CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

NOVEL APPROACHES ON THE CHARACTERIZATION OF THE WASTEWATER RESISTOME: POSSIBLE IMPLICATIONS ON HUMAN HEALTH AND WATER QUALITY MANAGEMENT

Thesis submitted to the Universidade Católica Portuguesa to attain

the degree of PhD in Biotechnology – with specialization in Microbiology

By

Jaqueline Maria Matias da Rocha

November 2021

CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

NOVEL APPROACHES ON THE CHARACTERIZATION OF THE WASTEWATER RESISTOME: POSSIBLE IMPLICATIONS ON HUMAN HEALTH AND WATER QUALITY MANAGEMENT

Thesis submitted to the Universidade Católica Portuguesa to attain

the degree of Ph.D. in Biotechnology with specialization in Microbiology

By

Jaqueline Maria Matias da Rocha

Under the supervision of Professor Célia Maria Manaia Rodrigues

Under the co-supervision of Professor Isabel da Silva Henriques

and Professor Margarita Gomila Ribas

November 2021

To my family and to Vitor

RESUMO

As bactérias resistentes a antibióticos e os seus genes de resistência a antibióticos são considerados contaminantes emergentes, hoje amplamente disseminados no meio ambiente. As estações de tratamento de águas residuais domésticas são importantes recetores e emissores destes contaminantes. Nestas estações, os esgotos são submetidos a diferentes tipos de tratamento que embora reduzam os níveis de bactérias resistentes a antibióticos e genes de resistência a antibióticos, não evitam que haja ainda fortes cargas emitidas para o meio recetor. Tendo em conta que a maioria das bactérias não é cultivável e que essa fração de bactérias pode conter genes de resistência a antibióticos, os métodos independentes de cultivo têm vindo a ser usados para avaliar a ocorrência de resistência aos antibióticos no ambiente. o método quantitativo de *Polymerase Chain Reaction* (qPCR) é um dos preferidos para quantificar genes de resistência a antibióticos em amostras ambientais. Contudo, é reconhecido que a harmonização de métodos será decisiva para que se possa comparar de modo confiável os resultados obtidos em diferentes partes do mundo.

Entre as bactérias resistentes a antibióticos emitidas por estações de tratamento de águas residuais, as da família *Enterobacteriaceae* representam uma fração importante e entre estas as da espécie *Klebsiella pneumoniae* merecem atenção especial. De facto, *K. pneumoniae* é uma espécie bacteriana que além da importância clínica quando associada aos humanos, também pode ser encontrada no meio ambiente, em solos, plantas, água e águas residuais. A capacidade de membros desta espécie de proliferar em diferentes ambientes e em humanos e animais pode constituir uma ameaça à saúde humana. No entanto, as características que podem ser mantidas ou perdidas durante a transição de *K. pneumoniae* por um contexto clínico ou ambiental ainda são desconhecidas.

Esta tese teve como objetivo 1) avançar no conhecimento relativamente ao uso de métodos analíticos harmonizados de qPCR que possam permitir comparações confiáveis da quantificação de genes; 2) desenvolver um padrão interno baseado em células que pudesse ser usado em diferentes laboratórios para normalizar resultados e avaliar as perdas durante o processamento das amostras de água, extração de DNA e quantificação por qPCR; e 3) contribuir para um melhor conhecimento da ecologia de *K. pneumoniae* e inferir sobre as possíveis dinâmicas entre nichos clínicos e ambientais, com especial foco na diversidade genética e estabilidade da resistência a antibióticos.

Para atingir o primeiro objetivo, determinantes genéticos que codificam resistência a sulfonamidas (*sul1* e *sul2*), quinolonas (*qnrS*) e β-lactâmicos (bla_{TEM}) e o gene 16S rRNA foram monitorizados em extratos de DNA fornecidos por parceiros que estavam a investigar processos de tratamento de águas residuais, à escala real ou em sistemas piloto. Paralelamente, foi estudada a influência do transporte de DNA, protocolos, padrões e equipamento de qPCR na comparação da quantificação de genes em ensaios inter-laboratoriais. Estes resultados e a literatura disponível justificaram os esforços para atingir o segundo objetivo. Um padrão interno, que consistia num fragmento de um gene que não se

encontra no ambiente, clonado em *Escherichia coli*, foi desenvolvido e testado. Este padrão interno deverá ser adicionado a amostras de água residual ou água com o objetivo de controlar as perdas de DNA. A emissão de bactérias resistentes a antibióticos por estações de tratamento de águas residuais é uma questão preocupante, contudo não é claro se estas bactérias sobreviverão e manterão as suas características uma vez libertadas no meio ambiente. Para investigar este tópico, *K. pneumoniae* foi usada como espécie-modelo e duas abordagens de investigação distintas foram utilizadas. Um grupo de isolados de *K. pneumoniae* resistentes a cefalosporinas de 3ª geração (25 de águas residuais; 34 clínicas) foi comparado com base em características fenotípicas, genotípicas e genómicas (n = 22). Em paralelo, um grupo mais amplo de genomas obtidos de uma base de dados pública (21 países; 61 ambientais, 78 clínicos) foi comparado com base na análise do pangenoma e nos perfis de características de resistência a antibióticos e metais, virulência, sistemas de efluxo, stresse oxidativo e *quorum sensing*.

De acordo com os resultados obtidos e sua interpretação, concluiu-se que a eficiência do tratamento e a qualidade de águas residuais, em relação à resistência a antibióticos, devem ser sempre medidas com base na abundância absoluta (por volume), e não na abundância relativa (por total de bactérias). A comparação inter-laboratorial pareceu ser confiável, embora a qualidade e estabilidade dos extratos de DNA durante o transporte, bem como as especificidades dos consumíveis e do equipamento, possam ser críticos para os resultados de monitorização. O uso de um padrão interno baseado em células poderá contribuir para superar estas limitações. Este padrão interno permitiu estimar o efeito de matriz da água, que esteve associado a uma subestimativa que variou de 0.1–0.9 log do número de cópias do gene por mililitro de amostra, independentemente do tipo de água. Os isolados clínicos e de águas residuais foram indistinguíveis com base na caracterização fenotípica e genotípica, embora os dados sugerissem que linhagens distintas possam prevalecer em contexto clínico ou ambiental. Os determinantes genéticos relacionados com funções de efluxo, stresse oxidativo e *quorum sensing* estavam presentes em isolados clínicos e de águas residuais, enquanto que a resistência a antibióticos e metais ou genes de virulência variavam de estirpe para estirpe, estando associados a elementos genéticos móveis, principalmente transposões, sequências de inserção ou elementos integrativos e conjugativos, respetivamente. A análise de um grupo maior e geograficamente mais diverso de genomas sugeriu que os alelos dos genes de virulência e resistência a antibióticos e metais são mais prevalentes e diversos em isolados clínicos, enquanto alguns alelos de determinantes genéticos relacionados com *quorum sensing*, sistemas de efluxo e stresse oxidativo são mais prevalentes e diversos em isolados ambientais.

Os estudos realizados são promissores relativamente à possibilidade de implementar esquemas inter-laboratoriais de monitorização de resistência a antibióticos que forneçam resultados comparáveis e confiáveis. A harmonização de alguns procedimentos e o uso de

padrões internos permitirão comparações de quantificação de genes de resistência a antibióticos em águas residuais e, portanto, melhorar e promover estudos de vigilância mundialmente. A dupla evidência de que as características de resistência a antibióticos observadas em *K. pneumoniae* clínicas são mantidas no meio ambiente e que é entre os isolados ambientais que as características de tolerância ao stresse parecem ser mais diversas, sugere a alta capacidade de *K. pneumoniae* para se disseminar a partir de águas residuais e no ambiente, aumentando os riscos de transmissão ao ser humano.

Palavras-chave: resistência a antibióticos, águas residuais, padrão interno, *Klebsiella pneumoniae*, genómica comparada

ABSTRACT

Antibiotic resistant bacteria and antibiotic resistance genes are considered contaminants of emerging concern, nowadays widely disseminated in the environment. Urban wastewater treatment plants are major recipients and reservoirs for these contaminants. In urban these plants, wastewater is subjected to different types of treatment that reduce the levels of antibiotic resistant bacteria and antibiotic resistance genes, although not completely. Considering that most of the bacteria are not culturable and that this fraction might harbor antibiotic resistance genes, culture-independent methods are currently used to assess antibiotic resistance in the environment. Among these methods, quantitative PCR is considered the gold standard used to quantify antibiotic resistance genes in environmental samples, although the lack of harmonized methods seriously limits the reliable comparison of results obtained in different laboratories.

Among the antibiotic resistant bacteria emitted by wastewater treatment plants, *Enterobacteriaceae* represent an important fraction and among these *Klebsiella pneumoniae* deserve special attention. Indeed, *K. pneumoniae* is a bacterial species that besides the clinical importance when associated to humans, can also be found in the environment, in soils, plants, water, and wastewater. The capacity of this bacterial species to thrive in different environments and in humans and animals increases its significance in terms of human health threat might constitute a human health threat. However, the traits that might be maintained or lost during the transit of *K. pneumoniae* through clinical and environmental contexts are still unknown.

This thesis aimed to 1) advance the knowledge regarding the use of harmonized analytical quantitative PCR methods that might enable reliable comparisons of genes quantification ; 2) design a cell-based internal standard that could be used in different laboratories to assess losses during water samples filtration, DNA extraction and quantitative PCR quantification; and 3) contribute to a better understanding of the ecology of *K. pneumoniae* and infer about possible dynamics between clinic and environmental niches, with special focus on genetic diversity and antibiotic resistance stability.

To tackle the first aim, genetic determinants encoding resistance to sulfonamides (*sul1* and *sul2*), quinolones (*qnrS*), and β-lactams (bla_{TEM}) and the 16S rRNA gene were monitored in DNA extracts supplied by partners who were investigating wastewater treatment processes, at full- or pilot-scale. In parallel, the influence on genes quantification of DNA shipment, quantitative PCR protocols, standards and equipment was studied in an interlaboratory comparison. These results and the literature available justified the efforts to meet the second aim. An internal standard, consisting in a cloned gene fragment not found in wastewater samples was designed and tested. This internal standard is to be used to spike wastewater or water samples aiming to control DNA losses during the processing of the sample and DNA extraction process. The emission of antibiotic resistant bacteria by wastewater treatment plants is an issue of concern, however it is not clear if these bacteria will survive and maintain

their features once in the environment. To investigate this topic, *K. pneumoniae* was used as a model species and two distinct research approaches were used. A group of 3rd generation cephalosporin-resistant *K. pneumoniae* isolates (25 wastewater; 34 clinical) was compared based on phenotypic, genotypic and genomic analyses (n=22) and a broader group of genomes collected from a public database (21 countries, 61 environmental; 78 clinical) was compared based on a core and pangenome approach and profiles of antibiotic and metal resistance, virulence, efflux systems, oxidative stress and quorum sensing traits.

According to the results obtained and their analysis, it was concluded that wastewater treatment efficiency and wastewater quality regarding antibiotic resistance emissions should always be measured based on absolute abundance (per volume), rather than in relative abundance (per total bacteria). The interlaboratory comparison seemed to be reliable, although DNA extract quality and stability during shipment, as well as consumables and equipment specificities, may be critical for the monitoring findings. The use of a cell-based internal standard may contribute to overcome those limitations. This internal standard permitted to estimate the water matrix effect which was associated with an underestimation that ranged 0.1–0.9 log gene copy number mL^{-1} of sample, irrespective of the water type. Clinical and wastewater isolates were indistinguishable based on phenotypic and genotypic characterization, although distinct lineages may prevail in clinical or environmental settings. Genetic determinants related to efflux, oxidative stress or quorum sensing functions were common to clinical and wastewater isolates, while antibiotic and metal resistance or virulence genes, were variable across the genomes and associated with mobile genetic elements, mostly transposons, insertion sequences or integrative and conjugative elements. The analysis of a larger and geographically more diverse group of genomes, suggested that antibiotic and metal resistance and virulence gene alleles were more prevalent and diverse in clinical isolates, while some quorum sensing, efflux systems and oxidative stress genetic determinants alleles were more prevalent and diverse in environmental isolates.

The studies performed unveil a promising opportunity to implement comparable and reliable antibiotic resistance monitoring schemes. The harmonization of some procedures and the use of internal standards will enable worldwide comparisons of antibiotic resistance genes in wastewaters, and therefore improve and promote surveillance studies worldwide. The double evidence that antibiotic resistance features observed in clinical *K. pneumoniae* are maintained in the environment and that it is among the environmental isolates that stress dwelling features seem to be more diverse, supports the high capacity of *K. pneumoniae* for spreading through wastewater or in the environment, enhancing the risks of transmission back to humans.

Keywords: antibiotic resistance, wastewater, internal standard, *Klebsiella pneumoniae*, comparative genomics

ACKNOWLEDGMENTS

To Universidade Católica Portuguesa – Escola Superior de Biotecnologia, for receiving me as a PhD student and for providing the necessary conditions for the development of my work.

To European Social Fund, under Programa Operacional Regional do Norte – Norte2020, for the grant awarded (NORTE-08-5369-FSE-000007) and for the financial support to attend international scientific symposiums. The work presented at this thesis was supported by the project NanoDiaBac – Nanofluidics for Ultrafast Diagnosis of Bacterial Infections (ENMED/0001/2014), StARE – Stopping Antibiotic Resistance Evolution (WaterJPI/0001/2013), OVERSEAS - Overseas Environmental Antibiotic Resistance Surveillance project (FLAD/NSF - "Programa de Redes de Investigação Portugal/EUA 2015" - Ciência Ambiental - Projecto 2015/298); Wastewater Reuse and Contaminants of Emerging Concern of NORMAN project, RISK.AR: Avaliação dos riscos associados a bactérias ambientais resistentes a antibióticos: propagação e transmissão aos humanos" (nº 28196), and by FEMS Research and Training Grant (FEMS GO 2017 008).

To my supervisor, Prof. Célia Manaia, for giving me the opportunity to integrate her research group and for believing in my capacities to develop this work. Thank you for your patience, for your support and for your guidance.

To my co-supervisor, Prof. Isabel Henriques, for the guidance and helpful discussions about the work performed. Thank you for your availability and for your support.

To my co-supervisor, Prof. Margarita Gomila for receiving me in her group at the Universitat de les Illes Balears, for the countless hours of discussions and for her help and patience. I am also grateful to all the research team for their contribution to improve my work and for being so friendly and nice.

To all my colleagues from the lab Catarina Ferreira, Diana Bogas, Ivone Moreira, Telma Fernandes, Nazareno Sccacia, Gianuário Fortunato, Joana Silva, João Magalhães and to the technicians Rui Magalhães, Susana Xis, Ana Martins at Escola Superior de Biotecnologia, who directly or indirectly supported me and contributed to the development of my work.

To Prof. Amy Pruden for receiving me at her lab at Virginia Tech University allowing me to perform part of the work of this thesis based on inter-laboratory comparisons. Also, to her research group which was enthusiastic and helped whenever it was needed.

To Doctor Arsénio Fialho for receiving me at his lab at Instituto Superior Técnico, in Lisbon, which allowed me to perform the experimental assays using the model organism *Galleria mellonella*. Also, to his research team for receiving me so well. A special thanks to Doctor

Dalila Mil-Homens for her availability, for her help and for her support on infection assays using the model organism *Galleria mellonella*.

To Terri and Lou for being so nice and friendly during my stay in US and for making me living the experience of a Thanksgiving Day, even if it was in September.

To Margarita Gomila and Tony Busquets for being my family too during my forced stay in Mallorca during the Covid pandemic situation. You are amazing and I will never forget the moments we lived and the support you gave me.

To my friends for their big support along this journey. A special thanks to some good friends: Sofia Amaro, Rita Inácio, Ana Rita Monforte, Catarina Ferreira, Diana Bogas, Flávio Arrais, Gonçalo Amaro for their unconditional friendship and support along this journey.

At last, but not the least, I would like to express my deep gratitude to my family, for her constant support, encouragement and patience through all these years. Sorry for the times I was overloaded with work and that I was not so available and patient as I should be. My dearest thanks go to Vitor, for his love, support and patience and for making me believe that life is good, and we should enjoy it always thinking and feeling positive.

Without all of you, this work would not have become possible.

TABLE OF CONTENTS

[1] These tables are provided in a link.

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

CHAPTER 3. MONITORING ANTIBIOTIC RESISTANCE GENES IN WASTEWATER: POTENTIALS AND LIMITATIONS OF QUANTITATIVE PCR

CHAPTER 4. CELL-BASED INTERNAL STANDARD FOR QPCR DETERMINATIONS OF ANTIBIOTIC RESISTANCE INDICATORS IN ENVIRONMENTAL WATER SAMPLES

CHAPTER 5. THIRD GENERATION CEPHALOSPORIN-RESISTANT *KLEBSIELLA PNEUMONIAE* THRIVING IN PATIENTS AND IN WASTEWATER: WHAT DO THEY HAVE IN COMMON?

CHAPTER 6. ENVIRONMENTAL AND CLINICAL *KLEBSIELLA PNEUMONIAE*: TWO SIDES OF THE SAME COIN

LIST OF TABLES

CHAPTER 1. INTRODUCTION

CHAPTER 3. MONITORING ANTIBIOTIC RESISTANCE GENES IN WASTEWATER: POTENTIALS AND LIMITATIONS OF QUANTITATIVE PCR

CHAPTER 4. CELL-BASED INTERNAL STANDARD FOR QPCR DETERMINATIONS OF ANTIBIOTIC RESISTANCE INDICATORS IN ENVIRONMENTAL WATER SAMPLES

CHAPTER 5. THIRD GENERATION CEPHALOSPORIN-RESISTANT *KLEBSIELLA PNEUMONIAE* THRIVING IN PATIENTS AND IN WASTEWATER: WHAT DO THEY HAVE IN COMMON?

CHAPTER 6. ENVIRONMENTAL AND CLINICAL *KLEBSIELLA PNEUMONIAE*: TWO SIDES OF THE SAME COIN

[1] These tables are provided in a link.

LIST OF ABBREVIATIONS

ANIb	Pairwise average nucleotide identity
ARB	Antibiotic resistant bacteria
ARGs	Antibiotic resistance genes
BLAST	Basic Local Alignment Search Tool
BLASTn	Basic Local Alignment Search Tool for Nucleotide sequences
BLASTp	Basic Local Alignment Search Tool for Protein sequences
CYP	Cyprus
ddPCR	Digital polymerase chain reaction
DNA	Deoxyribonucleic acid
ES	Spain
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp.
EU/EEA	European Union/ European Economic Area
G+C	Guanine and cytosine
GLASS	Global Antimicrobial Resistance and Use Surveillance System
ICE	Integrative and conjugative elements
MDR	Multidrug resistant

MicroBIGG-E	Pathogen Detection Microbial Browser for Identification of Genetic and Genomic Elements
MLST	Multi Locus Sequence Typing
NCBI	National Center for Biotechnology Information
PCR	Polymerase chain reaction
PFGE	Pulse field gel electrophoresis
PT	Portugal
qPCR	Quantitative polymerase chain reaction
rRNA	Ribosomal ribonucleic acid
RWW	Wastewater treatment plant influent
sTWW	Wastewater collected after secondary treatment
tTWW	Wastewater collected after tertiary treatment
US	United States of America
UV	Ultraviolet
WGS	Whole genome sequence
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1. Antibiotic resistance

Antibiotic resistance is the capacity that some bacteria have to survive and proliferate in the presence of antibiotics. This capacity is due to cell properties, specifically some cell structures and proteins, in particular enzymes. Some of these antibiotic resistance properties are encoded by the so-called antibiotic resistance genes that use different processes to avoid the action of antibiotics. Antibiotic resistance may be an intrinsic property of bacteria, being in that case common to most or all members of the same species. Intrinsic resistance may result from morphological or physiological properties, such as the absence of cell-wall, or to biochemical properties encoded by core genome genes, such as some types of beta-lactamase enzymes (common to all members of the same species) (Davies & Davies, 2010). Intrinsic antibiotic resistance contrasts with acquired antibiotic resistance because it is a property that emerges in a single bacterial cell, by mutation or by horizontal gene transfer (transformation, transduction, conjugation and vesiduction) (Figure 1) (Soler & Forterre, 2020).

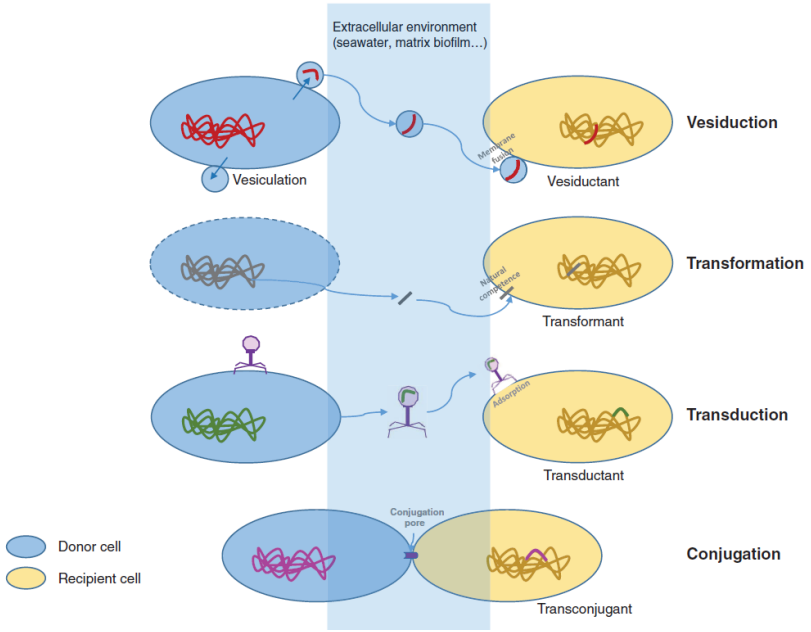


Figure 1 – Mechanisms of bacterial intercellular gene transfer, used also in the acquisition of antibiotic resistance (from Soler & Forterre, 2020).

Horizontal gene transfer has major relevance on the acquisition of antibiotic resistance, while gene mutation largely contributes to the diversity of those acquired genes (Álvarez-Martínez *et al.*, 2020; Woodford & Ellington, 2007). Some mobile genetic elements, particularly phages and plasmids, play a relevant role on the intercellular dissemination of antibiotic resistance, while others such as integrons and transposons are active on the intracellular mobilization of genes, for instance between plasmids and chromosomes (Partridge *et al.*, 2018). Both types of genetic element act synergically on the dissemination of resistance between cells and on the dynamics plasmid-chromosome. This dynamics is part of the capacity of bacteria to respond to environmental challenges and in ubiquitous bacteria, able to thrive in different types of environment, as a driver for gene acquisition, essential for adaptation. Antibiotics have been one of the drivers for the adaptation, emergence, and evolution of antibiotic resistant bacteria, mainly in clinical contexts, further disseminated in the environment (Andersson & Hughes, 2011; Andersson & Hughes, 2012).

Antibiotic resistance is a natural property of bacteria that became noticed mainly after the discovery and development of antibiotics as therapeutic agents to combat bacterial infections. The use of antibiotics in human medicine was a revolution that started in 1928 when Alexander Fleming discovered the penicillin produced by fungi of the species *Penicillium notatum* and insisted in its translation into a pharmaceutical product (Aminov, 2010) (Figure 2).

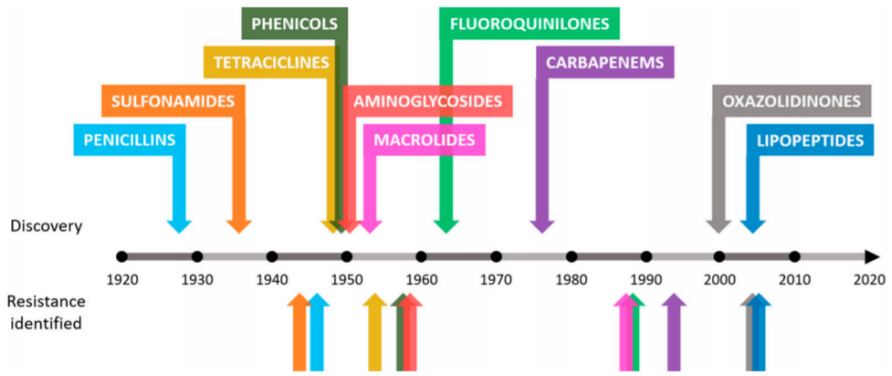


Figure 2 – Discovery of new classes of antibiotics and bacterial resistance identified (from Álvarez-Martinez *et al.*, 2020)

Nowadays, medicine uses antibiotics that are categorized into five major families, according to the chemical structure and mode of action: 1) inhibition of the cell wall synthesis (β-lactams), 2) disruption of the cell membrane (lipopeptides); 3) inhibition of protein synthesis (tetracyclines, aminoglycosides, macrolides, oxazolidinones, among others), 4) inhibition of

nucleic acids synthesis (fluoroquinolones), and 5) competitive inhibition of folic acid synthesis (sulfonamides) (Table 1) (Morar & Wright, 2010). Some of these are natural products of bacteria or fungi, while others are synthetic or semi-synthetic (Butler & Buss, 2006). Most antibiotics target the cell-wall, nucleic acids and protein synthesis, making use of structures that are specific of the bacterial cell (Butler & Buss, 2006). These targets justify the specificity of antibiotics against bacteria, while eukaryotic and archaea cells are not affected. Nonetheless, bacteria have developed mechanisms to overcome most of the antibacterial agents in use. Indeed, the increase of the rates of bacterial drug resistance has been faster than the development of new drugs, which compromises the success of antibiotic therapy leading to serious risks to human health.

Table 1 – Examples of major antibiotic classes, cellular action-targets and examples of resistance mechanisms.

Antibiotic class	Mechanism of action	Resistance mechanism
β-lactam	Cell wall synthesis inhibition	Enzymatic inactivation Efflux system Altered target
Sulfonamides	Interference with cell metabolism and growth arrest	Efflux system Altered target
Tetracyclines	Protein synthesis inhibition	Efflux systems Altered target
Aminoglycosides	Protein synthesis inhibition	Enzymatic inactivation Efflux system Altered target
Quinolones	Nucleic acids synthesis inhibition	Efflux systems Altered target
Lipopeptides	Cell membrane disruption	Altered target

Antibiotic resistance mechanisms can be categorized into the main groups, 1) prevention of access of the antibiotic to the target promoted by reduced membrane permeability or increased efflux, 2) antibiotic target modification and 3) direct modification of antibiotics through enzymatic inactivation (Blair *et al.*, 2015). Although organized in only these main mechanisms of resistance antibiotic resistance genes and variants present an impressive diversity and capacity to continuously evolve. To illustrate, the National Center for Biotechnology

Information (NCBI) Pathogen Detection Microbial Browser for Identification of Genetic and Genomic Elements (MicroBIGG-E) lists more than 5 million genes, gene variants, and housekeeping gene mutations (www.ncbi.nlm.nih.gov/pathogens/microbigge/#AMR, accessed in July, 20, 2021) related with antibiotic resistance in bacterial pathogens.

Besides the resistance mechanisms that each bacterial cell may acquire, there is also an important role for the interactions between species within a community. These mechanisms may offer a type of charity favoring the fitting of some individual cells even in the presence of antibiotics. Collective resistance refers to interactions within a community that favor the resistance to an antibiotic allowing the growth, even of susceptible populations, in its presence. This may be mediated by protection to the antibiotic exposure, for instance through drug inactivation by resistant community members, or through structural or physical protection (e.g. production of exopolysaccharides) (Meredith *et al.*, 2015; Vega & Gore, 2014). Many of these mechanisms are cell density-dependent providing protection to dense populations through the activation of quorum sensing and biofilm formation, and the inactivation of antibiotics (Evans *et al.*, 2018; Sorg *et al.*, 2016).

1.1. The environmental dimensions

Bacterial resistance to antibiotics may occur naturally in environmental bacteria. This has been documented in studies conducted in pristine areas, where human impacts are minimal or inexistent, such as Antarctic surface soils, and where antibiotic resistance genes have been detected (Van Goethem *et al.*, 2018). In the natural environment, antibiotic resistance genes might encode functions related with cell defense, biochemical signaling, or modulation of metabolic activity (Dantas *et al.,* 2008; Davies *et al.,* 2006; Martinez, 2009). The antibiotic resistance genes that nowadays spread among human pathogens and that are considered environmental contaminants, are believed to have origin in the natural resistome (i.e., the whole set of antibiotic resistance genes in a habitat) (Finley *et al.*, 2013). The arguments are that the intensive use of antibiotics in human and animal production and medicine and, in minor extent, in plants agriculture led to the spread of antibiotic residues, antibiotic resistant bacteria and antibiotic resistance genes through different compartments in the environment, such as water, soil or wildlife (Dantas *et al.*, 2008; D'Costa *et al.*, 2011; Rizzo *et al.*, 2013; Thaller *et al.*, 2010; Vredenburg *et al.*, 2014). In the environment, aquatic ecosystems offer suitable conditions for the spread of antibiotic resistance, forming a continuum between hospital effluents, urban wastewater and the receiving environments (Alves *et al.*, 2014; Varela *et al.*, 2014; Varela *et al.*, 2015a; Varela *et al.*, 2015b). For these reasons, the environmental dimensions of antibiotic resistance have gained increasing relevance, mainly when human

health is framed within the One-Health approach. Under this concept, antibiotic resistance and the human health risks must be approached assuming a continuum with the environmental and animal health and wellbeing (McEwen & Collignon, 2018). Indeed, human health is the major motivation to study and combat antibiotic resistance. For this reason the World Health Organization classified antimicrobial resistance as a major threat to public health in the twenty-first century, while recognized the high relevance of the natural environment as part of the solution and of the problem (WHO, 2015; WHO, 2020).

1.2. Wastewater: between humans and the environment

Antibiotic resistance is mainly disseminated by bacteria of human and animal origin, and any niche where these bacteria can proliferate are potential reservoirs. These comprise above all, the human and animal bodies, mainly the digestive tract, and more intensely when under antibiotic use pressure. Therefore, hospitals and health care facilities and animal production farms may represent important reservoirs (Berendonk *et al.*, 2015; Larsson *et al.*, 2018). While those reservoirs may be pivotal for the first emissions of antibiotic resistant bacteria and antibiotic resistance genes to the environment, there are other paths of dissemination that gained increasing relevance. One of these refers to the urban wastewater treatment plants that receive, treat, and deliver to the natural environment, the sewage produced by the served community. This is one of the parts of the urban water cycle, where sewage treatment represents a major barrier for the spread of chemical, physical and biological contaminants, assuring environment and human health protection. The other part of the urban water cycle comprises the production and distribution of water for human consumption, in a process that may include disinfection (Manaia *et al.*, 2016). This “clean” part of the urban water cycle may be threatened by the adverse impacts of the discharge of untreated or deficiently treated wastewater. Therefore, sewage treatment represents one of the critical control points where antibiotic resistance may be controlled (Manaia *et al.,* 2016; Bürgmann *et al*., 2018). Urban wastewater treatment plants receive high loads of nutrients, of human-commensal bacteria and of chemical contaminants, whose mixture may facilitate the survival or proliferation of antibiotic resistant bacteria (Qiu *et al.*, 2012; Paul *et al.*, 2019). These contaminants together with the high loads of antibiotics cause a suitable environment for antibiotic resistance survival (Singer *et al.*, 2016). While in some countries antibiotic consumption has been decreasing, in others it is still increasing (Klein *et al*., 2018). Overall, high amounts are still consumed, and it is estimated that 30 to 90% of the ingested antibiotics by humans and animals are excreted untransformed in urine and feces (Du & Liu, 2012). Therefore, the widely and, eventually, excessive consumption of antibiotics, the incomplete metabolism of drugs and its partial removal in urban wastewater treatment plants, are resulting in the ubiquitous occurrence of

antibiotic resistant bacteria and antibiotic resistance genes in aquatic environments (Alexander *et al.*, 2015; Cacace *et al.*, 2019).

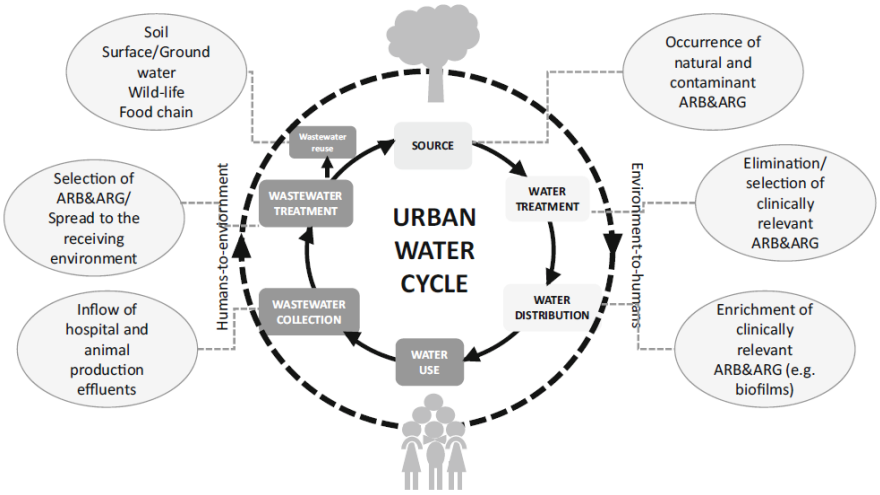


Figure 3 – Antibiotic resistance dissemination in different compartments of the urban water cycle (from Manaia *et al.*, 2016).

Urban wastewater treatment plants are important recipients and reservoirs of antibiotic resistant bacteria and antibiotic resistance genes in the environment (Figure 3). These receive mostly wastewater from domestic uses and pre-treated industrial and hospital effluents that may contain antibiotic residues, antibiotic resistant bacteria and antibiotic resistance genes, which will be subjected to different treatments and finally discharged into the environment, in aquatic systems (Manaia *et al.*, 2016; Rizzo *et al.*, 2013;). The wastewater environment is known to be a complex environment to bacteria where they will be mixed with detergents, metals, pharmaceutical compounds, chemicals and other products from anthropogenic activity that might create selective pressure and promote the proliferation and the spread of antibiotic resistance among the bacterial communities present (Rizzo *et al.*, 2013; Zhang *et al.*, 2012). The treatment processes used in urban wastewater treatment plants were designed to remove nutrients and pathogens, and although these objectives are attained, antibiotic resistant bacteria and antibiotic resistance genes are still emitted to the receiving environment (Pallares-Vega *et al.*, 2019; Pärnänen *et al.*, 2019; Rizzo *et al.*, 2013). Treated wastewater that underwent conventional treatment, consisting of primary and activated sludge secondary processes, is reported to contain loads of culturable bacteria reaching 10^5–10^7 colony-forming units/100 mL (Narciso-da-Rocha *et al.*, 2018; Novo & Manaia, 2010; Novo *et al.*, 2013; Rizzo *et al.*, 2013). Raw wastewater has been reported to contain predominantly bacteria of the phyla

Proteobacteria, *Actinobacteria* and *Firmicutes*, in which are included bacterial families reported as potential antibiotic resistance carriers and clinically-relevant groups, such as *Enterobacteriaceae*, *Enterococcaceae*, *Pseudomonadaceae*, among others (Alexander *et al*., 2016; Manaia, 2017; Mckinney & Pruden, 2012; Narciso-da-Rocha & Manaia, 2017; Sousa *et al*., 2017; Varela *et al*., 2015a, Varela *et al*., 2015b). Most of the wastewater bacteria are non-culturable, although they might harbor antibiotic resistance genes. Therefore, the impacts of wastewater discharges in receiving environments cannot be properly assessed based only on culturable bacteria. In a review comparing the final effluent of urban wastewater treatment plants in China, Estonia, Finland, Portugal, Spain and United States of America, Manaia *et al*. (2016) reported that 10^{14}-10^{18} copies of genes encoding resistance to tetracycline or β-lactams can be discharged per day into the environment. These values have been consistently confirmed in other studies (Pallares-Vega *et al*., 2019; Pärnänen *et al*., 2019). Metagenomics studies have unveiled an impressive diversity of antibiotic resistance genes in wastewater, with 100 to > 400 types being detected in the influent and 1-5 five times less types in the final effluent (An *et al*., 2018; Lira *et al*., 2020; Ng *et al*., 2019; Quintela-Baluja *et al*., 2019). Despite these reductions, these studies have reported also that some genes were enriched during treatment. Despite the high potential of metagenomic analysis to explore the wastewater resistome, this approach provides results whose interpretation in terms of treatment efficiency and risk assessment may be biased and complex. For this reason, some authors have preferred to assess treatment efficiency and impacts of wastewater treatment plants based on the use of targeted analysis of quantitative polymerase chain reaction (qPCR) (Cacace *et al*., 2019; Storteboom *et al*., 2010). This approach requires the selection of suitable markers, normally genes that have been reported in clinical isolates or with a wide distribution such as bla_{CTX}, bla_{KPC}, bla_{NDM}, bla_{IMP}, bla_{VIM} involved in β-lactams resistance, *vanA* involved in vancomycin resistance, *sul1* and *sul2* involved in sulfonamides resistance and *tet* genes, encoding resistance to tetracycline (Du *et al*., 2014; Narciso-da-Rocha *et al*., 2014; Rodriguez-Mozaz *et al*., 2015; Szczepanowski *et al*., 2009). Quantitative PCR, through the intra-polation to a calibration curve, allows the estimation of gene abundance per volume of sample, which is adequate to assess treatment efficiency and to risk assessment. While these determinations may be influenced by experimental conditions, the possibility of comparing data obtained in different laboratories has been regarded as an urgently need to enable antibiotic resistance monitoring (Berendonk *et al*., 2015; Manaia *et al*., 2016).

Water scarcity and water pollution have led to consider the use of treated wastewater an option, which is becoming mandatory in some situations. Around 1.2×10^9 people live in areas affected by serious water scarcity conditions (United Nations, 2014) and 1.8×10^9 are expected to be living in countries or regions affected by water scarcity by 2025 (FAO, 2012; United Nations, 2014). Wastewater reuse may be applied mainly for irrigation in agriculture and

landscaping (Drechsel *et al*., 2010). This practice is already in use in countries, such as France, Italy, Spain, Cyprus, Israel, USA, among others (EPA, 2012; Kalavrouziotis *et al*., 2015; Pedrero *et al*., 2010). However, such a practice may increase the risks of transmission of antibiotic resistant bacteria and antibiotic resistance genes from urban wastewater treatment plants to the environment and finally back to humans, via the food chain (Fuhrimann *et al*., 2016). At European level new guidelines for minimum requirements for water reuse have been recently announced (Alcalde-Sanz & Gawlik, 2017). It considers an integrated water management, with special focus on agricultural irrigation, and recommends that treated wastewater should fulfil minimum requirements regarding four parameters (*Escherichia coli*, biochemical oxygen demand, total suspended solids, turbidity), establishing different water quality requirements according to crop categories and irrigation mode. However, the release of antibiotics, antibiotic resistant bacteria or antibiotic resistance genes into the environment is not regulated so far (Rizzo *et al*., 2020). It is important to consider that even if the wastewater is not reused for irrigation, treated wastewater will end up in the environment in aquatic systems, used or connected to irrigation, leisure, or other ends, where may represent a risk to human health. Antibiotic resistant bacteria and their antibiotic resistance genes are nowadays recognized environmental contaminants of emerging concern that can pose serious risks to human health as they reduce antibiotic therapy efficacy (Pruden *et al*., 2006). Therefore, it is important to develop treatment methodologies that could have more efficacy on wastewater treatment and on the removal of antibiotic resistant bacteria and antibiotic resistance genes. High quality treated wastewater is crucial for ensuring adequate protection of the environment and human health.

1.2.1. Wastewater treatment

Sewage collected and pipped into a wastewater treatment plant (influent, raw wastewater) is subjected to different processes, which main objective is to remove debris, high organic loads and pathogens from sewage (Manaia *et al*., 2018).

