes* were inferred with Kaptive and Kleborate software (Lam et al., 2018; Wick, Heinz, Holt, & Wyres, 2018). Porin gene sequences were extracted from the RASTtk pipeline

annotation and translated with the Swiss institute of bioinformatics' translation tool (https://web.expasy.org/translate/) (Gasteiger et al., 2003).

2.8.5. PacBio sequencing and novel plasmid sequence

Long read sequencing on top 10 *E. coli* *bla*$_{NDM-1}$ transformant was performed as previously described (Papagiannitsis et al., 2019; Paskova et al., 2018). The plasmid sequence was extracted and compared to its nearest match with the blast option. Annotation of the plasmid sequence was manually curated, and the complete circularized sequence was visualized using the GenomeVx tool (Conant & Wolfe, 2008).

2.8.6. Pan genome and Recombination Analysis

GFF3 annotated files, for ECC and *K. aerogenes* isolates, served as input for the Roary pipeline (Page, 2015). The following command flags were used in pan-genome analysis: *roary -f outdir -e --mafft -v *.gff*.

Recombination hotspots were located through the Gubbins pipeline with recommended default parameters (Croucher, 2014). Pan-genome and recombination results were visualized using the web-based tool Phandango (Hadfield, 2017).

2.8.7. Core genome (cg) SNP phylogenetic analysis

MASH tool was used to identify the best *Enterobacter* sp. reference genome from the RefSeq NCBI databases through genome sketching and distance calculation (Ondov et al., 2016). Confirmatory species identification was performed though wgANI analyses (Rodriguez-R & Konstantinidis, 2016).

Phage identified sequences were removed from all genomes with PHASTER (Arndt et al., 2016). *Enterobacter* sp. CRENT-193

(NZ_CP024812.1) and *Klebsiella aerogenes* KCTC2190 (NC_015663.1) were used as references for read alignment within the Snippy pipeline (Seemann, 2015). Core SNP genome alignments were generated and cleaned with the snippy-clean option. SNPS were then extracted with the snp-sites option and input into gubbins for recombination hotspots removal. The resulting clean core alignment was used to infer the final phylogenetic tree with RAxML (Stamatakis, 2014) with 1000 iterations and bootstrap values support and the GTR-GAMMA model.

Chapter Three

Results

3.1. *In silico* primary identification

16S rDNA sequencing yielded inconsistent results for ECC isolates while correctly identifying *K. aerogenes* isolates (Table 4). However, based on *in silico* k-mer genome screening, 60% (n=9) of the isolates were identified as members of the ECC whereas the remaining 40% (n=6) were correctly identified as *K. aerogenes*. 53% (n=8) of the isolates were identified as *E. hormaechei* species (with varying entailing subspecies), with only one isolate (6%) being identified as *E. cloacae* (Table 4)

Table 4 16S rDNA Sequencing identification results versus the *in silico* KmerFinder tool output.

Isolate	Predicted lineage (kmer finder)	16S rDNA Sequencing
ENM-1	*Enterobacter cloacae subsp cloacae*	*E. cloacae*
ENM-2	*Enterobacter hormaechei subsp oharae*	*Shigella flexneri*
ENM-3	*Enterobacter hormaechei subsp oharae*	n/a
ENM-4	*Enterobacter hormaechei*	*E. xiangfangensis*
ENM-5	*Enterobacter hormaechei subsp. steigerwaltii*	*Shigella flexneri*
ENM-6	*Enterobacter hormaechei subsp. steigerwaltii*	*Shigella flexneri*
ENM-7	*Enterobacter hormaechei*	n/a
ENM-8	*Enterobacter hormaechei subsp. steigerwaltii*	*Shigella flexneri*
ENM-9	*Enterobacter hormaechei*	*E. xiangfangensis*
KAM-1	*Klebsiella aerogenes*	*Klebsiella aerogenes*
KAM-2	*Klebsiella aerogenes*	*Klebsiella aerogenes*
KAM-3	*Klebsiella aerogenes*	*Klebsiella aerogenes*
KAM-4	*Klebsiella aerogenes*	*Klebsiella aerogenes*
KAM-5	*Klebsiella aerogenes*	*Klebsiella aerogenes*

KAM-6	*Klebsiella aerogenes*	*Klebsiella aerogenes*

3.2. Susceptibility profiling

93% (n=14) of the studied isolates showed resistance to amoxicillin + clavulanic acid and 40% (n=6) were resistant to ticarcillin + tazobactam. Resistance to cephalosporins was also detected, with 60% (n=9) of the isolates being resistant to cefuroxime, 53% (n=8) to cefotaxime, 47% (n=7) to both ceftriaxone and ceftazidime. Only 27% (n=4) were resistant to cefepime. In regard to carabapenems, 27% (n=4) were resistant to ertapenem out of which three isolates (20%) were resistant to all the three tested carbapenems (Table 5). On the other hand, 53% (8/15) showed resistance to monobactams (aztreonam) while 20% (n=3) were resistant to other classes of antimicrobial agents such as tetracycline and quinolones. ENM-2 was the only isolate resistant to gentamicin and tobramycin while exhibiting susceptibility to amikacin. Accordingly, ENM-2 was classified as an extremely drug resistant (XDR) bacterium.

Table 5: Kirby-Bauer disk diffusion assay for the 18 antibiotics used in this study. The antimicrobial agents used belonged to seven classes of antimicrobial agents. S: Sensitive; R: Resistant; I: Intermediate Resistance.

Sample	ENM1	KAM1	KAM2	KAM3	KAM4	ENM2	ENM3	ENM4	ENM5	KAM5	ENM6	ENM7	KAM6	ENM8	ENM9
Antibiotic/Class	Penicillins														
Piperacillin 100 µg	R	R	I	S	R	R	S	S	S	R	R	S	R	S	S
Amoxicillin + clavulanic acid	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R
Piperacillin + Tazobactam	S	R	S	S	R	R	S	S	S	R	R	S	R	S	S
	Cephalosporins														
Cefuroxime CXM 30 µg	R	R	R	S	R	R	S	S	S	R	R	S	R	R	S
Ceftriaxone CRO 30 µg	R	R	I	S	R	R	S	S	S	R	R	S	R	S	S
Cefotaxime CTX 30 µg	R	R	R	S	R	R	S	S	I	R	R	S	R	S	S
Ceftazidime CAZ 30 µg	R	R	S	S	R	R	S	S	S	R	R	S	R	S	S
Cefepime CEF 30 µg	R	I	S	S	I	R	S	S	S	R	S	S	R	S	S
	Carbapenems														
Imipenem 10 µg	S	I	S	S	S	R	S	S	S	R	S	S	R	S	S
Meropenem 10 µg	S	S	S	S	S	R	S	S	S	R	S	S	R	S	S
Ertapenem 10 µg	S	I	S	S	I	R	S	S	S	R	R	S	R	I	S
	Monobactams														
Aztreonam 30 µg	R	R	R	S	R	R	S	S	S	R	R	S	R	S	S
	Aminoglycosides														
Amikacin 30 µg	S	I	S	S	I	S	S	S	S	S	S	S	S	S	S
Gentamicin 10 µg	S	S	S	S	S	R	S	S	S	S	S	S	S	S	S
Tobramycin 10 µg	S	S	S	S	S	R	S	S	S	S	S	S	S	S	S
	Tetracycline														
Tetracycline 30 µg	I	S	S	S	S	R	S	S	S	R	S	S	R	S	S
	Quinolones														
Ciprofloxacin 5 µg	R	S	S	S	R	R	S	S	S	S	S	S	S	S	S
	Sulfamides														
Trimethoprim + Sulfa 1.25µg/23.75µg	S	S	S	S	I	R	S	S	S	S	R	S	S	S	S

3.3. *hsp60* identification and clustering

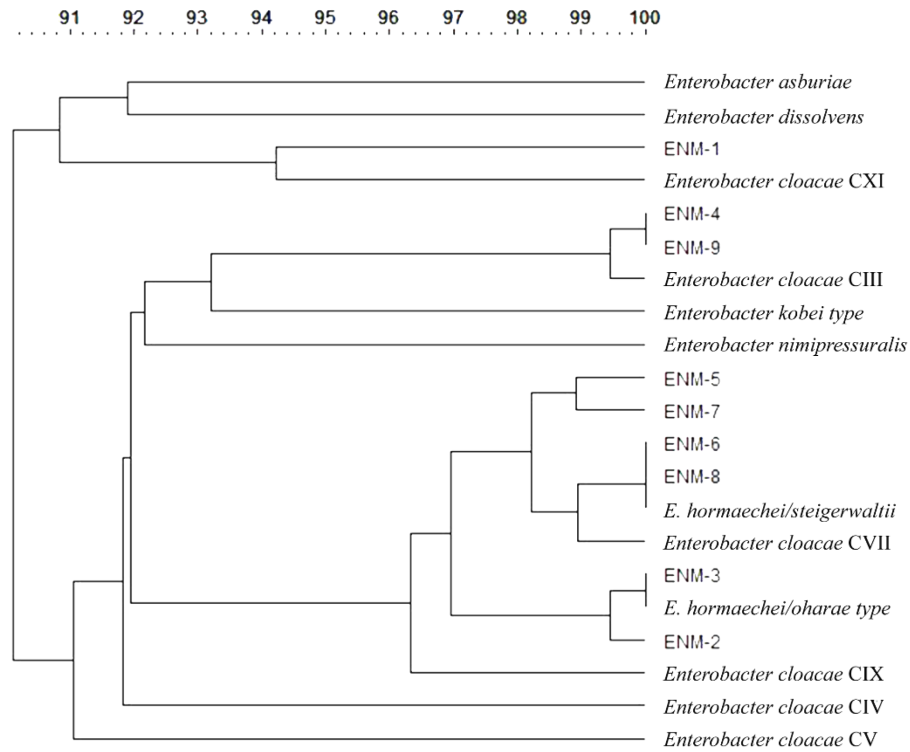

Figure 1: A rooted UPGMA tree based on the *hsp60* sequence analysis of all ECC (n=9) isolates against references matching (n=10) each Hoffmann cluster (Acession # AJ417108 to AJ417143). The tree was generated using clustal Omega multiple sequence alignment tool on EMBL-EBI.

Figure 1 shows that the *hsp60* sequencing results agreed with the k-mer based identification approach. ENM-1, identified as *E. cloacae* subsp. *cloacae*, clustered with 94% similarity to *E. cloacae* cluster XI within the first clade. All remaining isolates ENM-2 to ENM-9 clustered within the second clade. ENM4 and ENM9 shared identical *hsp60* sequences and mapped to cluster III (*E. horamechei* subsp. *hoffmanii*). ENM2 and ENM3 shared the highest similarity at more than 99% and were within cluster VI. ENM5, ENM6, ENM7 and ENM8 shared *hsp60* sequence homogeneity at 98%. The four isolates belonged to the cluster VIII along with *E. hormaechei* subsp. *steigerwaltii*.

3.4. wgANI identification

Figure 2 further confirmed the results observed with k-mer based identification and *hsp60* sequencing. The red regions in the figure denote high relatedness values that range from 97% to 100%; these values represented isolates that belonged to the same species (Rodriguez-R & Konstantinidis, 2016).

ENM1 matched with *E. cloacae* subsp. *cloacae* ATCC13047 with a 99% similarity, accordingly ENM1 belonged to the phylogenetic group G. ENM5, ENM6, ENM7 and ENM8 were closely related (99%) to *Enterobacter hormaechei* subsp. *steigerwaltii*, and as such were classified as members of the phylogenetic group B.