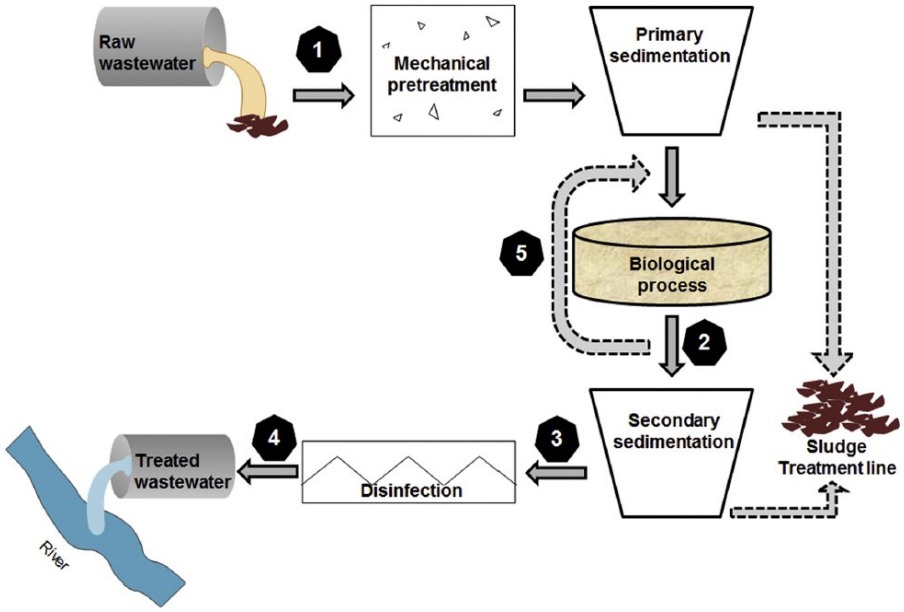


Figure 4 – Schematic presentation of the treatments applied in an urban wastewater treatment plant (from Manaia *et al*., 2018).

Most of the urban wastewater treatment plants apply a pre-treatment followed by primary treatment and a secondary biological process. The pre-treatment may be a mechanical process aiming to remove solids that might damage the downstream equipment. The influent may then undergo primary treatment aiming at settling the sedimentable solids, being ready for entering the secondary treatment area where biodegradable compounds are removed and microorganisms can be outcompeted or aggregated in flocs that will sediment (Figure 4) (Grady *et al*., 2011; Henze *et al*., 2008). This process is rather complex, not fully understood from the microbiological point of view, and reported as being subjected to a high degree of stochasticity (Bengtsson-Palme *et al*., 2016; Lira *et al*., 2020). Despite these fluctuations, it is generally estimated that a secondary effluent can have a load of up to 10^{12} antibiotic resistant bacteria per day or 10^{18} antibiotic resistance genes per day (Manaia *et al*., 2016; Vaz-Moreira *et al*., 2014). To further polish these effluents, some urban wastewater treatment plants also apply other processes that comprise tertiary treatment steps, and where disinfection may be considered an added value to remove microorganisms (Rizzo *et al*., 2020). Other aims of the tertiary treatment include increasing the removal of nutrients (nitrogen/phosphorus) and chemical micropollutants from secondary treatment effluent, increasing the final effluent water quality (Henze *et al*., 2008).

The most used disinfection method is chlorination, however ultraviolet (UV) radiation and ozonation are also employed. Some studies have compared the effect of these treatment processes on the removal of antibiotic resistant bacteria and antibiotic resistance genes. While some studies report that chlorination has shown to be more effective than UV or ozonation (Zhuang *et al*., 2015) and the sequential use of UV and ozonation revealed to be more effective on antibiotic resistance genes removal than UV irradiation only (Sharma *et al*., 2016), others report that chlorination did not prove to have significant contribution to ARGs removal (Di Cesare *et al*., 2016; Gao *et al*., 2012; Munir *et al*., 2011). Therefore, it seems that the efficiency of a treatment may be extremely dependent on the operation conditions employed. For example, the effect of chlorination on antibiotic resistance genes proved to be dependent on the dosage of chlorine and contact time, achieving a maximum inactivation of 1.30-1.49 log-units at 30 mg/min.L of chlorine (Sharma *et al*., 2016). Moreover, all bacteria, except sulfadiazine- and erythromycin-resistant bacteria were inactivated to levels below the detection limit by using 15 mg/min.L of chlorine (Yuan *et al*., 2015). Other treatment processes have been investigated or developed, although are only at the pilot- or bench-scale level, such as advanced oxidation processes based on the formation of hydroxyl radicals, nanomaterials, membrane filtration and coagulation, among others (Barancheshme & Munir, 2018). The development and implementation of treatment processes effective on the removal of these contaminants is crucial to avoid environmental pollution with antibiotic resistant bacteria and antibiotic resistance genes (Rizzo *et al*., 2020).

1.2.2. Assessment of wastewater treatment efficiency on antibiotic resistance reduction

Wastewater treatment efficiency is traditionally assessed based on some standardized physicochemical and microbiological parameters (Council Directive 91/271/EEC; Manaia *et al*., 2018). The microbiological parameters rely on the measurement of bacteria of enteric origin, such as enterococci, and mainly total and faecal coliforms and *Escherichia coli* (ISO 7899; ISO 9308). These or any other microbiological parameters are not indicated in the European legislation that defines minimal quality requirements of wastewater (Council Directive 91/271/EEC), although *Escherichia coli* is one of the quality and safety parameters recommended for wastewater reuse for irrigation (The European Parliament and the Council Regulation 2020/741). However, even in the water reuse minimum requirements, antibiotic resistance is not considered. In general, the occurrence of antibiotic resistant bacteria and antibiotic resistance genes in wastewater, although recognized worldwide, is currently not considered in any legislative framework. In part this is due to the difficulties in establishing

minimum requirements, justifying that further research efforts on the evaluation of advanced treatment processes and on monitoring methods are needed. Regular and harmonized monitoring will generate a body of information that can support the establishment of legislative policies. Different methodological approaches have been reported in the literature (Table 2) (Manaia *et al*., 2018).

Table 2 – Brief summary of strengths and weaknesses of culture-dependent and -independent methodologies used to assess antibiotic resistance in wastewater environments (adapted from Manaia *et al*., 2018).

	Culture-dependent methods	Culture-independent methods
Strengths	Allow phenotype identification Permit enumeration of viable cells Permit the identification and monitoring of clinically-relevant species Standardized methods Low cost	High specificity and sensitivity Targeted analyses Allow the determination of abundance and prevalence of specific ARGs Permit the simultaneous quantitative analysis of multiple ARGs and housekeeping genes in a single sample
Weaknesses	Laborious Time consuming Cultures are not representative of the whole community diversity Most bacteria are not culturable	DNA extraction may represent an important bias on the quantification The analysis is independent from the phenotype of the host The presence of PCR inhibitors in complex matrices may compromise the accuracy of ARGs quantification High cost No standardized methods

Culture-dependent methods rely on the cultivation of bacteria of interest to evaluate their growth, metabolism and phenotype, while the culture-independent methods are based on the analysis of nucleic acids (Hiller *et al*., 2019; Manaia *et al*., 2018; Rizzo *et al*., 2013). Most regulatory framework that include microbiological analysis establish culturing and numbers of culturable microorganisms as safety and quality criteria. Besides the major advantages that are the existence of standardized methods, the simplified translation of the results into

predefined quality criteria and the existence of data archived for decades, there are other potentials. These include the possibility of phenotypic characterization of isolates, the feasibility of implementation even with low technical resources, and potential for global data comparability. However, a major limitation refers to the fact that non-culturable populations, regardless their viability, are not considered when culture methods are used (Hiller *et al*., 2019; Manaia *et al*., 2018; Rizzo *et al*., 2013). Depending on the type of environment, the fraction of culturable bacteria may range from less than 0.1% to 10% (Vartoukian *et al*., 2010; Vaz-Moreira *et al*., 2013). This scenario created increasing interest on culture-independent methods that rely on the direct analysis of DNA, regardless the viability of the cells, growth requirements or cell injuries (Kim *et al*., 2013; Manaia *et al*., 2018). Culture-independent methods offer a comprehensive overview of the diversity and abundance of genes in a sample. The methods mostly used to investigate and monitor antibiotic resistance are quantitative PCR (qPCR), that enables the quantification of selected genes, and metagenomics that permit the search for a broad range of genes in samples with complex microbial communities (Aarestrup & Woolhouse, 2020; Hendriksen *et al*., 2019; Hutinel *et al*., 2019; Oulas *et al*., 2015; Pärnänen *et al*., 2019; Riquelme *et al*., 2021). The selection of the best method depends on the study aims. While metagenomics approaches are promising for non-target ARG profiling, the overall advantage of qPCR is the higher sensitivity and precision allowing the quantification of target genes even when these genes are at low-abundance values (Manaia *et al*., 2018). Therefore, one of the gold-standard methods currently used to quantify ARGs is the qPCR which is a targeted and accurate method that permits the detection and quantification of selected genes. This method has been used to quantify genes in clinical samples, groundwater, wastewater, manure, and soil (Böckelmann *et al*., 2009; Kim *et al*., 2013; Liotti *et al*., 2019; McKinney *et al*., 2018) presenting a high accuracy and specificity. Nevertheless, qPCR as well as the culture-dependent methods, is not exempt from biases. Quantitative PCR may be limited by the presence of detergents, humic acids and other compounds that might inhibit the qPCR reaction (Sidstedt *et al*., 2015; Smith & Osborn, 2009). Moreover, qPCR targets a limited number of genes and is dependent on qPCR primers design/specificity (Manaia *et al*., 2018). Another disadvantage might be the inability to directly discriminate between extracellular and intracellular DNA, although strategies have been used to try to solve this problem such as the use of propidium monoazide (Cangelosi & Meschke, 2014; Nocker *et al*., 2007).

1.2.3. Antibiotic resistance monitoring in wastewater – potentials and limitations

Antibiotic resistance is present in large amounts in domestic effluents, persist wastewater treatment and, consequently, urban wastewater treatment plants discharges may have adverse impact on the receiving environment (Alexander *et al*., 2015; Cacace *et al*., 2019). This fact has raised the interest on implementing feasible and reliable methods to monitor antibiotic resistant bacteria or antibiotic resistance genes in wastewaters. Monitoring antibiotic resistance in wastewater can have three main aims – i) assess the quality of the effluent and estimate the potential impacts in the receiving environment (e.g. Cacace *et al*., 2019); ii) assess the quality and efficacy of the treatment process (e.g. Pärnänen *et al*., 2019); iii) survey the profiles and abundance of antibiotic resistance in a population through the inspection of the respective sewage (Aarestrup & Woolhouse, 2020). The latter is out of the scope of the monitoring aims herein approached and will not be further discussed.

While in culture-dependent methods the protocols are usually established and standardised, they have limitations to stand alone to attain the aims mentioned above. However, before culture-independent methods can be implemented, some aspects need to be further explored, as it is recognized that the existing methods are not harmonized, and it is unknown if the results obtained worldwide are comparable (Berendonk *et al*., 2015; Manaia *et al*., 2016). For example, for culture-dependent methods, the procedures for determining antibiotic susceptibility of a given bacterial strain are well defined in the Clinical and Laboratory Standards Institute or European Committee on Antimicrobial Susceptibility Testing guidelines (CLSI, 2016; EUCAST, 2021). Although these guidelines are designed for the evaluation of resistance of clinical strains, they have been adapted to assess antimicrobial resistance of isolates obtained from environmental sources (Manaia *et al*., 2018). The harmonization of protocols and the definition of universal guidelines based on the use of culture-independent methods present different challenges. Determinations can be influenced by numerous variables, some of which may be difficult to control, making it challenging to compare results between different laboratories and regions of the world (Manaia *et al*., 2016; Manaia *et al*., 2018). Technical decisions such as the selection of the membranes used to concentrate the water biomass for further DNA extraction and the choice of the DNA extraction kit, may influence the final result. Also, the results may be affected by factors such as the primers, qPCR mixtures, qPCR protocol or equipment and qPCR standards, or the sequencing technique and depth, read length, assembly quality and reference database and criteria for annotation (Manaia *et al*., 2018). Another question refers to the sensitivity of the methods to use, since advanced wastewater treatment may lower antibiotic resistant bacteria and resistance genes to levels below the limits of detection (Becerra-Castro *et al*., 2016; Fortunato *et al*., 2018; Sousa *et al*., 2017). These are aspects that have been discussed and identified

as critical to advance the control of antibiotic resistance in wastewater environments (Di Cesare *et al*., 2020; Manaia *et al.,* 2018).

1.2.4. Wastewater as a bridge from clinics to the environment

In urban areas, domestic sewage represents the major human emission of pathogens and antibiotic resistant bacteria. Although wastewater treatment has a pivotal role on the removal of such microorganisms from sewage, an important fraction can survive, being discharged into the environment. The occurrence of pathogenic organisms in treated wastewater is mainly explained by human emissions and by the capacity of members of these species to endure treatment processes. The acronym ESKAPE has been used to list six groups of opportunistic pathogens that are leading causes of drug-resistant infections in hospitals, specifically *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp. (Boucher *et al*., 2009). Considering characteristics of these bacteria such as ubiquity, high incidence of clinical situations and severity of the infections, it is arguable that members of this list should be considered as priority targets in wastewater monitoring. In fact, all these organisms have been reported both in raw and treated wastewater (Akya *et al*., 2020; Higgins *et al*., 2018; Limayem *et al*., 2019; Mbanga *et al*., 2021; Surleac *et al*., 2020). Pathogens causing infections or thriving in transition environments, such as wastewater, irrigation or leisure waters, represent a human health threat. However, it is also possible that some of these bacteria cannot survive wastewater treatment and environmental challenges, or if they do, may tend to lose part of acquired traits. The effect of distinctive selective processes due to the stress imposed by the external conditions on the retention or loss of specific traits is still a major question. The evaluation of the stability and maintenance of virulence and resistance features in wastewater and in the environment is a suitable approach to study the effect that these transition environments might cause phenotypically and genotypically in bacteria (Biswal *et al*., 2014). Since these organisms can be present in humans and in distinct environments, the study about the alterations that transitions environments, such as wastewater, might promote in these organisms regarding antibiotic resistance and virulence might be of importance to evaluate the possible potential impact of these organisms on human health when discharged into the environment.

2. Relevance in clinics and in the environment of *Klebsiella pneumoniae*

The family *Enterobacteriaceae* is integrated in the class *Gammaproteobacteria* and includes genera and species with ubiquitous distribution in soil, water, plants, animals and humans (Brenner & Farmer, 2015). Members of this family stain Gram-negative, are non-spore forming rods, and are facultative anaerobes with chemoorganotrophic fermentative and respiratory metabolism (Brenner & Farmer, 2015). As the name hints, some members of the family *Enterobacteriaceae* are common enteric inhabitants of vertebrates, being *Escherichia coli* the most well-known species, also used as microbiological indicator of fecal contamination (ISO 9308). This family also comprises well-known pathogens, such as *Shigella dysenteriae*, some hemorrhagic lineages of *E. coli,* and *Klebsiella pneumoniae* (Berman, 2012). Ubiquity, virulence, and acquired antibiotic resistance, place some members of the family *Enterobacteriaceae* among the leading group of clinically-relevant bacteria with wide environmental distribution (Iredell *et al*., 2016; WHO, 2014).

K. pneumoniae is one of the species of the ESKAPE group that has called the attention of public health entities such as the World Health Organization, the Centers for Disease Control and Prevention (US) or the European Centre for Disease Prevention and Control, as a significant threat to global public health due to its high rates of antimicrobial resistance (ECDC, 2019; WHO, 2015). This species of the family *Enterobacteriaceae* has been reported in a wide range of sources, not only in association with humans and animals, but also in plants, soils, water, and wastewater (Navon-Venezia *et al*., 2017; Wyres & Holt, 2018). In humans, it can colonize and infect the respiratory, gastrointestinal and urinary tracts and the skin, being an opportunistic pathogen, mostly associated with urinary infections, pneumonia and wound infections (Navon-Venezia *et al*., 2017; Wyres & Holt, 2018). The threat posed by members of this species is illustrated by the fact that this species is one of the major causes of the bacterial infections caused by all Gram-negative bacteria (Navon-Venezia *et al*., 2017). The examination of 41 814 isolates of *Klebsiella pneumoniae* collected over the year of 2019 in 30 EU/EEA countries, demonstrated that more than a third (36.6%) of the isolates were resistant to at least one of the antimicrobial groups under surveillance (i.e., fluoroquinolones, 3rd generation cephalosporins, aminoglycosides and carbapenems) (ECDC, 2019). The highest EU/EEA population-weighted mean resistance percentage was reported for 3rd generation cephalosporins (31.3%), followed by fluoroquinolones (31.2%), aminoglycosides (22.3%) and carbapenems (7.9%). Single resistance was less commonly reported than resistance to two or more antimicrobial groups, with the most common resistance phenotype being the combined resistance to fluoroquinolones, 3rd generation cephalosporins and aminoglycosides (ECDC, 2019). Indeed, 3rd generation cephalosporin-resistant and carbapenem-resistant

Enterobacteriaceae species, including *Klebsiella pneumoniae*, are part of the WHO list of priority pathogens for research and development of new antibiotics, being classified in the critical level (Figure 5) (WHO, 2015).

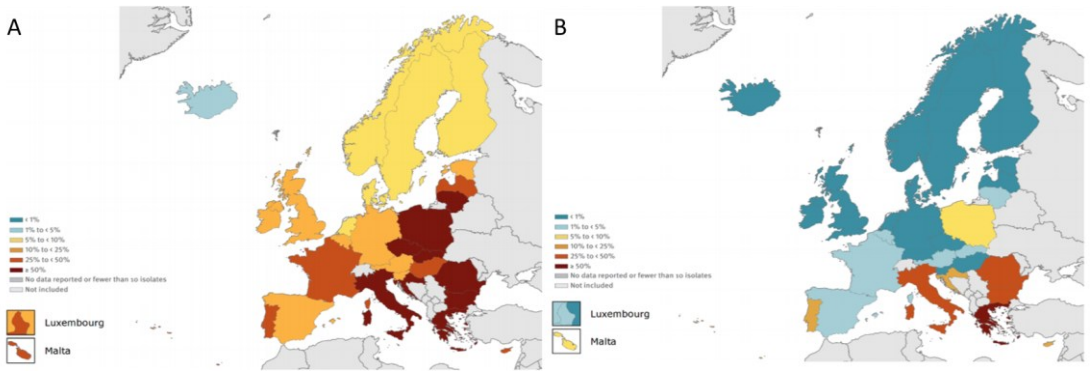


Figure 5 - Percentage of invasive *Klebsiella pneumoniae* isolates resistant to 3rd generation cephalosporins (A) and carbapenems (B) by country, EU/EEA, 2019.

The clinical relevance of this species is also related to its high genome plasticity, compared to other Gram-negative organisms (Wyres & Holt, 2018). In members of this species, plasmids and plasmid-related genes are frequently observed as extra- or intra-chromosome elements, hinted by genome regions with high variability in the guanine and cytosine (G+C) content, reflecting the intense horizontal gene transfer activity (Wyres & Holt, 2018). Paczosa and Mecsas (2016) described this species as “going in the offense with a strong defence” due to the wide range of clinically-relevant traits that can be observed in members of this bacterial group. Besides having clear advantages regarding antimicrobial resistance, this species has also a high infection capacity. This is promoted by traits related with virulence such as capsules, lipopolysaccharides, fimbriae but also traits related to metals and antibiotic resistance which confer additional advantage in aggressive environments (Paczosa & Mecsas, 2016). These features make of *Klebsiella pneumoniae* an interesting model organism to assess the ecology, persistence and stability of bacteria with acquired antibiotic resistance that are able to thrive in the environment. Such kind of study is increasingly feasible thanks to the possibility to explore and compare bacterial genomes.

3. Genomics and the study of antibiotic resistance ecology

Genomics is an area of knowledge and application that is developing fast, with crucial contributions from technology and bioinformatics, and pushed and pulled by numerous areas of intervention, where microbiology is only a tiny fraction (Figure 6). Therefore, numerous advances in microbial genomics are paced with other synchronous developments. Over the last decades, new DNA sequencing technologies, with decreasing costs have emerged (e.g., Roche 454, Illumina, Pacific Biosciences, Oxford Nanopore technology) (Gupta & Verma, 2019). While short read sequencing technologies such as Illumina have the limitation of missing and failing the highly frequent repetitive regions of bacterial genomes, these limitations have been compensated by the use of long read technologies such as Pacific Biosciences or Oxford Nanopore technologies (Gupta & Verma, 2019; Segerman, 2020).

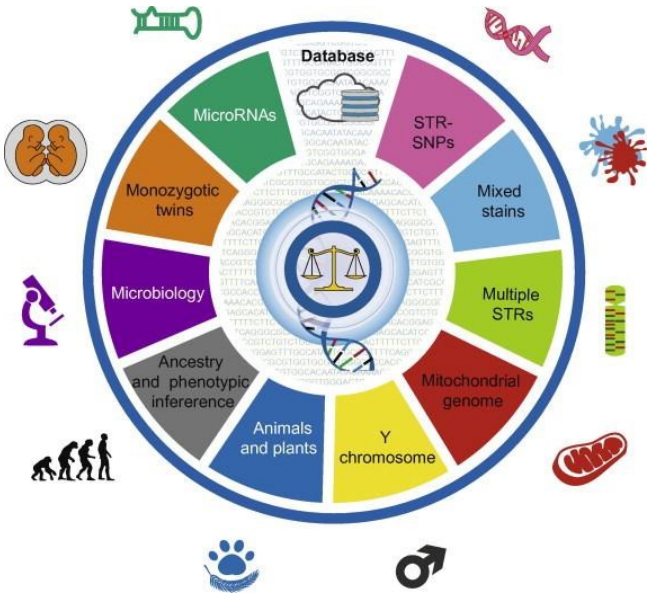


Figure 6 – Areas of intervention of genomics (from Yang *et al*., 2014).

Besides metagenomic approaches, briefly addressed above, next generation sequencing also leveraged whole-genome sequence (WGS) analyses, permitting an improved and advanced knowledge of functional capacities, evolution and epidemiology of microorganisms (Gerner-Smidt *et al*., 2019; Quainoo *et al*., 2017; Runcharoen *et al*., 2017). Either based on the examination of genomes of microbial isolates or genomes assembled from metagenomes, microbial genomics is now a still under exploitation field that promises to bring important

insights into the ecology and evolution of acquired antibiotic resistance (Baquero, 2012; Murray *et al*., 2020). One of the most important characteristics of genome sequence analysis is the relatively high reproducibility and the portability, which permits the public access to open data obtained worldwide. Public databases providing the whole genome sequence of pathogens and antibiotic resistant bacteria as well as filters and tools for preliminary analysis are nowadays essential for researchers and health professionals (Table 3).

Besides some fundamental research aspects, also practical implications are increasingly arising from the application of genomics studies. For instance, the Global Antimicrobial Resistance and Use Surveillance System (GLASS) promoted by WHO is exploring the use of whole-genome-sequencing for surveillance of antimicrobial resistance (WHO, 2020). According to the GLASS report (WHO, 2020), the use of WGS for global surveillance will promote the acquisition of information about the emergence and spread of antibiotic resistance and will help on the development of diagnostic tools to characterize antibiotic resistance in a shorter time, complementing phenotypic characterization (WHO, 2020). Research has demonstrated the usefulness of genomics to infer about diseases caused by transmission between humans and food, as part of an one-health approach (Gerner-Smidt *et al*., 2019), to track disease outbreaks (Quainoo *et al*., 2017), to track antibiotic resistant bacteria circulation in a specific environment (van Dorp *et al.,* 2019), to determine the location or antibiotic resistance determinants on core or accessory genome of an organism (Gorrie *et al*., 2018), to determine the geographic spread of a resistant strain to trace transmission networks (Nakano *et al*., 2018), among others. Genomics also supports comparative studies based on pangenome (core and accessory genes) analysis, providing information to estimate the genomic diversity of a dataset and bacterial evolution insights (Vernikos *et al*., 2015). Pangenome analysis revealed to be useful to provide insights into the potential ecological role and virulence (Wu *et al*., 2018; Guevarra *et al*., 2021), to complement culturomics in order to classify emerging bacteria or re-classify described ones (Caputo *et al*., 2019), as a tool for analysing pathogenic bacteria (Rouli *et al*., 2015) or to understand populations genomics and evolutionary events (Delgado-Blas, *et al*., 2021) across different bacterial species. These strategies and the availability of whole genome sequence data obtained worldwide will be crucial to monitor pathogen populations, to identify and track high-risk clones at the international level and to timely implement contention strategies. Moreover, understanding the mechanisms of disease, transmission and evolution, will help to advance in knowledge which is important to improve infection treatments globally.

Table 3 – Examples of public databases that provide, interpret and/or offer analysis tools to explore bacterial genomes.

Acronym	Description	Website
PATRIC	Pathosystems Resource Integration Center - provides integrated data and analysis tools to support biomedical research on bacterial infectious diseases.	https://www.patricbrc.org/
NCBI Pathogen Detection	NCBI Pathogen Detection - provides bacterial pathogen genomic sequences originating in environmental sources and clinical sources and identifies related sequences to uncover potential contamination sources.	https://www.ncbi.nlm.nih.gov/pathogens/
JSpeciesWS	JSpeciesWS - measures the probability of genomes belonging to the same species or not based on their complete or draft nucleotide sequence and give indication of genome size, G+C content, among others.	http://jspecies.ribohost.com/jspeciesws/
BIGSdb-Pasteur	BIGSdb-Pasteur – provides a collection of curated, open or private databases of genome sequences and genotypes based on multilocus sequence typing (MLST), whole genome based typing and supplementary schemes (in particular, antimicrobial resistance or virulence genes).	https://bigsdb.pasteur.fr/
CARD	Comprehensive Antibiotic Resistance Database – provides a database of resistance genes, their products and associated phenotypes.	https://card.mcmaster.ca/
AMRFinderPlus	AMRFinderPlus (NCBI) - identifies AMR genes, resistance-associated point mutations, and select other classes of genes using protein annotations and/or assembled nucleotide sequence.	https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/
VFDB	Virulence Factor Database – provides curated information about virulence factors of bacterial pathogens.	http://www.mgc.ac.cn/VFs/main.htm

Table 3 (cont.) - Examples of public databases that provide, interpret and/or offer analysis tools to explore bacterial genomes.

Acronym	Description	Website
CGE	Center for Genomic Epidemiology - provides access to bioinformatics resources such as: ResFinder (identification of acquired antibiotic resistance genes); PathogenFinder (prediction of bacteria's pathogenicity towards human hosts); VirulenceFinder (identification of acquired virulence genes); MLST (Multi Locus Sequence Typing (MLST) from an assembled genome or from a set of reads); PlasmidFinder (PlasmidFinder identifies plasmids in total or partial sequenced isolates of bacteria); SpeciesFinder (Prediction of bacterial species using the S16 ribosomal DNA sequence); Among others.	http://www.genomicepidemiology.org/
ISfinder	ISfinder - provides a basic framework for nomenclature and insertion sequences classification into related groups or families.	https://www-is.biotoul.fr/index.php
RAST	RAST – Rapid Annotation using Subsystem Technology - fully-automated service for annotating complete or nearly complete bacterial and archaeal genomes.	https://rast.nmpdr.org/rast.cgi

3.1. *Klebsiella pneumoniae* under the genomics lens

Specifically in the genus *Klebsiella* the use of genomics has contributed to elucidate biochemical and physiological properties of isolates recovered in clinical and research laboratories, and to assess relationships among closely related species that share 95-96% average nucleotide identity with *K. pneumoniae* (Holt *et al*., 2015; Long *et al*., 2017; Rodrigues *et al*., 2018). These approaches also unveiled the features associated with the core and accessory genome parts, highlighting clinically-relevant traits that might be associated to some critical features of the most spread or invasive *K. pneumoniae* members (Wyres *et al*., 2020). These approaches have also contributed to better understand aspects such as plasmid maintenance and dynamics in the hospital settings (Conlan *et al*., 2016; Gorrie *et al*., 2018; Martin *et al*., 2018). This kind of studies have evidenced how important these approaches can be to understand population structure, evolution and antibiotic resistance transmission. The same approaches promise to bring new insight into the dynamics of *K. pneumoniae* between humans, infection episodes and environmental lifestyles (Wyres *et al*., 2020). These studies will help to understand *K. pneumoniae* genomic epidemiology and evolution and therefore further and deeper studies are needed. However, whole genome sequence and genomics analyses have mostly relied on the study of clinical bacteria and clinically-relevant features. Some of the above listed databases (Table 3) are of this an example, being limited the number of genomes available of environmental origin and scarce and of poor quality the information provided. This upgrade is urgently needed to have an improved and more robust perspective of the relationship between the clinical and the environmental resistome.

4. Hypothesis and objectives of the thesis

Advances on the current knowledge of the wastewater resistome are essential to support the development of adequate legal frameworks and implementation of control measures as well as to unveil the human-health risks associated with the discharge of antibiotic resistant bacteria and antibiotic resistance genes in the environment.

The three hypothesis that oriented this thesis were:

1. That qPCR can be used in inter-laboratory assays for the quantification of antibiotic resistance genes in different types of wastewater to assess treatment efficiency and potential environmental impacts;

2. That DNA losses during extraction from wastewater and water samples could be controlled based on the use of a cell-based internal standard;

3. That clinical isolates of *K. pneumoniae,* used as a model bacterial species, may be lost or lose some antibiotic resistance and fitness features, once thriving in the environment.

The arguments to explore these three topics were:

1. The regular and integrated monitoring of antibiotic resistance in wastewater has been claimed as a priority on the combat of antibiotic resistance, being integrated surveillance one of the objectives of the World Health Organization global action plan on antimicrobial resistance (WHO, 2015). Moreover, numerous groups around the world have been developing improved and innovative wastewater treatment systems to reduce the resistance load of wastewater (Rizzo *et al*., 2020), being necessary to provide harmonized and comparable analytical methods that permit reliable comparisons.

2. Numerous studies have demonstrated that DNA extraction may be critical for the analysis of the resistome in environmental samples (Djurhuus *et al*., 2017; Hinlo *et al*., 2017; Li *et al*., 2018). A critical step may be related with cells – e.g., loss during filtration, failure to lyse, trapping in the sample debris and protein precipitates - and later on due to losses in the DNA target, mainly when qPCR is used. Therefore, the development of an internal standard that could be used in different laboratories, mainly in those with limited experience in DNA extraction, was considered a valid contribution.

3. Discussions about antibiotic resistance in wastewater environments may assume that antibiotic resistant bacteria are more resilient during treatment and in the environment than the susceptible counterparts, or, in opposition, that some lineages of multidrug resistant

bacteria may be lost or at least lose some relevant features during wastewater treatment or in the environment. The recognition of the lineages and features that survive wastewater treatment will contribute to improve the current knowledge about the potentials and limitations of wastewater treatment as well as to identify markers that may be useful for environmental monitoring.

CHAPTER 2

THESIS ROADMAP

The thesis is organized in 8 chapters, three of which are general – the Introduction in Chapter 1, General Discussion and Main Conclusions in Chapter 7, and Suggestions of Future Work in Chapter 8. Experimental work addressing each of three hypotheses, listed in section "Hypothesis and objectives of the thesis" in Chapter 1, is described in Chapter 3 to Chapter 6. The thesis includes work that is integrated in eight scientific articles, six published in peer-reviewed journals, between 2018 and 2020, one currently submitted and another one in preparation.

Chapter 3 was designed to address needs identified in ongoing collaborative studies. Once qPCR was identified as the method of choice to quantify antibiotic resistance genes in wastewater, as it allows the determination of abundance and prevalence values and supports the calculation of removal rates, the goal was to test distinct protocols targeting genes in different types of samples, processed in different laboratories. This study permitted the analysis of results obtained in distinct laboratories using the same protocols and standards, but different operators, batch of reagents and equipment. It also permitted the inference about the possible influence of DNA extract quality and shipment, and the use of distinct protocols. The promising results of this chapter raised the concern that DNA extraction may be poorly efficient in some laboratories (mainly researchers who are not microbiologist or molecular biologists) and that minor genes may be prone to higher deviations.

These considerations supported the design of Chapter 4. Previous experience has shown that cell resilience to lysis or cell losses due to different factors may affect DNA extraction yield and quality. Therefore, the aim was the design of an internal standard. The challenge was the identification of a gene that would not be found in wastewater environment. A rare gene encoding a molinate hydrolase from *Gulosibacter molinativorax* (Leite *et al*., 2015) was considered a good option, given its absence in the environment and in wastewater. A fragment of this gene was cloned in *Escherichia coli* to produce an internal standard. The use of this internal standard showed that the matrix effect, the major cause for cell loss or inadequate lysis during DNA extraction, had a moderate effect on DNA extraction. DNA extraction was neither affected by the operator nor by the kit used, being observed that one three times cheaper than the recommended was also suitable. The possibility of supplying an internal standard to groups with limited skills in DNA extraction or the recommendation of a low-cost DNA extraction kit or the demonstration that qPCR protocols performed in different laboratories were comparable were important achievements towards the systematization of antibiotic resistance monitoring processes.

Although it is consensual that antibiotic resistance monitoring in wastewater is important, the wide diversity of antibiotic resistance genes in these environments unveiled by metagenomics

studies (An *et al*., 2018; Lira *et al*., 2020; Ng *et al*., 2019), raises the question of what is indeed important for human health. A major question is which are the genetic determinants and bacteria lineages that may be responsible for future effects in humans, namely through infection. These were the questions addressed in Chapters 5 and 6. In Chapter 5, 34 clinical and 25 wastewater isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins, from Spain and Portugal, were compared based on clinically-relevant features, infection capacity (n=47), genetic lineage and genome analysis (n=22). These results were clear: although clinical isolates may be selected for higher virulence and antibiotic resistance, identical features persist in wastewater isolates. However, it was hypothesized that a larger number of isolates and isolation sources might bring a distinct perspective. This was the motivation for Chapter 6 study that included 139 genomes, retrieved from the public database National Center for Biotechnology Information (NCBI) and from the previous study (Chapter 5). These 139 genomes, included 78 of clinical and 61 of environmental origin, distributed by 21 countries. While some genetic determinants such as antibiotic and metal resistance genes, and virulence genes were more prevalent and diverse in clinical genomes, other such as efflux systems, quorum sensing, and oxidative stress were more abundant and diverse in the environmental genomes. These observations suggest that *K. pneumoniae* may evolve some fitness-relevant features in the environment, further specialized through mutation and gene acquisition in the clinical context. These conclusions are relevant starting points to further investigate if the antibiotic resistance genes or some key housekeeping features that may deserve to be monitored in the environment.

CHAPTER 3

MONITORING ANTIBIOTIC RESISTANCE GENES IN WASTEWATER: POTENTIALS AND LIMITATIONS OF QUANTITATIVE PCR

The work presented in this chapter is part of the following publications:

Rocha, J., Cacace, D., Kampouris, I., Guilloteau, H., Jäger, T., Marano, R.B.M., Karaolia, P, Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Berendonk, T.U., Schwartz, T. (2020). Inter-laboratory calibration of quantitative analyses of antibiotic resistance genes. Journal of Environmental Chemical Engineering, 8:102214. doi: 10.1016/j.jece.2018.02.022

Contribution: qPCR analysis and data compilation and analysis.

Narciso-da-Rocha, C.*, Rocha, J.*, Vaz-Moreira, I., Lira, F., Tamames, J., Henriques, I., Martinez, J., Manaia, C.M. (2018). Bacterial lineages with a major role on the dissemination of antibiotic resistance genes: a study in a full-scale urban wastewater treatment plant. Environment International 118:179-188. doi: 10.1016/j.envint.2018.05.040 (*equal contribution)

Contribution: qPCR analysis.

Rodríguez-Chueca, J., Varela, S., Rocha, J., Fernandes, T., Pablos, C., Encinas, A., Barceló, D., Rodríguez-Mozaz, S., Manaia, C. M., Encinas, A., Marugán, J. (2019). Assessment of full-scale tertiary wastewater treatment by UV-C based-AOPs: Removal or persistence of antibiotics and antibiotic resistance genes? Science of The Total Environment 652:1051-1061. doi: 10.1016/j.scitotenv.2018.10.223

Contribution: qPCR and data analysis.

Rocha, J.*, Fernandes, T.*, Riquelme, M. V., Zhu, N., Pruden, A., Manaia, C. M. (2019). Comparison of culture- and quantitative PCR-based indicators of antibiotic resistance in wastewater, recycled water, and tap water. International Journal of Environmental Research and Public Health, 16:4217. doi: 10.3390/ijerph16214217 (*equal contribution)

Contribution: microbiological and qPCR analysis and data compilation and analysis.

Michael, S.G., Kordatou, I.M., Nahim-Granados, S., Polo-López, M.I., Rocha, J., Martínez-Piernas, A.B., Fernández-Ibáñez, P., Agüera, A., Manaia, C.M., Fatta-Kassinos, D. (2020). Investigating the impact of UV-C- and solar-driven advanced oxidation on the removal of antibiotics, antibiotic resistance determinants and toxicity present in urban wastewater. Chemical Engineering Journal, 388:124383. doi: 10.1016/j.cej.2020.124383

Contribution: qPCR and data analysis.

Abstract

Urban wastewater treatment plants are major recipients of antibiotics, antibiotic resistant bacteria, and antibiotic resistance genes. The development of wastewater treatment processes for the efficient removal of these contaminants is a priority. However, monitoring antibiotic resistance in wastewater presents some challenges. This work focused on the suitability of using quantitative PCR (qPCR) to monitor wastewaters. The aims were to 1) assess the suitability of results expressed in gene abundance or prevalence to evaluate treatment efficiency and wastewater quality; 2) assess the degree of information provided by different antibiotic resistance genes as possible markers; 3) explore possibilities and critical factors for producing comparable data in different laboratories – e.g., the effect of DNA extract quality and shipment, equipment and gene-specific bias. Because it allows the detection of specific genetic determinants, and the results can be expressed as absolute (per volume) or relative (prevalence per total bacteria) abundance the method of choice was qPCR. The genetic determinants monitored were among the most predominant antibiotic resistance genes in wastewater encoding resistance to sulfonamides (*sul1* and *sul2*), quinolones (*qnrS)*, and β-lactams (bla_{TEM}) and the 16S rRNA gene as a proxy of total bacteria. The work was conducted in collaboration with partners responsible for wastewater treatment, in full-scale plants or pilot systems, who supplied the DNA extracts. The results showed that although wastewater quality can be expressed in terms of antibiotic resistance prevalence, treatment efficiency must be measured based on absolute abundance. The interlaboratory comparison of monitoring results seemed reliable, although DNA extract quality and stability during shipment, as well as consumables and equipment specificities, may be critical for the monitoring findings. It was also observed that some genetic determinants or PCR primers may be inadequate for interlaboratory monitoring. This work shows that harmonized protocols may contribute to enable worldwide comparisons of antibiotic resistance genes in wastewaters, for technological development or surveillance purposes.

1. Introduction

Antibiotic resistant bacteria and antibiotic resistance genes are recognized contaminants of emerging concern (Manaia *et al*., 2016). In urban areas, wastewater treatment plants are major recipients of these contaminants (Rizzo *et al*., 2013). In the urban wastewater treatment plants, the influent is subjected to a series of treatment processes frequently constituting a primary treatment for the removal of major solids and a secondary treatment for the removal of biological loads. Some urban wastewater treatment plants may also apply a tertiary treatment, that may consist or include a disinfection treatment, aiming the removal of micropollutants and the improvement of quality of the final effluent (Manaia *et al*., 2018). Urban wastewater treatment plants can achieve removal values of bacteria, ranging 1-3 log-units /mL, meaning that high loads of antibiotic resistant bacteria and antibiotic resistance genes are discharged in the final effluent (Manaia *et al*., 2016; Narciso-da-Rocha *et al*., 2014; Novo *et al*., 2013; Pallares-Vega *et al*., 2019). The need to assess the impacts caused in the receiving environments and to improve wastewater treatment technologies is consensual, although far from being designed or implemented (Berendonk *et al*., 2015; Di Cesare *et al*., 2020). The traditional enumeration of bacteria and *Escherichia coli* monitoring in water are methods implemented to assess water quality but with serious limitations to be used as an antibiotic resistance monitoring method (Manaia *et al*., 2018). These methods have the advantage of being comparable among different laboratories because are designed for routine- and directive-oriented procedures in which the water quality is evaluated based on the presence of enterococci, *Escherichia coli* or other bacteria (ISO 7899; ISO 9308). Although these methods are informative of the phenotype of bacteria, they present disadvantages. Besides being time consuming, these approaches do not consider the viable but not culturable bacteria and are not oriented for antibiotic resistant bacteria analysis (Manaia *et al*., 2018). Culture-independent methods are based on nucleic acids analysis and therefore enable the detection of the usually large fraction of non-culturable bacteria, including antibiotic resistant bacteria as well as, at least some, of their genetic determinants. For these reasons, and because of requiring a lower workload, culture-independent methods have been used to monitor antibiotic resistance worldwide (Cacace *et al*., 2019; Böckelmann *et al*., 2009; Kim *et al*., 2013; Klein, 2002; Nguyen *et al*., 2021; Valasek, 2005). However, environmental monitoring of antibiotic resistance genes has been made exclusively as a research activity. Metagenomics and the quantitative PCR (qPCR) are the most used techniques for this purpose. The first, metagenomics, is a non-targeted analysis which enables the detection of a wide range of antibiotic resistance genes but does not allow quantifications in relation to the volume or mass of sample. The second, qPCR, is a targeted method based on the use of specific primers, and which allows the quantification of antibiotic resistance genes per volume or mass of sample supported by a calibration curve (in the traditional real-time thermal

cyclers). As a targeted method, qPCR presents high specificity and sensitivity being among the gold-standard culture-independent methods to be used to monitor antibiotic resistance genes in environmental samples (Kim *et al*., 2013). Variations to first qPCR methods have been developed, being important upgrades the avoidance of calibration curve and the large numbers of samples that can be processed simultaneously, e.g. in the qPCR array, high throughput qPCR or in the ddPCR (An *et al*., 2018; Gao *et al*., 2018; Pärnänen *et al*., 2019). Although the use of qPCR to monitor antibiotic resistance is expanding and revealing to be a promising tool, it has been recognized that results can be affected by the sample processing (e.g. filtering membranes for liquid samples, DNA extraction kits, standards used, primers and master mixes, protocols or equipment (Djurhuus *et al*.,2017; Hinlo *et al*., 2017; Kim *et al*., 2013; Li *et al*., 2018; Rocha *et al*., 2020). The present study bridged different research projects (Table 1) where, for different reasons, antibiotic resistance genes were being monitored in wastewaters, aiming to explore the potentials and pitfalls of centralized versus interlaboratory analyses. Specifically, it was aimed to 1) assess the suitability of results expressed in gene abundance and prevalence to evaluate treatment efficiency and wastewater quality; 2) compare the information provided by different antibiotic resistance genes and assess their suitability as possible markers; 3) investigate critical factors and how these can be controlled for producing comparable data in different laboratories – e.g., the effect of DNA extract quality and shipment, equipment or gene-specific bias.

The findings of this work show that adequately implemented and harmonized procedures may offer reliable data of antibiotic resistance monitoring obtained worldwide in treatment efficiency or environmental surveillance.

2. Material and Methods

2.1. Samples and DNA extracts

This work was performed using DNA extracts obtained in our premises or by different partners, who shipped the DNA extracts to CBQF/UCP. All sample processing and DNA extraction were performed according to a harmonized procedure and protocol (Pärnänen *et al*., 2019; Rocha *et al*., 2020). Samples of raw/untreated and treated wastewater were collected by the partners of each laboratory, filtered and the total DNA was extracted using the DNeasy PowerWater kit (QIAGEN, Hilden, Germany) (or an alternative, Table 1) according to the recommended procedures (common to all partners). The filtering membranes and the DNA extraction kits used for each partner to filter the samples and extract the total bacterial DNA are indicated in Table 1. The DNA extracts were shipped in refrigerated conditions at 4 ºC to UCP, Portugal

which upon DNA extracts reception kept the DNA extracts refrigerated at -20 ºC until their use for qPCR analyses.

2.2. Quantitative PCR

The genes 16S rRNA, targeting total bacteria; *gadAB*, targeting *Escherichia coli*; and the genes encoding resistance to the antibiotic classes sulfonamides (*sul1* and *sul2*), quinolones (*qnrS*), and β-lactams (bla_{TEM}, and bla_{OXA-1}) were quantified using qPCR. These genes were selected based on their wide distribution in wastewater (bla_{TEM}, bla_{OXA}, *sul1*, *qnrS*), and as markers of bacterial groups (16S rRNA, *gadAB*, *Escherichia coli*) (Chen *et al*., 2006; Du *et al*., 2014; Narciso-da-Rocha *et al*., 2014; Varela *et al*., 2015). The quantifications were performed at least in duplicate for each sample, using the Standard Curve method as described in Brankatschk *et al*. (2012). The primers, standards and qPCR conditions are presented in Table 2. Possible qPCR inhibition was assessed by quantifying target genes using 10- and 100-fold diluted samples, as suggested by Bustin *et al*. (2009). Two real-time thermocyclers were used: StepOnePlus™ (Life Technologies, Carlsbad, USA), in UCP (Portugal), and Bio-Rad Real-Time PCR Analysis Software in Virginia Tech (US).

Table 1 - Aims and design of the studies supporting this chapter.

All samples were filtered through polycarbonate membranes with 0.22 µm of pore (Whatman or Merck), DNA was extracted with the DNeasy PowerWater kit; and all qPCR reactions were performed in a StepOne™ Real-Time PCR System (Life Technologies). Exceptions are indicated in the table.

Study	Aim	Study type	Type of wastewater	Wastewater Treatments	Genes analysed	Observations
Narciso-da-Rocha *et al*., 2018	Evaluate variations in bacterial community composition and antibiotic resistance genes load and explore possible relationships among them.	Full-scale	RWW, sTWW, tTWW, RE-tTWW	Primary, activated sludge and UV	16S rRNA, *intI*1, *bla*$_{TEM}$, *bla*$_{OXA-A}$, *bla*$_{SHV}$, *bla*$_{CTX-M}$, *sul1*, *sul2*, and *qnrS*	Membranes Whatman
Rodríguez-Chueca *et al*., 2019	Application of advanced oxidation processes for the removal of antibiotics and antibiotic resistance genes from a wastewater effluent of a treatment plant.	Full-scale	sTWW and tTWW	H_2O_2/UV-C, PMS/UV-C and PMS/Fe(II)/UV-C	16S rRNA, *intI*1, *bla*$_{TEM}$, *bla*$_{OXA-A}$, *sul1*, *sul2*, and *qnrS*	Membranes Isopore Merck
Rocha *et al*., 2019	Comparison of measurements of tetracycline-, sulphonamide-, and cefotaxime-resistant presumptive total and fecal coliforms and presumptive enterococci versus antibiotic resistance genes quantified by quantitative polymerase chain reaction across waste-, recycled-, tap-, and freshwater. Comparison of qPCR measurements among two laboratories.	Full-scale / Pilot	RWW, sTWW, tTWW	Primary, activated sludge and UV	16S rRNA, 23S rRNA, *uidA*, *gadAB*, *intI*1, *bla*$_{OXA-1}$, *bla*$_{CTX-M}$, *sul1*, *sul2*, *tet*(A) and *tet*(O)	PT: Membranes Merck US: Membranes Whatman US: DNA extraction; FastDNA SPIN KIT (MP Biomedicals LCC) US: qPCR- Bio-Rad Real-Time PCR Analysis Software (Biorad)
Rocha *et al*., 2020	Inter-laboratory calibration to assess the variability inherent to the qPCR procedures for quantification of antibiotic resistance genes.	Full-scale	sTWW	Membrane bioreactor and activated sludge	16S rRNA, *intI*1, *bla*$_{TEM}$, *bla*$_{CTXM-32}$, *sul1*, *qnrS*, and *vanA*	Membranes Whatman
Michael *et al*., 2020	Explore the impact of UV-C/H2O2 and sunlight/H2O2 processes on removing antibiotics, cultivable bacteria and genes.	Pilot scale	sTWW	Activated sludge, sunlight/H_2O_2, and UV-C/H_2O_2	16S rRNA, *bla*$_{OXA-A}$, *bla*$_{SHV}$, *bla*$_{CTX-M}$, *sul1*, *sul2*, *qnrS*, *tet*(M), *vanA*, and *mecA*	Membranes Merck

RWW – Raw wastewater; sTWW – effluent of the secondary treatment; tTWW – effluent of the tertiary treatment; RE-tTWW – effluent of the tertiary treatment incubated 3 days in the dark at 20 ºC; UV – ultraviolet; H_2O_2/UV-C – hydrogen peroxide/ UV-C; PMS/UV-C – Peroxymonosulfate/ UV-C; and PMS/Fe(II)/UV-C – Peroxymonosulfate/iron/UV-C.

Table 2 - Primers, standards and qPCR conditions used in the different studies.

Target Gene	Primers sequence (reference)	qPCR program and conditions	Report of the results and Limits of Quantification per PCR reaction (LOQ, copy number)
16S rRNA	331F - TCCTACGGGAGGCAGCAGT 518R - ATTACCGCGGCTGCTGG (Nadkarni *et al*., 2002 and Muyzer *et al*., 1993)	95 °C - 10 min (1 cycle); 95 °C - 15 s, 60 °C - 1 min (45 cycles) Mastermix: Power SYBR Green Primer concentration: 500 nM DNA: 10 µL / 25 µL reaction	Rocha *et al*., 2020 LOQ: 376
16S rRNA	1369F - CGGTGAATACGTTCYCGG 1492R - GGWTACCTTGTTACGACTT (Suzuki *et al*., 2000)	98 °C - 2 min (1 cycle), 98 °C - 5 s - 55 °C - 5 s (40 cycles) Mastermix: Evagreen Primer concentration: 400 nM DNA: 1 µL / 20 µL reaction	Rocha *et al*., 2019 LOQ: 100
16S rRNA	1114F - CGGCAACGAGCGCAACCC 1275R - CCATTGTAGCACGTGTGTAGCC (Denman *et al*., 2016)	95 °C - 10 min (1 cycle); 95 °C - 15 s, 55 °C - 20 s and 72 °C - 10 s (35 cycles) Mastermix: KAPA SYBR® FAST ABI Prism Primer concentration: 200 nM DNA: 2 µL / 20 µL reaction	Narciso-da-Rocha *et al*., 2018; Michael *et al*., 2020 LOQ: 385 Rodríguez-Chueca *et al*., 2019 LOQ: 330 Rocha *et al*., 2019 LOQ: 402
gadAB	gadrt-1 - GCGTTGCGTAAATATGGTTGCCGA gadrt-2 - CGTCACAGGCTTCAATCATGCGTT (Chen *et al*., 2006)	**US**: 95 °C - 2 min (1 cycle), 95 °C - 5 s, 69 °C - 5 s (40 cycles) Mastermix: Evagreen Primer concentration: 400 nM DNA: 1 µL / 20 µL reaction **PT**: 95 °C - 10 min (1 cycle), 95 °C - 15 s, 60 °C - 1 min (40 cycles) Mastermix: KAPA SYBR® FAST ABI Prism Primer concentration: 400 nM DNA: 2 µL / 20 µL reaction	Rocha *et al*., 2019 LOQ-US: 10 /LOQ-PT: 43

Table 2 (cont.) - Primers, standards and qPCR conditions used in the different studies.

Target Gene	Primers sequence (reference)	qPCR program and conditions	Report of the results and Limits of Quantification per PCR reaction (LOQ, copy number)
sul1	sul1F - CGCACCGGAAACATCGCTGCAC sul1R - TGAAGTTCCGCCGCAAGGCTCG (Pei *et al*., 2006)	**US**: 98 °C - 2 min (1 cycle), 98 °C - 5 s, 69 °C - 5 s (40 cycles) Mastermix: Evagreen Primer concentration: 400 nM DNA: 1 µL / 20 µL reaction **PT:** 95 °C - 5 min (1 cycle), 95 °C - 10 s, 60 °C - 30 s (35 cycles) Mastermix: FAST SYBR Primer concentration: 300nM DNA: 5 µL / 20 µL reaction (Rodríguez-Chueca *et al*., 2019; Rocha *et al*., 2020; Rocha *et al*., 2019; Michael *et al*., 2020; Narciso-da-Rocha *et al*., 2018)	Narciso-da-Rocha *et al*., 2018; Rocha *et al*., 2020; Michael *et al*., 2020 LOQ: 240 Rodríguez-Chueca *et al*., 2019 LOQ: 135 Rocha *et al*., 2019 LOQ-US: 100 / LOQ-PT: 96
sul2	sul2F - TCCGGTGGAGGCCGGTATCTGG sul2R - CGGGAATGCCATCTGCCTTGAG (Pei *et al*., 2006)	**US**: 98 °C - 2 min (1 cycle), 98 °C - 5 s, 67.5 °C - 5 s (40 cycles) Mastermix: Evagreen Primer concentration: 400 nM DNA: 1 µL / 20 µL reaction **PT:** 95 °C - 5 min (1 cycle), 95 °C - 15 s, 60 °C - 1 min (40 cycles) Mastermix: KAPA SYBR® FAST ABI Prism Primer concentration: 200 nM DNA: 2 µL / 20 µL reaction	Narciso-da-Rocha *et al*., 2018; Rodríguez-Chueca *et al*., 2019; Michael *et al*., 2020 LOQ: 47 Rocha *et al*., 2019 LOQ-US: 100/ LOQ-PT: 47

Table 2 (cont.) - Primers, standards and qPCR conditions used in the different studies.

Target Gene	Primers sequence (reference)	qPCR program and conditions	Report of the results and Limits of Quantification per PCR reaction (LOQ, copy number)
***bla*$_{OXA-1}$**	oxa1F - TATCTACAGCAGCGCCAGTG oxa1R - CGCATCAAATGCCATAAGTG (Yang *et al*., 2012)	**US**: 98 °C - 2 min (1 cycle), 98 °C - 5 s, 62 °C - 5 s (40 cycles) Mastermix: Evagreen Primer concentration: 400 nM DNA: 1 µL / 20 µL reaction **PT**: 95 °C - 10 min (1 cycle), 95 °C - 15 s, 60 °C - 1 min (40 cycles) Mastermix: KAPA SYBR® FAST ABI Prism Primer concentration: 200 nM DNA: 2 µL / 20 µL reaction	Rocha *et al*., 2019 LOQ-US: 10 /LOQ-PT: 38
***bla*$_{TEM}$**	blaTEMF - TTCCTGTTTTTGCTCACCCAG blaTEMR - CTCAAGGATCTTACCGCTGTTG (Bibbal *et al*., 2007)	95 °C - 10 min (1 cycle), 95 °C - 15 s, 60 °C - 1 min (40 cycles) Mastermix: SYBR® Select Master Mix Primer concentration: 200 nM DNA: 2 µL / 20 µL reaction	Narciso-da-Rocha *et al*., 2018; Rocha *et al*., 2020; Michael *et al*., 2020 LOQ: 75 Rodríguez-Chueca *et al*., 2019 LOQ: 45
qnrS	qnrSrtF- GACGTGCTAACTTGCGTGAT qnrSrtR - TGGCATTGTTGGAAACTTG (Marti *et al*., 2013)	95 °C - 10 min (1 cycle), 95 °C - 15 s, 60 °C - 1 min (40 cycles) Mastermix: SYBR® Select Master Mix Primer concentration: 600 nM DNA: 2 µL / 20 µL reaction	Narciso-da-Rocha *et al*., 2018 LOQ: 75 copies Rodríguez-Chueca *et al*., 2019 LOQ: 10 Rocha *et al*., 2020; Michael *et al*., 2020 LOQ: 96

2.3. Criteria used to analyse qPCR results

The qPCR results were analysed based on the following criteria: standard curve efficiency between 90 and 110%, correct melting temperature value, and unique melting peak. Amplifications in which the melting curves presented shoulders (increased signal in the baseline, e.g., due to primer dimers), multiple melting peaks (additional unspecific amplification) or incorrect melting temperatures (> ±1 °C from standards) were not considered. It was considered in each case that any gene amplification product that was below the limit of quantification (LOQ), would not be considered for the analysis.

2.4. Calculations and statistical analyses

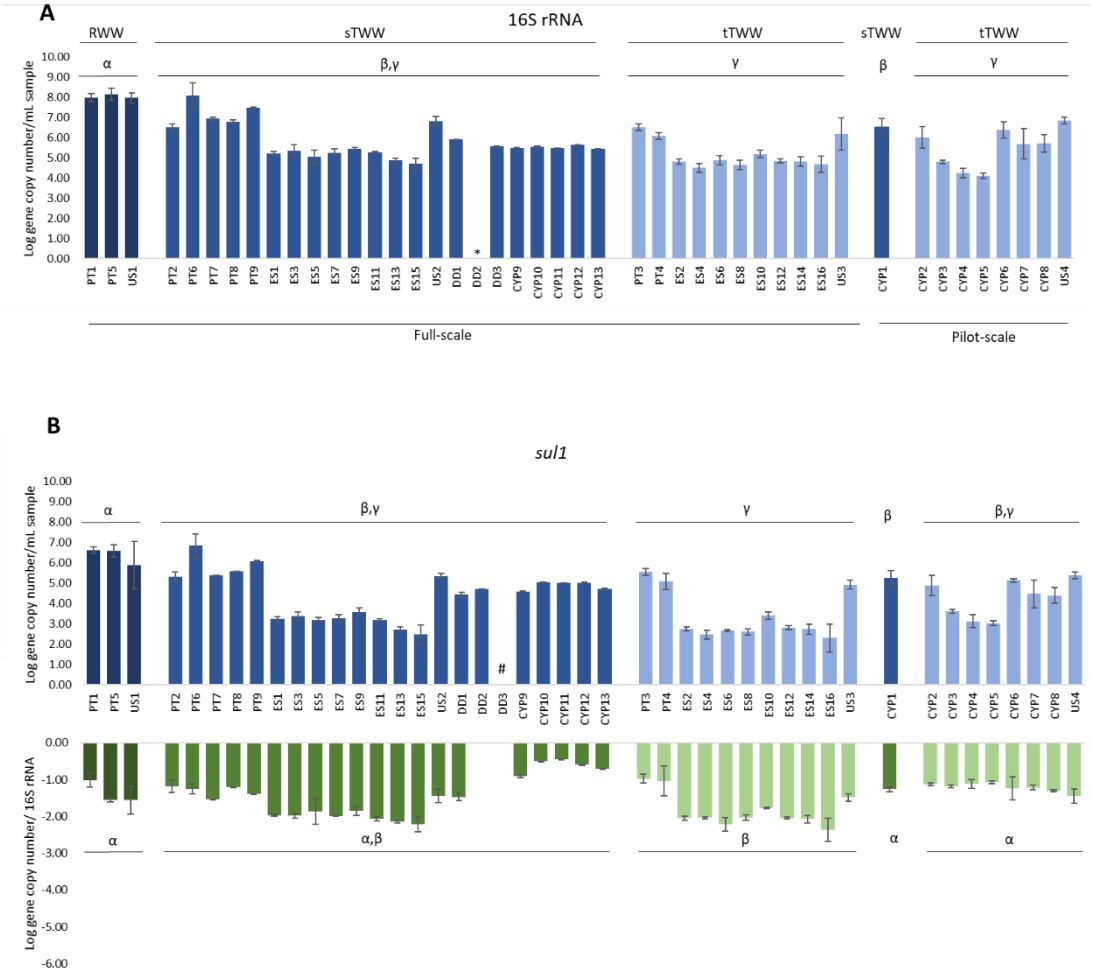
One-way analysis of variance (ANOVA) and Tukey's and Bonferroni post-hoc tests (SPSS Statistics for Windows v.27.0; IBM Corp., Armonk, NY, USA) were used to assess statistically significant differences ($p<0.05$) of prevalence and/or abundance of genes between types of wastewater. Log removal values of genes abundance (or prevalence) was calculated as: Log removal (log-units) = average of gene abundance (or prevalence) in the influent of the treatment – average of gene abundance (or prevalence) in the effluent of the treatment.

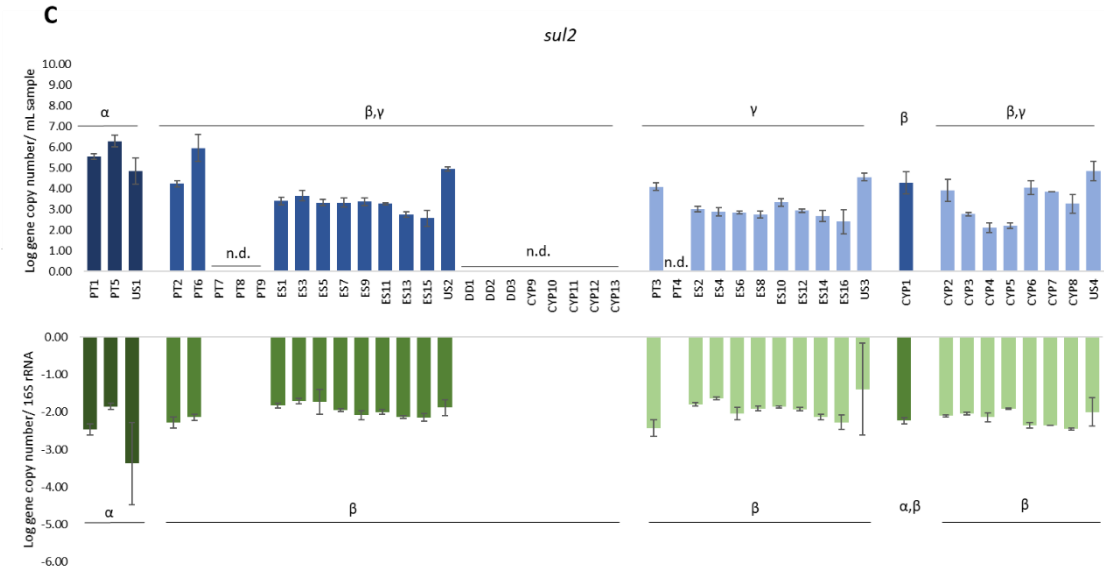
3. Results and Discussion

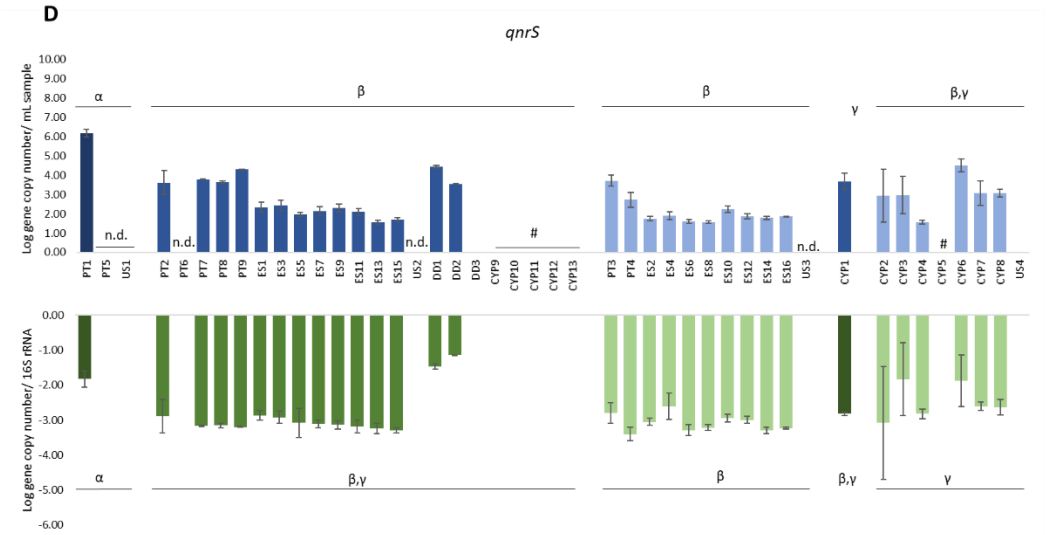
3.1. Gene abundance vs. gene prevalence

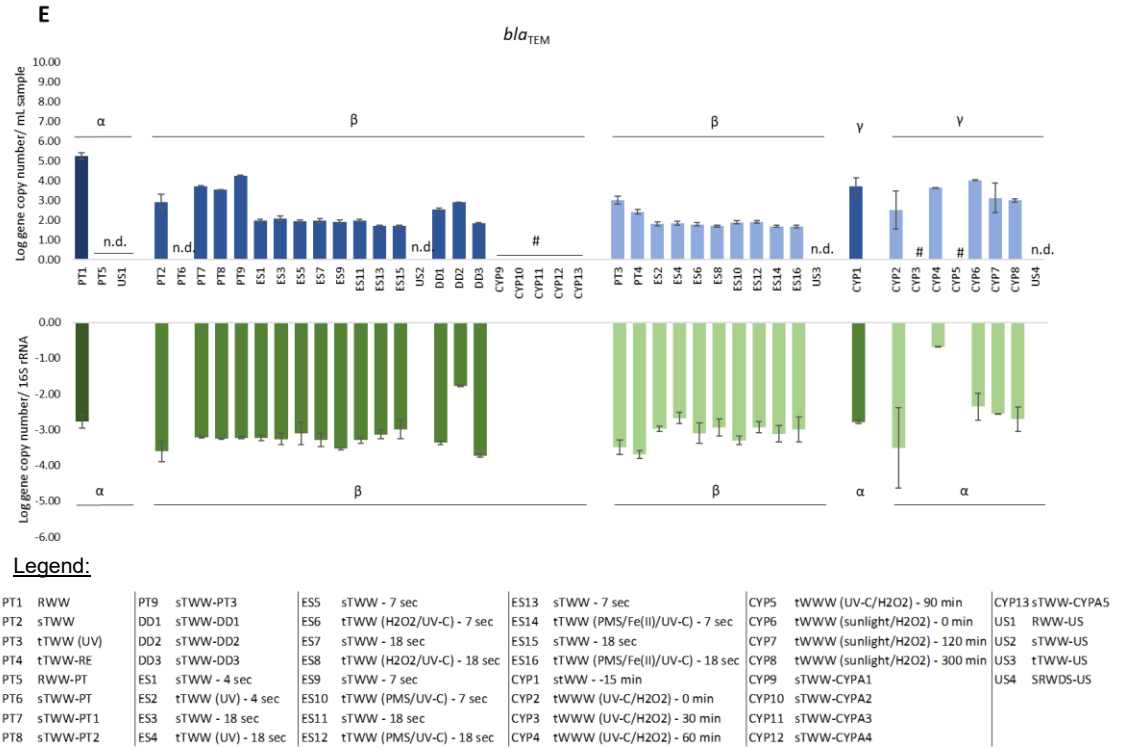
One of the objectives of this study was to interpret the results expressed in gene abundance or prevalence to evaluate treatment efficiency and wastewater quality. The results presented in Figure 1 show the qPCR quantifications for the genes 16S rRNA, *sul1*, *sul2*, *qnrS* and bla_{TEM}, expressed as abundance, in the upper part, or prevalence, in the lower part. Quantifications of the 16S rRNA gene express the abundance of bacteria in the water samples and is used also to estimate the prevalence values. The abundance of this gene was the highest in RWW (raw wastewater), ranging 7.97-8.15 log-units/mL, followed by sTWW, where it ranged 4.70-8.09 log-units/mL. Disinfection, tested at full- or pilot-scale (tTWW), led to 16S rRNA gene quantifications ranging 4.51-6.50 log-units/mL.

The quantifications of the 16S rRNA gene were not proportional to the DNA concentration in the extracts (Table 3). The effect was more notorious in samples of tertiary treatment processes in which DNA losses during extraction may be more severe (Table 3).




C
sul2
Log gene copy number/ mL sample
Log gene copy number/ 16S rRNA
α
β,γ
γ
β
β,γ
n.d.
PT1 PT5 US1 PT2 PT6 PT7 PT8 PT9 ES1 ES3 ES5 ES7 ES9 ES11 ES13 ES15 US2 DD1 DD2 DD3 CYP9 CYP10 CYP11 CYP12 CYP13 PT3 PT4 ES2 ES4 ES6 ES8 ES10 ES12 ES14 ES16 US3 CYP1 CYP2 CYP3 CYP4 CYP5 CYP6 CYP7 CYP8 US4
α
β
β
α,β
β


D
qnrS
Log gene copy number/ mL sample
Log gene copy number/ 16S rRNA
α
β
β
γ
β,γ
n.d.
#
PT1 PT5 US1 PT2 PT6 PT7 PT8 PT9 ES1 ES3 ES5 ES7 ES9 ES11 ES13 ES15 US2 DD1 DD2 DD3 CYP9 CYP10 CYP11 CYP12 CYP13 PT3 PT4 ES2 ES4 ES6 ES8 ES10 ES12 ES14 ES16 US3 CYP1 CYP2 CYP3 CYP4 CYP5 CYP6 CYP7 CYP8 US4
α
β,γ
β
β,γ
γ



Legend:

PT1	RWW	PT9	sTWW-PT3	ES5	sTWW - 7 sec	ES13	sTWW - 7 sec	CYP5	tWWW (UV-C/H2O2) - 90 min	CYP13	sTWW-CYPA5				
PT2	sTWW	DD1	sTWW-DD1	ES6	tTWW (H2O2/UV-C) - 7 sec	ES14	tTWW (PMS/Fe(II)/UV-C) - 7 sec	CYP6	tWWW (sunlight/H2O2) - 0 min	US1	RWW-US				
PT3	tTWW (UV)	DD2	sTWW-DD2	ES7	sTWW - 18 sec	ES15	sTWW - 18 sec	CYP7	tWWW (sunlight/H2O2) - 120 min	US2	sTWW-US				
PT4	tTWW-RE	DD3	sTWW-DD3	ES8	tTWW (H2O2/UV-C) - 18 sec	ES16	tTWW (PMS/Fe(II)/UV-C) - 18 sec	CYP8	tWWW (sunlight/H2O2) - 300 min	US3	tTWW-US				
PT5	RWW-PT	ES1	sTWW - 4 sec	ES9	sTWW - 7 sec	CYP1	stWW - -15 min	CYP9	sTWW-CYPA1	US4	SRWDS-US				
PT6	sTWW-PT	ES2	tTWW (UV) - 4 sec	ES10	tTWW (PMS/UV-C) - 7 sec	CYP2	tWWW (UV-C/H2O2) - 0 min	CYP10	sTWW-CYPA2						
PT7	sTWW-PT1	ES3	sTWW - 18 sec	ES11	sTWW - 18 sec	CYP3	tWWW (UV-C/H2O2) - 30 min	CYP11	sTWW-CYPA3						
PT8	sTWW-PT2	ES4	tTWW (UV) - 18 sec	ES12	tTWW (PMS/UV-C) - 18 sec	CYP4	tWWW (UV-C/H2O2) - 60 min	CYP12	sTWW-CYPA4						

Figure 1 - Gene abundance (gene copy number /mL of sample) (A-E, upper part) and genes prevalence (abundance of the gene/abundance of 16S rRNA gene) (A-E, lower part) in wastewater samples. RWW: wastewater treatment plant influent, sTWW: wastewater collected after secondary treatment, tTWW: wastewater collected after tertiary treatment, and SRWDS: simulated reclaimed water distribution system. For tertiary treatment samples within parenthesis is indicated the type of treatment tested and the period of time of treatment. PT refers to samples obtained in Portugal, DD to samples obtained in Germany, CYP to samples obtained in Cyprus, ES to samples obtained in Spain and US samples obtained in the USA. Error bars represent the standard deviation of at least two replicates. * refers to different melting temperature from the standards and # refers to determinations below the limit of quantification; n.d. refers to values not determined; α,β,γ refer to statistically significant differences ($p<0.05$).

Table 3 - Samples volume filtered, DNA extracts concentration and shipment.

	Sample	Volume filtered (mL)	Shipped (Yes/No)	Average DNA concentration in the DNA extract (ng/µL)	Average DNA concentration (mg /100 mL)	Average log 16S rRNA gene copy number / 100 mL sample
RWW (Full scale)	PT1	25	No	25.5	10.2	797
	PT5	50	No	52.9	10.6	815
	US1	100	Yes	231.7	23.2	797
sTWW (Full-scale)	PT2	250	No	56.3	2.3	650
	PT6	50	No	126.4	25.3	809
	PT7	150	Yes	24.5	1.6	694
	PT8	150	Yes	33.4	2.2	679
	PT9	150	Yes	72.0	4.8	749
	ES1	100	Yes	1.2	0.1	521
	ES3	100	Yes	2.1	0.2	535
	ES5	100	Yes	1.3	0.1	504
	ES7	100	Yes	1.3	0.1	526
	ES9	100	Yes	1.2	0.1	544
	ES11	100	Yes	1.2	0.1	527
	ES13	100	Yes	0.8	0.1	486
	ES15	100	Yes	0.6	0.1	470
	US2	100	Yes	9.4	0.9	681
	DD1	150	Yes	7.7	0.5	591
	DD2	150	Yes	13.1	0.9	*
	DD3	150	Yes	5.5	0.4	557
	CYP9	150	Yes	3.8	0.3	549
	CYP10	150	Yes	4.5	0.3	555
	CYP11	150	Yes	3.9	0.3	546
	CYP12	150	Yes	4.4	0.3	563
	CYP13	150	Yes	4.5	0.3	545
tTWW (Full-scale)	PT3	300	No	84.6	2.8	650
	PT4	300	No	58.7	2.0	609
	ES2	100	Yes	0.9	0.1	480
	ES4	100	Yes	0.7	0.1	451
	ES6	100	Yes	0.4	0.0	489
	ES8	100	Yes	0.5	0.0	464
	ES10	100	Yes	1.8	0.2	519
	ES12	100	Yes	0.8	0.1	485
	ES14	100	Yes	0.6	0.1	481
	ES16	100	Yes	0.5	0.1	468

Table 3 (cont.) - Samples volume filtered, DNA extracts concentration and shipment.

	Sample	Volume filtered (mL)	Shipped (Yes/No)	Average DNA concentration in the DNA extract (ng/µL)	Average DNA concentration (mg /100 mL)	Average log 16S rRNA gene copy number / 100 mL sample
	US3	100	Yes	6.8	0.7	618
sTWW (Pilot-scale)	CYP1	200-215	Yes	23.8	1.1-1.2	650
tTWW (Pilot-scale)	CYP2	200-250	Yes	16.2	0.6-0.8	601
	CYP3	200-250	Yes	4.6	0.2	480
	CYP4	200	Yes	0.9	0.0	425
	CYP5	200-250	Yes	0.6	0.0	411
	CYP6	100-150	Yes	26.9	1.8-2.7	639
	CYP7	100-150	Yes	22.8	1.5-2.3	568
	CYP8	150-200	Yes	26.4	1.3-1.8	571
	US4	100	Yes	3.1	0.3	683

* Different melting temperature compared to the standard.

Legend:

PT1 RWW	PT9 sTWW-PT3	ES5 sTWW - 7 sec	ES13 sTWW - 7 sec	CYP5 tWWW (UV-C/H2O2) - 90 min	CYP13 sTWW-CYPA5
PT2 sTWW	DD1 sTWW-DD1	ES6 tTWW (H2O2/UV-C) - 7 sec	ES14 tTWW (PMS/Fe(II)/UV-C) - 7 sec	CYP6 tWWW (sunlight/H2O2) - 0 min	US1 RWW-US
PT3 tTWW (UV)	DD2 sTWW-DD2	ES7 sTWW - 18 sec	ES15 sTWW - 18 sec	CYP7 tWWW (sunlight/H2O2) - 120 min	US2 sTWW-US
PT4 tTWW-RE	DD3 sTWW-DD3	ES8 tTWW (H2O2/UV-C) - 18 sec	ES16 tTWW (PMS/Fe(II)/UV-C) - 18 sec	CYP8 tWWW (sunlight/H2O2) - 300 min	US3 tTWW-US
PT5 RWW-PT	ES1 sTWW - 4 sec	ES9 sTWW - 7 sec	CYP1 stWW - -15 min	CYP9 sTWW-CYPA1	US4 SRWDS-US
PT6 sTWW-PT	ES2 tTWW (UV) - 4 sec	ES10 tTWW (PMS/UV-C) - 7 sec	CYP2 tWWW (UV-C/H2O2) - 0 min	CYP10 sTWW-CYPA2	
PT7 sTWW-PT1	ES3 sTWW - 18 sec	ES11 sTWW - 18 sec	CYP3 tWWW (UV-C/H2O2) - 30 min	CYP11 sTWW-CYPA3	
PT8 sTWW-PT2	ES4 tTWW (UV) - 18 sec	ES12 tTWW (PMS/UV-C) - 18 sec	CYP4 tWWW (UV-C/H2O2) - 60 min	CYP12 sTWW-CYPA4	

The observation of Table 3 suggests that even extracts with low DNA concentrations, may have enough DNA to serve as template to the PCR amplification of abundant genes as is the case of 16S rRNA gene. In general, the antibiotic resistance genes (*sul1*, *sul2*, *qnrS* and *bla*$_{TEM}$) presented abundance values of about 0.60-3.73 log-units/mL lower than the 16S rRNA gene. However, in some sTWW or tTWW samples the lowest gene quantifications ranged values of 1.3-1.7 log-units/mL (*bla*$_{TEM}$ and *qnrS),* which may be close to the limits of quantification. These lowest quantification values were observed in samples with low DNA concentrations and may be artefactual, due to the quality of DNA extracts. Hence, 16S rRNA gene quantification is always essential to assess the abundance of bacterial biomass in a sample, being recommended in any study to evaluate microbiological wastewater quality. The 16S rRNA gene is the most commonly used normalization parameter to assess the prevalence of antibiotic resistance genes (Cacace *et al*., 2019; Gao *et al*., 2012; Rodriguez-Mozaz *et al*., 2015). Such normalization informs about the fraction of bacteria that in a community harbours antibiotic resistance genes. The results presented in Table 4 show that, in general, the

reduction of the 16S rRNA gene and of the tested antibiotic resistance genes presented close values.

Table 4 – Values of log removal of the gene abundance and of the gene prevalence in the different treatment processes tested.