ENM4 showed 100% similarity to ENM9, and the two isolates showed 99% similarity to *Enterobacter hormaechei* subsp. *hoffmannii* and were classified as members of phylogenetic group D. ENM3 was a member of phylogenetic group C along with *Enterobacter hormaechei* subsp. *oharae.*

ENM2, on the other hand, had discrepancies between its *hsp60* identification result and ANI matrix clustering. It showed 99% similarity to *Enterobacter cloacae* subsp. *xiangfangensis* and 97% to E. *hormaechei* subsp. *oharae* reference genomes, and as such it belonged to phylogenetic group A.

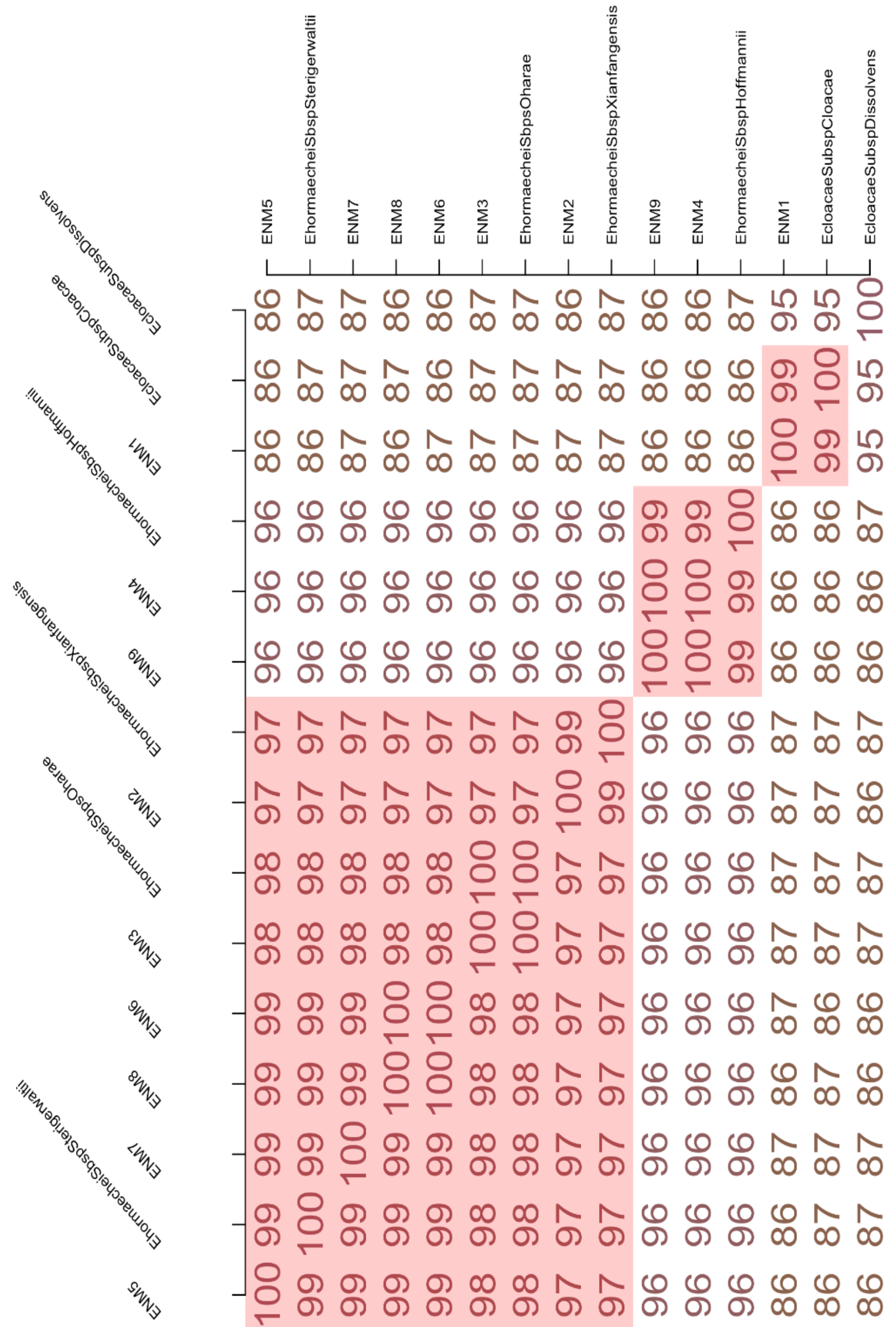

Figure 2 All vs All Average Nucleotide Identity (ANI) percentage relatedness matrix of ECC isolates and 6 reference genomes representing the five subspecies of *E. hormaechei* and *E. cloacae* subsp. *cloacae*. Red blocks represent isolate relatedness higher than 97% and white blocks indicate relatedness values lower than 96%.

3.5. Pan genome and recombination analysis

Pan genome and recombination analysis were performed for ECC and *K. aerogenes* isolates. The primary goal was to determine recombination hotspots

that drive molecular and phylogenetic shifts. Accordingly, our focus will be to highlight and discuss some of the genes that were identified within the accessory genomes, and also to reveal the advantageous discrepancies.

ECC isolates shared a core genome of 2030 coding sequences, with a total count of 10186 CDS. The observed output was unusual but was due to the significant differences between the two studied species under the ECC (Figure 3).

ENM2 was the only isolate positive for *tssK* and *vasK* genes, which are part of the type VI secretion system. The additional accessory genes were identified to be within mobile genetic elements including plasmids and prophages. ENM4 and ENM9 on the other hand, showed less genetic diversity in its accessory genome, carrying genes for fimbrial proteins, insertion sequences (*IS*4), and some type VI secretion system proteins such as *evpB*. ENM6 and ENM8 showed a slight difference only pertaining to prophage sequences, while in ENM3, the accessory genome component only mapped to hypothetical proteins with some phage protein remnants. Finally, ENM1 was the most diverse as compared to the others, with a 65 kbp unique fragment containing numerous paralogs. ENM1 carried genes encoding for curli amyloid fibers such as *csgD* and *csgA*.

K. aerogenes isolates shared a core genome of 4080 protein coding sequences with an additional 1959 shell genes attributed to mobile genetic elements. KAM2 and KAM3's accessory genomes showed hits to phage tail proteins in addition to a high amount of hypothetical proteins. Likewise, KAM1 and KAM4 displayed phage insertion, assembly and tail proteins (Figure 4). KAM5 and KAM6 carried two important virulence associated genes. The first was *virB*, a virulence regulon that regulates the virulence gene cascade in *S.*

flexneri. The second was HigB toxin, which inhibits protein synthesis in host cells (Hurley & Woychik, 2009; Kane & Dorman, 2012).

We plotted the recombination analysis data through the absence/presence matrices with its corresponding ML tree (Figure 5 & 6). Recombination blocks were plotted on the x axis, representing their genomic position relative to the Isolates. Metadata blocks describing total SNP count, total recombination blocks and total SNPs in recombination block were all studied. Figure 5 shows two major recombination blocks for the ECC isolates (blue circles). The first locus is sized at 240 Kbp and includes coding regions of significant importance such as the 16S ribosomal DNA gene, penicillin binding proteins, permeases, transporters and the housekeeping gene *dnaA*. The second block with a size of 100 Kbp that had ABC transporters and efflux pumps effector proteins highlighting its potential role in multi-drug resistance.

The alignment between KAM isolates yielded three minor recombination blocks. The first block is sized at 8 kbp and codes for efflux transport proteins. The second block was the largest being at 110 kbp and encodes the complete type VI secretion system machinery. Finally, the final 104 kbp block encoded for a lipid A biosynthesis protein along outer membrane protein suggesting *K. aerogenes* reliance on recombination to adopt a resistant phenotype.

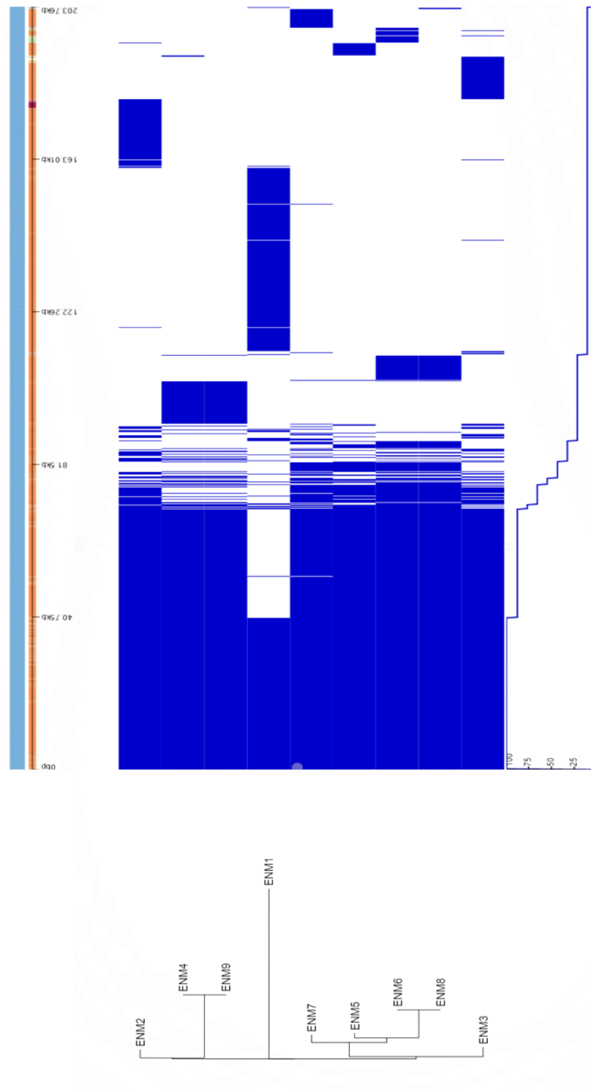

Figure 3 Representation of pan genome analysis denoting gene absence and presence (right) with its associated accessory binary genes tree (left) of ECC isolates (ENM1 → ENM9). Adjacent blue blocks on the left represent core genes shared between isolates whilst scattered block on the right represent cloud and shell accessory genes.

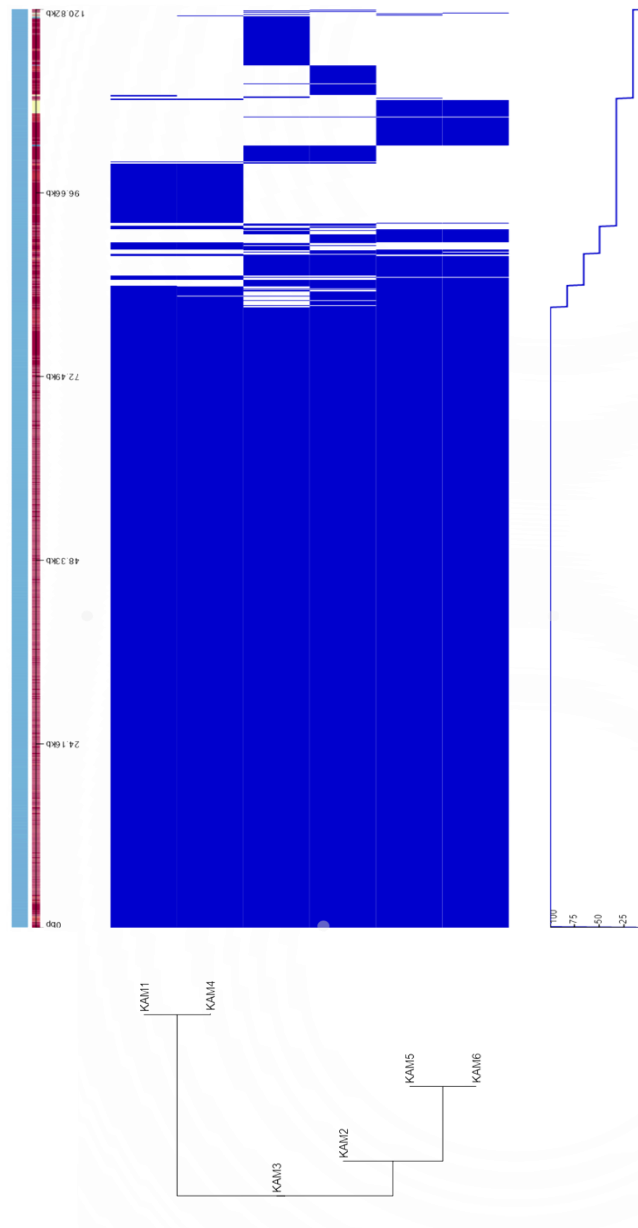

Figure 4 Representation of pan genome analysis denoting gene absence and presence (right) with its associated accessory binary genes tree (left) of *K. aerogenes* isolates (KAM1 → KAM6). Adjacent blue blocks on the left represent core genes shared between isolates whilst scattered block on the right represent cloud and shell accessory genes.

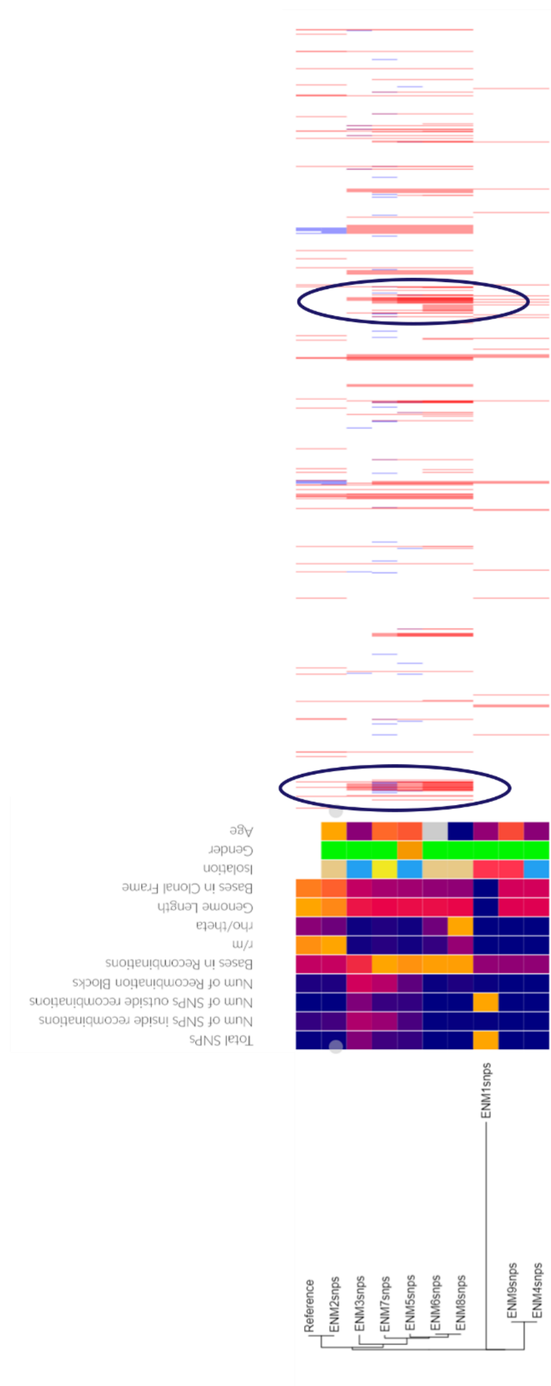

Figure 5 Recombination blocks within the core genome alignment of ECC (ENM1→ENM9) isolates with relative metadata heatmap. Red blocks denote shared recombination sites and blue block signal independent recombination sites. Major recombination hotspots are marked with dark blue circles (n=2) emphasizing presence/absence of shared recombination.

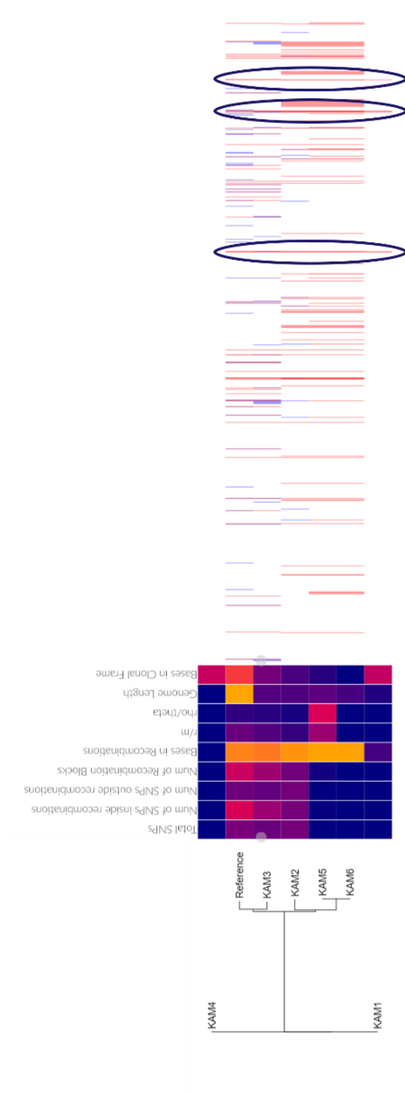

Figure 6 Recombination blocks within the core genome alignment of *K. aerogenes* (KAM1→KAM6) isolates with relative metadata heatmap. Red blocks denote shared recombination sites and blue block signal independent recombination sites. Major recombination hotspots are marked with dark blue circles (n=3) emphasizing presence/absence of shared recombination.

3.6. Core genome (cg) SNP phylogenetic analysis

cgSNP-based phylogenetic characterization reflected close homology with the isolates' identification results. In Figure 7, ENM1 formed a singleton clade with the greatest distance from the remaining isolates highlighting the heterogeneity within the polyphyletic ECC. ENM9 and ENM4 formed a

second clade while the remaining isolates 46% (n=7) clustered in the largest clade showcasing the small evolutionary distance between core genomes.

Figure 8 shows a similar topology as that seen in Figure 7. However, despite clustering in two different clades KAM1 and KAM4 show small-scale distance variations. Clade 3 revealed KAM5 and KAM6 as being almost identical with minimal SNP variation, while KAM3 was the closest to *K. aerogenes* KCTC2190.

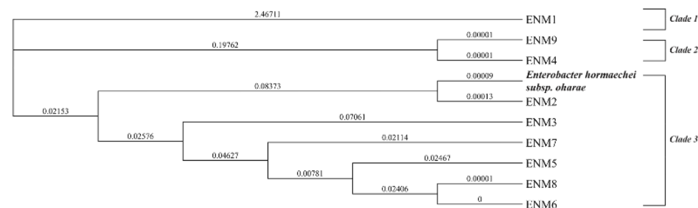

Figure 7 RAxML generated clean core maximum-likelihood tree with 1000 bootstrap support elucidating the phylogenetic relationship between ECC isolates (ENM1→ENM9) with one *E. hormaechei* subsp. *oharae* reference genome (accession # LRJW01000000). MGEs and recombination events were removed prior to generating the tree.

Figure 8 RAxML generated clean core maximum-likelihood tree with 1000 bootstrap support elucidating the phylogenetic relationship between KAM isolates (KAM1→KAM6) and the *K. aerogenes* NCBI reference genome (accession # NC_015663.1) without the interference of recombination and MGEs.

3.7. Virulence factors

In silico k-type analysis through captive identified two different K-types. KAM1 and KAM4 (33%) were identified as K442, while KAM2, KAM3, KAM5 and KAM6 (66%) were identified as K42. It is worthy to note that there were missing genes in the *cps* locus (Figure 9). Seven missing genes for KAM3 and five for all the other isolates.

Using Kleborate, only KAM2 and KAM3 were found to carry genes encoding for the siderophores yersiniabactin and colibactin, which are important virulence determinants. KAM2 harbored the *ybt*17 on the integrative conjugative element ICE*kp10* isoform. The locus' ST was 289 and had two allelic variants. KAM3 also had the *ybt17* with ST289 but didn't show any locus variation. Moreover, KAM2 and KAM3 harbored the *clb3* colibactin locus, with KAM2 showing no variation while KAM3 having two cbST17 variants. Finally, all the KAM isolates carried a chromosomal *ampC* gene.

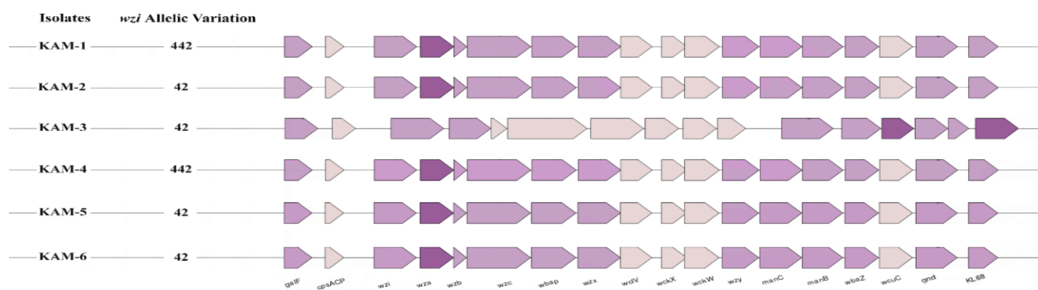

Figure 9 *wzi* loci alignment for all *K. aerogenes* isolates (n=6) with the locus allelic type shown. Purple color indicates genes that are present while beige indicates gene absence and deletions.

Table 6 Kleborate analysis output. Virulence and resistance systems detected along the yersiniabactin (*ybt)* and colibactin (*clb*) loci sequence type. A value of 0 indicates no hits, 1 = one identified system, 2 = two identified systems.

strain	virulence	Resistance	Yersiniabactin	YbST	Colibactin	CbST	*bla*
KAM1	0	1	-	-	-	-	*AmpC*
KAM2	2	1	*ybt* 17; ICEKp10	289-2LV	*clb*3	13	*AmpC*
KAM3	2	1	*ybt* 17; ICEKp10	289-1LV	*clb*3	17-2LV	*AmpC*
KAM4	0	1	-	-	-	-	*AmpC*
KAM5	0	1	-	-	-	-	*AmpC*
KAM6	0	1	-	-	-	-	*AmpC*

3.8. PFGE

Pulse-Field gel electrophoresis was performed to investigate the clonal relationship between the isolates (Figure 10A). Members of the ECC did not show any significant clonal relatedness.Based on the banding patterns and clustering, seven

distinct pulsotypes were detected. ENM4 and ENM9 had identical pulsotypes, whereas ENM6 and ENM8 showed 95% similarity.

In the case of the *K. aerogenes* isolates, KAM-1 and KAM-4 shared a high degree of clonal relatedness and clustered separately from KAM-2 and KAM-3 (Figure 10B). KAM-5 and KAM-6 remained untypable using both *Xba*I and *Sme*I as a secondary enzyme.