	Values of log removal of genes abundance (white) or prevalence (grey)								
	16S rRNA	***sul1***	***sul1***	***sul2***	***sul2***	***qnrS***	***qnrS***	***bla*$_{TEM}$**	***bla*$_{TEM}$**
PT1-PT2	1.48	1.29	0.17	1.31	-0.19	2.57	1.07	2.33	0.84
PT2-PT3	-0.01	-0.21	-0.21	0.15	0.15	-0.10	-0.09	-0.11	-0.12
PT3-PT4	0.41	0.45	0.06	n.d.	n.d.	0.99	0.60	0.60	0.20
ES1-ES2	0.41	0.49	0.07	0.39	-0.03	0.60	0.19	0.16	-0.26
ES3-ES4	0.85	0.90	0.05	0.77	-0.07	0.54	-0.31	0.25	-0.60
ES5-ES6	0.16	0.51	0.35	0.47	0.31	0.36	0.21	0.15	-0.01
ES7-ES8	0.61	0.66	0.05	0.57	-0.04	0.57	0.10	0.27	-0.35
ES9-ES10	0.25	0.18	-0.07	0.04	-0.21	0.07	-0.18	0.02	-0.23
ES11-ES12	0.41	0.39	-0.03	0.34	-0.08	0.23	-0.18	0.05	-0.36
ES13-ES14	0.05	-0.01	-0.07	0.06	0.00	-0.23	0.06	0.04	-0.02
ES15-ES16	0.03	0.19	0.16	0.16	0.14	-0.15	-0.05	0.03	0.00
PT5-PT6	0.05	-0.25	-0.30	0.34	0.28	n.d.	n.d.	n.d.	n.d.
US1-US2	1.16	0.52	-0.11	-0.09	-1.50	n.d.	n.d.	n.d.	n.d.
US2-US3	0.63	0.43	0.04	0.37	-0.49	n.d.	n.d.	n.d.	n.d.
US3-US4	-0.66	-0.45	-0.04	-0.28	0.61	n.d.	n.d.	n.d.	n.d.
CYP1-CYP2	0.49	0.36	-0.13	0.36	-0.13	0.75	0.26	1.21	0.72
CYP2-CYP3	1.21	1.26	0.05	1.15	-0.06	-0.03	-1.25	<LOQ	<LOQ
CYP3-CYP4	0.55	0.49	-0.07	0.65	0.10	1.40	0.99	<LOQ	<LOQ
CYP4-CYP5	0.14	0.10	-0.04	-0.09	-0.23	<LOQ	<LOQ	<LOQ	<LOQ
CYP6-CYP7	0.70	0.67	-0.03	0.18	0.00	1.43	0.73	0.90	0.20
CYP7-CYP8	-0.03	0.07	0.10	0.59	0.09	0.00	0.03	0.12	0.14
TOTAL AVERAGE VALUE	0.42	0.38	0	0.37	-0.07	0.56	0.13	0.43	0.01

Legend:

PT1	RWW	PT9	sTWW-PT3	ES5	sTWW - 7 sec	ES13	sTWW - 7 sec	CYP5	tWWW (UV-C/H2O2) - 90 min	CYP13	sTWW-CYPA5
PT2	sTWW	DD1	sTWW-DD1	ES6	tTWW (H2O2/UV-C) - 7 sec	ES14	tTWW (PMS/Fe(II)/UV-C) - 7 sec	CYP6	tWWW (sunlight/H2O2) - 0 min	US1	RWW-US
PT3	tTWW (UV)	DD2	sTWW-DD2	ES7	sTWW - 18 sec	ES15	sTWW - 18 sec	CYP7	tWWW (sunlight/H2O2) - 120 min	US2	sTWW-US
PT4	tTWW-RE	DD3	sTWW-DD3	ES8	tTWW (H2O2/UV-C) - 18 sec	ES16	tTWW (PMS/Fe(II)/UV-C) - 18 sec	CYP8	tWWW (sunlight/H2O2) - 300 min	US3	tTWW-US
PT5	RWW-PT	ES1	sTWW - 4 sec	ES9	sTWW - 7 sec	CYP1	stWW - -15 min	CYP9	sTWW-CYPA1	US4	SRWDS-US
PT6	sTWW-PT	ES2	tTWW (UV) - 4 sec	ES10	tTWW (PMS/UV-C) - 7 sec	CYP2	tWWW (UV-C/H2O2) - 0 min	CYP10	sTWW-CYPA2		
PT7	sTWW-PT1	ES3	sTWW - 18 sec	ES11	sTWW - 18 sec	CYP3	tWWW (UV-C/H2O2) - 30 min	CYP11	sTWW-CYPA3		
PT8	sTWW-PT2	ES4	tTWW (UV) - 18 sec	ES12	tTWW (PMS/UV-C) - 18 sec	CYP4	tWWW (UV-C/H2O2) - 60 min	CYP12	sTWW-CYPA4		

Also, in Figure 1 it is observed that for the same gene in the same water sample, prevalence values tend to be identical irrespective of the treatment type, while abundance (per volume) varies (e.g *sul1* and *sul2* pilot-scale treatment) vary. Indeed, the prevalence of the genetic determinants analysed was, in most of the cases, non-significantly different in RWW, sTWW and/or tTWW ($p>0.05$). Moreover, calculation of prevalence values may be influenced by cumulative standard deviations of two determinations. And may be more problematic for low

abundance genes. This was observed, for example, in the sample CYP2 for the genes *qnrS* and *bla*$_{TEM}$ where a high standard deviation in the gene abundance promoted a high standard deviation in the prevalence values estimated for that sample. These observations lead to the conclusion that the assessment of water treatment efficiency can be reliably measured based on the reduction of the 16S rRNA gene and, if that is the option, other genes, but the results should be always expressed per volume (absolute abundance). However, although the gene abundance (16S rRNA or antibiotic resistance genes) informs about bacterial load, it provides limited insight about wastewater quality regarding antibiotic resistance. In this case, it is the ratio between the antibiotic resistance gene and the 16S rRNA gene that better characterizes the samples. For instance, in Figure 1 the pilot-scale treated wastewater samples, irrespective of treatment, presented always similar prevalence values. This observation suggests that the prevalence of antibiotic resistance in a wastewater may be fairly stable, no matter the treatment efforts that are applied. The consequent conclusion is that removal of bacteria and dilution with water or soils with lower antibiotic resistance prevalence may be favourable to reduce the resistance prevalence.

3.2. Possible antibiotic resistance markers

A second objective of this study was to compare the tested genetic determinants regarding the degree of information they can provide about wastewater treatment efficiency. In the pilot- as in the full-scale systems samples the abundance of genes could be ranked as 16S rRNA>*sul1*>*sul2*>*qnrS*>*bla*$_{TEM}$ (Figure1A-E, upper part). These results are in line with literature in which these genes are widely distributed in wastewater and they may persist even after wastewater treatment (Cacace *et al*., 2019; Fouz *et al*., 2020). The same hierarchization was observed for prevalence values - *sul1*>*sul2*>*qnrS*>*bla*$_{TEM}$ (Figure 1B-E, lower part). The results showed that the abundance of all the examined genetic determinants was lower in sTWW and tTWW compared to RWW ($p<0.05$), and no statistically significant differences were observed between sTWW and tTWW ($p>0.05$) (Figure 1A-E, upper part). These results suggest that the secondary treatment is the one causing the sharpest variations on genes abundance, which was also previously observed (Narciso-da-Rocha *et al*., 2018).

The selection of possible markers is useful to assess wastewater treatment efficiency. A suitable marker should be present in all samples and in high abundance to be determined even after treatment. These criteria were the basis for the selection of the genetic determinants that are being compared. The analysis of the results showed, however, that only 16S rRNA and *sul1* fulfilled these criteria as the other genes were not determined in some samples, probably due to poor DNA quality/concentration. The average removal values expressed in

log gene copy number / mL ranged 0.37 to 0.56 (Table 4), which are values extremely low when compared to what is described in the literature, normally above 1.5 (Narciso da Rocha *et al*., 2018; Pallares-Vega *et al*., 2019). This is probably because disinfection, that corresponded to most of the samples examined in this study, had a mild effect on antibiotic resistance genes elimination. However, disinfection may have jeopardized the quality of DNA, which concentration was very low in some samples (Table 3). And, consequently, genes that are in low relative abundance values may seem to be completely eliminated. This assumption may be misleading. Hence, the recommendation is that treatment efficiency and antibiotic resistance removal should be assessed based on at least 2-3 high abundance genes, such as *sul1* or the class 1 integron integrase gene *intI*1, a proxy of antibiotic resistance (Gillings *et al*., 2015).

3.3. Interlaboratory antibiotic resistance determinations

A third objective of this study was the assessment of the feasibility of comparing data produced in different laboratories and to identify critical factors that could bias the results. The set of data obtained allowed this discussion, because the samples and DNA extracts examined were obtained and analysed in different laboratories based on common protocols (PT7-9, DD1-3, CYP9-13) or the same sample was examined with own laboratory protocols (US1-2). Part of the DNA extracts were shipped (Table 1).

In this case study, samples from US were analysed in both laboratories.The samples followed the sequence – sample collection, processing, DNA extraction, and qPCR analysis in US. Then the DNA extracts (stored at – 80 ºC in the meanwhile) were shipped to Portugal and qPCR analysis of the same determinants were performed (Figure 2) and compared to the obtained in the US. The operator was the same and the protocols were those in use in each laboratory. Quantifications performed in both laboratories differed in values that ranged 0.0 log-units/mL for *sul1* gene in RWW1A sample, to 2.7 log-units for $bla_{\text{OXA-1}}$ gene quantification in RWW3A sample (Figure 2). In 22/60 determinations it was observed that gene quantifications were lower in PT than in US, possibly due to DNA degradation during shipment. Indeed, the quantifications of DNA in some shipped extracts were lower when measured after storage and shipment, mainly in raw wastewater where DNAses are more likely to be abundant and active (Table 5).

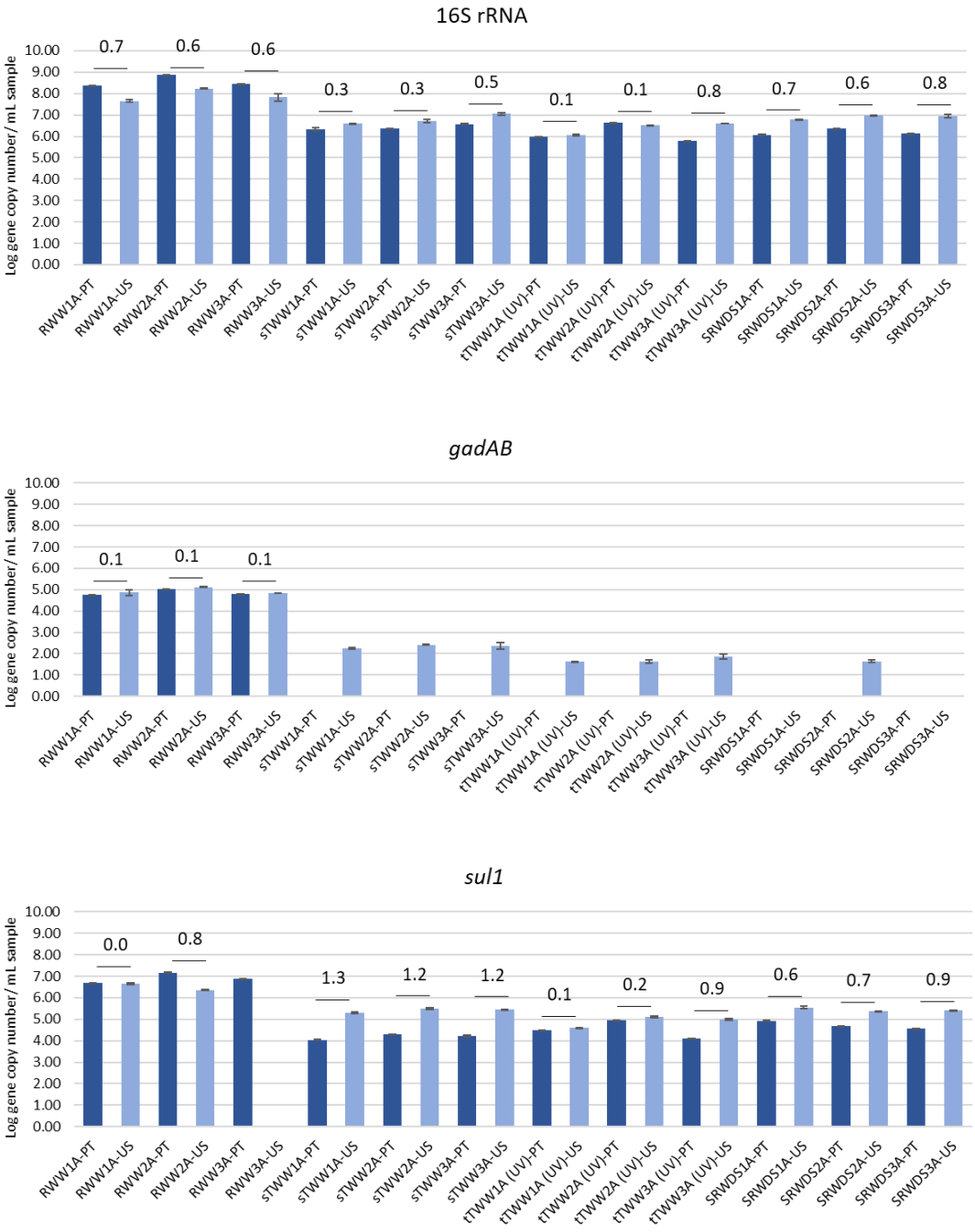
Table 5 – Wastewater samples types and the possible effect of the shipment on the DNA extracts concentration. The quantifications performed in both countries were done with Qubit.

Sample type	Replicates	DNA concentration determined in US (ng/µL)	DNA concentration determined in UCP (ng/µL)
RWW	1A	92.5	90.0
RWW	2A	315.0	210.0
RWW	3A	202.0	100.0
sTWW	1A	9.1	8.9
sTWW	2A	9.7	9.0
sTWW	3A	8.7	8.7
tTWW	1A	7.5	8.1
tTWW	2A	6.6*	20.2*
tTWW	3A	3.2	3.1
SRWDS	1A	1.5	1.7
SRWDS	2A	5.3	3.6
SRWDS	3A	2.4	2.5

*, possibly due to DNA solubilization

Seven of these situations (PT samples in *gadAB* gene quantification) led to the impossibility of detecting some genes that were below LOQ in PT while were measurable in US. The most evident situations were observed for the gene *gadAB* and *sul2* (Figure 2). The opposite, i.e. higher determinations in Portugal were also observed, specifically for the genes 16S rRNA gene, *sul1* or $bla_{\text{OXA-1}}$ in RWW. The fact that in some cases higher quantifications were observed after shipment might be related to the calibration curve or the qPCR equipment sensitivity. Distinct limits of quantification (per qPCR reaction) might contribute for some of the differences observed (e.g., 10 gene copies in the US and 43 gene copies in PT for *gadAB*

gene (Table 2). DNA degradation during shipment may also explain some of these results (Table 5).



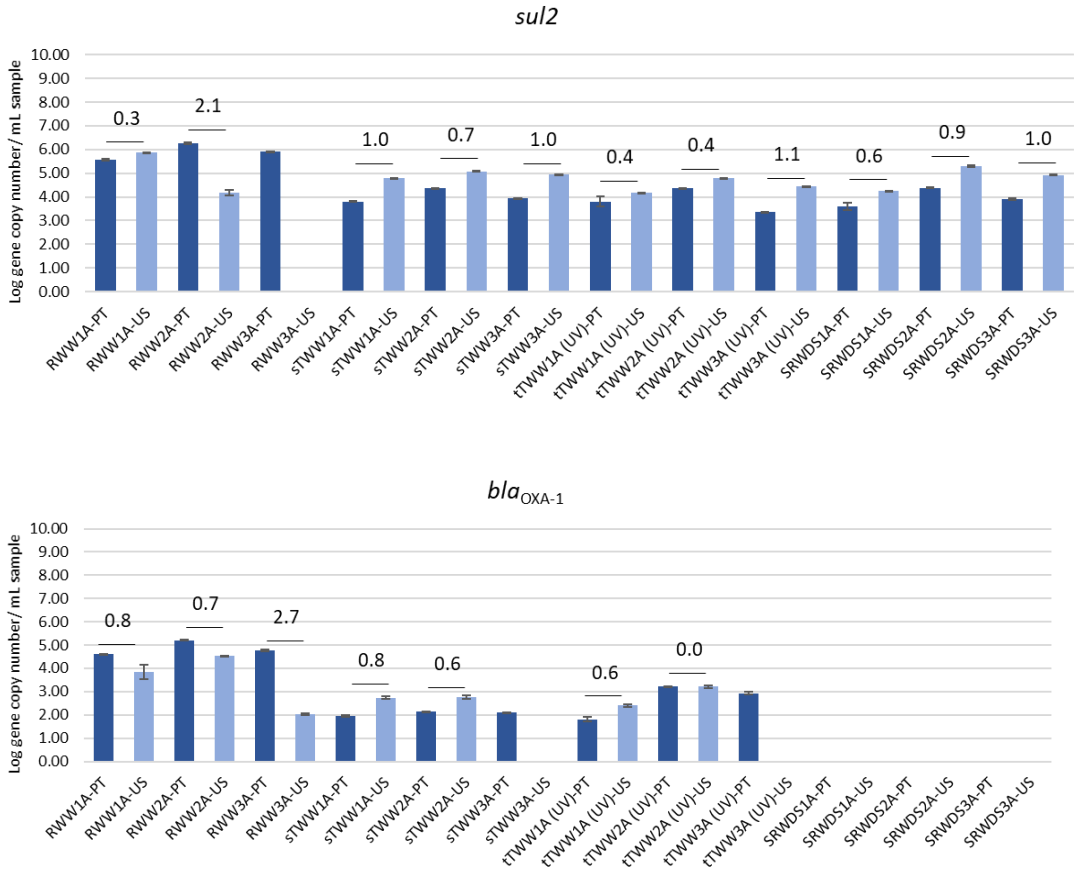


Figure 2 - Comparison of gene abundance (gene copy number/mL of sample) values obtained in Portugal and in the US in wastewater samples. RWW: wastewater treatment plant influent, sTWW: wastewater collected after secondary treatment, tTWW: wastewater collected after UV disinfection, and SRWDS: simulated reclaimed water distribution system. PT refers to quantification performed in Portugal, and US to data obtained in the US. The absolute difference values in the mean quantification of technical triplicates are indicated above the bars, whenever the target was quantifiable in both laboratories. The lack of gene quantification performed in Portugal for some samples was due to quantification below the limit of detection and the lack of gene quantification performed in US for some samples was due to quantification below the limit of detection (*bla*$_{OXA-1}$ and *sul2*) or due to quantification below the limit of quantification (*gadAB*). The gene copy numbers are listed in the Table 2.

In another study by Rocha *et al*. (2020) five partners located in different countries compared the variability of qPCR procedures for the quantification of antibiotic resistance genes using the same DNA extracts and qPCR protocols between different laboratories and concluded that

the equipment might be a factor introducing some discrepancy in the results. The present study aimed to take this comparison further, since it used distinct qPCR standards, protocols, and equipment, as would be the case in published literature from different labs. These results suggest that although cooperation between groups is always valuable, there are conditions to perform and compare data produced in different laboratories, as long as the margins of discrepancy are taken into account, as well as the attenuation of variables that may influence most the results. The thermal cycler and consumables may be critical in this aspect. Other factors may include the calibration curve, a difficulty that may be overcome by using a synthetic reference, and the DNA quality.

The results summarized in Table 3 show that DNA extracts obtained in laboratories that are not equipped/trained for molecular biology experiments (CYP and ES) and/or that DNA extracts that after tertiary treatment may have lower concentration. This fact may impose a serious bias in the analysis, mainly for the analysis of rare genes or when treatment efficiency is being analysed. Moreover, it can be aggravated by the DNA shipment, during which DNA degradation may occur. However, the comparison of the DNA concentration and of the gene quantification did not provide a clear evidence about degradation, at least in treated wastewater samples. DNA quality is of major relevance to assess treatment efficiency and poor-quality extracts will not affect the quantifications of abundant genes but may be critical for rarer genes (Table 6). All these observations suggest treatment efficiency may be overestimated in some cases. It is recommended that a cell-based internal standard is used to control losses during DNA extraction.

Table 6 - Simulation of the impact of DNA losses on gene quantification.

WATER (100 mL)	Losses (DNA extraction, shipment) (%)	DNA extract (100 µL)	qPCR reaction (2 µL)
1×10^6 gene copy	0	1×10^6	2×10^4
	50	5×10^5	1×10^4
	90	1×10^5	2×10^3
1×10^3 gene copy	0	1×10^3	2×10^1
	50	5×10^2	1×10^1
	90	1×10^2	2×10^0

4. Conclusions

It was concluded that the assessment of wastewater treatment efficiency can be reliably measured based on the reduction of the 16S rRNA gene and, if antibiotic resistance is under evaluation, also of some of the respective genes. Reasons to express the results per volume (absolute abundance) were provided and supported. Moreover, the 16S rRNA gene quantification was observed to be essential to assess the abundance of bacterial biomass in a sample, being recommended in any study to evaluate microbiological wastewater quality. Of note is the fact that antibiotic resistance prevalence, in contrast to abundance, presents modest variation, regardless the type of treatment used. Consequently, it is argued that total bacterial removal and dilution with microbial communities with lower antibiotic resistance prevalence may be favourable to reduce the resistance prevalence.

The selection of gene markers should accomplish specific requirements. Low abundance genes may be inadequate to assess treatment efficiency, although may be useful to assess wastewater quality. It is recommended that treatment efficiency and antibiotic resistance removal is assessed based on at least 2-3 high abundance genes, where a gene such as *sul1*, analysed in this study, should be included.

The comparison of data produced in different laboratories is a priority in the field of wastewater and antibiotic resistance. The results showed that such aim is feasible for assessing treatment efficiency and wastewater quality. However, the margins of discrepancy that are observed must be considered. The variables that may generate discrepancies must be controlled and attenuated. For instance, DNA extracts quality is critical, and the use of a cell-based internal standard might be useful to control losses in DNA extraction.

5. References

An, X.-L., Su, J.-Q., Li, B., Ouyang, W.-Y., Zhao, Y., Chen, Q.-L., Cui, L., Chen, H., Gillings, M. R., Zhang, T., & Zhu, Y.-G. (2018). Tracking antibiotic resistome during wastewater treatment using high throughput quantitative PCR. *Environment International*, 117, 146–153. https://doi.org/10.1016/j.envint.2018.05.011

Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M. N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., & Martinez, J. L. (2015). Tackling antibiotic resistance: The environmental framework. *Nature Reviews Microbiology*, 13(5), 310–317. https://doi.org/10.1038/nrmicro3439

Bibbal, D., Dupouy, V., Ferré, J. P., Toutain, P. L., Fayet, O., Prère, M. F., & Bousquet-Mélou, A. (2007). Impact of three ampicillin dosage regimens on selection of ampicillin resistance in *Enterobacteriaceae* and excretion of bla_{TEM} genes in swine feces. *Applied and Environmental Microbiology*, 73(15), 4785–4790. https://doi.org/10.1128/AEM.00252-07

Böckelmann, U., Dörries, H. H., Ayuso-Gabella, M. N., De Marçay, M. S., Tandoi, V., Levantesi, C., Masciopinto, C., Houtte, E. Van, Szewzyk, U., Wintgens, T., & Grohmann, E. (2009). Quantitative PCR monitoring of antibiotic resistance genes and bacterial pathogens in three European artificial groundwater recharge systems. *Applied and Environmental Microbiology*, 75(1), 154–163. https://doi.org/10.1128/AEM.01649-08

Brankatschk, R., Bodenhausen, N., Zeyer, J., & Burgmann, H. (2012). Simple absolute quantification method correcting for quantitative PCR efficiency variations for microbial community samples. *Applied and Environmental Microbiology*, 78(12), 4481–4489. https://doi.org/10.1128/AEM.07878-11

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Cacace, D., Fatta-Kassinos, D., Manaia, C. M., Cytryn, E., Kreuzinger, N., Rizzo, L., Karaolia, P., Schwartz, T., Alexander, J., Merlin, C., Garelick, H., Schmitt, H., de Vries, D., Schwermer, C. U., Meric, S., Ozkal, C. B., Pons, M. N., Kneis, D., & Berendonk, T. U. (2019). Antibiotic resistance genes in treated wastewater and in the receiving water

bodies: A pan-European survey of urban settings. *Water Research*, 162, 320–330. https://doi.org/10.1016/j.watres.2019.06.039

Chen, Y. C., Higgins, M. J., Maas, N. A., & Murthy, S. N. (2006). DNA extraction and *Escherichia coli* quantification of anaerobically digested biosolids using the competitive touchdown PCR method. *Water Research*, 40(16), 3037–3044. https://doi.org/10.1016/j.watres.2006.06.020

Denman, S. E., & McSweeney, C. S. (2006). Development of a real-time PCR assay for monitoring anaerobic fungal and cellulolytic bacterial populations within the rumen. *FEMS Microbiology Ecology*, 58(3), 572–582. https://doi.org/10.1111/j.1574-6941.2006.00190.x

Di Cesare, A., Corno, G., Manaia, C. M., & Rizzo, L. (2020). Impact of disinfection processes on bacterial community in urban wastewater: Should we rethink microbial assessment methods? *Journal of Environmental Chemical Engineering*, 8(5), 1–12. https://doi.org/10.1016/j.jece.2020.104393

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., Goldsmith, D. B., Michisaki, R., Breitbart, M., Boehm, A. B., & Chavez, F. P. (2017). Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Frontiers in Marine Science*, 4, 1–11. https://doi.org/10.3389/fmars.2017.00314

Du, J., Ren, H., Geng, J., Zhang, Y., Xu, K., & Ding, L. (2014). Occurrence and abundance of tetracycline, sulfonamide resistance genes, and class 1 integron in five wastewater treatment plants. *Environmental Science and Pollution Research*, 21(12), 7276–7284. https://doi.org/10.1007/s11356-014-2613-5

Fouz, N., Pangesti, K. N. A., Yasir, M., Al-Malki, A. L., Azhar, E. I., Hill-Cawthorne, G. A., & El Ghany, M. A. (2020). The contribution of wastewater to the transmission of antimicrobial resistance in the environment: Implications of mass gathering settings. *Tropical Medicine and Infectious Disease*, 5(1), 1–25. https://doi.org/10.3390/tropicalmed5010033

Gao, M., Qiu, T., Sun, Y., & Wang, X. (2018). The abundance and diversity of antibiotic resistance genes in the atmospheric environment of composting plants. *Environment International*, 116(9), 229–238. https://doi.org/10.1016/j.envint.2018.04.028

Gao, P., Munir, M., & Xagoraraki, I. (2012). Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional

municipal wastewater treatment plant. *Science of the Total Environment*, 421–422, 173–183. https://doi.org/10.1016/j.scitotenv.2012.01.061

Gillings, M. R., Gaze, W. H., Pruden, A., Smalla, K., Tiedje, J. M., & Zhu, Y. G. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME Journal*, 9(6), 1269–1279. https://doi.org/10.1038/ismej.2014.226

Hinlo, R., Gleeson, D., Lintermans, M., & Furlan, E. (2017). Methods to maximise recovery of environmental DNA from water samples. *PLoS ONE*, 12(6), 1–22. https://doi.org/10.1371/journal.pone.0179251

ISO 7899. (2000). Water quality – Detection and enumeration of intestinal Enterococci. International Organization for Standardization, Geneva, Switzerland.

ISO 9308. (2014). Water quality – Enumeration of *Escherichia coli* and coliform bacteria. International Organization for Standardization, Geneva, Switzerland.

Kim, J., Lim, J., & Lee, C. (2013). Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. *Biotechnology Advances*, 31(8), 1358–1373. https://doi.org/10.1016/j.biotechadv.2013.05.010

Klein, D. (2002). Quantification using real-time PCR technology: applications and limitations. *Trends in Molecular Medicine*, 8(6), 257–260. https://doi.org/10.1016/S1471-4914(02)02355-9

Li, A. D., Metch, J. W., Wang, Y., Garner, E., Zhang, A. N., Riquelme, M. V., Vikesland, P. J., Pruden, A., & Zhang, T. (2018). Effects of sample preservation and DNA extraction on enumeration of antibiotic resistance genes in wastewater. *FEMS Microbiology Ecology*, 94(2), 1–11. https://doi.org/10.1093/femsec/fix189

Manaia, C. M., Macedo, G., Fatta-Kassinos, D., & Nunes, O. C. (2016). Antibiotic resistance in urban aquatic environments: can it be controlled? *Applied Microbiology and Biotechnology*, 100(4), 1543–1557. https://doi.org/10.1007/s00253-015-7202-0

Manaia, C. M., Rocha, J., Scaccia, N., Marano, R., Radu, E., Biancullo, F., Cerqueira, F., Fortunato, G., Iakovides, I. C., Zammit, I., Kampouris, I., Vaz-Moreira, I., & Nunes, O. C. (2018). Antibiotic resistance in wastewater treatment plants: Tackling the black box. *Environment International*, 115, 312–324. https://doi.org/10.1016/j.envint.2018.03.044

Marti, E., & Balcázar, J. L. (2013). Real-time PCR assays for quantification of *qnr* genes in environmental water samples and chicken feces. *Applied and Environmental Microbiology*, 79(5), 1743–1745. https://doi.org/10.1128/AEM.03409-12

Michael, S. G., Michael-Kordatou, I., Nahim-Granados, S., Polo-López, M. I., Rocha, J., Martínez-Piernas, A. B., Fernández-Ibáñez, P., Agüera, A., Manaia, C. M., & Fatta-Kassinos, D. (2020). Investigating the impact of UV-C/H_2O_2 and sunlight/H_2O_2 on the removal of antibiotics, antibiotic resistance determinants and toxicity present in urban wastewater. *Chemical Engineering Journal*, 388, 124383. https://doi.org/10.1016/j.cej.2020.124383

Muyzer, G., De Waal, E. C., & Uitterlinden, A. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*, 59(3), 695–700.

Nadkarni, M. A., Martin, F. E., Jacques, N. A., & Hunter, N. (2002). Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology*, 148(1), 257–266. https://doi.org/10.1099/00221287-148-1-257

Narciso-da-Rocha, C., Rocha, J., Vaz-Moreira, I., Lira, F., Tamames, J., Henriques, I., Martinez, J. L., & Manaia, C. M. (2018). Bacterial lineages putatively associated with the dissemination of antibiotic resistance genes in a full-scale urban wastewater treatment plant. *Environment International*, 118, 179–188. https://doi.org/10.1016/j.envint.2018.05.040

Narciso-Da-Rocha, C., Varela, A. R., Schwartz, T., Nunes, O. C., & Manaia, C. M. (2014). *bla*$_{TEM}$ and *vanA* as indicator genes of antibiotic resistance contamination in a hospital-urban wastewater treatment plant system. *Journal of Global Antimicrobial Resistance*, 2(4), 309–315. https://doi.org/10.1016/j.jgar.2014.10.001

Nguyen, A. Q., Vu, H. P., Nguyen, L. N., Wang, Q., Djordjevic, S. P., Donner, E., Yin, H., & Nghiem, L. D. (2021). Monitoring antibiotic resistance genes in wastewater treatment: Current strategies and future challenges. *Science of the Total Environment*, 783, 146964. https://doi.org/10.1016/j.scitotenv.2021.146964

Novo, A., André, S., Viana, P., Nunes, O. C., & Manaia, C. M. (2013). Antibiotic resistance, antimicrobial residues and bacterial community composition in urban wastewater. *Water Research*, 47(5), 1875–1887. https://doi.org/10.1016/j.watres.2013.01.010

Pallares-Vega, R., Blaak, H., van der Plaats, R., de Roda Husman, A. M., Hernandez Leal, L., van Loosdrecht, M. C. M., Weissbrodt, D. G., & Schmitt, H. (2019). Determinants of presence and removal of antibiotic resistance genes during WWTP treatment: A cross-sectional study. *Water Research*, 161, 319–328. https://doi.org/10.1016/j.watres.2019.05.100

Pärnänen, K. M. M., Narciso-Da-Rocha, C., Kneis, D., Berendonk, T. U., Cacace, D., Do, T. T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J. L., Michael, S. G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., … Manaia, C. M. (2019). Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances*, 5(eaau9124), 1–10. https://doi.org/10.1126/sciadv.aau9124

Pei, R., Kim, S. C., Carlson, K. H., & Pruden, A. (2006). Effect of river landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Research*, 40(12), 2427–2435. https://doi.org/10.1016/j.watres.2006.04.017

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., Michael, I., & Fatta-Kassinos, D. (2013). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of the Total Environment*, 447, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Rocha, J., Cacace, D., Kampouris, I., Guilloteau, H., Jäger, T., Marano, R. B. M., Karaolia, P., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Berendonk, T. U., & Schwartz, T. (2020). Inter-laboratory calibration of quantitative analyses of antibiotic resistance genes. *Journal of Environmental Chemical Engineering*, 8(1), 102214. https://doi.org/10.1016/j.jece.2018.02.022

Rocha, J., Fernandes, T., Riquelme, M. V., Zhu, N., Pruden, A., & Manaia, C. M. (2019). Comparison of culture-and quantitative PCR-based indicators of antibiotic resistance in wastewater, recycled water, and tap water. *International Journal of Environmental Research and Public Health*, 16(21). https://doi.org/10.3390/ijerph16214217

Rodríguez-Chueca, J., Varella della Giustina, S., Rocha, J., Fernandes, T., Pablos, C., Encinas, Á., Barceló, D., Rodríguez-Mozaz, S., Manaia, C. M., & Marugán, J. (2019). Assessment of full-scale tertiary wastewater treatment by UV-C based-AOPs: Removal or persistence of antibiotics and antibiotic resistance genes? *Science of the Total Environment*, 652, 1051–1061. https://doi.org/10.1016/j.scitotenv.2018.10.223

Rodriguez-Mozaz, S., Chamorro, S., Marti, E., Huerta, B., Gros, M., Sànchez-Melsió, A., Borrego, C. M., Barceló, D., & Balcázar, J. L. (2015). Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Research*, 69, 234–242. https://doi.org/10.1016/j.watres.2014.11.021

Suzuki, M. T., Taylor, L. T., & DeLong, E. F. (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Applied and Environmental Microbiology*, 66(11), 4605–4614. https://doi.org/10.1128/AEM.66.11.4605-4614.2000

Valasek, M. A., & Repa, J. J. (2005). The power of real-time PCR. *American Journal of Physiology - Advances in Physiology Education*, 29(3), 151–159. https://doi.org/10.1152/advan.00019.2005

Varela, A. R., Macedo, G. N., Nunes, O. C., & Manaia, C. M. (2015). Genetic characterization of fluoroquinolone resistant *Escherichia coli* from urban streams and municipal and hospital effluents. *FEMS Microbiology Ecology*, 91(5), 1–12. https://doi.org/10.1093/femsec/fiv015

Yang, Y., Zhang, T., Zhang, X. X., Liang, D. W., Zhang, M., Gao, D. W., Zhu, H. G., Huang, Q. G., & Fang, H. H. P. (2012). Quantification and characterization of β-lactam resistance genes in 15 sewage treatment plants from East Asia and North America. *Applied Microbiology and Biotechnology*, 95(5), 1351–1358. https://doi.org/10.1007/s00253-011-3810-5

CHAPTER 4

CELL-BASED INTERNAL STANDARD FOR QPCR DETERMINATIONS OF ANTIBIOTIC RESISTANCE INDICATORS IN ENVIRONMENTAL WATER SAMPLES

Ecological Indicators 113 (2020) 106194

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Cell-based internal standard for qPCR determinations of antibiotic resistance indicators in environmental water samples

Jaqueline Rocha, Célia M. Manaia*

Universidade Católica Portuguesa, CBQF – Centro de Biotecnologia e Química Fina – Laboratório Associado, Escola Superior de Biotecnologia, Rua de Diogo Botelho 1327, 4169-005 Porto, Portugal

ARTICLE INFO

Keywords:
Quantitative PCR
DNA extraction
Internal standard
Matrix effect

ABSTRACT

Quantitative PCR (qPCR) has been used to quantify antibiotic resistance genes (ARGs) in water, wastewater, soil, sediment and tissue samples. Concerns regarding the comparability of data obtained in different laboratories has been a major bottleneck to incentivize the compilation of publicly available of ARGs quantifications gathered from different reports. In this study, the influence of the DNA extraction kits (NZY Tissue gDNA Isolation kit or DNeasy PowerWater kit) and of the operator on the DNA extraction yield and on qPCR genes quantification was assessed. Since in wastewater and water samples the matrix effect can affect the DNA recovery and, therefore, gene quantification, an internal standard, consisting in a cloned gene not found in environmental samples, was tested. The aim was to assess how qPCR determinations in wastewater and water samples can be affected by the matrix effect. The results show that the DNA extraction operator did not significantly influence DNA yield. The use of distinct kits resulted in qPCR gene quantifications that did not differ in more than 1 log-unit mL^{-1}. The matrix effect, assessed based on the use of an internal standard, was associated with an underestimation that ranged 0.1–0.9 log gene copy number mL^{-1} of sample, irrespective of the water type.

The reliability on the use of a DNA extraction kit that costs about 3 times less than the most commonly used can be an incentive for the use of DNA based analyses of ARGs in environmental waters. Moreover, the fact that both the DNA extraction operator and the reduced matrix effect have little influence on the final results, are good news, encouraging the compilation of data produced in distinct laboratories. Nevertheless, harmonization efforts are still necessary to minimize bias that may be due associated with other conditions, such as equipment.

1. Introduction

The environmental contamination with antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARGs) is widely recognized and known to assume concerning proportions in aquatic environments (Allen et al., 2010; Fatta-Kassinos et al., 2011; Kolpin et al., 2002; Kümmerer, 2009; Martinez, 2009). This situation represents a public health issue, requiring urgent measures that allow the combination of monitoring efforts and implementation of control processes (Berendonk et al., 2015; Manaia et al., 2016). The monitoring of ARB and their ARGs in the environment can be performed based on culture-dependent or culture-independent methods (Manaia et al., 2018, Manaia et al., 2016). While culture-dependent methods have the advantage of being comparable when involve the use of routine and directive-oriented procedures (ISO 7899; ISO 9308; Drinking water directive, (98/83/EC)), they have the limitations that are not designed for ARB enumeration and leave non-culturable bacteria aside the analyses. Therefore, culture-independent methods are considered an essential complement or alternative to assess the quality and safety of water environments in terms of antibiotic resistance occurrence (Manaia et al., 2018, 2016; Vartoukian et al., 2010). Quantitative PCR (qPCR) has been, in this aspect, one of the methods of choice (Kim et al., 2013; Klein, 2002; Valasek, 2005). As for other quantitative methods, an adequate implementation of qPCR involves the use of identical conditions in all assays, an objective that may be difficult to reach, given the diversity of operators, samples to analyse, reagents used, among other.

Abbreviations: ARGs, antibiotic resistance genes; ASLP, activated sludge from a municipal wastewater treatment plant; *EcmolA* +, *Escherichia coli* JM109 cloned with the plasmid containing the internal standard (*molA*); HE, hospital effluent; NZY, NZY Tissue gDNA Isolation kit; PW, DNeasy PowerWater kit; qPCR, quantitative PCR; RWWA, influent wastewater from an airport wastewater treatment plant; RWWP, influent wastewater from a municipal wastewater treatment plant; sTWWA, secondary treatment effluent from an airport wastewater treatment plant; sTWWP, secondary treatment effluent from a municipal wastewater treatment plant; UP, ultrapure water

* Corresponding author.
E-mail address: cmanaia@porto.ucp.pt (C.M. Manaia).

https://doi.org/10.1016/j.ecolind.2020.106194
Received 30 October 2019; Received in revised form 2 February 2020; Accepted 5 February 2020

In a previous study using common DNA extracts in a ring test involving five laboratories, Rocha et al. (2018) demonstrated that ARGs quantification in wastewater samples varied up to 28%, which could be attributed to a combination of the real-time thermal cycler, the respective operator or the reagents batch. In the current study, we aimed at assessing other potential influential variables. For example, the DNA extraction method, which can influence the yield or the occurrence of potential qPCR interfering agents (e.g. humic acids, heavy metals, phenolic compounds) (Foerstner et al., 2005; Venter et al., 2004), may have implications on the accurate DNA amplification and gene quantification (Bessetti, 2007). These and other potential biases may hamper the comparison of data worldwide or even in the same laboratory at different time scales (Manaia et al., 2016; Smith and Osborn, 2009).