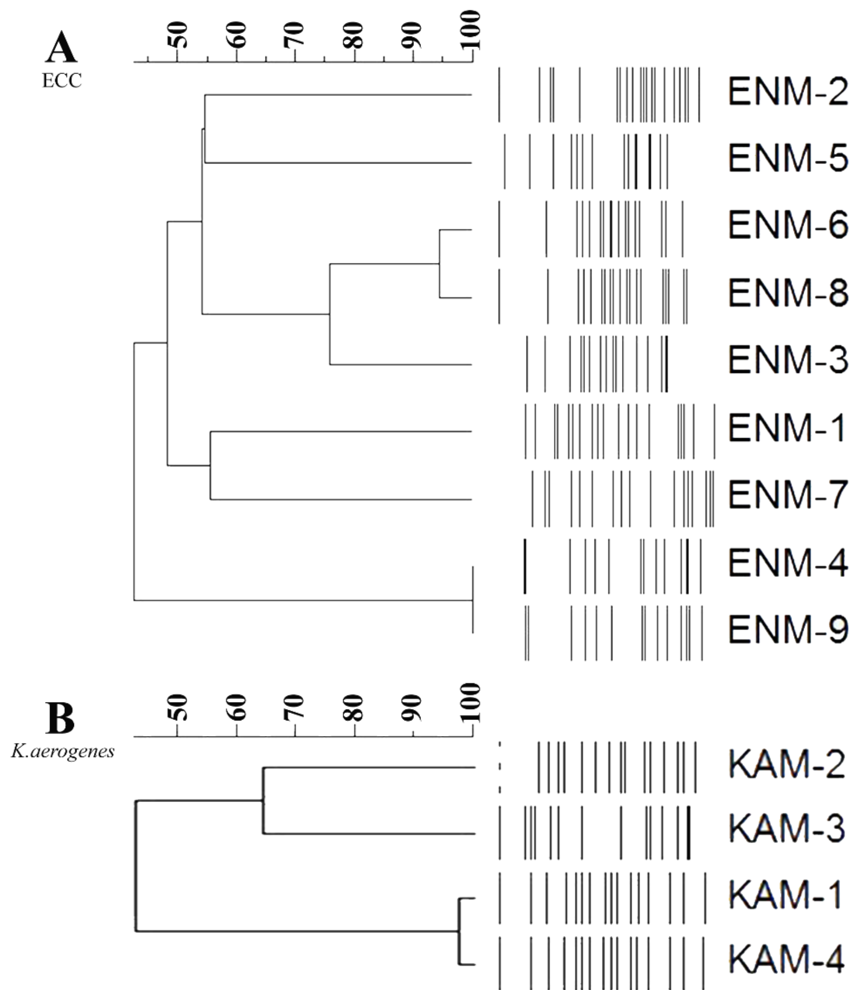

Figure 10: Pulse-field gel electrophoresis elucidating the clonality of isolates based on their banding pattern obtained through *Xba*I restriction digestion. Dendograms were generated using BioNumerics v7.6.1. A) PFGE-based dendogram for ECC isolates (n=9) showcasing 7 different pulsotypes. B) PFGE-based dendogram for *K.aerogenes* isolates (n=6) demonstrating 3 different pulsotypes.

3.9. Sequence type characterization and molecular resistance profiling

In silico (MLST) analysis was performed using seven housekeeping genes (*dnaA, fusA, gyrB, leuS, pyrG, rplB and rpoB*). In addition, MLST PCR assay was used to complement missing alleles (Miyoshi-Akiyama, 2013).

ENM-6 and ENM-8 were assigned to ST113 and ENM-2 was assigned to ST114. ENM-3 and ENM-7 were assigned to STs 198 and 50, respectively. ENM1, ENM-4, ENM-5 and ENM-9's allelic profiles differed by one for four of the alleles from the closest ST profiles present in the database (table 7).

KAM-2 and KAM-3 were assigned STs 93 and 4, respectively. We couldn't determine the ST's of KAM-1, KAM-4, KAM-5 and KAM-6 (table 7).

On the other hand, all the studied isolates (n=15) harbored an *ampC* family β-lactamase gene ($bla_{\text{CMH-3}}$, $bla_{\text{ACT-7/14/15/15}}$ and $bla_{DHA\text{-}1}$) (figure 13). *Enterobacter* sp. has intrinsic resistance to penicillins and early generation cephalosporins due to chromosomal copies of *ampC* genes which confirmed the results from *in silico* resistance gene screening. $bla_{\text{CTX-M-15}}$ was associated with ENM1 and ENM2. Using PBRT, plasmidfinder and PLACNETw we couldn't detect any plasmids in ENM1. This suggested that $bla_{\text{CTX-M-15}}$ was chromosomally encoded. In fact, blasting the sequence on NCBI and ISfinder, $bla_{\text{CTX-M-15}}$ was associated with IS*Ecp1* and a Tn2 like structures in both ENM1 and ENM2 (Figure 11).

Additionally, ENM-2 was the only isolate positive for $bla_{\text{NDM-1}}$, $bla_{\text{OXA-1}}$, $bla_{\text{CTX-M-15}}$, $bla_{\text{SHV-12}}$ and $bla_{\text{ACT-16}}$ (figure 13). Also, aminoglycoside determinants resistance *aac(6')Ib-cr, aadA1, aac(3)-Iia* were detected in 27% (n=4) of the

isolates, and 20% (n=3) additionally harbored the fluoroquinolone resistance determinant *qnrS1*.

KAM5 and KAM6 both showed phenotypic resistance to all tested carbapenems. *In silico* analysis showed no carbapenemase encoding genes despite being positive for an IncF plasmid. Comparative analysis of the outer membrane protein (omp36) showed a truncation event only in KAM5 and KAM6. DNA sequences for both loci were extracted from the RASTtk and translated. Non-synonymous mutation was detected through a single nucleotide polymorphism (C→T) in a glutamine residue (CAG) introducing an amber (TAG) stop codon leading to a non-sense mutation in the omp36 locus (Figure 12).

Table 7 MLST allelic variations for all housekeeping genes loci for all isolates (N=15). Allelic combinations with no match were termed as novel and submitted to the database curators.

Isolate	**Loci alleles**							**ST**
	dnaA	*fusA*	*gyrB*	*leuS*	*pyrG*	*rplB*	*rpoB*	
ENM-1	192	44	65	214	36	22	90	Novel
ENM-2	53	35	20	44	45	4	6	114
ENM-3	68	8	75	63	65	34	6	198
ENM-4	59	9	12	172	67	6	6	Novel
ENM-5	9	4	15	13	43	30	24	Novel
ENM-6	4	22	68	69	37	4	24	113
ENM-7	4	4	4	6	37	4	25	50
ENM-8	4	22	68	69	37	4	24	113
ENM-9	59	9	12	172	67	6	6	Novel
KAM-1					Novel			
KAM-2	3	3	2	18	3	1	10	93
KAM-3	3	3	1	4	3	1	2	4
KAM-4					Novel			
KAM-5					Novel			
KAM-6					Novel			

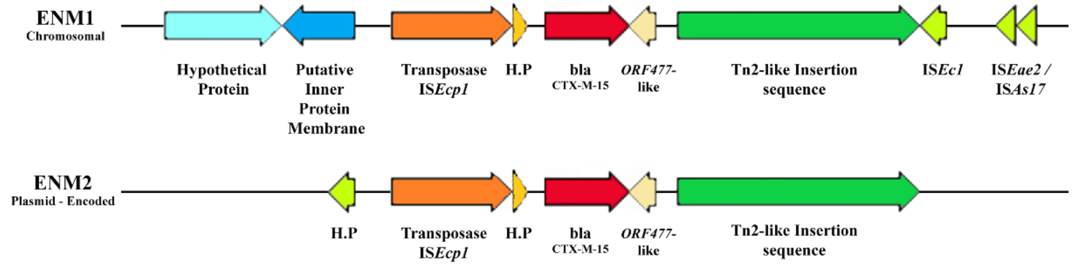

Figure 11 *bla*$_{\text{CTX-m-15}}$ genetic environment comparison between ENM1 and ENM2. Both chromosomal and plasmid encoded copies of *bla*$_{\text{CTX-m-15}}$ are carried by an *IS*Ecp1-*bla*$_{CTX\text{-}M\text{-}15}$-*orf*477*Δ* transposition unit.

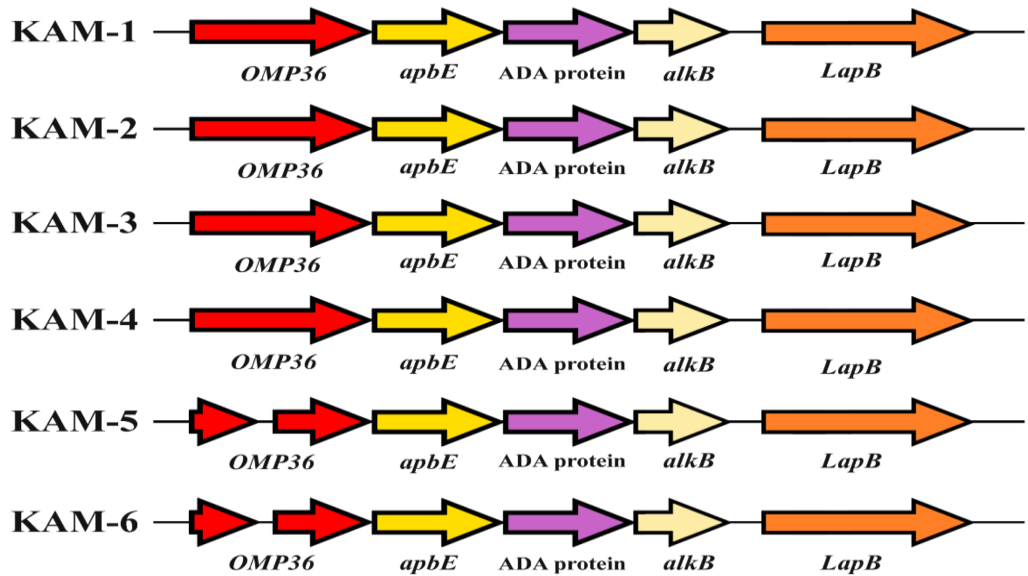

Figure 12 Alignment of outer membrane protein (omp36) encoding genes and their genetic environment. KAM1 → KAM4 show a functional *omp36* gene sequence (red arrow) while KAM5 and KAM6 carry a truncated form of the gene (truncated red arrows).

3.10. Plasmid typing

46% (n=7) of isolates were positive for the IncF (IncFII, IncFIA, IncFIB) incompatibility group. ENM-2 was also positive for the IncH and IncN2 groups. ENM-3 and ENM-7 were positive for the IncR and IncI2 incompatibility replicons, respectively. These results were confirmed by plasmid-based replicon typing (PBRT), and it additionally revealed that 33% (n=4) of the isolates were positive for IncX3 (Table 8).

Table 8 Plasmid based replicon typing and *in silico* based plasmid incompatibility profiling of study isolates (n=15). IncF isoforms were identified in 46% (n=7) of isolates.