Studies comparing the effect of filtering membrane (Djurhuus et al., 2017; Hinlo et al., 2017); sample preservation (Hinlo et al., 2017; Li et al., 2017); DNA extraction methods (Djurhuus et al., 2017; Hinlo et al., 2017; Li et al., 2017); among others, have contributed to better understand the influence of water sample processing on the final results. Nevertheless, the sample processing and qPCR protocols used vary considerably worldwide (Rocha et al., 2018). The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin, 2010; Bustin et al., 2009), aim at improving qPCR reliability, reproducibility and comparability of data. However, MIQE guidelines are focused on qPCR analysis and not on sample processing. In this context, this work was designed to assess the influence of DNA extraction kits and operators on DNA yield and on gene qPCR quantification and also to quantify the bias imposed by the matrix effect of the water sample, by using an internal standard. To our knowledge this is the first study that reports the use of a bacterial cell-based internal standard for the determination of the matrix effect of water samples on DNA extraction and qPCR determination. The evaluation of the usefulness of the internal standard to monitor ARGs in water and wastewater samples was motivated by the internationally claimed need of surveillance of antibiotic resistance to be implemented at critical control points, which include wastewater treatment plants, untreated wastewater sources (hospital effluents, areas of poor sanitation infrastructure), and water bodies impacted by these sources (Berendonk et al. 2015; Manaia et al., 2016; Huijbers et al., 2019). The results of this study aim at contributing to increase the body of knowledge that may facilitate and support the comparison of data obtained under distinct conditions in different laboratories or time scales.

2. Material and methods

2.1. Samples

Seven types of water and wastewater samples were tested in this study (Table S1). Samples were collected from the influent and secondary wastewater treatment effluent (RWWA and sTWWA) from an airport wastewater treatment plant and from the influent, secondary wastewater treatment effluent and activated sludge samples (RWWP, sTWWP and ASLP) from a municipal a wastewater treatment plant. Both wastewater treatment plants operated with primary and conventional activated sludge secondary treatments. Other samples comprised the hospital effluent and river water. All samples were collected in the Northern region of Portugal. The samples were collected in sterilized bottles, transported to the lab in refrigerated conditions and immediately processed.

2.2. DNA extraction

Total DNA was extracted from sample volumes varying from 25 mL to 100 mL, processed in triplicate. Samples were filtered through polycarbonate membranes (0.22 μm porosity, Whatman, UK) and stored at −80 °C until DNA extraction. For DNA kit extraction comparison were used the NZY Tissue gDNA Isolation kit (Nzytech, Lisbon, Portugal) and the DNeasy PowerWater kit (QIAGEN, Hilden, Germany) that at the time of the work performance was known as PowerWater DNA Isolation Kit and was commercialized by MO BIO Laboratories Inc., CA, USA. For simplicity, from this point forward, the most recent designation will be as synonymous of the previous. The approximate costs per reaction of the DNeasy PowerWater kit and the NZY Tissue gDNA Isolation kit, were 9 € and 3 €, respectively. For this comparison, four environmental and four ultrapure water-internal standard spiked samples, extracted by two different operators, were analysed. To assess the matrix effect, using the internal standard described below, a total of 14 environmental and 14 spiked ultrapure water were extracted using the DNeasy PowerWater kit) (Table S1). Both kits were used according to manufacturer instructions with the following exceptions: in the DNeasy PowerWater kit the lysis period of time was increased from 5 min (as recommended) to 15 min; an extra centrifugation of 30 s was performed before DNA elution, and DNA was eluted twice with 50 μL of elution buffer warmed at 55 °C, to increase DNA recovery; in the NZY Tissue gDNA Isolation kit, before DNA extraction procedure the membranes were inserted into a 2 mL tube to which were added 360 μL of NT1 buffer and 50 μL of proteinase K, the volume of these two reagents was increased to ensure that the polycarbonate membrane used to concentrate the biomass in the water sample was covered by the lysis solutions, and DNA was also eluted twice with 50 μL of elution buffer warmed at 55 °C. The concentration of the DNA extracts was determined using Qubit (Thermo Fisher Scientific, USA). DNA extracts were preserved at −20 °C until their use for quantitative PCR analyses.

2.3. Genes quantification using qPCR

The quantitative PCR assays targeting five chromosomal/housekeeping genes (16S rRNA, *marA*, *rpoB*, *uidA* and *ecf*), eight ARGs ($bla_{\text{CTX-M}}$, $bla_{\text{OXA-A}}$, bla_{SHV}, bla_{TEM}, bla_{IMP}, bla_{VIM}, *qnrS*, *sul1*), one gene encoding the class 1 integrons integrase related with horizontal gene transfer (*intI1*) and one internal standard (*molA*) gene used the primers and conditions listed in Table S2. The chromosomal/housekeeping genes were selected as a measurement of total bacteria (16S rRNA); *Enterobacteriaceae* (*marA*), *Escherichia coli* (*uidA*); *Pseudomonas aeruginosa* (*ecf*); and *Acinetobacter* spp. (*rpoB*) due to their association with water environments, including wastewater, and with humans (Atrouni et al., 2016; Castiglioni et al., 2008; Jang et al., 2017; Mena and Gerba, 2009). The ARGs encoding resistance to the β-lactams ($bla_{\text{CTX-M}}$, $bla_{\text{OXA-A}}$, bla_{SHV}, bla_{TEM}, bla_{IMP}, bla_{VIM}), fluoroquinolones (*qnrS*) and sulphonamides (*sul1*) and the mobile genetic elements-related encoding gene (*intI1*) were selected based on their common occurrence in domestic wastewater and widespread in environmental compartments, correlation with anthropogenic pollution, and on the fact they have been reported as clinically relevant genes (Du et al., 2014; Gillings et al., 2015; Narciso-da-Rocha et al., 2018; Narciso-Da-Rocha et al., 2014; Szczepanowski et al., 2009; Varela et al., 2016; Walsh et al., 2005; Zhang et al., 2009). Quantifications were made based on the Standard Curve method as described in Brankatschk et al. (2012) using a StepOne™ Real-Time PCR System (Life Technologies, Carlsbad, CA, USA). The qPCR results, based on three independent DNA extracts for each sample, were analysed according to the quality and acceptability criteria described in Rocha et al. (2018). Briefly, the criteria used were: standard curve efficiency between 90 and 110%; Ct values in the test samples could be interpolated in the standard curve; identical amplicon melting temperature in reference (used for the standard curve) and test samples; observation of a single and correct melting point; and absence of shoulders (increased signal in the baseline, e.g. due to primer dimers).

2.4. Internal standard

The criteria to select an internal standard to assess the matrix effect was the use of a bacterial species that might mimic the behaviour of

Table 1
DNA yield obtained with distinct DNA extraction kits or operators. DNA concentration yield obtained by two operators using the DNA extraction kits NZY Tissue gDNA Isolation kit (NZY) or DNeasy PowerWater kit (PW). UP – ultrapure water; HE – hospital effluent. All samples were spiked with *EcmolA+* .

		NZY kit	PW kit
Sample	Sample volume (mL)	DNA concentration (ng μL^{-1}) (Average ± Standard deviation)	DNA concentration (ng μL^{-1}) (Average ± Standard deviation)
UP1 – Operator 1 (n = 3)	50	3.0 ± 0.1	2.0 ± 0.8
UP1 – Operator 2 (n = 3)		3.3 ± 0.1	2.2 ± 0.4
UP2 – Operator 1 (n = 3)		4.4 ± 0.6	4.5 ± 3.3
UP2 – Operator 2 (n = 3)		4.6 ± 0.8	3.3 ± 1.6
UP3 – Operator 1 (n = 3)		3.6 ± 0.6	3.6 ± 0.9
UP3 – Operator 2 (n = 3)		4.1 ± 1.8	1.8 ± 0.1
UP4 – Operator 1 (n = 3)		4.2 ± 0.6	2.7 ± 0.4
UP4 – Operator 2 (n = 3)		4.3 ± 0.9	4.1 ± 3.9
River1 – Operator 1 (n = 3)		4.2 ± 1.0*	16.0 ± 2.6*
River1 – Operator 2 (n = 3)		4.8 ± 1.3*	11.3 ± 1.7*
River3 – Operator 1 (n = 3)		4.4 ± 0.8*	13.4 ± 0.8*
River3 – Operator 2 (n = 3)		5.9 ± 1.2	9.6 ± 1.9
HE2 – Operator 1 (n = 3)		17.4 ± 3.7	16.5 ± 4.5
HE2 – Operator 2 (n = 3)		13.2 ± 4.9	42.3 ± 10.8
HE4 – Operator 1 (n = 3)		11.6 ± 5.7	36.7 ± 16.3
HE4 – Operator 2 (n = 3)		28.1 ± 20.9	30.9 ± 20.3

*Represent statistically significant differences observed ($p < 0.01$) for the same operator between kits based on independent-samples T-test.

other water and wastewater bacteria during DNA extraction and the presence of a gene that is not expected to be found in these environments. To fulfil these criteria, was selected the *Escherichia coli* strain JM109 (NZYTech, Lisbon, Portugal), transformed with the internal standard gene *molA*. The gene *molA* (accession no. FN985594) of *Gulosibacter molinativorax* ON4^T strain encodes a molinate hydrolase, which is involved in the degradation of the herbicide molinate, used for the control of barnyard grass in paddy fields (Duarte et al., 2011; Lopes et al., 2013; Nunes et al., 2013). A conventional PCR targeting a *molA* amplicon with 1069 bp was performed using the primers F10 (5′-ACG ATCGCGATTGTCGGCGG-3′) and R1070 (5′–GGAGTTCACCCTGGGAC ATA −3′). The amplification was performed in a reaction volume of 50 µL with 2x KCl buffer, 2x $(NH_4)_2SO_4$ buffer, 0.25 mM $MgCl_2$, 2.5 mM dNTP mix (Thermo Scientific, USA), 2.5 µL dimethylsulfoxide (Applichem, Germany), 2.5 µM each primer, 7.5 U of Taq polymerase (Thermo Scientific, USA) and 4 µL of template DNA. The PCR conditions were 5 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 55 °C, 1.45 min at 72 °C, and a final extension of 10 min at 72 °C. A negative control reaction, without template DNA was included. The amplicon obtained was cloned with the plasmid pTZ57R/T using the InsTAclone PCR Cloning Kit #K1214 (Thermo Scientific, CA, USA), in the commercial strain *Escherichia coli* JM109, from this point forward called *EcmolA+*. *EcmolA+* was grown in Luria Broth medium (Invitrogen, CA, USA) supplemented with ampicillin (50 mg mL^{-1}), at 37 °C overnight and plasmid DNA with the *molA* insert was extracted using the commercial kit GeneJET Plasmid Miniprep K0503 (Thermo Scientific, CA, USA) and purified for further use as qPCR standard.

To determine the optimal dose of *EcmolA+* to be used as internal standard, 6, 7, 8, 9 and 10 log-units of *EcmolA+* fresh culture suspensions were prepared. A volume of 1 mL of each suspension was spiked into 1 L of ultrapure water and environmental water samples from river and secondary treated effluent wastewater. The final doses of *EcmolA+* in UP and in the samples were of 3, 4, 5, 6 and 7 log-units of *EcmolA+* mL^{-1}. Based on this preliminary test, an optimal final dose of 6 log-units *EcmolA+* mL^{-1} was selected and the same procedure was applied to a total of 14 environmental water samples (Table S1). This internal standard dose was determined since higher doses of *EcmolA+* (7 log-units *EcmolA+* mL^{-1}) would result in an overestimation of the 16S rRNA gene in samples. Lower doses (3 to 5 log-units *EcmolA+* mL^{-1}) could lead to internal standard (*molA*) abundance below the limit of quantification or to unspecific amplification of the gene, with non-compliance with the quality criteria (e.g. amplicon melting curves with additional melting peaks or shoulders). Ultrapure water was spiked with the same dose of *EcmolA+* used in environmental samples. Viable *EcmolA+* spiked in ultrapure water were enumerated on Plate Count Agar (PCA, Liofilchem, Roseto degli Abruzzi, Italy).

The gene *molA* was analysed in spiked and non-spiked with the internal standard. The abundance of the *molA* gene recovered in ultrapure water was compared with the abundance of the *molA* gene recovered in field water samples spiked and used to determine the internal standard losses due to sample matrix effect, which were calculated as follows: A = log (*molA* gene copy number/mL of spiked ultrapure water) – log (*molA* gene copy number/mL of spiked water sample) and B = A/log (*molA* gene copy number/mL of spiked ultrapure water) × 100.

2.5. Statistical analyses

Independent-samples T-test was used to assess statistically significant differences between 1) the yield of DNA obtained using two different DNA extraction kits and between two operators ($p < 0.01$), 2) the copy number of target genes using DNA extracts obtained by two operators ($p < 0.01$), 3) the matrix effect between both kits, 4) the original and the corrected copy number of target genes ($p < 0.01$). One-way analysis of variance (ANOVA) and Tukey's and Bonferroni post-hoc tests were used to assess statistically significant differences ($p < 0.01$) between the *molA* losses due to matrix effect in different water samples. These analyses were performed with the aid of the SPSS Statistics (for Windows v.24.0; IBM Corp., Armonk, NY, USA).

3. Results

3.1. DNA extraction kits effect on DNA yield and on genes quantification

Two DNA extraction kits were compared in terms of DNA extraction yield, gene quantification and inter-operator reproducibility. One was the NZY Tissue gDNA Isolation kit, recommended for a variety of matrices, including animal cells and tissues, Gram-positive and Gram-negative bacteria, mouse tails, yeast, forensic samples and clinical samples and the other was the DNeasy PowerWater kit, recommended for filtered water samples, even water containing heavy amounts of contaminants. These comparative assays involved spiked-ultrapure water, river and hospital effluent samples. The DNA concentrations in the extracts obtained using the DNeasy PowerWater kit and the NZY Tissue gDNA Isolation kit are summarized in Table 1. In most samples, the DNA concentration yield was not statistically different for both DNA

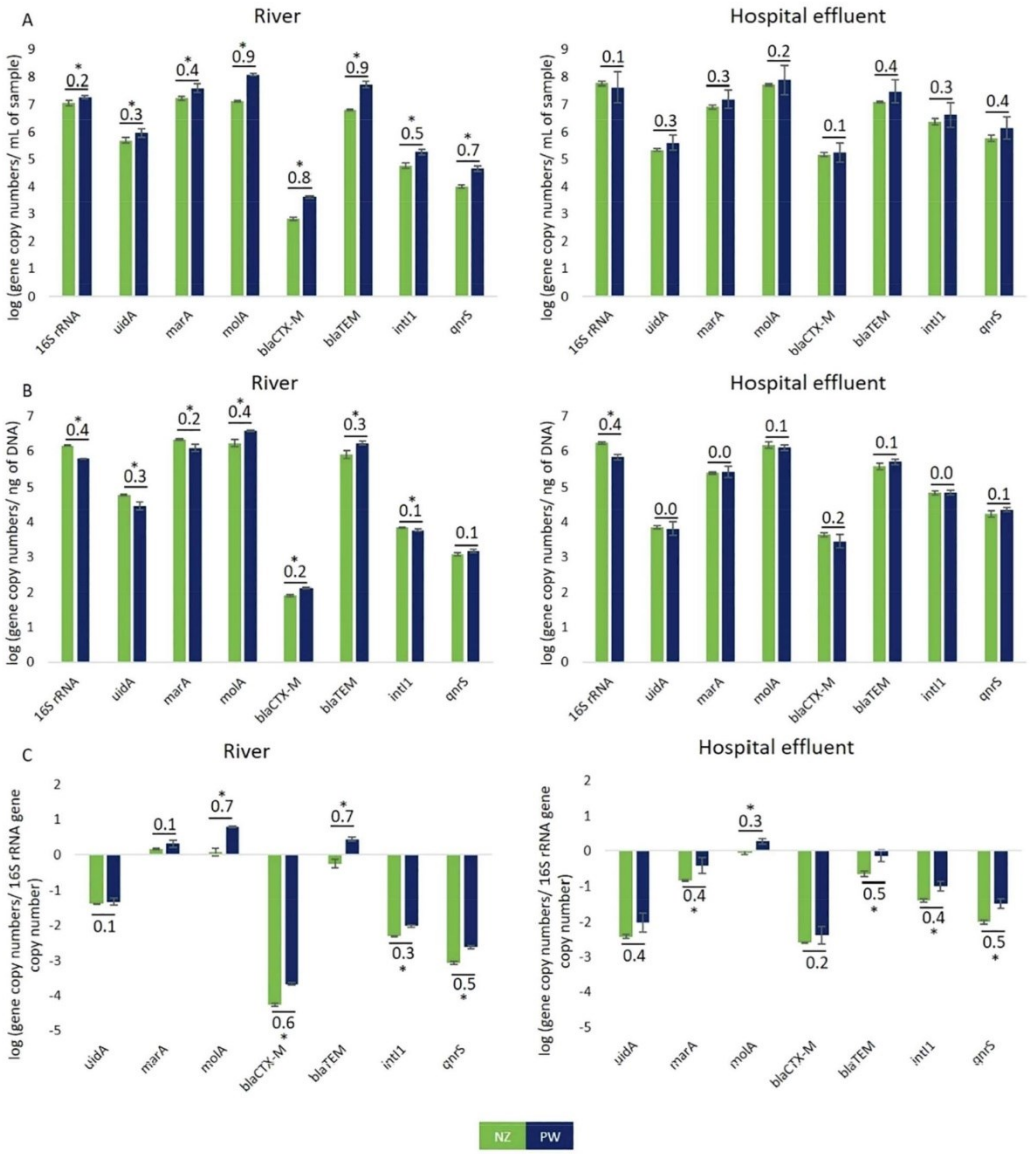


Fig. 1. Comparison of gene quantification in DNA extracts prepared with two different kits. Chromosomal (16S rRNA gene, *uidA*, *marA*), putatively plasmid associated (bla_{CTX-M}, bla_{TEM}, *intI*1, and *qnrS*), and internal standard (*molA*) genes quantification in DNA extracts prepared with the NZY Tissue gDNA Isolation kit (NZ) and DNeasy PowerWater kit (PW). These results refer to environmental water samples that were spiked with the internal standard *molA* gene. The internal standard *molA* gene was also monitored in non-spiked environmental water samples, and it was confirmed its absence in environmental samples. Results are expressed as logarithm transformed values of gene copy numbers obtained per (A) mL of sample, (B) ng of DNA or (C) 16S rRNA gene copy number. Values above the bars represent the difference between logarithmic values of genes quantification obtained after DNA extraction with both kits. * indicate statistically significant differences between both DNA extraction kits ($p < 0.01$).

extraction kits ($p < 0.01$) (Table 1). Exceptions were observed for River1 (operators 1 and 2) and River3 (operator 1) samples, with higher yields in DNeasy PowerWater kit extracts. The vulnerability to operator variations, assessed by the use of the same samples by two operators extracting DNA simultaneously, demonstrated that, also in this aspect, NZY Tissue gDNA Isolation kit and DNeasy PowerWater kit were not significantly different ($p < 0.01$).

Given the consistency of results between both operators, further assays on qPCR gene determinations used the DNA extract set obtained by operator 1. Even though the DNA yield obtained by both extraction

kits was not significantly different, it was hypothesized that differences could be observed in the gene quantification. In hospital effluent samples, no significant differences were observed for the quantification of genes abundance in extracts obtained with both kits. In contrast, in river samples, most of the genes were quantified either per water volume or per ng of DNA at significantly higher amounts in the DNeasy PowerWater kit than in the NZY Tissue gDNA Isolation kit extracts (Fig. 1). However, these differences were never higher than 0.9 log gene copies mL^{-1} of sample or 0.4 log gene copies ng^{-1} of DNA (Fig. 1). As demonstrated using the internal standard, the matrix effect for the NZY Tissue gDNA Isolation kit estimated for river and hospital effluent samples were in average 0.3 and 5.5%, respectively, not significantly different of what was observed for DNeasy PowerWater kit. Considering that both kits did not show an overwhelming performance difference, it was taken the decision to proceed with the internal standard assays using the DNeasy PowerWater kit, since it is recommended by the manufacturer for water samples.

3.2. Matrix effect in different water types

The internal standard, used to measure the matrix effect, consisted in a culture suspension of the *Escherichia coli* strain JM109 *molA* clone, *EcmolA*+, spiked simultaneously in water or wastewater and in ultrapure water, where the matrix effect was assumed to be null. In preliminary assays, different doses of *EcmolA*+ (between 6 and 10 log-units *EcmolA*+ mL^{-1} sample) were spiked in ultrapure water and environmental water samples to determine the adequate dose to use in further experiments. The value determined as most suitable was 9 log-units *EcmolA*+ mL^{-1} to reach a final abundance of 6 log-units *EcmolA*+ mL^{-1} of spiked ultrapure water or sample. The matrix effect, estimated based on the quantification of the gene *molA* in control spiked ultrapure water and in samples is shown in Table 2. Although the quantification of the *molA* gene in spiked samples was, in average, underestimated in 0.1 to 0.9 log-units, occasionally, it was overestimated in 0.05 and 0.23 log-units, in sTWWP and in river samples, respectively. In average, the matrix effect ranged of 0.1–0.5 log-units, corresponding to 1.2% in sTWWP, to 6.7% in sTWWA and RWWP. Interestingly, it was concluded that the matrix effect was not dependent on the origin of the sample ($p < 0.01$) (Table 2). The observed overestimation in two samples compared to the respective ultrapure water samples was not related with the presence of the *molA* gene in non-spiked environmental samples, since it was confirmed the absence of the gene in environmental samples. In two situations the quantification of the internal standard *molA* was higher in environmental water samples than in spiked ultrapure water, although all the quality criteria cited above were observed. This might be due to the fact that the internal standard (*molA*) could be adsorbed onto samples particles, hampering an adequate homogenization of the internal standard prior to their filtration. This is sometimes observed, with distinct replicates of the same sample presenting peaks for some genes. This effect is not uncommon and can be attributed to adsorption onto samples particles, which may influence the amount of the gene recovered after DNA extraction in these samples. This effect would never be observed in ultrapure water, explaining the apparent unexpected results.

Table 2
Matrix effect estimated based on the internal standard gene (*molA*). sTWWA and sTWWP – secondary treated wastewater effluent from sampling site A and P; ASLP – activated sludge treated effluent from sampling site P; RWWA and RWWP – influent wastewater from sampling site A and P and HE – hospital effluent. The internal standard *molA* gene was also analysed in non-spiked environmental water samples, and it was confirmed its absence in environmental samples.

DNA extraction matrices	Matrix effect	
	Average Log-units ± Standard Error (A)	Average ± Standard Error in % (B)
River (n = 6)	0.3 ± 0.2	4.5 ± 2.8
sTWWA (n = 6)	0.5 ± 0.1	6.7 ± 1.3
sTWWP (n = 6)	0.1 ± 0.1	1.2 ± 1.7
ASLP (n = 6)	0.4 ± 0.0	5.3 ± 0.5
RWWA (n = 6)	0.2 ± 0.1	2.0 ± 1.0
RWWP (n = 6)	0.5 ± 0.1	6.7 ± 1.7
HE (n = 6)	0.4 ± 0.1	4.4 ± 1.0

A = log (*molA* gene copy number/mL of spiked ultrapure water) – log (*molA* gene copy number/mL of spiked water sample).
B = A/log (*molA* gene copy number/mL of spiked ultrapure water) × 100.

3.3. Gene quantification and internal standard correction

A total of 14 samples of water (river, n = 2) and municipal and airport wastewater treatment plants (influent, secondary effluent and activated sludge, n = 10) and hospital effluent (n = 2) (Table S1) were used. These samples were characterized for the abundance (gene copy number/volume of sample) and prevalence (gene copy number/16S rRNA gene copy number) of five chromosomal genes and nine putatively plasmid-associated genes. The quantifications were corrected according to the matrix effect percentage values (Table 2 and Fig. 2). Except for samples from secondary wastewater treatment effluent (sTWWA and sTWWP), where *uidA* gene abundance per volume of water was higher than *rpoB*, the chromosomal genes abundance could be ranked as 16S rRNA > *marA* > *rpoB* > *uidA* > *ecf*. The same hierarchy was observed for genes prevalence (normalized by 16S rRNA gene abundance) (Fig. 2). Regarding the putatively-associated plasmid genes, the bla_{TEM}, *intI*1 and *sul1* were the most abundant genes in the majority of samples ranging 7–8, 6–8 and 6–7 log (gene copies mL^{-1} of sample), respectively. These genes were in higher abundance than *marA* and *uidA* genes, specific for *Enterobacteriaceae* and *Escherichia coli*, respectively (Castiglioni et al., 2008; Chern et al., 2009) and, than the genes encoding β-lactamases (bla_{CTX-M}, bla_{OXA-A}, bla_{SHV}, bla_{IMP} and bla_{VIM}), widespread in Gram-negative bacteria (Rawat and Nair, 2010). These results are in agreement with the proposal that have been made that they are widespread and that *intI*1 is considered an indicator of anthropogenic contamination (Du et al., 2014; Gillings et al., 2015; Narciso-Da-Rocha et al., 2014).

Also, the carbapenems encoding genes bla_{VIM} and bla_{IMP} presented higher abundance (6 log-units each) in hospital effluent samples compared to other samples, in which abundance ranged from 2 to 4 log-units. This might be related to the fact that these ARGs are associated with last resort antibiotics that are only administrated in hospitals (Meletis, 2016). A major question of this study was if due to matrix effect the abundance of genes could be underestimated in environmental water samples. Underestimation ranged 0.0–0.6 log-units in abundance (per volume of water) and 0.0–0.8 log-units in prevalence (per 16S rRNA gene). The significant differences on the genes abundance due to the matrix effect were more frequently observed in chromosomal than in plasmid-associated genes ($p < 0.01$) (Fig. 2). The matrix effect correction, estimated based on the internal standard *molA* associated to a plasmid which was inserted into a cell, led to significant differences with the same frequency in the prevalence of either plasmid-associated or chromosome associated genes.

4. Discussion

DNA extraction is a critical step on the genetic analyses of environmental samples. For water samples, processing usually involves a step of biomass concentration, frequently through filtration, DNA extraction, and the genetic sample characterization using targeted methods, such as qPCR or non-targeted approaches, such as metagenomics (Manaia et al., 2018). Aware of the importance of this step on the quality of genetic analyses of water samples, manufacturers have

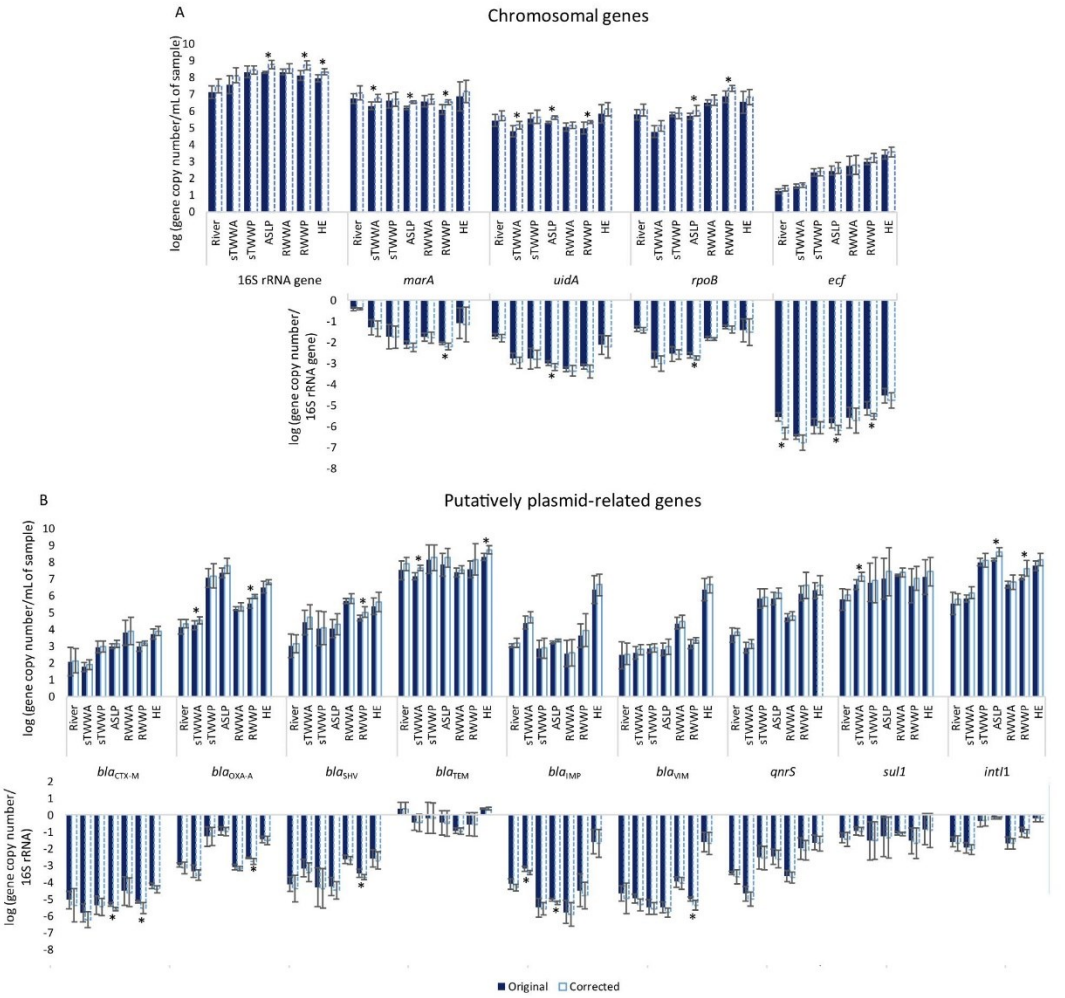


Fig. 2. Matrix effect on genes quantification assessed based on the use of the internal standard *EcmolA+*. (A) Chromosomal genes and (B) putatively plasmid-associated ARGs abundance (gene copy number/mL of sample) and prevalence (gene copy number/16S rRNA gene copy number) were determined in 7 environmental water types from different origins: river, secondary treated wastewater effluent from sampling site A and P (sTWWA and sTWWP), activated sludge treated effluent from sampling site P (ASLP), influent wastewater from sampling site A and P (RWWA and RWWP) and hospital effluent (HE). The internal standard *molA* gene was also analysed in non-spiked environmental water samples, and it was confirmed its absence in environmental samples. The results express the direct gene quantification and respective corrected value due to matrix effect percentage, estimated based on the internal standard losses. * indicate statistically significant differences, $p < 0.01$.

developed DNA extraction kits, which are designed to process filtering membranes containing the water sample biomass and to avoid the matrix effect (e.g. DNeasy PowerWater kit). Highly convenient and popular (Cacace et al., 2019; Pärnänen et al., 2019), these water-specific solutions are expensive and, naturally, cannot completely overcome all kind of matrix effect that can exist in water samples. Therefore, this work aimed to assess if a generalist DNA extraction kit could be adapted for the analysis of water samples, both at the level of sample processing and gene quantification. The matrix effect, which may influence the genetic analyses results, is in chemical analyses overcome through the use of an internal standard (Skoog et al., 2017). Therefore, a cell-based internal standard was designed for this study to measure matrix effects. To our knowledge, this is the first time a cell-based internal standard was designed and implemented for analysis of genes in water samples.

The influence of DNA extraction procedure on DNA yields has been discussed in previous studies, with distinct outcomes. For instance, Hinlo et al. (2017) did not observe differences on the DNA yield obtained using Qiagen's DNeasy Blood and Tissue kit and DNeasy PowerWater kit; Li et al. (2017) observed that the DNA yield was higher using FastDNA SPIN Kit for Soil, compared to PowerSoil DNA Isolation Kit and ZR Fecal DNA MiniPrep; and Djurhuus et al. (2017) observed

the DNA yield was extremely variable between DNeasy Blood and Tissue kit, the DNeasy PowerWater kit, and standard phenol/chloroform methods. In the current study, the fact that statistically significant differences were obtained for DNA extraction from river using DNeasy PowerWater kit compared with NZY Tissue gDNA Isolation kit may be due to the fact that DNeasy PowerWater kit is described as being optimized to remove humic acids, heavy metals, polysaccharides, among other substances in water samples. The differences observed for river sample in genes abundance (genes copy number per mL of sample or ng of DNA) might also be related with the fact that in river samples higher DNA yield was obtained with DNeasy PowerWater kit. Riediger et al. (2016) also observed a higher gene quantification using DNeasy PowerWater kit in river and ultrapure water samples, than using QIAamp DNA mini and PowerSoil DNA Isolation kits. The lack of statistically significant differences observed for DNA extraction between operators using both kits is also interesting and in agreement with the report of Li and colleagues (Li et al., 2017), in which no statistically significant differences (p greater than 0.05) (DNA yield differences up to 12 μg) were observed for DNA yield obtained by different operators. Overall, if DNA extraction kits comparison is analysed in a cost benefit perspective, the choice deserves weighting. The DNeasy PowerWater kit cost around 9 € and the NZY Tissue gDNA Isolation kit costs around 3 € per reaction. If convenience characteristics are taken into account, such as being 1) cheap, 2) user friendly or 3) fast, a simple kit as NZY Tissue gDNA Isolation kit can be adopted to encourage the survey of environmental samples.

The matrix effect, an obligatory issue when water chemical analyses are in discussion (Carbajo et al., 2015; Van De Steene and Lambert, 2008; Zhou and Kang, 2013), is an issue poorly explored in the microbiological analyses domain. The matrix effect refers to particles and organic matter that may impede the extraction of a given analyte (Schrader et al., 2012). Environmental samples matrix might be composed by substances such as detergents, phenolic compounds, humic acids, heavy metals and other contaminants that might influence DNA extraction efficiency and qPCR performance, due to, for example, bacterial interaction with samples matrix or inhibition of qPCR reactions (Kim et al., 2013; Wilson, 1997). The use of external controls spiked into DNA extracts (Cloud et al., 2003; Volkmann et al., 2007) or of internal controls added to the samples before the DNA extraction (Burggraf and Olgemöller, 2004) were reported as a way to understand the reliability of qPCR results. However, these approaches used cell-free DNA, which did not consider the losses due to filtration, an obligatory step in water sample processing, and to deficient cell lysis, a crucial step during DNA extraction which might influence the amount of DNA and genes recovered and quantified. Cell-based internal standards were used for the detection of *Helicobacter pylori* in drinking water, expectedly at low abundance (Sen et al., 2007). Therefore, to our knowledge, this is the first report of a cell-based internal standard used to assess the effect of samples matrix effect on DNA extraction and genes quantification. It was also insightful to observe that the matrix effect is not dependent on the water type, with river water samples presenting matrix effects in the same range of values as raw wastewater. Also the importance of water sample heterogeneity, with particles onto which biological analytes may adsorb, was suggested in this study.

In summary, these results contribute to debate the common assumption that quantitative PCR may have limited value for the comparison of data obtained in distinct laboratories, due to different types of technical biases. In this work, using these two kits, it was observed that the DNA extraction kit, the operator, or the samples matrix effect had a limited impact on the final genes quantification. Moreover, it is suggested that the simplification of DNA extraction procedures, through the adoption of DNA extraction kits that are cheaper and/or user friendly, may not create important bias on the comparison of the quantification of ARGs in DNA extracts. These are encouraging findings that may incentivize labs worldwide, even with low resources, to collaborate in ARGs surveillance studies in the environment.

CRediT authorship contribution statement

Jaqueline Rocha: Conceptualization, Data curation, Investigation, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing. **Célia M. Manaia:** Funding acquisition, Project administration, Resources, Investigation, Conceptualization, Supervision, Methodology, Validation, Data curation, Writing - review & editing.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the support of the staff who provided the wastewater samples from hospital and treatment plants; the colleagues R. Lopes and O. Nunes from Laboratório de Engenharia de Processos, Ambiente, Biotecnologia e Energia, Faculdade de Engenharia, Universidade do Porto who kindly provided the *Gulosibacter molinativorax* ON4^T strain for this study; the control DNA extraction operators T. Fernandes and C. Ferreira and Gianuário. Fortunato for supplying the qPCR protocols for *rpoB*, *ecf*, bla_{IMP} and bla_{VIM} genes. The work was possible thanks to the support of the host research Centre funded by the FCT project UID/Multi/50016/2019.

Funding

This work was supported by Fundação para a Ciência e Tecnologia WaterJPI/0001/2013 STARE – "Stopping Antibiotic Resistance Evolution" and by European Regional Development Fund (FEDER), through the North Regional Operational Program (North PO), under the project DEPCAT: Demonstration of new equipment involving integrated catalytic processes for the treatment of organic pollutants and disinfection of waters (NORTE-01-0247-FEDER-033330). J. Rocha was supported by the International PhD Programme in Biotechnology – BIOTECH.DOC, NORTE-08-5369-FSE-000007.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2020.106194.

References

Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K.A., Davies, J., Handelsman, J., 2010. Call of the wild: antibiotic resistance genes in natural environments. Nat. Rev. Microbiol. 8, 251–259. https://doi.org/10.1038/nrmicro2312.

Atrouni, A.A., Joly-Guillou, M.L., Hamze, M., Kempf, M., 2016. Reservoirs of non-*baumannii Acinetobacter* species. Front. Microbiol. 7, 1–12. https://doi.org/10.3389/fmicb.2016.00049.

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M.-N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., Martinez, J.L., 2015. Tackling antibiotic resistance: the environmental framework. Nat. Rev. Microbiol. 13, 310–317. https://doi.org/10.1038/nrmicro3439.

Bessetti, B.J., 2007. An introduction to PCR inhibitors. Promega Notes 9–10. https://doi.org/10.1161/STROKEAHA.109.559013.

Brankatschk, R., Bodenhausen, N., Zeyer, J., Bürgmann, H., 2012. Simple absolute quantification method correcting for quantitative PCR efficiency variations for microbial community samples. Appl. Environ. Microbiol. 78, 4481–4489. https://doi.

org/10.1128/AEM.07878-11.
Burggraf, S., Olgemöller, B., 2004. Simple technique for internal control of real-time amplification assays. Clin. Chem. 50, 819–825. https://doi.org/10.1373/clinchem.2003.027961.
Bustin, S.A., 2010. Why the need for qPCR publication guidelines?-The case for MIQE. Methods 50, 217–226. https://doi.org/10.1016/j.ymeth.2009.12.006.
Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: Minimum Information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622. https://doi.org/10.1373/clinchem.2008.112797.
Cacace, D., Fatta-Kassinos, D., Manaia, C.M., Cytryn, E., Kreuzinger, N., Rizzo, L., Karaolia, P., Schwartz, T., Alexander, J., Merlin, C., Garelick, H., Schmitt, H., de Vries, D., Schwermer, C.U., Meric, S., Ozkal, C.B., Pons, M.N., Kneis, D., Berendonk, T.U., 2019. Antibiotic resistance genes in treated wastewater and in the receiving water bodies: A pan-European survey of urban settings. Water Res. 162, 320–330. https://doi.org/10.1016/j.watres.2019.06.039.
Carbajo, J.B., Petre, A.L., Rosal, R., Herrera, S., Letón, P., García-calvo, E., Fernández-alba, A.R., Perdigón-melón, J.A., 2015. Continuous ozonation treatment of ofloxacin: Transformation products, water matrix effect and aquatic toxicity. J. Hazard. Mater. 292, 34–43. https://doi.org/10.1016/j.jhazmat.2015.02.075.
Castiglioni, S., Pomati, F., Miller, K., Burns, B.P., Zuccato, E., Calamari, D., Neilan, B.A., 2008. Novel homologs of the multiple resistance regulator *marA* in antibiotic-contaminated environments. Water Res. 42, 4271–4280. https://doi.org/10.1016/j.watres.2008.07.004.
Chern, E.C., Brenner, K.P., Wymer, L., Haugland, R.A., 2009. Comparison of fecal indicator bacteria densities in marine recreational waters by QPCR. Water Qual. Expo. Heal. (1), 203–214. https://doi.org/10.1007/s12403-009-0019-2.
Cloud, J.L., Hymas, W.C., Turlak, A., Croft, A., Reischl, U., Daly, J.A., Carroll, K.C., 2003. Description of a multiplex *Bordetella pertussis* and *Bordetella parapertussis* LightCycler® PCR assay with inhibition control. Diagn. Microbiol. Infect. Dis. 46, 189–195. https://doi.org/10.1016/S0732-8893(03)00045-2.
Djurhuus, A., Port, J., Closek, C.J., Yamahara, K.M., Romero-Maraccini, O., Walz, K.R., Goldsmith, D.B., Michisaki, R., Breitbart, M., Boehm, A.B., Chavez, F.P., 2017. Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. Front. Mar. Sci. 4. https://doi.org/10.3389/fmars.2017.00314.
Drinking water directive (98/83/EC) on the quality of water intended for human consumption. European Commission. Official J. L 330, 05/12/1998 P. 0032 – 0054. Brussels.
Du, J., Ren, H., Geng, J., Zhang, Y., Xu, K., Ding, L., 2014. Occurrence and abundance of tetracycline, sulfonamide resistance genes, and class 1 integron in five wastewater treatment plants. Environ. Sci. Pollut. Res. 21, 7276–7284. https://doi.org/10.1007/s11356-014-2613-5.
Duarte, M., Ferreira-da-Silva, F., Lünsdorf, H., Junca, H., Gales, L., Pieper, D.H., Nunes, O.C., 2011. *Gulosibacter molinativorax* ON4^{T} molinate hydrolase, a novel cobalt-dependent amidohydrolase. J. Bacteriol. 193, 5810–5816. https://doi.org/10.1128/JB.05054-11.
Fatta-Kassinos, D., Meric, S., Nikolaou, A., 2011. Pharmaceutical residues in environmental waters and wastewater: Current state of knowledge and future research. Anal. Bioanal. Chem. https://doi.org/10.1007/s00216-010-4300-9.
Foerstner, K.U., von Mering, C., Hooper, S.D., Bork, P., 2005. Environments shape the nucleotide composition of genomes. EMBO Rep. 6, 1208–1213. https://doi.org/10.1038/sj.embor.7400538.
Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., Zhu, Y.G., 2015. Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. ISME J. 9, 1269–1279. https://doi.org/10.1038/ismej.2014.226.
Hinlo, R., Gleeson, D., Lintermans, M., Furlan, E., 2017. Methods to maximise recovery of environmental DNA from water samples. PLoS One 12. https://doi.org/10.1371/journal.pone.0179251.
Huijbers, P.M.C., Flach, C.F., Larsson, D.G.J., 2019. A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. Environ. Int. 130, 104880. https://doi.org/10.1016/j.envint.2019.05.074.
ISO 7899, 2000. Water quality – Detection and enumeration of intestinal Enterococci. International Organization for Standardization, Geneva, Switzerland.
ISO 9308, 2014. Water quality – Enumeration of *Escherichia coli* and coliform bacteria. International Organization for Standardization, Geneva, Switzerland.
Jang, J., Hur, H.G., Sadowsky, M.J., Byappanahalli, M.N., Yan, T., Ishii, S., 2017. Environmental *Escherichia coli*: ecology and public health implications—a review. J. Appl. Microbiol. 123, 570–581. https://doi.org/10.1111/jam.13468.
Kim, J., Lim, J., Lee, C., 2013. Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2013.05.010.
Klein, D., 2002. Quantification using real-time PCR technology: applications and limitations. Trends Mol. Med. 8, 257–260. https://doi.org/10.1016/S1471-4914(02)02355-9.
Kolpin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., Buxton, H.T., 2002. Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: A national reconnaissance. Environ. Sci. Technol. 36, 1202–1211. https://doi.org/10.1021/es011055j.
Kümmerer, K., 2009. Antibiotics in the aquatic environment – A review – Part II. Chemosphere 75, 435–441. https://doi.org/10.1016/j.chemosphere.2008.12.006.
Li, A.D., Metch, J.W., Wang, Y., Garner, E., Zhang, A.N., Riquelme, M.V., Vikesland, P.J., Pruden, A., Zhang, T., 2017. Effects of sample preservation and DNA extraction on enumeration of antibiotic resistance genes in wastewater. FEMS Microbiol. Ecol. 94. https://doi.org/10.1093/femsec/fix189.
Lopes, A.R., Danko, A.S., Manaia, C.M., Nunes, O.C., 2013. Molinate biodegradation in soils: Natural attenuation versus bioaugmentation. Appl. Microbiol. Biotechnol. 97, 2691–2700. https://doi.org/10.1007/s00253-012-4096-y.
Manaia, C.M., Macedo, G., Fatta-Kassinos, D., Nunes, O.C., 2016. Antibiotic resistance in urban aquatic environments: can it be controlled? Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-015-7202-0.
Manaia, C.M., Rocha, J., Scaccia, N., Marano, R., Radu, E., Biancullo, F., Cerqueira, F., Fortunato, G., Iakovides, I.C., Zammit, I., Kampouris, I., Vaz-Moreira, I., Nunes, O.C., 2018. Antibiotic resistance in wastewater treatment plants: Tackling the black box. Environ. Int. 115, 312–324. https://doi.org/10.1016/j.envint.2018.03.044.
Martinez, J.L., 2009. Environmental pollution by antibiotics and by antibiotic resistance determinants. Environ. Pollut. 157, 2893–2902. https://doi.org/10.1016/j.envpol.2009.05.051.
Meletis, G., 2016. Carbapenem resistance: overview of the problem and future perspectives. Ther. Adv. Infect. Dis. 3, 15–21. https://doi.org/10.1177/2049936115621709.
Mena, K.D., Gerba, C.P., 2009. Risk assessment of *Pseudomonas aeruginosa* in water. Rev. Environ. Contam. Toxicol. 201, 71–115. https://doi.org/10.1007/978-1-4419-0032-6_3.
Narciso-da-Rocha, C., Rocha, J., Vaz-Moreira, I., Lira, F., Tamames, J., Henriques, I., Martinez, J.L., Manaia, C.M., 2018. Bacterial lineages putatively associated with the dissemination of antibiotic resistance genes in a full-scale urban wastewater treatment plant. Environ. Int. 118, 179–188. https://doi.org/10.1016/j.envint.2018.05.040.
Narciso-Da-Rocha, C., Varela, A.R., Schwartz, T., Nunes, O.C., Manaia, C.M., 2014. Bla_{TEM} and *vanA* as indicator genes of antibiotic resistance contamination in a hospital-urban wastewater treatment plant system. J. Glob. Antimicrob. Resist. 2, 309–315. https://doi.org/10.1016/j.jgar.2014.10.001.
Nunes, O.C., Lopes, A.R., Manaia, C.M., 2013. Microbial degradation of the herbicide molinate by defined cultures and in the environment. Appl. Microbiol. Biotechnol. 97, 10275–10291. https://doi.org/10.1007/s00253-013-5316-9.
Pärnänen, K.M.M., Narciso-da-Rocha, C., Kneis, D., Berendonk, T.U., Cacace, D., Do, T.T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J.L., Michael, S.G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., Stedtfeld, R.D., Tiedje, J.M., Della Giustina, S.V., Walsh, F., Vaz-Moreira, I., Virta, M., Manaia, C.M., 2019. Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. Sci. Adv. 5, eaau9124. https://doi.org/10.1126/sciadv.aau9124.
Rawat, D., Nair, D., 2010. Extended-spectrum β-lactamases in Gram negative bacteria. J. Glob. Infect. Dis. 2, 263–274. https://doi.org/10.4103/0974-777x.68531.
Riediger, I.N., Hoffmaster, A.R., Casanovas-Massana, A., Biondo, A.W., Ko, A.I., Stoddard, R.A., 2016. An optimized method for quantification of pathogenic *Leptospira* in environmental water samples. PLoS One 11, 1–12. https://doi.org/10.1371/journal.pone.0160523.
Rocha, J., Cacace, D., Kampouris, I., Guilloteau, H., Jäger, T., Marano, R.B.M., Karaolia, P., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Berendonk, T.U., Schwartz, T., 2018. Inter-laboratory calibration of quantitative analyses of antibiotic resistance genes. J. Environ. Chem. Eng. (in press). https://doi.org/10.1016/j.jece.2018.02.022.
Schrader, C., Schielke, A., Ellerbroek, L., Johne, R., 2012. PCR inhibitors – Occurrence, properties and removal. J. Appl. Microbiol. 113, 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x.
Sen, K., Schable, N.A., Lye, D.J., 2007. Development of an internal control for evaluation and standardization of a quantitative PCR assay for detection of *Helicobacter pylori* in drinking water. Appl. Environ. Microbiol. 73, 7380–7387. https://doi.org/10.1128/AEM.00687-07.
Skoog, D.A., Holler, F.J., Crouch, S.R., 2017. Principles of instrumental analysis. 7th Edition, Cengage learning. Australia.
Smith, C.J., Osborn, A.M., 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS Microbiol. Ecol. 67, 6–20. https://doi.org/10.1111/j.1574-6941.2008.00629.x.
Szczepanowski, R., Linke, B., Krahn, I., Gartemann, K.H., Gützkow, T., Eichler, W., Pühler, A., Schlüter, A., 2009. Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. Microbiology 155, 2306–2319. https://doi.org/10.1099/mic.0.028233-0.
Valasek, M.A., 2005. The power of real-time PCR. Adv. Physiol. Educ. 29, 151–159. https://doi.org/10.1152/advan.00019.2005.
Van De Steene, J.C., Lambert, W.E., 2008. Comparison of matrix effects in HPLC-MS/MS and UPLC-MS/MS analysis of nine basic pharmaceuticals in surface waters. J. Am. Soc. Mass Spectrom. 19, 773–1718. https://doi.org/10.1016/j.jasms.2008.01.013.
Varela, A.R., Nunes, O.C., Manaia, C.M., 2016. Quinolone resistant *Aeromonas* spp. as carriers and potential tracers of acquired antibiotic resistance in hospital and municipal wastewater. Sci. Total Environ. 542, 665–671. https://doi.org/10.1016/j.scitotenv.2015.10.124.
Vartoukian, S.R., Palmer, R.M., Wade, W.G., 2010. Strategies for culture of "uncultur-able" bacteria. FEMS Microbiol. Lett. 309, 1–7. https://doi.org/10.1111/j.1574-6968.2010.02000.x.
Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H.,

Pfannkoch, C., Rogers, Y.H., Smith, H.O., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science 304, 66–74. https://doi.org/10.1126/science.1093857.

Volkmann, H., Schwartz, T., Kirchen, S., Stofer, C., Obst, U., 2007. Evaluation of inhibition and cross-reaction effects on real-time PCR applied to the total DNA of wastewater samples for the quantification of bacterial antibiotic resistance genes and taxon-specific targets. Mol. Cell. Probes 21, 125–133. https://doi.org/10.1016/j.mcp.2006.08.009.

Walsh, T.R., Toleman, M.A., Poirel, L., Nordmann, P., 2005. Metallo-β-lactamases: the quiet before the storm? Clin. Microbiol. Rev. 18, 306–325. https://doi.org/10.1128/CMR.18.2.306-325.2005.

Wilson, I.G., 1997. Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. 63, 3741–3751.

Zhang, X.X., Zhang, T., Fang, H.H.P., 2009. Antibiotic resistance genes in water environment. Appl. Microbiol. Biotechnol. 82, 397–414. https://doi.org/10.1007/s00253-008-1829-z.

Zhou, J.L., Kang, Y., 2013. Matrix effect in high-performance liquid chromatography-tandem mass spectrometry analysis of antibiotics in environmental water samples. J. Sep. Sci. 36, 564–571. https://doi.org/10.1002/jssc.201200750.

CHAPTER 5

THIRD GENERATION CEPHALOSPORIN-RESISTANT *KLEBSIELLA PNEUMONIAE* THRIVING IN PATIENTS AND IN WASTEWATER: WHAT DO THEY HAVE IN COMMON?

The work presented in this chapter was submitted to the journal *BMC Genomics* and is currently under review:

Rocha, J., Ferreira, C., Mil-Homens, D., Busquets, A., Fialho, A.M., Henriques, I., Gomila, M., Manaia, C.M. (under review). Third generation cephalosporin-resistant *Klebsiella pneumoniae* thriving in patients and in wastewater: what do they have in common? BMC Genomics.

Contribution:

Contributor Role	JR	CF	DM	AB	AF	IH	MG	CM
Conceptualisation								x
Methodology	x						x	x
Validation							x	x
Investigation	x	x	x	x			x	x
Data Curation	x						x	
Writing – Original Draft Preparation	x						x	x
Writing – Review and Editing	x	x	x	x	x	x	x	x
Supervision							x	x
Project Administration								x
Funding							x	x

Abstract

Klebsiella pneumoniae are ubiquitous bacteria and recognized multidrug-resistant opportunistic pathogens that can be released into the environment, mainly through sewage, where they can survive even after wastewater treatment. A major question is if once released into wastewater, the selection of lineages missing clinically-relevant traits may occur. Wastewater (n=25) and clinical (n=34) 3rd generation cephalosporin-resistant *K. pneumoniae* isolates were compared based on phenotypic, genotypic and genomic analyses. Clinical and wastewater isolates were indistinguishable based on phenotypic and genotypic characterization. The analysis of whole genome sequences of 22 isolates showed that antibiotic and metal resistance or virulence genes, were associated with mobile genetic elements, mostly transposons, insertion sequences or integrative and conjugative elements. These features were variable among isolates, according to the respective genetic lineage rather than the origin. It is suggested that once acquired, clinically relevant features of *K. pneumoniae* may be preserved in wastewater, even after treatment. This evidence highlights the high capacity of *K. pneumoniae* for spreading through wastewater, enhancing the risks of transmission back to humans.

1. Introduction

The species *Klebsiella pneumoniae,* within the family *Enterobacteriaceae,* include opportunistic pathogens, with ubiquitous distribution (Wyres & Holt, 2018; Pendleton *et al*., 2013). The ubiquity and clinical relevance of *K. pneumoniae* is due, in part, to the genome plasticity, in which genes acquisition, such as those encoding antibiotic resistance, is a major driver (Wyres & Holt, 2018; Paczosa & Mecsas, 2016; Beceiro *et al*., 2013). Indeed, genes acquired by horizontal gene transfer, encoding resistance against aminoglycosides, 3rd generation cephalosporins, carbapenems and fluoroquinolones (Wyres *et al*., 2020; Navon-Venezia *et al*., 2017) or metals such as arsenic, copper, tellurium and mercury are frequent in *K. pneumoniae* (Bialek-Davenet *et al*., 2014). The ubiquity and clinical relevance of *K. pneumoniae* is also due to a wide array of genes that encode functions related with adhesion, protection (capsules) or siderophore production (Paczosa & Mecsas, 2016). The combination of these features and ubiquitous distribution make *K. pneumoniae* an important opportunistic pathogen, responsible for one third of the hospital infections caused by Gram-negative bacteria (Navon-Venezia *et al*., 2017).

The ubiquity of *K. pneumoniae* is illustrated by its occurrence in healthy humans and animals, and in plants, soil, water and wastewater (Bagley, 1985; Podschun & Ullmann, 1998; Wyres & Holt, 2018), suggesting that it may circulate among distinct compartments. In urban areas, domestic sewage represents the major human emission of pathogens and antibiotic resistant bacteria. Although wastewater treatment has a pivotal role for the removal of such microorganisms from sewage, an important fraction can survive, being discharged into the environment (Manaia *et al*., 2016; Vaz-Moreira *et al*., 2014; Rizzo *et al*., 2013).

The occurrence of *K. pneumoniae* in treated wastewater is mainly explained by human emissions and by the capacity of members of this species to endure treatment processes. In fact, the survival or even proliferation of virulent and multidrug resistant *K. pneumoniae* discharged through sewage in the environment has been reported (Fouz *et al*., 2020; Runcharoen *et al*., 2017). These facts explain why identical *K. pneumoniae* sequence types (e.g. ST11, ST15, ST17, ST258 or ST147) have been reported in both clinical settings and wastewater (Wyres & Holt, 2018; Wyres *et al*., 2020).

Pathogens causing infection or thriving in the environment are supposedly exposed to distinct challenges, which may hypothetically be associated with the retention or loss of specific traits and/or responsible for distinctive selective processes. A major question is if *K. pneumoniae* found in the environment retain the features observed in clinical isolates, or if some properties, like antibiotic resistance or virulence, can be lost. In accordance, this work aimed to investigate if phenotypic and genotypic traits and genome characteristics are shared by clinical and

wastewater isolates. It was also aimed to compare both groups in terms of genome features associated with horizontal gene transfer. Our hypothesis was addressed using a set of isolates resistant to 3rd generation cephalosporins, as this is an increasingly frequent phenotype in clinical and environmental isolates (Navon-Venezia *et al*., 2017; Müller *et al*., 2018). A group of 59 isolates (25 from wastewater, 34 clinical) was characterized phenotypically and genotypically targeting clinically-relevant traits and a subset of these isolates (7 from wastewater, 15 clinical – 11 from patients and 4 from clinical environment) was further compared based on genome analyses. The results suggested that phylogeny, more than strains origin, may explain the profile of acquired traits.

2. Material and methods

2.1. Study structure and bacterial strains

Fifty-nine *K. pneumoniae* isolates, identified based on the 16S rRNA gene sequence and exhibiting resistance to 3rd generation cephalosporins (cefotaxime and ceftazidime) were selected for this study. The selection of these isolates for this study was due to the clinically relevance of *K. pneumoniae* isolates resistant to 3rd generation cephalosporins and due to their ubiquity across clinical and environmental niches. The bacterial isolates were from wastewater (n=25): 3 from hospital effluent, 12 from raw wastewater and 10 from treated wastewater and clinical (n=34): 30 from patients (26 from Portugal, 4 from Spain) and 4 from the respective clinical environment (Spain) (Table S1). Wastewater isolates were recovered in independent events in the Northern region of Portugal between 2011 and 2016 and from a laboratory collection of 49 isolates were selected those that were resistant to 3rd generation cephalosporins, a feature common to all clinical isolates. The clinical isolates were obtained from urine, faeces or blood samples, among others, of hospitalized patients, collected over a period of 18 months, from 2014 to 2016 in Porto, Portugal. Clinical samples, including urine, faeces, blood and haemoculture samples were processed at the hospital in accordance with the manual for good laboratory practices implemented in this health unit. There was no treatment of personal data and the biological samples are not related to any data that allows the identification of individuals. The 8 clinical isolates recovered in Spain were collected during a hospital outbreak in the Balearic Islands from patients (n=4) and from drains and surface (n=4). As it was hypothesized that during infection the isolates are exposed to specific conditions, which are different from the natural environment, the inclusion of the Spanish outbreak isolates was considered interesting to assess whether genome variation could present a distinct pattern when compared to Portuguese clinical isolates. The isolates were

classified in 3 categories: clinical when isolated from patients, clinical environment when isolated from hospital settings, and wastewater (Table S1). The choice for wastewater isolates was based on the fact that this comprises a highly competitive and stressful environment, where pathogens are supposed to be eliminated. The phenotypic characterization of the isolates included clinically-relevant traits such as antibiotic resistance phenotype analyses, trans-species conjugation assays, biofilm formation and infection capacity. Genotypic characterization involved the detection of antibiotic resistance genes and the determination of the number and size of plasmids.

2.2. Antibiotic resistance phenotype and genotype

Resistance phenotypes were determined based on the disk diffusion method, incubated for 24 h at 37 ºC, as recommended by the Clinical Laboratory Standards Institute (CLSI, 2016) for antibiotics belonging to 5 different classes: β-lactams (AMC, amoxicillin with clavulanic acid, 20/10 µg; AML, amoxicillin, 25 µg; ATM, aztreonam, 30 µg; MEM, meropenem, 10 µg); aminoglycosides (CN, gentamicin, 10 µg; AK, amikacin, 30 µg); quinolones (CIP, ciprofloxacin, 5 µg); sulfonamides (RL, sulfamethoxazole, 25 µg); and tetracyclines (TE, tetracycline, 30 µg), the combination sulfamethoxazole/trimethoprim (SXT, sulfamethoxazole/trimethoprim, 1.25/23.75 µg) and trimethoprim (W, trimethoprim, 5 µg) was also tested. Bacterial isolates were classified as resistant or susceptible according to the inhibition zone diameters recommended by CLSI, 2016 guidelines for *Enterobacteriaceae* (CLSI, 2016). The reference strain *Escherichia coli* ATCC® 25922 was included in each assay as quality control. The genes *bla*$_{CTX}$, *bla*$_{IMP}$, *bla*$_{KPC}$, *bla*$_{OXA}$, *bla*$_{SHV}$, *bla*$_{TEM}$, *bla*$_{VIM}$, and *mcr* were screened by PCR using the primers, annealing temperatures and the conditions described in the literature indicated in Table S2. Positive and negative controls were included in each reaction and amplicons were randomly confirmed based on DNA sequence analysis.

2.3. Plasmid analyses

The plasmid replicon types were screened in total DNA extracted from each isolate using the PCR conditions recommended by Carattoli *et al*., 2005. Positive and negative controls were included in each reaction and amplicon sequencing for authenticity confirmation. In addition, the number and size of plasmids were determined by pulse field gel electrophoresis (PFGE), as described by Ferreira *et al*., 2019. Briefly, cell suspensions were prepared in a cell suspension buffer (100 mM Tris-HCl, pH 8; 100 mM EDTA, pH 8) reaching a final turbidity of

1.4-1.5 at 610 nm and cells were lysed in solidified plugs (1% melted SeaKem Gold agarose (Lonza, Switzerland)) using cell lysis buffer (50 mM Tris-HCl, pH 8; 50 mM EDTA, pH 8; 1% N-lauroylsarcosine sodium salt) and proteinase K (20 mg/mL) for 2 h at 55 °C in an incubator with constant agitation (150 rpm). Plugs were washed twice with sterile ultrapure water at 55 °C for 15 min and four times with TE buffer at 55 °C for 15 min with constant agitation and stored at 4 °C until further analyses. DNA in the plugs was digested with S1 nuclease (per plug: 12.5 µL S1 buffer; 50 U S1 nuclease; total volume of 125 µL)(Thermo Scientific, USA) for 30 min at 37 °C. The PFGE was performed using a 1% SeaKem Gold agarose gel run in CHEF III DR System (Bio-Rad, Laboratories, Hercules, CA, United States) with 0.5 X TBE (45 mM Tris-HCl, pH 8.0; 45 mM boric acid; 1 mM EDTA) for 18 h at 14 ºC, with an initial switch time of 6.8 s and final switch time of 35.4 s, 6 V, with a 120º angle. Gels were stained with ethidium bromide (1 mg/mL) for 20 min and distained twice with water for 20 min. The number and size of plasmids harbored by each isolate were determined based on comparisons of the profiles obtained with the molecular weight of bands obtained for the control *Salmonella enterica* serovar Braenderup H9812 digested with XbaI (Thermo Scientific, USA) (Magalhães *et al*., 2015).

2.4. Conjugation assays

The trans-species conjugation assays used *Escherichia coli* J53 as recipient strain. *E. coli* J53 was selected as recipient strain as it was considered a better discriminating feature while normalized possible bias due to intraspecies genetic proximity between donor and recipient. Donors and recipient were cultured in Luria-Bertani (LB) broth overnight at 30 ºC with agitation. Conjugation was performed using a 1:3 ratio of donor and recipient cells inoculated in LB medium supplemented with ceftazidime (2 mg/L) and incubated for 20 h at 28 ºC. Putative transconjugants were selected on LB agar supplemented with sodium azide (100 mg/L) and ceftazidime (2 mg/L), incubated overnight at 37 ºC. The transconjugants were characterized based on antibiotic resistance phenotypes and genes as described above.

2.5. Biofilm formation

The capacity to form biofilm was tested on flat bottom 96-well polystyrene microtiter plates (Orange Scientific, Belgium), in LB (10 g tryptone, 5 g yeast extract, 10 g NaCl), modified LB (mLB) (5 g tryptone, 2.5 g yeast extract, 1.0 g NaCl) and mLB supplemented with cefotaxime (2 mg/L) media. A volume of 100 µL of cultures grown at 37 ºC during 18 h with an optical

density of 0.1 -0.3 at 610 nm was inoculated in a microtiter plate and incubated at 37 ºC during 24 h. After measuring the cultures turbidity at 595 nm, planktonic cells were removed by washing the wells 3 times with of 1X phosphate-buffered saline solution. The sediment, presumable biofilm, was air dried for 15 min and was fixed with 98% (v/v) methanol during 15 min, to be stained with 0.1% of crystal violet solution for 10 min at room temperature. The excess of stain was removed under low running tap water and the dye bound to the adherent cells was suspended in 100 µL of ethanol 95% (v/v), during 15 min at room temperature. The optical density of this solution was measured at 570 nm. All assays were made in triplicate. Biofilm formation capacity was classified into 4 categories: negative ($OD \leq OD_c$), weak ($OD_c < OD \leq 2xOD_c$), moderate ($2xOD_c < OD \leq 4xOD_c$) and strong ($4xOD_c < OD$). OD refers to the optical density measured in each well and OD_c = average OD value for the negative control + 3 x standard deviation of average OD values of the negative control (Stepanovic *et al*., 2000).

2.6. Infection capacity

Infection capacity assays used *Galleria mellonella* as model organism. A group of 47 isolates (23 wastewater, 20 clinical and 4 clinical environment) selected based on the capacity to form biofilm (highest scores), the capacity for trans-species conjugation, and/or the detection of 2 or more plasmids by PFGE, were tested as previously described (Mil-Homens *et al*., 2010). Plate Count Agar (PCA) overnight cultures, at 37 °C, were suspended in saline solution (0.85% NaCl) at a density of ~$1x10^5$ CFUs/µL. A volume of 5 µL of that suspension was injected into each *G. mellonella* larva, in 10 larvae replicates that were incubated at 37 °C in the dark for 72 h. Test larvae were starved overnight at 37 ºC, in the dark, were approximately 2–3 cm in length and had no signs of darkening. The injection site (the hindmost left proleg) was disinfected with ethanol and the injection was performed using a micrometer adapted to control the injection volume onto a micro-syringe (Mil-Homens *et al*., 2016). At least three independent experiments for each isolate were performed, therefore for each isolate were tested 30 larvae. As negative control, 10 larvae were injected with 0.85% NaCl in each experiment. After the incubation period, selected because after 72 h may be observed pupa formation (Harding *et al*., 2012), the injected larvae were individually examined for survival, movement, cocoon formation and melanization, according to Tsai *et al*., 2016. The *G. mellonella* Health Index Scoring System was used to assess the larvae health status where a score ≥9 represents a healthy larva and a score <9 represents an infected larva (Loh *et al*., 2013). From 0-8 scores, the lowest the score, the stronger the infection capacity of the bacterial isolate and consequently more severe was the effect on larva health.

2.7. Genome analysis

A group of 22 isolates (7 of wastewater, 11 clinical, 4 of clinical environment) were selected for whole genome sequencing based on the phenotypic and genotypic characterization performed (Table S3). Representative genomes from the groups determined (Figure 1) and presenting clinically relevant traits such as moderate/strong biofilm producers, conjugative, presenting multidrug-resistance, among others characteristics were selected for genomic analysis. The genomes sequences were determined using the paired-end Illumina HiSeq, the quality of the reads obtained was checked with the FastQC v0.11.8 software and the genomes were assembled using SPAdes v3.11.1 software. The resulting contigs with low coverage (<2%), or with a size below 500 bp were removed and the quality of the genomes was assessed using the CheckM method (Parks *et al*., 2015).

The coverage of genomes was determined based on the formula C=NxL/G (C, coverage; N, number of reads; L, average read length; G, genome size). The whole genomes shotgun sequences obtained in this study have been deposited in the DDBJ/ENA/GenBank database. Accession numbers are indicated in Table S3. The version described in this paper are the first version. Genomes average nucleotide identity based on the blast algorithm (ANIb) and the percentage in GC was determined using JSpeciesWS online service (http://jspecies.ribohost.com/jspeciesws/#analyse).

For genome comparison whole genomes were annotated using PROKKA version 1.12 (Seemann, 2014) and the amino acid sequences obtained for all the functional categories for each isolate were compared using the criteria of 70% similarity over 50% of coverage alignment using the GET_HOMOLOGUES software (Contreras-Moreira & Vinuesa, 2013). Average identity matrixes were calculated with BLASTp scores among the protein sequences of the genomes and a dendrogram representing the degree of similarity of the genomes based on the amino acid sequences presence or absence was obtained.

The internal fragments of a group of 7 housekeeping genes (*gapA*, *infB*, *mdh*, *pgi*, *phoE*, *rpoB* and *tonB*) from the *K. pneumoniae* Multi Locus Sequence Type (MLST) scheme (Diancourt *et al*., 2005) were extracted from each *K. pneumoniae* genome to determine the MLST using the Institute Pasteur database. Concatenated sequences were aligned using MEGA7 and a phylogenetic tree was constructed using the Maximum Likelihood method with a bootstrap of 1000 replicates. Clinically relevant genome features such as the genes annotated as encoding for antibiotic and metal resistance, virulence, quorum sensing, and oxidative stress related functions and plasmid replicon type were screened on the *K. pneumoniae* whole genomes sequences. A total of 127 clinically relevant genes were downloaded from publicly available databases and used to perform BLASTn searches against a database created in house

constituted by the genomes in analyses. Metal resistance (n=35), virulence (n=26) and efflux systems (n= 17) related genes sequences were downloaded from the Institute Pasteur database (https://bigsdb.pasteur.fr/cgibin/bigsdb/bigsdb.pl?db=pubmlst_klebsiella_seqdef&page=downloadAlleles), antibiotic resistance genes (n= 21) from ResFinder database (Zankari *et al*., 2012) and cross-checked with CARD database (https://card.mcmaster.ca/), plasmids replicon type genes (n=12) were downloaded from PlasmidFinder 2.1 tool from Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/PlasmidFinder/). Quorum sensing (n=7) and oxidative stress (n=9) related genes were searched in Uniprot database using the terms "*Klebsiella pneumoniae* quorum sensing" and "*Klebsiella pneumoniae* oxidative stress" and downloaded from NCBI database. Considering that the presence of genes encoding metals resistance and yersiniabactin virulence was variable among closely phylogenetically related strains, the genetic context of those genes was explored seeking to unveil possible hints of genes acquisition. The genetic linkage between antibiotic resistance encoding genes was investigated aiming to assess acquisition patterns.

2.8. Statistical analysis

The phenotypic and genotypic characteristics of the 59 *K. pneumoniae* isolates were compared based on the Jaccard similarity index using the software Primer & Permanova v6 (Primer-e, New Zealand) and expressed as a dendrogram obtained with UPGMA algorithm. The data was organized in a 1/0 (presence1; absence =0) table for the following characteristics: 1) PCR gene screening; 2) antibiotic resistance phenotype; 3) number of plasmids; 4) size of plasmids (<150 Kbp=0; ≥150 Kbp=1); 5) plasmids replicon types; 6) conjugation with *E. coli* J53; 7) ARGs acquired in transconjugants (1 gene=0; ≥2 genes=1); 8) multidrug resistance acquired by transconjugants (≤2 antibiotic classes=0; ≥3 antibiotic classes=1) and 9) biofilm formation in LB and mLB (negative/weak=0; moderate/strong=1). Fisher's exact test, adequate to compare small samples with low frequency values (Kim, 2017), was used to evaluate statistically significant differences between clinical and environmental isolates phenotypic and genotypic characteristics using a p-value ≤ 0.05.

3. Results

3.1. Preliminary characterization based on phenotype and selected genetic traits

The characterization of the isolates (Table S1) based on phenotypic and genetic features evidenced identical profiles shared by clinical and wastewater isolates (Figure 1). In general, the prevalence values observed for each of the tested characteristics were not significantly different. An exception was the trans-species conjugation (with *E. coli* J53). A total of 73% of the clinical isolates (22/30, the clinical environment was not included in this analysis) was capable of conjugating with *E. coli*, while only 40% (10/25) of wastewater had such a capacity (p-value=0.016; Fisher's exact test) (Figure 1). Transconjugant *E. coli* J53 with acquired antibiotic resistance genes (ARGs) and multidrug resistant (MDR) phenotypes (i.e. resistance to antibiotics belonging to 3 or more classes) were more frequent for clinical than for wastewater donors (ARGs 53% - 16/30 *vs.* 36% -9/30; MDR 50% - 15/30 *vs.* 24% - 6/25, respectively), although not statistically significantly different (ARGS – p-value=0.278; MDR – p-value=0.057; Fisher's exact test) (Table S4). The capacity to form biofilm in LB medium was significantly more frequent in wastewater (72%) than in clinical (30%) isolates (p-value=0.003; Fisher's exact test). In diluted LB, with half the nutrients concentration and 10 times less sodium chloride (mLB), more isolates tested positive for biofilm formation (wastewater, 84% vs. clinical, 70%) and significant differences were not observed (p-value=0.341; Fisher's exact test) (Figure 1). The addition of cefotaxime (2 mg/L) to culture media did not affect the biofilm forming capacity. Genotypic features presented, in general, identical prevalence values in both groups. Genes encoding resistance against carbapenems and other β-lactams (bla_{CTX}, bla_{OXA}, bla_{SHV}, bla_{TEM} and bla_{KPC}) presented identical prevalence in clinical and in wastewater isolates (0.056<p-value<1.00; Fisher's exact test). Also, the PFGE analysis revealed that irrespective of the origin most isolates harbored 1-3 plasmids. Plasmids larger than 150 Kbp were present in 87% of clinical and 88% of wastewater isolates, while smaller plasmids (<150 Kbp) were more frequent in wastewater isolates (72% vs. 53% in clinical), although not statistically significant (>150 Kbp – p-value=1.00; <150 Kbp – p-value=0.177; Fisher's Exact test) (Figure 1). The *G. mellonella* Health Index scores varied between 0 and 9 (strong to no-infection capacity), being 0 and 1 more frequent in clinical (30%) than in wastewater isolates (4%) (p-value=0.038; Fisher's exact test). While the score 9 ranged 20 to 22% in both groups, scores 2-8 were more frequent in wastewater (74%) than in clinical isolates (50%), though not statistically significant (p-value=0.127; Fisher's exact test) (Figure 1).

The pheno- and genotypic characterization was integrated based on a numerical taxonomy approach (59 strains, 27 characteristics) (Figure 1). The resultant dendrogram permitted the

ad hoc definition of 6 groups (A to F) (cut-off ~60%). Clinical and wastewater isolates clustered together, and 6 isolates were unclustered (Figure 1 and Table S5). These results suggest that wastewater isolates, regardless the type of wastewater, retained clinically relevant features.

Aiming to better explore these results, 22 isolates included in clusters A, B, D, E and F (Figure 1) were selected for whole genome sequence analysis (Table S5). These were 7 of wastewater (Portugal, 4 RWW, 2 tTWW and 1 hospital effluent), 11 of patients (7 Portugal; 4 Spain) and 4 of clinical environment (Spain). This selection comprised isolates with clinically relevant traits, such as trans-species conjugation and MDR transfer, meropenem resistance, biofilm formation capacity and distinct *G. mellonella* infection indices (0-9). Group C, composed of three isolates unable to conjugate with *E. coli* J53 and not yielding any of the plasmid replicon types tested, were not included in this analysis, mainly interested in assessing clinically-relevant features occurring in distinct lineages.

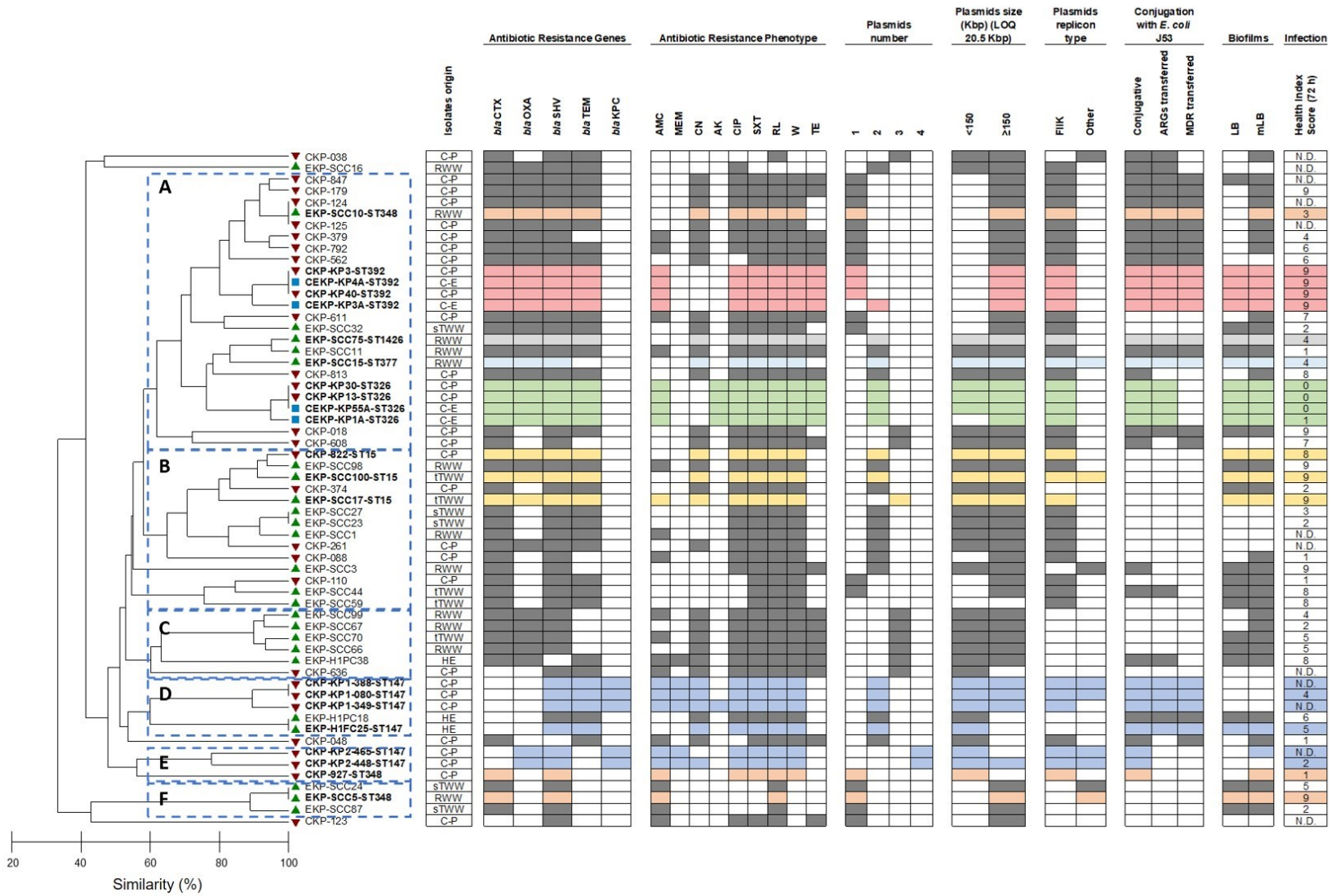

Antibiotic Resistance Genes
Antibiotic Resistance Phenotype
Plasmids number
Plasmids size (Kbp) (LOQ 20.5 Kbp)
Plasmids replicon type
Conjugation with E. coli J53
Biofilms
Infection
Isolates origin
blaCTX
blaOXA
blaSHV
blaTEM
blaKPC
AMC
MEM
CN
AK
CIP
SXT
RL
W
TE
1
2
3
4
<150
≥150
FIIK
Other
Conjugative
ARGs transferred
MDR transferred
LB
mLB
Health Index Score (72 h)
A
B
C
D
E
F
CKP-038
EKP-SCC16
CKP-847
CKP-179
CKP-124
EKP-SCC10-ST348
CKP-125
CKP-379
CKP-792
CKP-562
CKP-KP3-ST392
CEKP-KP4A-ST392
CKP-KP40-ST392
CEKP-KP3A-ST392
CKP-611
EKP-SCC32
EKP-SCC75-ST1426
EKP-SCC11
EKP-SCC15-ST377
CKP-813
CKP-KP30-ST326
CKP-KP13-ST326
CEKP-KP55A-ST326
CEKP-KP1A-ST326
CKP-018
CKP-608
CKP-822-ST15
EKP-SCC98
EKP-SCC100-ST15
CKP-374
EKP-SCC17-ST15
EKP-SCC27
EKP-SCC23
EKP-SCC1
CKP-261
CKP-088
EKP-SCC3
CKP-110
EKP-SCC44
EKP-SCC59
EKP-SCC99
EKP-SCC67
EKP-SCC70
EKP-SCC66
EKP-H1PC38
CKP-636
CKP-KP1-388-ST147
CKP-KP1-080-ST147
CKP-KP1-349-ST147
EKP-H1PC18
EKP-H1FC25-ST147
CKP-048
CKP-KP2-465-ST147
CKP-KP2-448-ST147
CKP-927-ST348
EKP-SCC24
EKP-SCC5-ST348
EKP-SCC87
CKP-123
20
40
60
80
100
Similarity (%)

Figure 1 - Clustering analysis using Jaccard similarity based on distinctive traits of the 3rd generation cephalosporin-resistant *K. pneumoniae* isolates. Antibiotic classes tested: β-lactams and carbapenems (AMC-Amoxicillin+Clavulanate (20/10 µg); MEM-Meropenem (10 µg)); aminoglycosides (CN-Gentamicin (10 µg)); AK-Amikacin (30 µg)); quinolones (CIP-Ciprofloxacin (5 µg)); sulfonamides (SXT-Sulfamethoxazole/Trimethoprim (1.25/23.75 µg); RL-Sulfamethoxazole (25 µg); W–Trimethoprim (5 µg)); and tetracyclines (TE-Tetracycline(30 µg)). ARGs *bla*$_{IMP}$, *bla*$_{VIM}$ and *mcr* were also screened and were not detected in none of the isolates. All the isolates showed a resistance phenotype against amoxicillin and aztreonam. Red, green and blue symbols in the dendrogram refer to isolates obtained from patients, wastewater and clinical settings, respectively. Colors in the table refer to the trait presence and different colors refer to different sequence types. A to F letters in the dendrogram refer to the main clusters identified. C-P – clinical isolate obtained from patients; C-E – isolate obtained from clinical environment; HE – hospital effluent; RWW – raw influent wastewater; sTWW – secondary treatment effluent wastewater; tTWW – tertiary treatment effluent wastewater; N.D. - not determined; LB - Luria-Bertani; mLB - modified Luria-Bertani; ST – sequence type; CKP – clinical K. pneumoniae; EKP – environmental K. pneumoniae; CEKP – clinical environment K. pneumoniae.