Isolate	Plasmid Inc Types
ENM-1	No hit
ENM-2	IncFII, IncH1A, IncH1B, IncN2
ENM-3	IncFII(K), IncR
ENM-4	No hit
ENM-5	IncX3
ENM-6	IncFIB, IncFII
ENM-7	IncI2
ENM-8	IncFIB, IncFII, IncX3
ENM-9	No hit
KAM-1	IncX3
KAM-2	IncFIIK, IncX3
KAM-3	No hit
KAM-4	No hit
KAM-5	IncFIA, IncFII
KAM-6	IncFIA, IncFII

Isolates	Subspecies	*hsp60* Identification	Sequence Type (ST)	Plasmidic Profile
ENM-1	*Enterobacter cloacae subsp cloacae*	Cluster XI	New	x
ENM-2	*Enterobacter hormaechei subsp xiangfangensis*	Cluster VI	114	IncFII, IncH1A, IncH1B, IncN2
ENM-3	*Enterobacter hormaechei subsp oharae*	Cluster VI	198	IncFII(K), IncR
ENM-4	*Enterobacter hormaechei subsp hoffmanii*	Cluster III	New	x
ENM-5	*Enterobacter hormaechei subsp. steigerwaltii*	Cluster VIII	New	IncX3
ENM-6	*Enterobacter hormaechei subsp. steigerwaltii*	Cluster VIII	113	IncFIB, IncFII
ENM-7	*Enterobacter hormaechei subsp. steigerwaltii*	Cluster VIII	50	IncI2
ENM-8	*Enterobacter hormaechei subsp. steigerwaltii*	Cluster VIII	113	IncFIB , IncFII, IncX3
ENM-9	*Enterobacter hormaechei subsp hoffmanii*	Cluster III	New	x
KAM-1	*Klebsiella aerogenes*	x	New	IncX3
KAM-2	*Klebsiella aerogenes*	x	93	IncFIIK, IncX3
KAM-3	*Klebsiella aerogenes*	x	4	x
KAM-4	*Klebsiella aerogenes*	x	New	x
KAM-5	*Klebsiella aerogenes*	x	New	IncFIA IncFII
KAM-6	*Klebsiella aerogenes*	x	New	IncFIA IncFII

Antibacterial Determinants	
beta-lactamases	*CTX-M-15*, *SHV-12*, *ACT*, *CMH*, *OXA-1*, *NDM-1*, *ampC*
Aminoglycosides	*aac(3)-I*, *aadA1*, *aac(6')Ib-cr*
Fluoroquinolone	*QnrS1*, *QnrB*
Tetracycline	*tet(A)*
Trimethoprim	*dfrA*

Figure 13 Molecular characterization and attributes of all study isolates (n=15). Correct nomenclature and relative subspecies. 14 sequence types (including novel STs) and their molecular characteristics such as plasmid carriage and different categories of resistance determinants.

3.11. Novel *bla*$_{NDM-1}$ plasmid

Plasmid sequences were extracted during assembly of ENM2 raw reads using the --plasmid option in SPAdes. Long read sequencing with the PacBio SMRT system yielded three contigs. The first, a 5Mbp contig representing genomic DNA, the second was a 213 kbp contig representing the IncHIA and IncHIB multi replicon plasmid. Finally, the third contig showed a 104 kbp fragment. Alignment of the 104 kbp fragment with the extracted plasmid

sequences showed a 100% coverage. Overlapping regions were manually removed, automatically annotated, manually curated and circularized (Figure14).

pNDM1_ENM2 sequence was mapped to the NCBI database with BLASTn. Its nearest match was a 216378 bp plasmid designated as *E. coli* strain AR_0128 plasmid (CP021720) (Unpublished data). Query coverage was 100% with 0 E value and 99.98% identity coverage. pNDM1_ENM2 is an IncFII plasmid carrying only a $bla_{\text{NDM-1}}$ gene within a Tn125 isoform (figure 15)

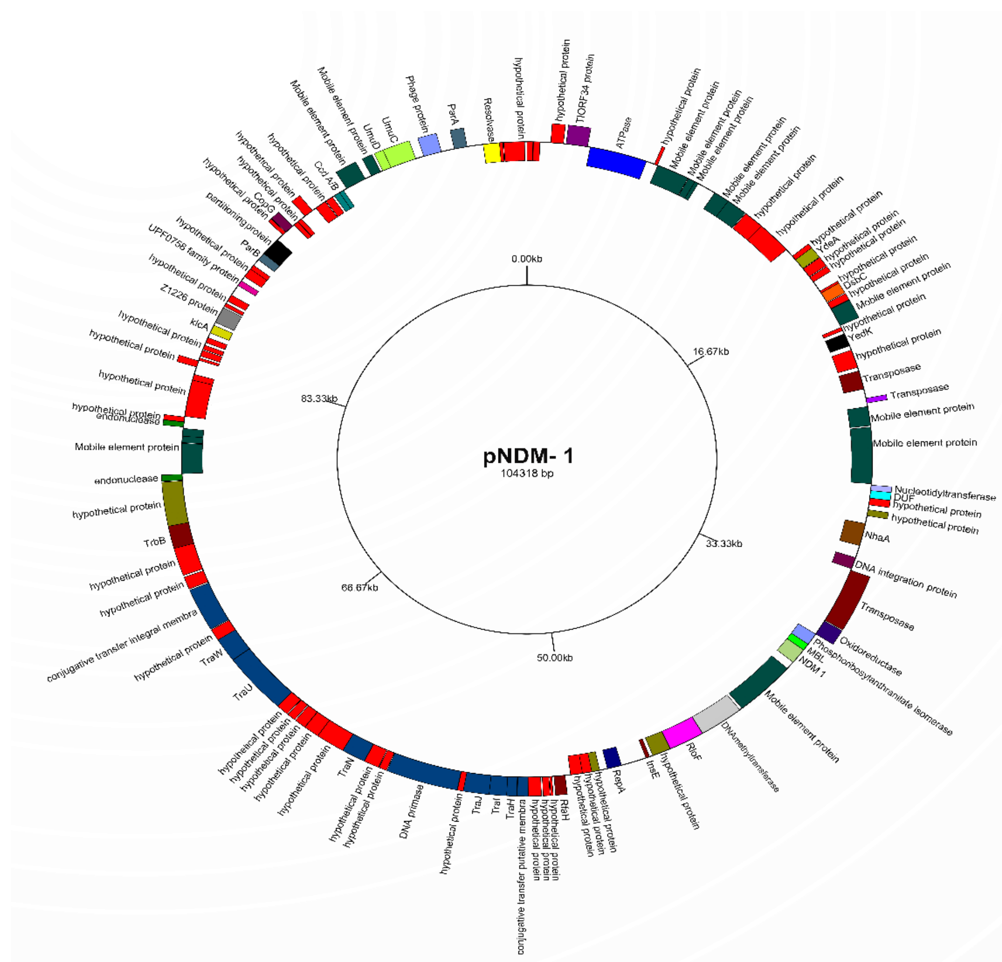

Figure 14 Overview of plasmid pNDM1_ENM2. Complete sequence of the 104318 bp plasmid, fully annotated. *tra* genes are colored in blue, *parA* and *parB* genes colored in teal and IncFII *repA* gene colored in navy blue. $bla_{\text{NDM-1}}$ is flanked a bleomycin resistance gene (*mbl)* downstream and an *IS*Aba125 upstream.

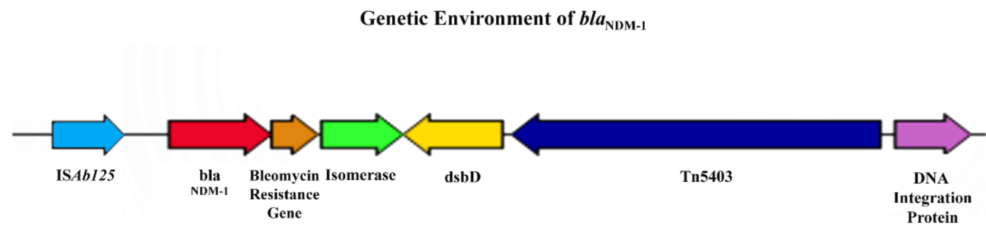

Figure 15 Genetic environment of *bla*$_{NDM-1}$ allocated on the pDNM1_ENM2 plasmid. IS*Ab125* (cyan) is found upstream of *bla*$_{NDM-1}$ (red), while a bleomycin resistance gene (orange) and a Tn*5403* (navy blue) are found downstream.

Chapter Four

Discussion

Members of the *Enterobacter cloacae* complex and *Klebsiella aerogenes* are opportunistic pathogens. As our treatment options are depleting in an alarming manner both bacteria keep amassing and devising new mechanisms of resistance coupled with their unlimited ability to undergo perpetual phylogenomic renewal.

In this study, ECC isolates demonstrated the heterogeneity of the complex as 8/15 (53%) of study isolates were identified as native to the *E. hormaechei* species (Hoffman clusters III, VI and VIII). Surprisingly, only one isolate was recognized as an *E. cloacae* (Hoffman cluster XI). As a result of erroneous and outdated identification practices, *E. hormaechei* carriage tends to be underestimated in molecular epidemiological studies related to *Enterobacter* spp. (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019). Local studies investigating *Enterobacteriaceae* spread in the region fail to correctly identify *Enterobacter* isolates (Herte, 2012; Moubareck et al., 2005). However, in studies specifically aimed at characterizing the predominant clinical species, *E. hormaechei* was found to be the most clinically relevant even in colonizing uncommon hospital apparel and prosthetic implants (Morand et al., 2009). With advances in sequencing technologies, population level analysis becomes more feasible and stringent suggesting a much-needed switch to WGS identification

methods especially when tackling highly heterogenic phyla such as the ECC (Chavda et al., 2016; Kluytmans-van den Bergh et al., 2016).

Core genomic recombination is one of the key driving forces that aid in bacterial differentiation and adaptation, other than inter and intraspecies MGE exchange (Marttinen, Croucher, Gutmann, Corander, & Hanage, 2015). ECC isolates boast a significant recombination ratio at 1.04 where such forces are responsible for the complex' heterogeneity (Paauw et al., 2008). Figure 5 shows two major recombination blocks shared by the ECC isolates within this study. Our results highlighted potential shared recombination events between all ECC isolates within the 16S rRNA gene. This is in accordance with the consensus that 16S sequencing proves inconsistent in the phylogenetic typing of ECC members (Chavda et al., 2016). In fact, evidence suggests that within the ECC the 16S rRNA genes were modeled overtime through recurrent recombination events (Sato & Miyazaki, 2017). Moreover, potential recombination events span penicillin binding proteins and membrane permeases suggesting a role in conferring non-enzymatic drug resistance.

In parallel, *K. aerogenes* favors recombination events skewed towards improving its invasiveness, resistance and survivability. Recombination blocks were detected in this study in loci coding for efflux transport proteins, outer membrane porins and lipid A biosynthesis (Figure 6). We detected a major recombination hotspot in type VI secretion system encoding loci indicating the potential in acquiring putative genes that aim at enhancing the effectiveness of the type VI secretion system in host

infection scenarios (Kirchberger, Unterweger, Provenzano, Pukatzki, & Boucher, 2017).

Analyzing the presence/absence pan genome block of ECC revealed great disparity between ENM1 and the rest (Figure 3). Within the large putative accessory block in ENM1, type IV secretion encoding loci were detected. Despite the importance of the type IV secretion system, there is scarce data about its presence and effect in *E. cloacae* isolates (Mezzatesta, Gona, & Stefani, 2012). More efforts have been directed at elucidating the presence of the type III secretion system within member of the ECC and its cytotoxic effect (Krzymińska, Mokracka, Koczura, & Kaznowski, 2009). The curli gene was also detected within the accessory block of ENM1. This gene encodes for curli amyloid fibers which function in facilitating cell to cell adhesion and is thought to have a role in biofilm formation (Mezzatesta, Gona, & Stefani, 2012). ENM2, ENM4 and ENM9 were all found to harbor a type VI secretion system assembly loci within the accessory genome. Although, these isolates are members of the same overall phylogenetic cluster they display large variation in their virulence determinants and ability to acquire such factors. This hints at the potential in finding selective virulence marker for the different opportunistic pathogens within the ECC, which could aid in treatment and identification in an outbreak scenario (Paauw et al., 2009).