3.2. Comparative genome analyses

The 22 isolates were affiliated to 7 multi-locus sequence types (MLST), 3 of which included wastewater and clinical isolates (ST348, ST15, ST147). Two were represented by single wastewater isolates (ST377 and ST1426). The other two corresponded to the Spanish outbreak isolates (ST326 and ST392), each with two clinical and two environmental isolates (Figure 2A).

The dendrogram produced based on the average nucleotide identity values (98-100%) (blast algorithm, ANIb) supported the affiliation to 7 MLST groups (Figure 2B). Also, the comparison of the genomes based on the presence/absence of annotated deduced amino acid sequences supported the formation of the same groups (Figure 2C). The three types of analyses evidenced the closest relationship between ST15 and ST326, and ST147 and ST392. This organization did not coincide with that reported in Figure 1, with isolates of the ST147 and ST348 divided by different groups (D and E; and A, E and F, respectively), while others were clustered in groups A or B (ST326, ST392 and ST15, respectively).

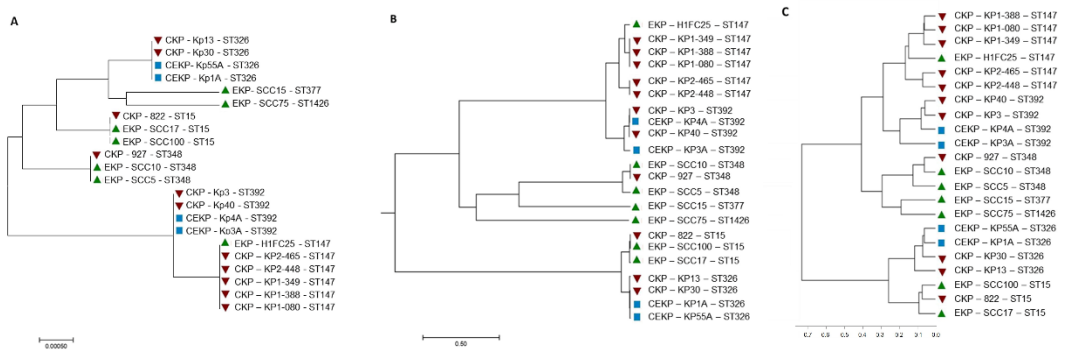


Figure 2 - Genome-based phylogenetic analysis of the 22 3rd generation cephalosporin-resistant *K. pneumoniae* isolates. (A) Maximum Likelihood Tree based on the concatenated MLST gene sequences (*gapA*, *infB*, *mdh*, *pgi*, *phoE*, *rpoB*, and *tonB*); (B) UPGMA dendrogram based on ANIb pairwise values comparisons among the genomes, and (C) UPGMA dendrogram representing the degree of similarity of the genomes based on the amino acid sequences presence or absence. Red, green and blue symbols in the dendrogram refer to isolates obtained from patients, wastewater and clinical environment, respectively.

A further analysis focused on genetic determinants related to efflux systems, oxidative stress, quorum sensing, virulence, plasmid replicon types and antibiotic and metal resistance genes, as these may be associated to opportunistic pathogens or acquired traits. The genetic determinants related to efflux systems, oxidative stress, and quorum sensing detected were common to all the examined genomes (Table S6), irrespective of the origin or genetic lineage (Figure 3). Also, *mrk* genes, encoding Type 3 fimbriae involved in bacterial adhesion, and *wzi* and *wzc* genes, involved in bacterial capsule production, and the ARG *bla*$_{SHV}$ were observed in the 22 genomes (Figure 3). However, these were sometimes represented by different variants (Table S6). The presence of genetic determinants encoding antibiotic resistance, plasmid replicon types, metal resistance, and virulence was variable among the 22 genomes. Antibiotic resistance genes rarely observed were *aadB*1 (in ST392, 1 clinical environment), *ermB* (in ST147, 1 hospital effluent), and *arr-3* (in ST147, 2 clinical). In contrast, the genes *bla*$_{OXA}$, *dfrA*, *fosA* and *aac(6´)Ib-cr* were common to most of the 22 isolates, with a few exceptions observed in 1 ST348 RWW isolate where *bla*$_{OXA}$, *dfrA* and *aac(6´)Ib-cr* were not detected, in 1 ST147 hospital effluent isolate lacking *bla*$_{OXA}$ and *aac(6´)Ib-cr*, and in 1 ST392 clinical isolate that did not yield the *fosA* gene (Figure 3 and Table S6).

The most common plasmid replicon type was IncFIB (K), curiously not detected in 4 genomes of the ST147 (1 hospital effluent and 3 clinical) and in one ST348. The plasmid replicon type

IncFIA was observed in more than half of the isolates (n=13), although in any the ST392, ST1426, ST348, or ST377. Two uncommon plasmid replicon types were colRNAI and IncX4, observed in a single hospital effluent ST147 isolate that lacked the IncFIB (K). Two other uncommon replicon types were IncHI2 and IncHI2A detected in the ST392 isolate of the clinical environment in the Spanish outbreak.

Metal resistance genes were associated with tellurium, mercury, arsenic, copper and silver. Tellurium resistance genes, presumably organized in an operon, were observed in the genomes of 6 isolates, 5 of which of Portugal (2 of wastewater, ST377 and ST348 and 3 clinical, ST147). These genes were not present in all isolates of the same ST, suggesting their acquired character. This was confirmed by the occurrence of *ter* genes in a clinical environment isolate of the ST392 recovered during the outbreak. Curiously, this isolate harbored the unique IncHI2 and IncHI2A plasmid replicon types, referred above, although *ter* and these Inc groups were not in the same contig. The genes putatively constituting the mercury operon were observed in 9 genomes, all (n=4) of the ST326 (clinical and clinical environment), 3 of wastewater (2, ST15 and 1, ST147) and 1 clinical (ST15) and 1 of clinical environment (ST392). It was suggested that *mer* genes were acquired by ST147 and ST392 isolates in clinical context (hospital effluent and clinical environment, respectively). Genes associated with the copper and silver operon were detected in all ST326 and ST392 isolates, being variable among the ST15 (1 clinical and 1 wastewater isolates) and the ST147 (2 clinical isolates). Arsenic-resistance related genes, described as part of the arsenic operon, were detected in all isolates that also presented the copper and silver resistance genes, being a wastewater and clinical ST348 isolates the exception. The other ST348 isolate with origin in wastewater only yielded *ter* genes, and curiously it was the only one that lacked the IncFIB(K) plasmid replicon (Figure 3). These observations suggest that in some cases the metal resistance genes were acquired, as they varied within the same genetic lineage and were associated with plasmid replicon types. However, these genes were not lost in wastewater, not even after treatment.

Virulence genes related with iron transport, including siderophore production and capsular serotype K2 were variable among the examined genomes, although these variations were mainly associated with phylogenetic lineages. The only exceptions were observed for isolates of ST15 and ST147. The virulence genes *kfu*, related to iron transport system, were detected in all the isolates (n=7) of the ST15 and ST326. The presence of the genes *fyuA, irp* and *ybt* associated with yersiniabactin siderophore production, detected in 6 isolates, was variable among and within lineages, suggesting its acquired character. These genes were observed in all ST348 isolates, in 1 (out of 3) ST15 isolates, and in 2 (out of 6) ST147 isolates. Interestingly, these 2 ST147 isolates were those yielding the arsenic, cooper and silver resistance genes, while lacked mercury or tellurium resistance genes, present in other genomes of the same

ST147. The genes *kpiA* were observed in all isolates (n=15) of ST392, ST1426, ST348, ST147, and ST377 (Figure 3).

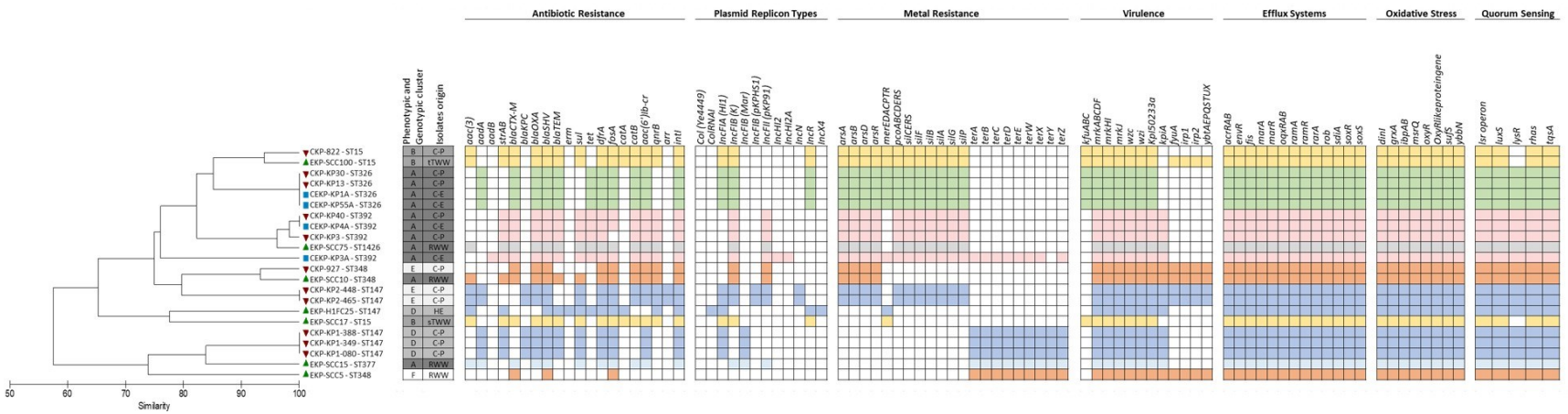


Figure 3 - Clustering analysis of 22 3rd generation cephalosporin-resistant *K. pneumoniae* based on selected clinically relevant features (genes encoding antibiotic resistance, metal resistance, virulence, efflux systems, oxidative stress, quorum sensing and plasmid replicon types). A presence/absence table supported the Jaccard similarity index estimation and UPGMA clustering analysis. Colors in the table refer to the trait presence and different colors refer to different isolates MLST. Red, green, and blue symbols in the dendrogram refer to isolates obtained from patients, wastewater, and clinical settings, respectively. C-P – clinical isolate obtained from patients; C-E – isolate obtained from clinical environment; HE – hospital effluent; RWW – raw influent wastewater; sTWW – secondary treatment effluent wastewater; tTWW – tertiary treatment effluent wastewater.

3.3. Antibiotic and metal resistance and virulence genes genetic context

The genetic context of acquired features, whose presence spanned distinct genetic lineages or varied within a single lineage was investigated. These comprised antibiotic (i.e. *bla*$_{CTX}$ e *bla*$_{KPC}$) and metal (*ter*, *mer*, *sil*, *pco*) resistance, and yersiniabactin-related virulence genes. The tellurium-related genes were associated to insertion sequences of the families IS*66* and IS*NCY* in ST377 (1 RWW), of the family IS*NCY* in ST147 (3 clinical) and in ST348 (1 RWW) and of the family IS*256* in ST392 (1 clinical environment) (Figure 4A, Figure S1). All the *ter* genes examined presented 99-100% sequence identity with other genomes of *K. pneumoniae* available in public databases, except in the clinical environment isolate (KP3A) recovered during the hospital Spanish outbreak (ST392). In this case, the *ter* genes presented 65-83% nucleotide sequence identity with the others analysed in this study. Moreover, this was the only isolate were the *ter* genes were associated to the Tn*3* and IS*256*. Based on BLASTn search these tellurium operon genes hinted high sequence identity with similar genes observed in the genomes of bacteria of the genera *Citrobacter*, *Enterobacter* and *Salmonella*. In the 6 genomes containing the *ter* genes were also identified *tra* genes, reported as necessary for bacterial conjugation (Virolle *et al*., 2020). These *tra* genes were *traI*, *traF*, *traG* in the 5 genomes with the *K. pneumoniae ter*-type, different from the *traN* and *traU* observed in the genome with the atypical *ter* genes (Figure 4A). Moreover, only in this latter atypical genome, other metal resistance genes (mercury, arsenic, copper, and silver) were observed beside the *ter* genes (Figure 3, Figure 4A, 4B and 4C).

Mercury-related genes were observed in distinct contexts. In the 3 ST15 isolates and in 1 out of 3 ST392 isolates, those genes were flanked by transposases. In the single ST147 isolate where this gene was detected, it was flanked by a recombinase and a gene involved in conjugation (*traC*), also observed to flank *mer* genes in the 4 ST326 isolates (Figure 4B). Therefore, depending on the sequence type of the isolates, the mercury-related genes were flanked by different mobilization genes. The genes related to the metals arsenic, copper, and silver were linked in the same contig in most of the isolates (11 out of 13 isolates) and their acquisition through horizontal gene transfer was suggested by their association to transposases in ST147, ST15, ST326, ST392 (Figure 4C, Figure S2). In the ST1426 isolate, these genes were not in the same contig (Figure 4C, Figure S2). In those 11 genomes, the silver-related genes were flanked by a transposase and the copper-related genes by an IS*L3* transposase. The arsenic-related genes were flanked either by both IS*NCY* and IS*L3* transposases (n= 8) or only by IS*L3* transposases (n=3). Two out of the 3 ST348 isolates

presented arsenic-, but not copper- or silver- related genes, and in 1 out of 2 isolates these genes were associated in the same contig in an insertion sequence (Figure 4C, Figure S2).

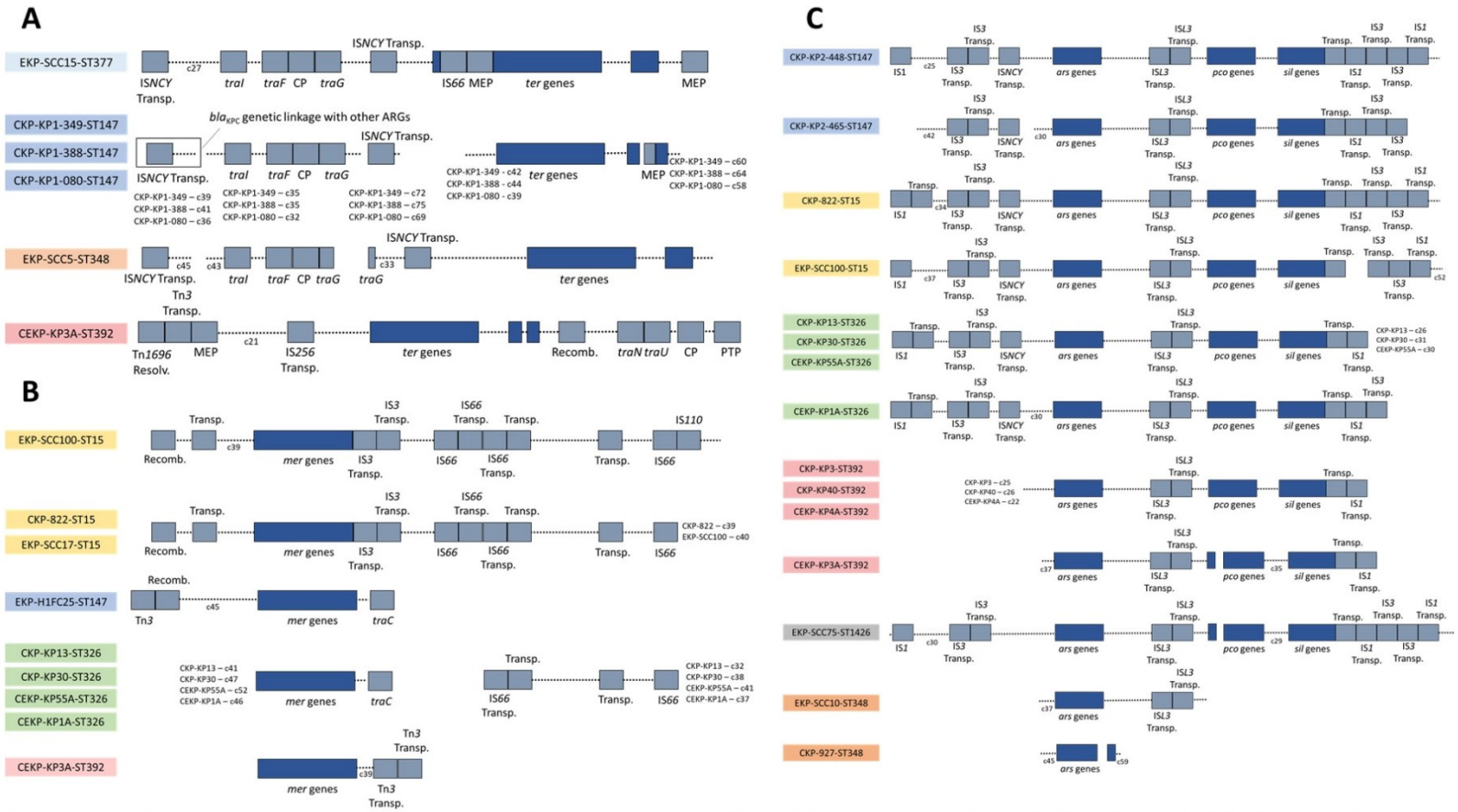


Figure 4 - Schematic presentation of the genetic environment of A) tellurium (*ter*), B) mercury (*mer*) and C) arsenic (*ars*), copper (*pco*) and silver (*sil*) resistance-related genes. The contig number (c) is indicated for each isolate close to the schematic presentations.

The yersiniabactin *locus*, involved in siderophore production (*fyuA, irp1, irp2, ybtA, ybtE, ybtP, ybtQ, ybtS, ybtT, ybtU* and *ybtX*), was observed to be linked to a prophage integrase and to P-type conjugative transfer protein and *traM* in ST15, ST147 and ST348 (Figure 5, Figure S4). The yersiniabactin siderophore production related genes were inserted in integrative and conjugative elements (ICE), whose organization varied according to the sequence type of the isolates. These genes were inserted in ICEkp4 (99.97% of identity with ICEkp4) in the ST15 isolate, in ICEkp5 (99.49% of identity with ICEkp5) in the ST348 isolates, and in ICEkp12 (99.99% of identity with ICEkp12), in the ST147 isolates.

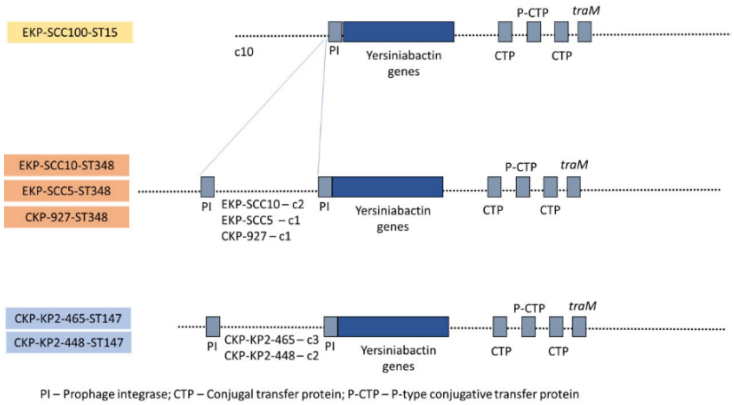


Figure 5 - Schematic presentation of yersiniabactin virulence factor genetic environment. The contig number (c) is indicated for each isolate in the cases where more than one isolate shares the same yersiniabactin genetic environment (e.g EKP-SCC10 – c2).

In general, as expected, the most variable genome features investigated within each phylogenetic group were the plasmid replicon types and the antibiotic resistance genes profile (Figure 3). The genetic linkage of antibiotic resistance genes was investigated aiming to find hints of distinct acquisition paths. The *bla*$_{\text{CTX-M-15}}$ gene, encoding antibiotic resistance to cephalosporins, was observed to be flanked by insertion sequences and/or transposases in all the genomes where this gene was detected (16/22). The *bla*$_{\text{CTX-M-15}}$ gene was associated to genes encoding resistance to β-lactams (*bla*$_{\text{TEM}}$), aminoglycosides (*strB*, *strA*), and sulfonamides (*sul*) in ST392, ST15, ST348 and, ST1426 (Figure 6). The gene *bla*$_{\text{KPC}}$, encoding resistance to carbapenems was observed to be flanked by transposases in the clinical isolates of the ST147 (n=5). The *bla*$_{\text{KPC-3}}$ gene in 2/5 clinical isolates was also associated in the same contig with genes involved in conjugation (Figure 6), as has been reported before (Ferreira *et*

al., 2021). In the other 3/5 clinical isolates the *bla*$_{KPC-3}$ gene was associated with genes encoding resistance to quinolones (*aac(6´)-Ib-cr*), aminoglycosides (*strA* and *strB*), β-lactams (*bla*$_{OXA-1}$/ *bla*$_{OXA-9}$, *bla*$_{TEM-1A}$) and sulfonamides (*sul1*/ *sul2*) (Figure 6 and Table S6).

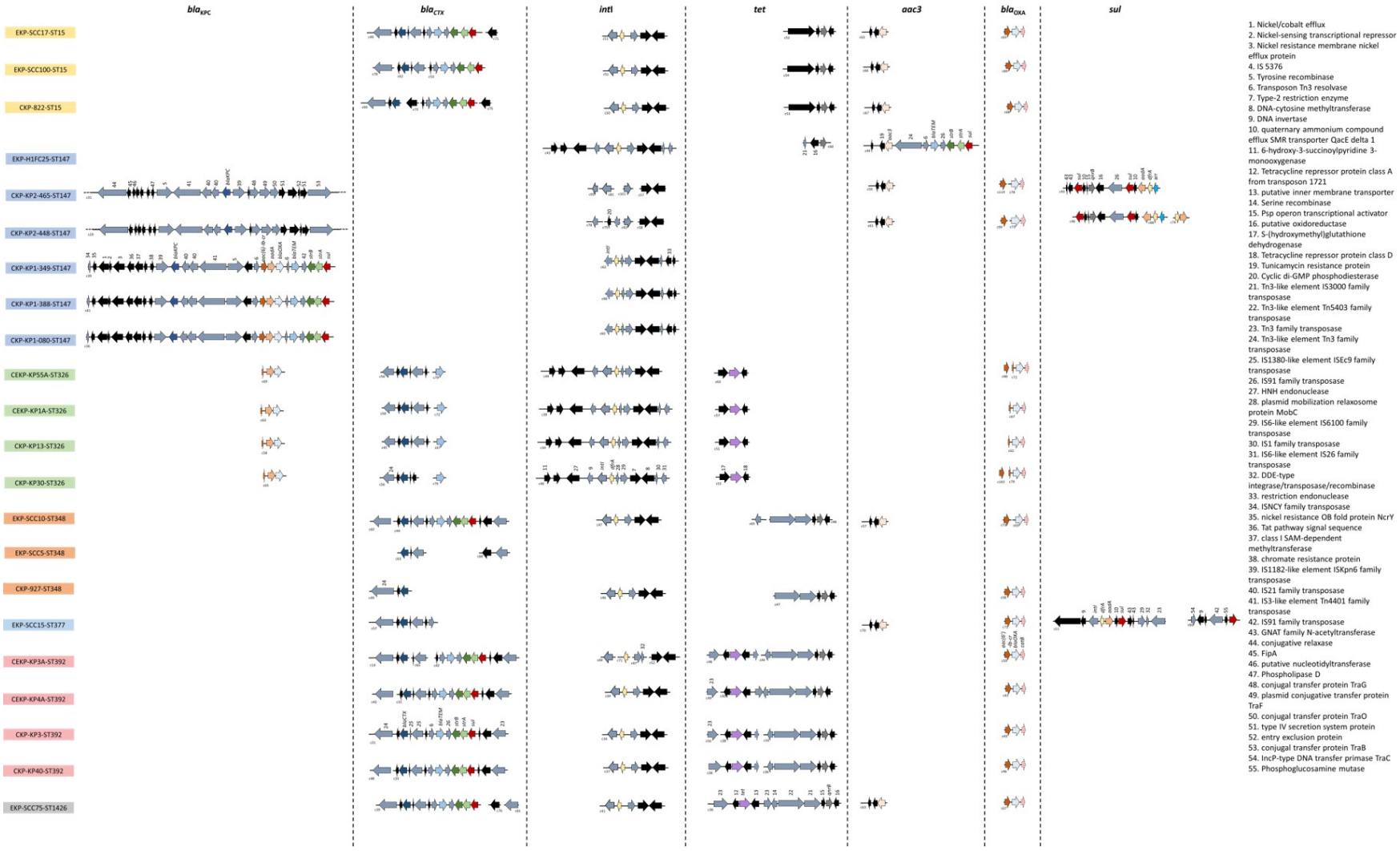

blaKPC
blaCTX
intI
tet
aac3
blaOXA
sul
EKP-SCC17-ST15
EKP-SCC100-ST15
CKP-822-ST15
EKP-H1FC25-ST147
CKP-KP2-465-ST147
CKP-KP2-448-ST147
CKP-KP1-349-ST147
CKP-KP1-388-ST147
CKP-KP1-080-ST147
CEKP-KP55A-ST326
CEKP-KP1A-ST326
CKP-KP13-ST326
CKP-KP30-ST326
EKP-SCC10-ST348
EKP-SCC5-ST348
CKP-927-ST348
EKP-SCC15-ST377
CEKP-KP3A-ST392
CEKP-KP4A-ST392
CKP-KP3-ST392
CKP-KP40-ST392
EKP-SCC75-ST1426
1. Nickel/cobalt efflux
2. Nickel-sensing transcriptional repressor
3. Nickel resistance membrane nickel efflux protein
4. IS 5376
5. Tyrosine recombinase
6. Transposon Tn3 resolvase
7. Type-2 restriction enzyme
8. DNA-cytosine methyltransferase
9. DNA invertase
10. quaternary ammonium compound efflux SMR transporter QacE delta 1
11. 6-hydroxy-3-succinoylpyridine 3-monooxygenase
12. Tetracycline repressor protein class A from transposon 1721
13. putative inner membrane transporter
14. Serine recombinase
15. Psp operon transcriptional activator
16. putative oxidoreductase
17. S-(hydroxymethyl)glutathione dehydrogenase
18. Tetracycline repressor protein class D
19. Tunicamycin resistance protein
20. Cyclic di-GMP phosphodiesterase
21. Tn3-like element IS3000 family transposase
22. Tn3-like element Tn5403 family transposase
23. Tn3 family transposase
24. Tn3-like element Tn3 family transposase
25. IS1380-like element ISEc9 family transposase
26. IS91 family transposase
27. HNH endonuclease
28. plasmid mobilization relaxosome protein MobC
29. IS6-like element IS6100 family transposase
30. IS1 family transposase
31. IS6-like element IS26 family transposase
32. DDE-type integrase/transposase/recombinase
33. restriction endonuclease
34. ISNCY family transposase
35. nickel resistance OB fold protein NcrY
36. Tat pathway signal sequence
37. class I SAM-dependent methyltransferase
38. chromate resistance protein
39. IS1182-like element ISKpn6 family transposase
40. IS21 family transposase
41. IS3-like element Tn4401 family transposase
42. IS91 family transposase
43. GNAT family N-acetyltransferase
44. conjugative relaxase
45. FipA
46. putative nucleotidyltransferase
47. Phospholipase D
48. conjugal transfer protein TraG
49. plasmid conjugative transfer protein TraF
50. conjugal transfer protein TraO
51. type IV secretion system protein
52. entry exclusion protein
53. conjugal transfer protein TraB
54. IncP-type DNA transfer primase TraC
55. Phosphoglucosamine mutase

Figure 6 - Genetic environment of the genes encoding resistance to carpanemens (*bla*$_{KPC}$), cephalosporins (*bla*$_{CTX}$), tetracyclines (*tet*), aminoglycosides (*aac3*), β-lactams (*bla*$_{TEM}$), and sulfonamides (*sul*). The genetic environment of class I integron (*int*) encoding genes is also presented. The contig number (c) is indicated for each isolate next to the contig that is represented.

4. Discussion

As pathogens causing an infection or thriving in the environment, bacteria can be exposed to distinct conditions, which hypothetically may be associated with the retention or loss of specific traits either phenotypic or genotypic (Martin & Bachman, 2018). This hypothesis was behind the comparative characterization made in this study. The phenotypic and genotypic characterization organized the isolates according to the antibiotic resistance traits, being resistance to tetracycline, meropenem, sulfonamides and amoxicillin with clavulanic acid the most differentiating. Other differentiating traits were the capacity to conjugate with *E. coli* J53 and transfer antibiotic resistance genes and the biofilm forming capacity. All the features that differentiated the groups have been observed in clinical and in environmental strains (Barati *et al*., 2016; Obasi *et al*., 2017). The comparative analysis revealed that the origin of the isolates was not determinant for group organization, although most of the clinical isolates clustered in group A (Figure 1). This group (A) was characterized by moderate/strong biofilm forming capacity, presence of β-lactamase encoding genes (*bla*$_{CTX}$, *bla*$_{OXA}$, *bla*$_{SHV}$ and *bla*$_{TEM}$), resistance to tetracycline, presence of a single plasmid of high molecular weight and of the replicon type FIIK, as well as the capacity to transfer ARGs and MDR to *E. coli* J53. Also, characteristics such as the capacity of the isolates to conjugate with *E. coli* J53 or to form biofilm were significantly more frequent in clinical or in wastewater isolates, respectively. These observations may suggest a certain degree of habitat adaptation. For instance, the trans-species conjugative capacity described in *K. pneumoniae* (Suzuki *et al*., 2019; Yi *et al*., 2010), may be favored in habitats shared by different species, as is the case of *E. coli* in the gut of patients under antibiotherapy (Wyres & Holt, 2018; Stanley *et al*., 2018). In turn, the capacity to form biofilm may represent an advantage for nutrient capture or stress protection favoring wastewater bacteria with those properties (Paczosa *et al*., 2016). Trans-species conjugation and biofilm formation were apparently widespread over the distinct genetic lineages, although with curious exceptions - ST15 strains were not observed to conjugate with *E. coli* J53 and some clinical ST147 strains did not form biofilm. However, these differences are probably not associated with the origin of those strains, whose ubiquity is suggested in the literature, with the ST15, ST147 and ST348 described in clinical and environmental sources

(Diestra *et al*., 2010; Fasciana *et al*., 2019; Navon-Venezia *et al*., 2017; Wyres & Holt, 2016; Marques *et al*., 2019; Trigo Da Roza *et al*., 2019). The ST326, ST392, ST377 and ST1426 have been reported in other studies as being associated to the clinical settings (Fursova *et al*., 2020; Shelenkov *et al*., 2020; Zenati *et al*., 2017). Indeed, ST326 and ST392 were the lineages recovered from an outbreak in a hospital, and the capacity of a clinical environment ST392 isolate to acquire metal resistance genes was strongly suggested in this study. Also, to our knowledge, ST377 and ST1426 were never reported in wastewater.

Genes related to efflux, oxidative stress or quorum sensing were detected in all the 22 isolates examined. This observation agrees with the literature that identify some genes related with efflux and quorum sensing as part of the *K. pneumoniae* core genome (Wyres *et al*., 2020; Lery *et al*., 2014). An exception was the regulator *lysR* gene not detected in isolates of the ST15, which is involved in quorum sensing, oxidative stress response, and has also been associated to the regulation of virulence factors, mainly to the expression of adhesins in early stages of the biofilm formation, being important in the process of infection (Hennequin & Forestier, 2009; Reading & Sperandio, 2005). This observation is in agreement with the ongoing discussion about the truncated nature of this regulator in members of this lineage (Machuca *et al*., 2018; Naha *et al*., 2021). Interestingly, these isolates were moderate/strong biofilm producers and were not able to infect *G. mellonella*, suggesting the importance of *lysR* for infection. The genes *mrk*, *wzc*, *wzi* and bla_{SHV} were also observed in the 22 genomes analysed, which is in agreement with previous observations due to its localization in the chromosome (Holt *et al*., 2015; Paczosa & Mecsas, 2016; Wyres *et al*., 2020).

Genes whose detection varied among the examined genomes were related with antibiotic and metal resistance or virulence. Antibiotic resistance genes bla_{KPC} and bla_{CTX} were genetically linked to other ARGs and were observed to be mainly associated with transposons of the types Tn4401 and Tn3-like. According to the literature, these transposons are widespread in *K. pneumoniae* and are often related with acquisition of bla_{KPC} and bla_{CTX} (Rodrigues *et al*., 2016; Mbelle *et al*., 2020). Metal resistance-related genes were associated with the insertion sequences IS66, ISNCY, ISL3, IS3 or IS1, in a pattern shared by isolates of different lineages or origins, except one isolate of the ST392 in which was detected the IS256. The fact that these metal-related genes were flanked by insertion sequences suggests the potential for mobilization (Hendrickx *et al*., 2020; Håkonsholm *et al*., 2020). One ST392 isolate was the only Spanish outbreak strain where the *ter* genes were detected, and it presented a unique sequence and context when compared to the other isolates included in this study. Since this was an outbreak isolate it may suggest different paths of gene acquisition from other species, due to selective pressures, as is hinted by the high sequence similarity with homologous genetic elements reported in other genera.

Among the virulence genes detected, genes related to the production of yersiniabactin siderophores, which enable iron acquisition from the host to survive and propagate during the infection process (Paczosa *et al*., 2016), were observed in some ST15 and ST147 and in all ST348 isolates. The yersiniabactin *loci* were associated with integrative and conjugative elements (ICE), as described by Lam and colleagues (2018). Among these, ICEkp4, ICEkp5 and ICEkp12 were specific of the sequence type of the isolates, considering that the first was observed in ST15, the second in ST348 and the third in ST147. This is in agreement with the literature, although in the same sequence types other ICE associated with virulence genes have been reported (Lam *et al*., 2018). Nevertheless, the results suggested that the phylogenetic lineage, more than the origin of the isolates, might explain the paths of acquisition of virulence genes. ICEkp are typically constituted by a P4-like integrase gene in the left end, followed by the yersiniabactin locus, and by a mobilization module constituted by the *xis* excisionase, *virB*-type 4 secretion system (T4SS), *ori*T transfer origin and *mobBC* proteins (responsible for mobilization) (Lam *et al*., 2018). In some ICE structures it is also found a zinc and manganese metabolism module (Lam *et al*., 2018), observed in our study in ICEkp4, ST15 isolates.

In this study it was observed that genes related to antibiotic or metals resistance were flanked by insertion sequences, transposases or genes involved in bacterial conjugation. Moreover, genes related to virulence were flanked by ICEs. This observation suggests that the mobilization of the three types of genes uses different mechanisms. Moreover, the genetic analysis of the mobilization structures suggests that the genetic lineage, rather than the source of isolation, are determinant for the genotype and phenotype of the strain. Evidences that isolates of the ST392, ST147, ST15 and ST348 have the potential to acquire or lose metal resistance genes and that such a capacity is related with some plasmid replicon types was evidenced in this study despite the limited number of genomes examined.

Plasmids are present in almost all *K. pneumoniae* isolates with a broad range of replicon types associated (Navon-Venezia *et al*., 2017). Among them, *bla*$_{\text{CTX-M-15}}$ is commonly associated to IncFII plasmids that simultaneously carry other antibiotic resistance genes (Wyres & Holt, 2016). Indeed, the isolates of the ST377, ST392, ST348 and ST147 that harbored the gene *bla*$_{\text{CTX-M}}$, all presented the *bla*$_{\text{CTX-M-15}}$ variant and tested positive for the replicon type IncFII. Also, the *bla*$_{\text{KPC-3}}$ observed in clinical isolates of the ST147 was previously described as being associated to the replicon type IncN (Ferreira *et al.,* 2021), although this and other replicon types are reported as vectors for these carbapenemase encoding genes (IncX3, IncR, IncHI1 and IncI2) (Navon-Venezia *et al*., 2017).

A major question of this study was whether the isolation habitat, clinical or wastewater, could influence *K. pneumoniae* isolates features. The phenotypic and genomic studies did not

evidence features that highlighted specialization to the isolation habitat. The results suggest that clinical isolates once in wastewater may retain clinically relevant traits, even those that were acquired through horizontal gene transfer and were associated with transposons, insertion sequences or integrative and conjugative elements. Moreover, it is suggested that phylogeny, more than the isolates origin, may explain the profile of acquired traits, although genetic variation may occur within the same genetic lineage.

5. References

Bagley, S. T. (1985). Habitat association of *Klebsiella* species. *Infection control, 6*(2), 52–58. https://doi:10.1017/s0195941700062603.

Barati, A., Ghaderpour, A., Chew, L. L., Bong, C. W., Thong, K. L., Chong, V. C., & Chai, L. C. (2016). Isolation and characterization of aquatic-borne *Klebsiella pneumoniae* from tropical estuaries in Malaysia. *International Journal of Environmental Research and Public Health*, *13*(426), 1-16. https://doi.org/10.3390/ijerph13040426

Beceiro, A., Tomás, M., & Bou, G. (2013). Antimicrobial resistance and virulence: A successful or deleterious association in the bacterial world? *Clinical Microbiology Reviews*, *26*(2), 185–230. https://doi.org/10.1128/CMR.00059-12

Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Jones, L., Delannoy-Vieillard, A. S., Garin, B., Hello, S. Le, Arlet, G., Nicolas-Chanoine, M. H., Decré, D., & Brisse, S. (2014). Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerging Infectious Diseases*, *20*(11), 1812–1820. https://doi.org/10.3201/eid2011.140206

Bisiklis, A., Papageorgiou, F., Frantzidou, F.,& Alexiou-Daniel, S., (2007). Specific detection of *bla*$_{VIM}$ and *bla*$_{IMP}$ metallo-β-lactamase genes in a single real-time PCR. Clinical Microbiology and Infection, *13*(12), 1201–1203. https://doi.org/10.1111/j.1469-0691.2007.01832.x

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

CLSI. (2016). *Clinical and Laboratory Standards Institute: Performance Standards for Antimicrobial Susceptibility Testing Supplement M100S*.