The yersiniabactin and colibactin siderophores were detected in 2/15 (13%) of study isolates (Table 6). KAM2 and KAM3 were found to carry a chancy isoform ICE*Kp10* that coharbors the yersinbactin (*ybt)* and colibactin (*clb*) siderophores. ICE*Kp10* is generally associated with *ybt17* locus and the *clb3*

locus in the cargo region (Lam et al., 2018). A study by Arabaghian et al., presented a diverse set of ICE*Kp* isoforms from multiple carbapenem resistant *K. pneumoniae* isolates recovered from clinical settings in Lebanon. However, none coharbored the *ybt* and *clb* suggesting positive selection for the isoform in *K. aerogenes* which enhance iron sequestration in hostile host environments and promotes an aggressive invasive phenotype (hypervirulence) in the absence of multi drug resistance (Arabaghian et al., 2019; Lam et al., 2018).

Two *wzi* (K-types) were identified within the *K. aerogenes* isolates. KL42 4/15 (26%) and KL442 2/15 (13%). Both K-types were not identified previously in the region (Arabaghian et al., 2019). Additionally, both K loci present novel deleterious structures KL42D1 and KL442D1(Figure 9). Such deletions from the K locus genes can be caused by recombination events within the central regions of the locus or CDS protein disruption by IS integration (Wyres et al., 2016).

ECC and *K. aerogenes* demonstrated high clonal diversity in this study. Multi-locus sequence types identified for the ECC included but are not limited to: ST113 2/15 (13%), ST50 1/15 (6%), ST198 1/15 (6%) and ST114 1/15 (6%) (Table7). ST114 has been identified as an epidemic clone and clonal complex CC114. This ST is generally associated with carriage of $bla_{\text{CTX-M-15}}$ and $bla_{\text{VIM-1}}$ and recently an NDM-1 *Enterobacter hormaechei* subsp. *xiangfangensis* epidemic clone has been detected from 37 countries (Izdebski et al., 2015; Peirano et al., 2018). In contrast ST198 was only associated with ESBL carriage, mainly $bla_{\text{CTX-M-9}}$ variant and was not classified as CREC in any outbreak report (Izdebski et al., 2015). Four

novel STs were identified in this study with their allelic variation detailed in table 7. This revealed a gap in the data related to the clonality of ECC members in the region.

Despite *K. aerogenes*'s role as a successful opportunistic pathogen and its ability to integrate MGEs and acquire extreme drug resistance, data on its clonal diversity and geographical dissemination is currently scarce (Malek et al., 2019). KAM2 and KAM3 sequence types were identified as ST93and ST4, respectively (table 7). ST93 and ST4, currently represent the dominant international clones that are involved in outbreak of HAI scenarios (Male et al., 2019). KAM1, KAM4, KAM5 and KAM6 all represent novel STs confirming the lack in data about species diversity, global distribution and epidemic clone identification.

Under the selective positive pressure of treatment *K. aerogenes* has been shown to introduce such mutations into outer membrane porins. A combination of such mutations with overexpression of chromosomal *ampC* leads to pan-carbapenem resistance (Davin-Regli & Pagès, 2015). KAM5 and KAM6 displayed carbapenemase independent resistant phenotype despite harboring incFII and incFIA plasmids (table 8). *K. aerogenes*' main carbapenem resistance mechanism involves porin alterations and overproduction of chromosomal *ampC* gene (table 6). As shown in figure 11, a non-sense mutation was induced in the *omp36* gene, thus reducing potential uptake of harmful molecules. The most alarming issue is the fact that this is a reversible phenotype which is only governed by treatment with carbapenems. Therefore, *K. aerogenes* has a highly volatile adaptability that

can be summoned at will and under the correct selective forces (Davin-Regli & Pagès, 2015; Malek et al., 2019; Sanders & Sanders, 1997).

In contrast, ECC members are becoming one of the key drivers of enzyme dependent carbapenem resistance within the *Enterobacteriaceae* family (Peirano et al., 2018). CREC is imposing a massive burden on healthcare system around the globe, especially in the USA where it has become the second most widespread carbapenem resistant *Enterobacteriaceae* (Annavajhala, Gomez-Simmonds, & Uhlemann, 2019; Chavda et al., 2016). ENM1 showed no plasmid carriage, yet its genome analysis showed that it harbors $bla_{\text{CTX-M-15}}$ ESBL gene. This hints at the potential integration of the CTX-M-15 transposition unit in the chromosome within a recombination hotspot. Blast analysis of the IS*Ecp1* and a Tn2 like transposition unit (Guenther et al., 2017) showed similarity with the $bla_{\text{CTX-M-15}}$ allocated on the incHIA/incHIB plasmid in ENM2 (figure 10). The observed resistance genotype in ENM1 can only be a testament to the high genome plasticity of *E. cloacae* and can serve as a countdown for the development of pan-resistant super bugs through vertical gene transfer (Kluytmans-van den Bergh et al., 2016).

pNDM1_ENM2, a novel incFII plasmid was identified as a 104311 bp $bla_{\text{NDM-1}}$ bearing plasmid from *Enterobacter hormaechei* subsp. *xiangfangensis* ST114 (Figure 13). $bla_{\text{NDM-1}}$'s genetic environment is almost identical to that of a Tn*125* where the gene is flanked by two IS*Aba125* upstream and downstream. However, the $bla_{\text{NDM-1}}$-ble_{MBL} genes were flanked by a downstream IS*Aba125* and an upstream Tn*5403* (Rinkel, Hubert, Roux, & Lett, 1994) which is from the family of Tn*3*-like

transposons (Figure 14) (Arabaghian et al., 2019). NDM-1's transposon genetic shared 98% homology to an NDM locus allocated on an incFII plasmid carried by an *E. hormaechei* subsp. *oharae* in Brazil (Carvalho-Assef et al., 2014) suggesting global epidemic spread of *bla*$_{NDM-1}$-incFII associated plasmids.

Recently antimicrobial susceptibility data from 13 hospitals in Lebanon were published temporarily substituting for the lack official national AMR data (Moghnieh et al., 2019). Surprisingly, no data was present about the carriage and dissemination of AMR in ECC members and *K. aerogenes* (Moghnieh et al., 2019). Also, local studies have tackled the topic of AMR was tackled only by performing phenotypic based screening and PCR assays to assess ESBL and carbapenem resistance gene carriage in Enterobacteriaceae (Daoud et al., 2017). To our knowledge, this is the first study that performs epidemiological and comprehensive WGS based end to end bioinformatics data analysis on members of the *E. cloacae* complex and *K. aerogenes* isolates, in Lebanon and the region.

References

Andrews. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved June 10, 2019, from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Annavajhala, M. K., Gomez-Simmonds, A., & Uhlemann, A.-C. (2019). Multidrug-Resistant Enterobacter cloacae Complex Emerging as a Global, Diversifying Threat. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.00044

Arabaghian, H., Salloum, T., Alousi, S., Panossian, B., Araj, G. F., & Tokajian, S. (2019). Molecular Characterization of Carbapenem Resistant Klebsiella pneumoniae and Klebsiella quasipneumoniae Isolated from Lebanon. *Scientific Reports*, *9*(1), 531. https://doi.org/10.1038/s41598-018-36554-2

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*(W1), W16–W21. https://doi.org/10.1093/nar/gkw387

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bennett, P M. (2008). Plasmid encoded antibiotic resistance: Acquisition and transfer of antibiotic resistance genes in bacteria. *British Journal of Pharmacology*, *153*(Suppl 1), S347–S357. https://doi.org/10.1038/sj.bjp.0707607

Bennett, Peter M. (2004). Genome plasticity: Insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods in Molecular Biology (Clifton, N.J.)*, *266*, 71–113. https://doi.org/10.1385/1-59259-763-7:071

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bornet, C., Davin-Regli, A., Bosi, C., Pages, J. M., & Bollet, C. (2000). Imipenem resistance of enterobacter aerogenes mediated by outer membrane permeability. *Journal of Clinical Microbiology*, *38*(3), 1048–1052.

Brady, C., Cleenwerck, I., Venter, S., Coutinho, T., & De Vos, P. (2013). Taxonomic evaluation of the genus Enterobacter based on multilocus sequence analysis (MLSA): Proposal to reclassify E. nimipressuralis and E. amnigenus into Lelliottia gen. nov. as Lelliottia nimipressuralis comb. nov. and Lelliottia amnigena comb. nov., respectively, E. gergoviae and E. pyrinus into Pluralibacter gen. nov. as Pluralibacter gergoviae comb. nov. and Pluralibacter pyrinus comb. nov., respectively, E. cowanii, E. radicincitans, E. oryzae and E. arachidis into Kosakonia gen. nov. as Kosakonia cowanii comb. nov., Kosakonia radicincitans comb. nov., Kosakonia oryzae comb. nov. and Kosakonia arachidis comb. nov., respectively, and E. turicensis, E. helveticus and E. pulveris into Cronobacter as Cronobacter zurichensis nom. nov., Cronobacter helveticus comb. nov. and Cronobacter pulveris comb. nov., respectively, and emended description of the genera Enterobacter and Cronobacter. *Systematic and Applied Microbiology*, *36*(5), 309–319. https://doi.org/10.1016/j.syapm.2013.03.005

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., … Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, *5*, 8365. https://doi.org/10.1038/srep08365

Cabral, A. B., Maciel, M. A. V., Barros, J. F., Antunes, M. M., Barbosa de Castro, C. M. M., & Lopes, A. C. S. (2017). Clonal spread and accumulation of β-lactam resistance determinants in Enterobacter aerogenes and Enterobacter cloacae complex isolates from infection and colonization in patients at a public hospital in Recife, Pernambuco, Brazil. *Journal of Medical Microbiology*, *66*(1), 70–77. https://doi.org/10.1099/jmm.0.000398

Campos, L., Lobianco, L. F., Seki, L., Santos, R. M. R., & D Asensi, M. (2007). Outbreak of Enterobacter hormaechei septicaemia in newborns caused by contaminated parenteral nutrition in Brazil. *The Journal of Hospital Infection*, *66*, 95–97. https://doi.org/10.1016/j.jhin.2007.02.013

Cantón, R., González-Alba, J. M., & Galán, J. C. (2012). CTX-M Enzymes: Origin and Diffusion. *Frontiers in Microbiology*, *3*. https://doi.org/10.3389/fmicb.2012.00110

Carattoli, A. (2009). Resistance plasmid families in Enterobacteriaceae. *Antimicrobial Agents and Chemotherapy*, *53*(6), 2227–2238. https://doi.org/10.1128/AAC.01707-08

Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L., & Threlfall, E. J. (2005). Identification of plasmids by PCR-based replicon typing. *Journal of Microbiological Methods*, *63*(3), 219–228. https://doi.org/10.1016/j.mimet.2005.03.018

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., … Hasman, H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, *58*(7), 3895–3903. https://doi.org/10.1128/AAC.02412-14

Carvalho-Assef, A. P. D., Pereira, P. S., Albano, R. M., Berião, G. C., Tavares, C. P., Chagas, T. P. G., … Asensi, M. D. (2014). Detection of NDM-1-, CTX-M-15-, and qnrB4-producing Enterobacter hormaechei isolates in Brazil. *Antimicrobial Agents and Chemotherapy*, *58*(4), 2475–2476. https://doi.org/10.1128/AAC.02804-13

CDC. (2019, March 20). What Exactly is Antibiotic Resistance? Retrieved June 8, 2019, from Centers for Disease Control and Prevention website: https://www.cdc.gov/drugresistance/about.html

Chavda, K. D., Chen, L., Fouts, D. E., Sutton, G., Brinkac, L., Jenkins, S. G., … Kreiswirth, B. N. (2016). Comprehensive Genome Analysis of Carbapenemase-Producing Enterobacter spp.: New Insights into Phylogeny, Population Structure, and Resistance Mechanisms. *MBio*, *7*(6), e02093-16. https://doi.org/10.1128/mBio.02093-16

Conant, G. C., & Wolfe, K. H. (2008). GenomeVx: Simple web-based creation of editable circular chromosome maps. *Bioinformatics (Oxford, England)*, *24*(6), 861–862. https://doi.org/10.1093/bioinformatics/btm598

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., … Harris, S. R. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Research*, *43*(3), e15–e15. https://doi.org/10.1093/nar/gku1196

Dagher, C., Salloum, T., Alousi, S., Arabaghian, H., Araj, G. F., & Tokajian, S. (2018). Molecular characterization of Carbapenem resistant Escherichia coli recovered from a tertiary hospital in Lebanon. *PLOS ONE*, *13*(9), e0203323. https://doi.org/10.1371/journal.pone.0203323

Dalben, M., Varkulja, G., Basso, M., Krebs, V. L. J., Gibelli, M. A., van der Heijden, I., … Costa, S. F. (2008). Investigation of an outbreak of Enterobacter cloacae in a neonatal unit and review of the literature. *The Journal of Hospital Infection*, *70*(1), 7–14. https://doi.org/10.1016/j.jhin.2008.05.003

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*, 218431. https://doi.org/10.1155/2011/218431

Daoud, Z., Farah, J., Sokhn, E. S., El Kfoury, K., Dahdouh, E., Masri, K., … Matar, G. M. (2017). Multidrug-Resistant Enterobacteriaceae in Lebanese Hospital Wastewater: Implication in the One Health Concept. *Microbial Drug Resistance*, *24*(2), 166–174. https://doi.org/10.1089/mdr.2017.0090

Davin-Regli, A., & Pagès, J.-M. (2015). Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00392

Diene, S. M., Merhej, V., Henry, M., El Filali, A., Roux, V., Robert, C., … Rolain, J.-M. (2013). The rhizome of the multidrug-resistant Enterobacter aerogenes genome reveals how new “killer bugs” are created because of a sympatric lifestyle. *Molecular Biology and Evolution*, *30*(2), 369–383. https://doi.org/10.1093/molbev/mss236

Evans, B. A., & Amyes, S. G. B. (2014). OXA β-Lactamases. *Clinical Microbiology Reviews*, *27*(2), 241–263. https://doi.org/10.1128/CMR.00117-13

Federhen, S. (2012). The NCBI Taxonomy database. *Nucleic Acids Research*, *40*(Database issue), D136–D143. https://doi.org/10.1093/nar/gkr1178

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784–3788. https://doi.org/10.1093/nar/gkg563

Gest, H. (2004). The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes and Records of the Royal Society of London*, *58*(2), 187–201. https://doi.org/10.1098/rsnr.2004.0055

Guenther, S., Semmler, T., Stubbe, A., Stubbe, M., Wieler, L. H., & Schaufler, K. (2017). Chromosomally encoded ESBL genes in Escherichia coli of ST38

from Mongolian wild birds. *The Journal of Antimicrobial Chemotherapy*, *72*(5), 1310–1313. https://doi.org/10.1093/jac/dkx006

Hadfield, J., Croucher, N. J., Goater, R. J., Abudahab, K., Aanensen, D. M., & Harris, S. R. (2018). Phandango: An interactive viewer for bacterial population genomics. *Bioinformatics*, *34*(2), 292–293. https://doi.org/10.1093/bioinformatics/btx610

Haertl, R., & Bandlow, G. (1993). Epidemiological fingerprinting of Enterobacter cloacae by small-fragment restriction endonuclease analysis and pulsed-field gel electrophoresis of genomic restriction fragments. *Journal of Clinical Microbiology*, *31*(1), 128–133.

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Møller, N., & Aarestrup, F. M. (2014). Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *Journal of Clinical Microbiology*, *52*(1), 139–146. https://doi.org/10.1128/JCM.02452-13

Herte, R. (2012). The threat of carbapenem-resistant Enterobacteriaceae in Lebanon: An update on the regional and local epidemiology | Elsevier Enhanced Reader. https://doi.org/10.1016/j.jiph.2012.02.003

Hoffmann, H., & Roggenkamp, A. (2003). Population genetics of the nomenspecies Enterobacter cloacae. *Applied and Environmental Microbiology*, *69*(9), 5306–5318. https://doi.org/10.1128/aem.69.9.5306-5318.2003

Hoffmann, H., Stindl, S., Ludwig, W., Stumpf, A., Mehlen, A., Heesemann, J., … Roggenkamp, A. (2005). Reassignment of enterobacter dissolvens to Enterobacter cloacae as E. cloacae subspecies dissolvens comb. nov. and emended description of Enterobacter asburiae and Enterobacter kobei. *Systematic and Applied Microbiology*, *28*(3), 196–205. https://doi.org/10.1016/j.syapm.2004.12.010

Hoffmann, H., Stindl, S., Ludwig, W., Stumpf, A., Mehlen, A., Monget, D., … Schleifer, K. H. (2005). Enterobacter hormaechei subsp. oharae subsp. nov., E. hormaechei subsp. hormaechei comb. nov., and E. hormaechei subsp. steigerwaltii subsp. nov., three new subspecies of clinical importance. *Journal of Clinical Microbiology*, *43*(7), 3297–3303. https://doi.org/10.1128/JCM.43.7.3297-3303.2005

Hoffmann, H., Stindl, S., Stumpf, A., Mehlen, A., Monget, D., Heesemann, J., … Roggenkamp, A. (2005). Description of Enterobacter ludwigii sp. nov., a

novel Enterobacter species of clinical relevance. *Systematic and Applied Microbiology*, *28*(3), 206–212. https://doi.org/10.1016/j.syapm.2004.12.009

Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., … Thomson, N. R. (2015). Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. *Proceedings of the National Academy of Sciences*, *112*(27), E3574–E3581. https://doi.org/10.1073/pnas.1501049112

Hormaeche, E., & Edwards, P. R. (1960). A Proposed Genus Enterobacter. *International Journal of Systematic and Evolutionary Microbiology*, *10*(2), 71–74. https://doi.org/10.1099/0096266X-10-2-71

Hurley, J. M., & Woychik, N. A. (2009). Bacterial Toxin HigB Associates with Ribosomes and Mediates Translation-dependent mRNA Cleavage at A-rich Sites. *Journal of Biological Chemistry*, *284*(28), 18605–18613. https://doi.org/10.1074/jbc.M109.008763

Izdebski, R., Baraniak, A., Herda, M., Fiett, J., Bonten, M. J. M., Carmeli, Y., … MOSAR WP2, WP3 and WP5 Study Groups. (2015). MLST reveals potentially high-risk international clones of Enterobacter cloacae. *The Journal of Antimicrobial Chemotherapy*, *70*(1), 48–56. https://doi.org/10.1093/jac/dku359

Kane, K. A., & Dorman, C. J. (2012). VirB-Mediated Positive Feedback Control of the Virulence Gene Regulatory Cascade of Shigella flexneri. *Journal of Bacteriology*, *194*(19), 5264–5273. https://doi.org/10.1128/JB.00800-12

Khan, A. U., Maryam, L., & Zarrilli, R. (2017). Structure, Genetics and Worldwide Spread of New Delhi Metallo-β-lactamase (NDM): A threat to public health. *BMC Microbiology*, *17*. https://doi.org/10.1186/s12866-017-1012-8

Kirchberger, P. C., Unterweger, D., Provenzano, D., Pukatzki, S., & Boucher, Y. (2017). Sequential displacement of Type VI Secretion System effector genes leads to evolution of diverse immunity gene arrays in *Vibrio cholerae*. *Scientific Reports*, *7*, 45133. https://doi.org/10.1038/srep45133

Klebsiella aerogenes MLST Home Page. (n.d.). Retrieved June 7, 2019, from https://pubmlst.org/kaerogenes/

Kluytmans-van den Bergh, M. F. Q., Rossen, J. W. A., Bruijning-Verhagen, P. C. J., Bonten, M. J. M., Friedrich, A. W., Vandenbroucke-Grauls, C. M. J. E., … Kluytmans, J. A. J. W. (2016). Whole-Genome Multilocus Sequence Typing

of Extended-Spectrum-Beta-Lactamase-Producing Enterobacteriaceae. *Journal of Clinical Microbiology*, *54*(12), 2919–2927. https://doi.org/10.1128/JCM.01648-16

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., … Peacock, S. J. (2012). Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. *PLOS Pathogens*, *8*(8), e1002824. https://doi.org/10.1371/journal.ppat.1002824

Krzymińska, S., Mokracka, J., Koczura, R., & Kaznowski, A. (2009). Cytotoxic activity of Enterobacter cloacae human isolates. *FEMS Immunology and Medical Microbiology*, *56*, 248–252. https://doi.org/10.1111/j.1574-695X.2009.00572.x

Lam, M. M. C., Wick, R. R., Wyres, K. L., Gorrie, C. L., Judd, L. M., Jenney, A. W. J., … Holt, K. E. (2018). Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in Klebsiella pneumoniae populations. *Microbial Genomics*, *4*(9). https://doi.org/10.1099/mgen.0.000196

Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M. R., Ahn, T.-H., … Ussery, D. W. (2015). Insights from 20 years of bacterial genome sequencing. *Functional & Integrative Genomics*, *15*(2), 141–161. https://doi.org/10.1007/s10142-015-0433-4

Lavigne, J.-P., Sotto, A., Nicolas-Chanoine, M.-H., Bouziges, N., Pagès, J.-M., & Davin-Regli, A. (2013). An adaptive response of Enterobacter aerogenes to imipenem: Regulation of porin balance in clinical isolates. *International Journal of Antimicrobial Agents*, *41*(2), 130–136. https://doi.org/10.1016/j.ijantimicag.2012.10.010

Letunic, I., & Bork, P. (n.d.). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, 2019. https://doi.org/10.1093/nar/gkz239

Lixandru, B. E., Cotar, A. I., Straut, M., Usein, C. R., Cristea, D., Ciontea, S., … Damian, M. (2015). Carbapenemase-Producing Klebsiella pneumoniae in Romania: A Six-Month Survey. *PLOS ONE*, *10*(11), e0143214. https://doi.org/10.1371/journal.pone.0143214

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkz268

Malek, A., McGlynn, K., Taffner, S., Fine, L., Tesini, B., Wang, J., … Pecora, N. (2019). Next-Generation-Sequencing-Based Hospital Outbreak Investigation Yields Insight into Klebsiella aerogenes Population Structure and Determinants of Carbapenem Resistance and Pathogenicity. *Antimicrobial Agents and Chemotherapy*, *63*(6). https://doi.org/10.1128/AAC.02577-18

Marttinen, P., Croucher, N. J., Gutmann, M. U., Corander, J., & Hanage, W. P. (2015). Recombination produces coherent bacterial species clusters in both core and accessory genomes. *Microbial Genomics*, *1*(5). https://doi.org/10.1099/mgen.0.000038

Matar, G. M., Cuzon, G., Araj, G. F., Naas, T., Corkill, J., Kattar, M. M., & Nordmann, P. (2008). Oxacillinase-mediated resistance to carbapenems in Klebsiella pneumoniae from Lebanon. *Clinical Microbiology and Infection*, *14*(9), 887–888. https://doi.org/10.1111/j.1469-0691.2008.02059.x

Mezzatesta, M. L., Gona, F., & Stefani, S. (2012). Enterobacter cloacae complex: Clinical impact and emerging antibiotic resistance. *Future Microbiology*, *7*(7), 887–902. https://doi.org/10.2217/fmb.12.61

Miyoshi-Akiyama, T., Hayakawa, K., Ohmagari, N., Shimojima, M., & Kirikae, T. (2013). Multilocus sequence typing (MLST) for characterization of Enterobacter cloacae. *PloS One*, *8*(6), e66358. https://doi.org/10.1371/journal.pone.0066358

Moghnieh, R., Araj, G. F., Awad, L., Daoud, Z., Mokhbat, J. E., Jisr, T., … Husni, R. (2019). A compilation of antimicrobial susceptibility data from a network of 13 Lebanese hospitals reflecting the national situation during 2015–2016. *Antimicrobial Resistance & Infection Control*, *8*(1), 41. https://doi.org/10.1186/s13756-019-0487-5

Morand, P. C., Billoet, A., Rottman, M., Sivadon-Tardy, V., Eyrolle, L., Jeanne, L., … Dumaine, V. (2009). Specific distribution within the Enterobacter cloacae complex of strains isolated from infected orthopedic implants. *Journal of Clinical Microbiology*, *47*(8), 2489–2495. https://doi.org/10.1128/JCM.00290-09

Moubareck, C., Daoud, Z., Hakime, N. I., Hamze, M., Mangeney, N., Matta, H., … Doucet-Populaire, F. (2005). Countrywide Spread of Community- and Hospital-Acquired Extended-Spectrum -Lactamase (CTX-M-15)-Producing Enterobacteriaceae in Lebanon. *Journal of Clinical Microbiology*, *43*(7), 3309–3313. https://doi.org/10.1128/JCM.43.7.3309-3313.2005

Musil, I., Jensen, V., Schilling, J., Ashdown, B., & Kent, T. (2010). Enterobacter cloacae infection of an expanded polytetrafluoroethylene femoral-popliteal bypass graft: A case report. *Journal of Medical Case Reports*, *4*, 131. https://doi.org/10.1186/1752-1947-4-131

Nordmann, P., Dortet, L., & Poirel, L. (2012). Carbapenem resistance in Enterobacteriaceae: Here is the storm! *Trends in Molecular Medicine*, *18*(5), 263–272. https://doi.org/10.1016/j.molmed.2012.03.003

Nordmann, P., Naas, T., & Poirel, L. (2011). Global Spread of Carbapenemase-producing Enterobacteriaceae. *Emerging Infectious Diseases*, *17*(10), 1791–1798. https://doi.org/10.3201/eid1710.110655

O'Hara, C. M., Steigerwalt, A. G., Hill, B. C., Farmer, J. J., Fanning, G. R., & Brenner, D. J. (1989). Enterobacter hormaechei, a new species of the family Enterobacteriaceae formerly known as enteric group 75. *Journal of Clinical Microbiology*, *27*(9), 2046–2049.

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., & Phillippy, A. M. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biology*, *17*(1), 132. https://doi.org/10.1186/s13059-016-0997-x

Paauw, A., Caspers, M. P. M., Leverstein-van Hall, M. A., Schuren, F. H. J., Montijn, R. C., Verhoef, J., & Fluit, A. C. (2009). Identification of resistance and virulence factors in an epidemic Enterobacter hormaechei outbreak strain. *Microbiology (Reading, England)*, *155*(Pt 5), 1478–1488. https://doi.org/10.1099/mic.0.024828-0

Paauw, A., Caspers, M. P. M., Schuren, F. H. J., Leverstein-van Hall, M. A., Delétoile, A., Montijn, R. C., … Fluit, A. C. (2008). Genomic diversity within the Enterobacter cloacae complex. *PloS One*, *3*(8), e3018. https://doi.org/10.1371/journal.pone.0003018

Papagiannitsis, C. C., Bitar, I., Malli, E., Tsilipounidaki, K., Hrabak, J., & Petinaki, E. (2019). IncC blaKPC-2-positive plasmid characterized from ST648 Escherichia coli. *Journal of Global Antimicrobial Resistance*. https://doi.org/10.1016/j.jgar.2019.05.001

Paskova, V., Medvecky, M., Skalova, A., Chudejova, K., Bitar, I., Jakubu, V., … Hrabak, J. (2018). Characterization of NDM-Encoding Plasmids From Enterobacteriaceae Recovered From Czech Hospitals. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.01549

Peirano, G., Matsumura, Y., Adams, M. D., Bradford, P., Motyl, M., Chen, L., … Pitout, J. D. D. (2018). Genomic epidemiology of global carbapenemase-producing enterobacter spp., 2008–2014. *Emerging Infectious Diseases*, *24*(6), 1010–1019. https://doi.org/10.3201/eid2406.171648

Podschun, R., & Ullmann, U. (1998). Klebsiella spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clinical Microbiology Reviews*, *11*(4), 589–603.

Rinkel, M., Hubert, J.-C., Roux, B., & Lett, M.-C. (1994). Transposon Tn5403, a mobilization-helper element: Complete nucleotide sequence and distribution in aquatic strains. *FEMS Microbiology Ecology*, *15*(1–2), 89–95. https://doi.org/10.1111/j.1574-6941.1994.tb00233.x

Rodriguez-R, L. M., & Konstantinidis, K. T. (2016). *The enveomics collection: A toolbox for specialized analyses of microbial genomes and metagenomes* (No. e1900v1). https://doi.org/10.7287/peerj.preprints.1900v1

Rozwandowicz, M., Brouwer, M. S. M., Fischer, J., Wagenaar, J. A., Gonzalez-Zorn, B., Guerra, B., … Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *Journal of Antimicrobial Chemotherapy*, *73*(5), 1121–1137. https://doi.org/10.1093/jac/dkx488

Sanders, W. E., & Sanders, C. C. (1997). Enterobacter spp.: Pathogens poised to flourish at the turn of the century. *Clinical Microbiology Reviews*, *10*(2), 220–241.

Sato, M., & Miyazaki, K. (2017). Phylogenetic Network Analysis Revealed the Occurrence of Horizontal Gene Transfer of 16S rRNA in the Genus Enterobacter. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02225

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seemann, T. (2015). *Snippy: fast bacterial variant calling from NGS reads*. Retrieved from https://github.com/tseemann/snippy

Shon, A. S., Bajwa, R. P. S., & Russo, T. A. (2013). Hypervirulent (hypermucoviscous) Klebsiella pneumoniae: A new and dangerous breed. *Virulence*, *4*(2), 107–118. https://doi.org/10.4161/viru.22718

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., & Chandler, M. (2006). ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Research*, *34*(Database issue), D32-36. https://doi.org/10.1093/nar/gkj014

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

STANDARD OPERATING PROCEDURE FOR PULSENET PFGE OF ESCHERICHIA COLI O157:H7, ESCHERICHIA COLI NON-O157 (STEC), SALMONELLA SEROTYPES, SHIGELLA SONNEI AND SHIGELLA FLEXNERI. (n.d.). 16.

Sutton, G. G., Brinkac, L. M., Clarke, T. H., & Fouts, D. E. (2018). Enterobacter hormaechei subsp. hoffmannii subsp. nov., Enterobacter hormaechei subsp. xiangfangensis comb. nov., Enterobacter roggenkampii sp. nov., and Enterobacter muelleri is a later heterotypic synonym of Enterobacter asburiae based on computational analysis of sequenced Enterobacter genomes. *F1000Research*, *7*. https://doi.org/10.12688/f1000research.14566.2

Tangcharoensathien, V., Sattayawutthipong, W., Kanjanapimai, S., Kanpravidth, W., Brown, R., & Sommanustweechai, A. (2017). Antimicrobial resistance: From global agenda to national strategic plan, Thailand. *Bulletin of the World Health Organization*, *95*(8), 599–603. https://doi.org/10.2471/BLT.16.179648

Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H., & Swaminathan, B. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: Criteria for bacterial strain typing. *Journal of Clinical Microbiology*, *33*(9), 2233–2239.

Tindall, B. J., Sutton, G., & Garrity, G. M. (2017). Enterobacter aerogenes Hormaeche and Edwards 1960 (Approved Lists 1980) and Klebsiella mobilis Bascomb et al. 1971 (Approved Lists 1980) share the same nomenclatural type (ATCC 13048) on the Approved Lists and are homotypic synonyms, with consequences for the name Klebsiella mobilis Bascomb et al. 1971 (Approved Lists 1980). *International Journal of Systematic and Evolutionary Microbiology*, *67*(2), 502–504. https://doi.org/10.1099/ijsem.0.001572

Vielva, L., de Toro, M., Lanza, V. F., & de la Cruz, F. (2017). PLACNETw: A web-based tool for plasmid reconstruction from bacterial genomes. *Bioinformatics (Oxford, England)*, *33*(23), 3796–3798. https://doi.org/10.1093/bioinformatics/btx462

WEINSTEIN, M. P. (2018). *M100-performance standards for antimicrobial susceptibility testing, 28th edition.* Place of publication not identified: CLINICAL AND LABORATORY.

Wenger, P. N., Tokars, J. I., Brennan, P., Samel, C., Bland, L., Miller, M., … Jarvis, W. (1997). An outbreak of Enterobacter hormaechei infection and colonization in an intensive care nursery. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *24*(6), 1243–1244. https://doi.org/10.1086/513650

WHO. (n.d.). WHO publishes list of bacteria for which new antibiotics are urgently needed. Retrieved June 4, 2019, from https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

Wick, R. R., Heinz, E., Holt, K. E., & Wyres, K. L. (2018). Kaptive Web: User-Friendly Capsule and Lipopolysaccharide Serotype Prediction for Klebsiella Genomes. *Journal of Clinical Microbiology*, *56*(6), e00197-18. https://doi.org/10.1128/JCM.00197-18

Wyres, K. L., Wick, R. R., Gorrie, C., Jenney, A., Follador, R., Thomson, N. R., & Holt, K. E. (2016). Identification of Klebsiella capsule synthesis loci from whole genome data. *Microbial Genomics*, *2*(12). https://doi.org/10.1099/mgen.0.000102

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., … Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *The Journal of Antimicrobial Chemotherapy*, *67*(11), 2640–2644. https://doi.org/10.1093/jac/dks261

Zhu, B., Lou, M.-M., Xie, G.-L., Wang, G.-F., Zhou, Q., Wang, F., … Duan, Y.-P. (2011). Enterobacter mori sp. nov., associated with bacterial wilt on Morus alba L. *International Journal of Systematic and Evolutionary Microbiology*, *61*(Pt 11), 2769–2774. https://doi.org/10.1099/ijs.0.028613-0