Contreras-Moreira, B., & Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Applied and Environmental Microbiology*, *79*(24), 7696–7701. https://doi.org/10.1128/AEM.02411-13

Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. D., & Brisse, S. (2005). Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *Journal of Clinical Microbiology, 43*(8), 4178–4182. https://doi.org/10.1128/JCM.43.8.4178-4182.2005

Diestra, K., Miró, E., Martí, C., Navarro, D., Cuquet, J., Coll, P., & Navarro, F. (2010).

Multiclonal epidemic of *Klebsiella pneumoniae* isolates producing DHA-1 in a Spanish hospital. *Clinical Microbiology and Infection*, *17*(7), 1032–1036. https://doi.org/10.1111/j.1469-0691.2010.03319.x

Dipersio, J.R., Deshpande, L.M., Biedenbach, D.J., Toleman, M.A., Walsh, T.R., & Jones, R.N. (2005). Evolution and dissemination of extended-spectrum β-lactamase-producing *Klebsiella pneumoniae*: Epidemiology and molecular report from the SENTRY Antimicrobial Surveillance Program (1997-2003). Diagnostic Microbiology and Infectious Disease, *51*(1), 1–7. https://doi.org/10.1016/j.diagmicrobio.2004.08.001

Fasciana, T., Gentile, B., Aquilina, M., Ciammaruconi, A., Mascarella, C., Anselmo, A., Fortunato, A., Fillo, S., Petralito, G., Lista, F., & Giammanco, A. (2019). Co-existence of virulence factors and antibiotic resistance in new *Klebsiella pneumoniae* clones emerging in south of Italy. *BMC Infectious Diseases*, *19*(1), 1–10. https://doi.org/10.1186/s12879-019-4565-3

Ferreira, C., Bikkarolla, S. K., Frykholm, K., Pohjanen, S., Brito, M., Lameiras, C., Nunes, O. C., Westerlund, F., & Manaia, C. M. (2021). Polyphasic characterization of carbapenem resistant *Klebsiella pneumoniae* clinical isolates suggests vertical transmission of the *bla*$_{KPC-3}$ gene. PLoS ONE, *16*(2), 1–18. https://doi.org/10.1371/journal.pone.0247058

Ferreira, C., Bogas, D., Bikarolla, S. K., Varela, A. R., Frykholm, K., Linheiro, R., Nunes, O. C., Westerlund, F., & Manaia, C. M. (2019). Genetic variation in the conjugative plasmidome of a hospital effluent multidrug resistant *Escherichia coli* strain. *Chemosphere*, *220*, 748–759. https://doi.org/10.1016/j.chemosphere.2018.12.130

Fouz, N., Pangesti, K. N. A., Yasir, M., Al-Malki, A. L., Azhar, E. I., Hill-Cawthorne, G. A., & El Ghany, M. A. (2020). The contribution of wastewater to the transmission of antimicrobial resistance in the environment: Implications of mass gathering settings. *Tropical Medicine and Infectious Disease*, *5*(1), 1-25. https://doi.org/10.3390/tropicalmed5010033

Fursova, N. K., Astashkin, E. I., Gabrielyan, N. I., Novikova, T. S., Fedyukina, G. N., Kubanova, M. K., Esenova, N. M., Sharapchenko, S. O., & Volozhantsev, N. V. (2020). Emergence of five genetic lines ST395NDM-1, ST13OXA-48, ST3346OXA-48, ST39CTX-M-14, and Novel ST3551OXA-48 of multidrug-resistant clinical *Klebsiella pneumoniae* in Russia. *Microbial Drug Resistance*, *26*(8), 924–933. https://doi.org/10.1089/mdr.2019.0289

Gootz, T.D., Lescoe, M.K., Dib-Hajj, F., Dougherty, B.A., He, W., Della-Latta, P., & Huard, R.C. (2009). Genetic organization of transposase regions surrounding *bla*$_{KPC}$ carbapenemase genes on plasmids from *Klebsiella strains* isolated in a New York City

Hospital. Antimicrobial Agents and Chemotherapy, *53*(5), 1998–2004. https://doi.org/10.1128/AAC.01355-08

Håkonsholm, F., Hetland, M. A. K., Svanevik, C. S., Sundsfjord, A., Lunestad, B. T., & Marathe, N. P. (2020). Antibiotic sensitivity screening of *Klebsiella* spp. and *Raoultella* spp. isolated from marine bivalve molluscs reveal presence of CTX-M-producing *K. pneumoniae*. *Microorganisms*, *8*(12), 1–15. https://doi.org/10.3390/microorganisms8121909

Harding, C. R., Schroeder, G. N., Reynolds, S., Kosta, A., Collins, J. W., Mousnier, A& Frankel, G. (2012). *Legionella pneumophila* pathogenesis in the *Galleria mellonella* infection model. *Infection and Immunity*, *80*(8), 2780–2790. https://doi.org/10.1128/IAI.00510-12

Hendrickx, A. P. A., Landman, F., de Haan, A., Borst, D., Witteveen, S., van Santen-Verheuvel, M. G., van der Heide, H. G. J., Schouls, L. M., Halaby, T., Steingrover, R., Cohen Stuart, J. W. T., Melles, D. C., van Dijk, K., Spijkerman, I. J. B., Notermans, D. W., Oudbier, J. H., van Ogtrop, M. L., van Dam, A., den Reijer, M., … Paltansing, S. (2020). Plasmid diversity among genetically related *Klebsiella pneumoniae* blaKPC-2 and blaKPC-3 isolates collected in the Dutch national surveillance. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-73440-2

Hennequin, C., & Forestier, C. (2009). oxyR, a LysR-type regulator involved in *Klebsiella pneumoniae* mucosal and abiotic colonization. *Infection and Immunity*, *77*(12), 5449–5457. https://doi.org/10.1128/IAI.00837-09

Henriques, I.S., Fonseca, F., Alves, A., Saavedra, M.J., & Correia, A., (2006). Occurrence and diversity of integrons and β-lactamase genes among ampicillin-resistant isolates from estuarine waters. *Research in Microbiology*, *157*(10), 938–947. https://doi.org/10.1016/j.resmic.2006.09.003

Holt, E.H., Wertheim, H., Zadoks, R.N., Baker, S., Whitehouse, C.A., Dance, D., Jenney, A., Connor, T. R., Hsu, L. Y., Severin, J., Brisse, S., Cao, H., Wilksch, J., Gorrie, C., Schultz, M. B., Edwards, D. J., Van Nguyen, K., Nguyen, T. V., Dao, T. T., … Thomson, N. R. (2015). Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public healt. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(27), E3574-E3581. https://doi.org/10.1073/pnas.1501049112

Kim, H.-Y., (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endoddonticts*. *42*(2), 152-155.

https://doi.org/10.5395/rde.2017.42.2.152

Lam, M. M. C., Wick, R. R., Wyres, K. L., Gorrie, C. L., Judd, L. M., Jenney, A. W. J., Brisse, S., & 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

Lery, L. M. S., Frangeul, L., Tomas, A., Passet, V., Almeida, A. S., Bialek-Davenet, S., Barbe, V., Bengoechea, J. A., Sansonetti, P., Brisse, S., & Tournebize, R. (2014). Comparative analysis of *Klebsiella pneumoniae* genomes identifies a phospholipase D family protein as a novel virulence factor. *BMC Biology*, *12*, 15–18. https://doi.org/10.1186/1741-7007-12-41

Loh, J. M. ., Adenwalla, N., Wiles, S., & Proft, T. (2013) The use of *Galleria mellonella* as an infection model for group A streptococcus. *Virulence*, *4*(5), 419–428. https://doi.org/10.4161/viru.24930

Machuca, J., López-Cerero, L., Fernández-Cuenca, F., Mora-Navas, L., Mediavilla-Gradolph, C., López-Rodríguez, I., & Pascual, Á. (2019). OXA-48-Like *Klebsiella pneumoniae* in Southern Spain in 2014-2015. *Antimicrobial Agents and Chemoterapy*, *63*(1), eo1396-18. https://doi.org/10.1128/AAC.01396-18

Magalhães, R., Almeida, G., Ferreira, V., Santos, I., Silva, J., Mendes, M. M., Pita, J., Mariano, G., Mâncio, I., Sousa, M. M., Farber, J., Pagotto, F., & Teixeira, P. (2015). Cheese-related listeriosis outbreak, Portugal, march 2009 to february 2012. *Eurosurveillance*. https://doi.org/10.2807/1560-7917.ES2015.20.17.21104

Manaia, C. M., Macedo, G., Fatta-Kassinos, D., & Nunes, O. C. (2016). Antibiotic resistance in urban aquatic environments: can it be controlled? *Applied Microbiology and Biotechnology*, *100*(4), 1543–1557. https://doi.org/10.1007/s00253-015-7202-0

Marques, C., Menezes, J., Belas, A., Aboim, C., Cavaco-Silva, P., Trigueiro, G., Gama, L. T., & Pomba, C. (2019). *Klebsiella pneumoniae* causing urinary tract infections in companion animals and humans: Population structure, antimicrobial resistance and virulence genes. *Journal of Antimicrobial Chemotherapy*, *74*(3), 594–602. https://doi.org/10.1093/jac/dky499

Martin, R. M., & Bachman, M. A. (2018). Colonization, infection, and the accessory genome of *Klebsiella pneumoniae*. *Frontiers in Cellular and Infection Microbiology*, *8*, 1–15. https://doi.org/10.3389/fcimb.2018.00004

Mbelle, N. M., Feldman, C., Sekyere, J. O., Maningi, N. E., Modipane, L., & Essack, S. Y. (2020). Pathogenomics and evolutionary epidemiology of multi-drug resistant clinical *Klebsiella pneumoniae* isolated from Pretoria, South Africa. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-58012-8

Mendes, R.E., Kiyota, K.A., Monteiro, J., Castanheira, M., Andrade, S.S., Gales, A.C., Pignatari, A.C.C., & Tufik, S., (2007). Rapid detection and identification of metallo-β-lactamase-encoding genes by multiplex real-time PCR assay and melt curve analysis. *Journal of Clinical Microbiology*, *45*(2), 544–547. https://doi.org/10.1128/JCM.01728-06

Mil-Homens, D., Ferreira-Dias, S., & Fialho, A. M. (2016). Fish oils against *Burkholderia* and *Pseudomonas aeruginosa*: In vitro efficacy and their therapeutic and prophylactic effects on infected *Galleria mellonella* larvae. *Journal of Applied Microbiology*, *120*(6), 1509–1519. https://doi.org/10.1111/jam.13145

Mil-Homens, D., Rocha, E. P. C., & Fialho, A. M. (2010). Genome-wide analysis of DNA repeats in *Burkholderia cenocepacia* J2315 identifies a novel adhesin-like gene unique to epidemic-associated strains of the ET-12 lineage. *Microbiology*, *156*(4), 1084-1096. https://doi.org/10.1099/mic.0.032623-0

Naha, S., Sands, K., Mukherjee, S., Saha, B., & Dutta, Shanta, Basu, S. (2021). OXA-181-Like Carbapenemases in *Klebsiella pneumoniae* ST14, ST15, ST23, ST48, and ST231 from Septicemic Neonates: Coexistence with NDM-5, resistome, transmissibility, and genome diversity. *MSphere*, *6*(1), e01156-20. https://doi:10.1128/mSphere.01156-20

Navon-Venezia, S., Kondratyeva, K., & Carattoli, A. (2017). *Klebsiella pneumoniae*: A major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiology Reviews*, *41*(3), 252–275. https://doi.org/10.1093/femsre/fux013

Obasi, A., Nwachukwu, S., Ugoji, E., Kohler, C., Göhler, A., Balau, V., Pfeifer, Y., & Steinmetz, I. (2017). Extended-Spectrum β-Lactamase-producing *Klebsiella pneumoniae* from pharmaceutical wastewaters in South-Western Nigeria. *Microbial Drug Resistance*, *23*(8), 1013–1018. https://doi.org/10.1089/mdr.2016.0269

Paczosa, M. K., & Mecsas, J. (2016). *Klebsiella pneumoniae*: going on the offense with a strong defense. *Microbiology and Molecular Biology Reviews*, *80*(3), 629–661. https://doi.org/10.1128/mmbr.00078-15

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and

metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Pendleton, J. N., Gorman, S. P., & Gilmore, B. F. (2013). Clinical relevance of the ESKAPE pathogens. *Expert Review of Anti-Infective Therapy*, *11*(3), 297–308. https://doi.org/10.1586/eri.13.12

Podschun, R., & Ullmann, U. (1998). *Klebsiella* spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clinical Microbiology Reviews*, *11*(4), 589–603. https://doi.org/10.1128/cmr.11.4.589

Reading, N. C., & Sperandio, V. (2005). Quorum sensing: The many languages of bacteria. *FEMS Microbiology Letters*, *254*, 1–11. https://doi.org/10.1111/j.1574-6968.2005.00001.x

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., Michael, I., & Fatta-Kassinos, D. (2013). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of the Total Environment*, *447*, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Rodrigues, C., Bavlovic, J., Machado, E., Amorim, J., Peixe, L., & Novais, Â. (2016). KPC-3-producing *Klebsiella pneumoniae* in Portugal linked to previously circulating non-CG258 lineages and uncommon genetic platforms (Tn4401d-IncFIA and Tn4401d-IncN). *Frontiers in Microbiology*, *7*, 1-8. https://doi.org/10.3389/fmicb.2016.01000

Runcharoen, C., Moradigaravand, D., Blane, B., Paksanont, S., Thammachote, J., Anun, S., Parkhill, J., Chantratita, N., & Peacock, S. J. (2017). Whole genome sequencing reveals high-resolution epidemiological links between clinical and environmental *Klebsiella pneumoniae*. *Genome Medicine*, *9*(1), 1–10. https://doi.org/10.1186/s13073-017-0397-1

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

Shelenkov, A., Mikhaylova, Y., Yanushevich, Y., Samoilov, A., Petrova, L., Fomina, V Gusarov, V., Zamyatin, M., Shagin, D., & Akimkin, V. (2020). Molecular typing, characterization of antimicrobial resistance, virulence profiling and analysis of whole-genome sequence of clinical *Klebsiella pneumoniae* isolates. *Antibiotics*, *9*(5). https://doi.org/10.3390/antibiotics9050261

Stanley, I. J., Kajumbula, H., Bazira, J., Kansiime, C., Rwego, I. B., & Asiimwe, B. B. (2018). Multidrug resistance among *Escherichia coli* and *Klebsiella pneumoniae* carried in the

gut of out-patients from pastoralist communities of Kasese district, Uganda. *PLoS ONE*, *13*(7), 1–12. https://doi.org/10.1371/journal.pone.0200093

Stepanović, S., Vuković, D., Dakić, I., Savić, B., & Švabić-Vlahović, M. (2000). A modified microtiter-plate test for quantification of staphylococcal biofilm formation. *Journal of Microbiological Methods*, *40*(2), 175–179. https://doi.org/10.1016/S0167-7012(00)00122-6

Suzuki, Y., Nazareno, P. J., Nakano, R., Mondoy, M., Nakano, A., Bugayong, M. P., Bilar, J., Perez, M. V., Medina, E. J., Saito-Obata, M., Saito, M., Nakashima, K., Oshitani, H., & Yano, H. (2020). Environmental presence and genetic characteristics of carbapenemase-producing enterobacteriaceae from hospital sewage and river water in the philippines. *Applied and Environmental Microbiology*, *86*(2). https://doi.org/10.1128/AEM.01906-19

Trigo Da Roza, F. T., Couto, N., Carneiro, C., Cunha, E., Rosa, T., Magalhães, M., Tavares, L., Novais, Â., Peixe, L., Rossen, J. W., Lamas, L. P., & Oliveira, M. (2019). Commonality of Multidrug-Resistant *Klebsiella pneumoniae* ST348 Isolates in Horses and Humans in Portugal. *Frontiers in Microbiology*, *10*, (1-9). https://doi.org/10.3389/fmicb.2019.01657

Tsai, C. J. Y., Loh, J. M. S., & Proft, T. (2016). *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. *Virulence*, *7*(3), 214–229. https://doi.org/10.1080/21505594.2015.1135289

Vaz-Moreira, I., Nunes, O. C., & Manaia, C. M. (2014) Bacterial diversity and antibiotic resistance in water habitats: Searching the links with the human microbiome. *FEMS Microbiology Reviews*, *38*(4), 761–778. https://doi.org/10.1111/1574-6976.12062

Virolle, C., Goldust, K., Djermoun, S., Bigot, S., & Lesterlin, C. (2020). Plasmid transfer by conjugation in Gram-negative bacteria: from the cellular to the community level. *Genes*, *11*(11), 1239. https://doi.org/10.3390/genes11111239

Weill, F.-X., Lailler, R., Praud, K., Kérouanton, A., Fabre, L., Brisabois, A., Grimont, P.A.D., & Cloeckaert, A., (2004). Emergence of Extended-Spectrum-β-Lactamase (CTX-M-9)-Producing Multiresistant Strains of *Salmonella enterica* Serotype Virchow in Poultry and Humans in France. Journal of Clinical Microbiology *42*(12), 5767–5773. https://doi.org/10.1128/JCM.42.12.5767–5773.2004

Wyres, K. L., & Holt, K. E. (2016). *Klebsiella pneumoniae* population genomics and antimicrobial-resistant clones. *Trends in Microbiology*, *24*(12), 944–956.

https://doi.org/10.1016/j.tim.2016.09.007

Wyres, K. L., & Holt, K. E. (2018). *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Current Opinion in Microbiology*, *45*, 131–139. https://doi.org/10.1016/j.mib.2018.04.004

Wyres, K. L., Lam, M. M. C., & Holt, K. E. (2020). Population genomics of *Klebsiella pneumoniae*. *Nature Reviews Microbiology*, *18*(6), 344–359. https://doi.org/10.1038/s41579-019-0315-1

Wyres, K. L., Nguyen, T. N. T., Lam, M. M. C., Judd, L. M., van Vinh Chau, N., Dance, D. A. B., Ip, M., Karkey, A., Ling, C. L., Miliya, T., Newton, P. N., Nguyen, L., Sengduangphachanh, A., Turner, P., Veeraraghavan, B., Vinh, P. V., Vongsouvath, M., Thomson, N. R., Baker, S., & Holt, K. E. (2019). Genomic surveillance for hypervirulence and multi-drug resistance in invasive *Klebsiella pneumoniae* from south and southeast Asia. *BioRxiv*, 1–16. https://doi.org/10.1101/557785

Xavier, B.B., Lammens, C., Ruhal, R., Malhotra-Kumar, Surbhi, Butaye, P., Goossens, H., & Malhotra-Kumar, S., (2016). Identification of a novel plasmid-mediated colistin resistance gene, *mcr-2*, in *Escherichia coli*, Belgium, June 2016. Eurosurveillance *21*, 6–11. https://doi.org/10.2807/1560-7917.ES.2016.21.27.30280

Yi, H., Xi, Y., Liu, J., Wang, J., Wu, J., Xu, T., Chen, W., Chen, B., Lin, M., Wang, H., Zhou, M., Li, J., Xu, Z., Jin, S., & Bao, Q. (2010). Sequence analysis of pKF3-70 in *Klebsiella pneumoniae*: Probable origin from R100-like plasmid of *Escherichia coli*. *PLoS ONE*, *5*(1), 1–9. https://doi.org/10.1371/journal.pone.0008601

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

Zenati, K., Sahli, F., Garcia, V., Bakour, S., Belhadi, D., Rolain, J. M., & Touati, A. (2017). Occurrence and clonal diversity of multidrug-resistant *Klebsiella pneumoniae* recovered from inanimate surfaces in Algerian hospital environment: First report of *armA*, *qnrB* and *aac(6')-Ib-cr* genes. *Journal of Global Antimicrobial Resistance*, *10*(2010), 148–153. https://doi.org/10.1016/j.jgar.2017.05.015

6. Supplementary files

6.1. Supplementary Figures

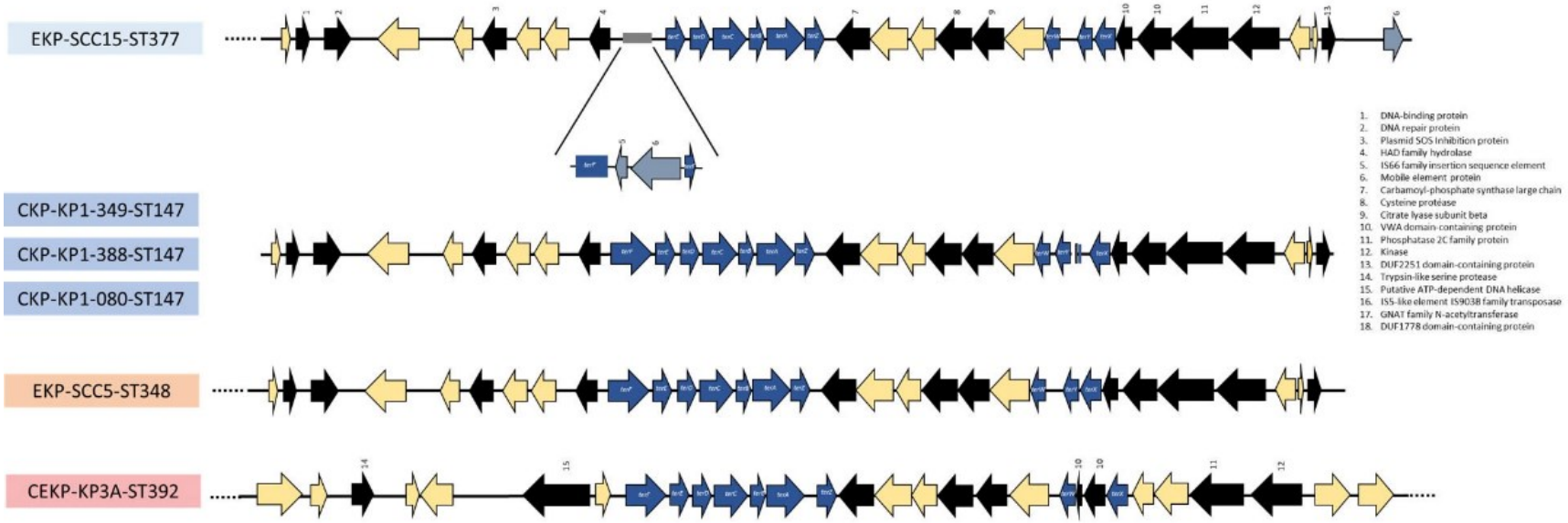


Figure S1 - Genetic environment of tellurium (*ter*) resistance-related genes. Blue arrows indicate the genes related to tellurium resistance and yellow arrows refer to hypothetical proteins.

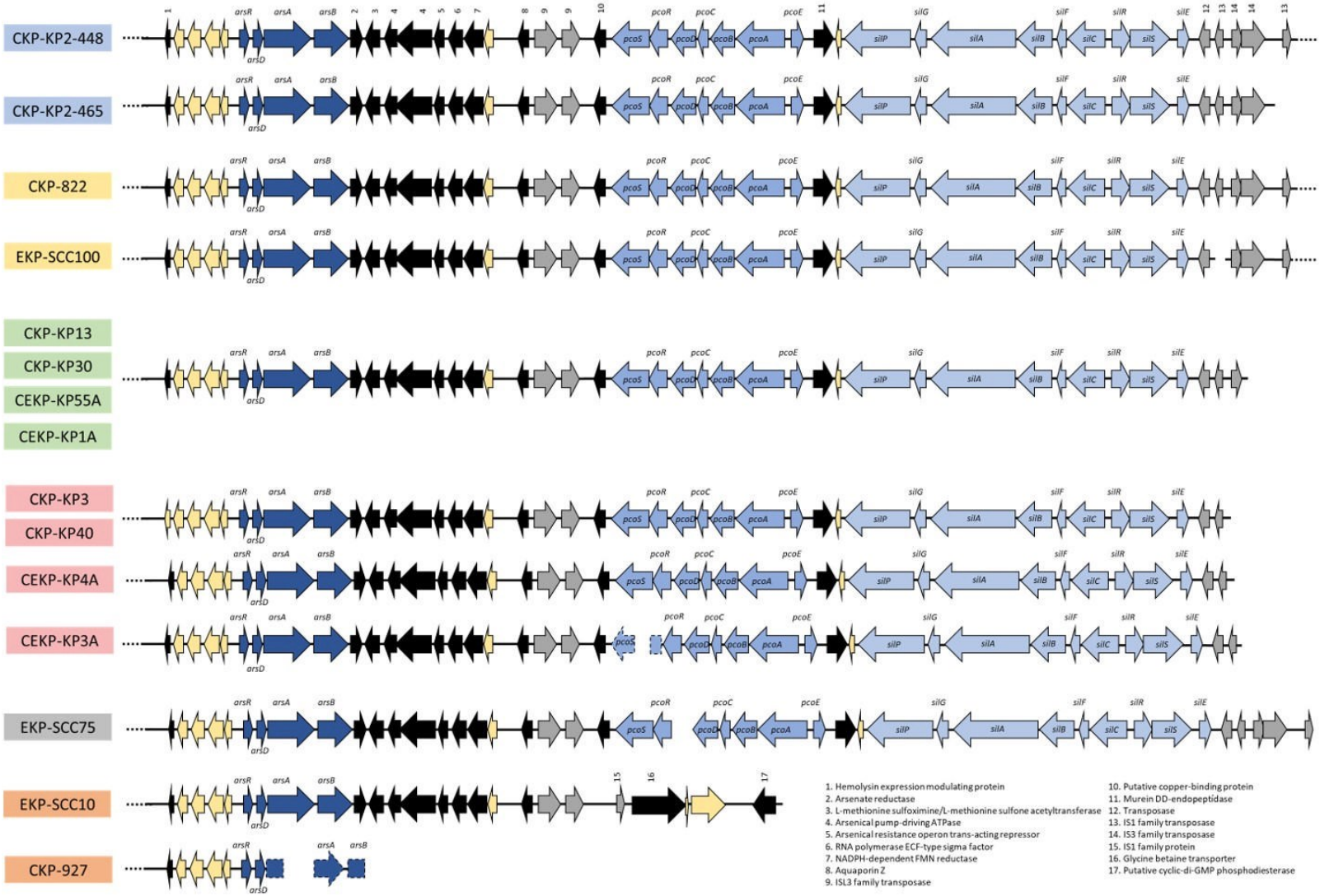

CKP-KP2-448
CKP-KP2-465
CKP-822
EKP-SCC100
CKP-KP13
CKP-KP30
CEKP-KP55A
CEKP-KP1A
CKP-KP3
CKP-KP40
CEKP-KP4A
CEKP-KP3A
EKP-SCC75
EKP-SCC10
CKP-927
1. Hemolysin expression modulating protein
2. Arsenate reductase
3. L-methionine sulfoximine/L-methionine sulfone acetyltransferase
4. Arsenical pump-driving ATPase
5. Arsenical resistance operon trans-acting repressor
6. RNA polymerase ECF-type sigma factor
7. NADPH-dependent FMN reductase
8. Aquaporin Z
9. ISL3 family transposase
10. Putative copper-binding protein
11. Murein DD-endopeptidase
12. Transposase
13. IS1 family transposase
14. IS3 family transposase
15. IS1 family protein
16. Glycine betaine transporter
17. Putative cyclic-di-GMP phosphodiesterase

Figure S2 - Genetic environment of arsenic (*ars*), copper (*pco*) and silver (*sil*) resistance-related genes. Blue arrows indicate genes related to resistance to arsenic, copper and arsenic metals, yellow arrows refer to hypothetical proteins.

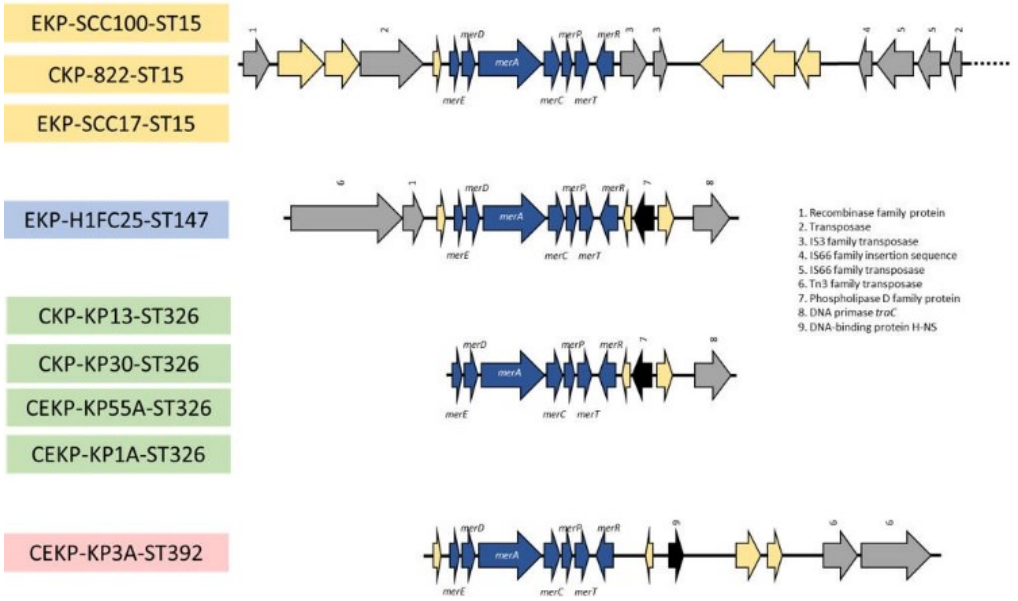


Figure S3 - Genetic context of mercury (*mer*) resistance-related genes. Blue arrows indicate the genes related to mercury resistance and yellow arrows refer to hypothetical proteins.

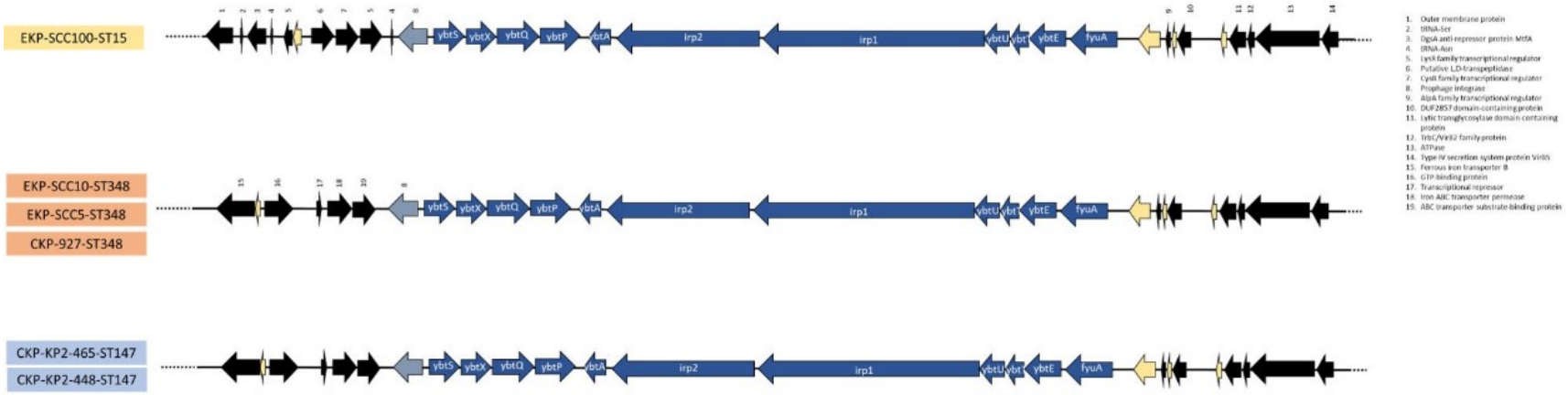


Figure S4 - Genetic environment of yersiniabactin virulence locus. Blue arrows indicate genes related to yersiniabactin virulence and yellow arrows refer to hypothetical proteins.

6.2. Supplementary Tables

Table S1 – 3rd generation cephalosporin-resistant *K. pneumoniae* isolates characterized based on β-lactam and carbapenem encoding genes, antimicrobial resistance phenotype, plasmid number, size and replicon types, conjugation properties and biofilm formation capacity. A total of 25 isolates were obtained from wastewater, 30 from patients and 4 from clinical settings.

Isolate name	Isolate origin	Isolation conditions	Isolate Category Classification
EKP-SCC1	Raw wastewater	membrane-Fecal Coliform medium + Cefotaxime (8 mg/L)	Environmental
EKP-SCC3	Raw wastewater		Environmental
*EKP-SCC5	Raw wastewater		Environmental
*EKP-SCC10	Raw wastewater		Environmental
EKP-SCC11	Raw wastewater		Environmental
*EKP-SCC15	Raw wastewater		Environmental
EKP-SCC16	Raw wastewater		Environmental
*EKP-SCC17	Secondary treatment effluent wastewater		Environmental
EKP-SCC23	Secondary treatment effluent wastewater		Environmental
EKP-SCC24	Secondary treatment effluent wastewater		Environmental
EKP-SCC27	Secondary treatment effluent wastewater		Environmental
EKP-SCC32	Secondary treatment effluent wastewater		Environmental
EKP-SCC44	Tertiary treatment effluent wastewater		Environmental
EKP-SCC59	Tertiary treatment effluent wastewater		Environmental
EKP-SCC66	Raw wastewater		Environmental
EKP-SCC67	Raw wastewater		Environmental
EKP-SCC70	Tertiary treatment effluent wastewater		Environmental

Table S1 (cont.) - 3rd generation cephalosporin-resistant *K. pneumoniae* isolates characterized based on β-lactam and carbapenem encoding genes, antimicrobial resistance phenotype, plasmid number, size and replicon types, conjugation properties and biofilm formation capacity.

Isolate name	Isolate origin	Isolation conditions	Isolate Category Classification
*EKP-SCC75	Raw wastewater		Environmental
EKP-SCC87	Secondary treatment effluent wastewater		Environmental
EKP-SCC98	Raw wastewater		Environmental
EKP-SCC99	Raw wastewater		Environmental
*EKP-SCC100	Tertiary treatment effluent wastewater		Environmental
*EKP-H1FC25	Hospital effluent	membrane-Fecal Coliform medium + Ciprofloxacin (4 mg/L)	Environmental
EKP-H1PC18	Hospital effluent	Plate Count Agar + Ciprofloxacin (4 mg/L)	Environmental
EKP-H1PC38	Hospital effluent		Environmental
*CKP-KP2-448	Urine	Cystine Lactose Eletrolyte Deficient/ Columbia agar + 5% sheep blood	Clinical
CKP-110	Urine		Clinical
*CKP-822	Urine		Clinical
CKP-847	Urine		Clinical
CKP-038	Urine		Clinical
CKP-088	Urine		Clinical
CKP-048	Urine		Clinical
CKP-123	Urine		Clinical
CKP-124	Urine		Clinical
*CKP-KP1-388	Urine		Clinical
*CKP-KP2-465	Urine		Clinical
CKP-636	Faeces	MacConkey agar	Clinical
*CKP-KP1-349	Faeces		Clinical
CKP-608	Liquid		Clinical

Table S1 (cont.) - 3rd generation cephalosporin-resistant *K. pneumoniae* isolates characterized based on β-lactam and carbapenem encoding genes, antimicrobial resistance phenotype, plasmid number, size and replicon types, conjugation properties and biofilm formation capacity.

Isolate name	Isolate origin	Isolation conditions	Isolate Category Classification
CKP-813	Liquid		Clinical
CKP-179	Blood		Clinical
*CKP-KP1-080	Blood		Clinical
CKP-125	Blood		Clinical
CKP-261	Blood		Clinical
CKP-018	Swab		Clinical
CKP-792	Swab		Clinical
CKP-562	Swab		Clinical
CKP-374	Expectoration		Clinical
*CKP-927	Expectoration		Clinical
CKP-379	Bronchial Secretions		Clinical
CKP-611	Bronchial Secretions		Clinical
*CKP-KP13	Blood	Blood agar	Clinical
*CKP-KP30	Blood		Clinical
*CKP-KP3	Blood		Clinical
*CKP-KP40	Blood		Clinical
*CEKP-KP4A	Hospital drain		Clinical environment
*CEKP-KP55A	Hospital surface		Clinical environment
*CEKP-KP3A	Hospital basin drain		Clinical environment
*CEKP-KP1A	Hospital bathroom drain		Clinical environment

Table S2 - Primer sequences and PCR amplification conditions used in this study.

Target gene	Primer name	Sequence	Annealing temperature (°C)	Amplicon size (bp)	References
***bla*$_{CTX-M}$**	blaCTX-M (fw)	CRATGTGCAGYACCAGTAA	53	540	Weill et al., 2004
	blaCTX-M (rev)	CGCRATATCRTTGGTGGTG			
***bla*$_{OXA-A}$**	blaOXA-A (fw)	ACACAATACATATCAACTTCGC	53	814	Henriques et al., 2006
	blaOXA-A (rev)	AGTGTGTTTAGAATGGTGATC			
***bla*$_{SHV}$**	blaSHV (fw)	TCGGGCCGCGTAGGCATGAT	62	626	DiPersio et al., 2005
	blaSHV (rev)	AGCAGGGCGACAATCCCGCG			
***bla*$_{TEM}$**	blaTEM (fw)	ATAAAATTCTTGAAGACGAAA	45	1080	DiPersio et al., 2005
	blaTEM (rev)	GACAGTTACCAATGCTTAATCA			
***bla*$_{KPC}$**	blaKPC (fw)	TGTCACTGTATCGCCGTCTAG	52	880	Gootz et al., 2009
	blaKPC (rev)	TTACTGCCCGTTGACGCCCAATCC			
***bla*$_{VIM}$**	q_blaVIM (fw)	GTACGCATCACCGTCGACAC	48	178	Bisiklis et al., 2007
	q_blaVIMspec (rev)	AGACGGGACGTACACAACTAAG			
***bla*$_{IMP}$**	q_blaIMP (fw)	GAATAGRRTGGCTTAAYTCTC	46	188	Mendes et al., 2007
	q_blaIMP (rev)	CCAAACYACTASGTTATC			
mcr	MCR2 (fw)	TGGTACAGCCCCTTTATT	55	1617	Xavier at al., 2016
	MCR2 (rev)	GCTTGAGATTGGGTTATGA			

Table S3 – 3rd generation cephalosporin-resistant *K. pneumoniae* isolates genomes used for comparative genomic analysis.

Isolate name	Isolate origin	Contigs number	Genome coverage	Genome size (Mbp)	G+C content (%)	Accession number
EKP-H1FC25	Hospital effluent	74	99x	5.36	57.3	JAGSXY000000000
EKP-SCC10	Raw wastewater	84	97x	5.46	57.1	JAGSZX000000000
EKP-SCC15	Raw wastewater	113	93x	5.67	56.9	JAGSZY000000000
EKP-SCC75	Raw wastewater	95	97x	5.44	57.3	JAGSZR000000000
CKP-822	Urine	90	97x	5.44	57.2	JAGSZS000000000
CKP-KP2-465	Urine	112	163x	5.75	56.9	JAGSZZ000000000
CKP-KP2-448	Urine	108	214x	5.76	56.9	JAGTAA000000000
CKP-KP1-349	Faeces	100	183x	5.65	56.9	JAGTAB000000000
CKP-KP1-388	Urine	107	201x	5.65	56.7	JAGTAC000000000
CKP-KP1-080	Blood	99	257x	5.65	56.9	JAGTAD000000000
CKP-KP3	Blood	56	16x	5.48	57.2	JAGTPB000000000
CKP-KP13	Blood	78	72X	5.72	56.9	JAGTOX000000000
CKP-KP30	Blood	100	41x	5.66	57.0	JAGTOV000000000
CKP-KP40	Blood	59	47x	5.82	56.7	JAGTOS000000000
CEKP-KP55A	Hospital surface	98	33x	5.70	56.9	JAGTOO000000000
CEKP-KP1A	Drain bathroom from hospital	91	40x	5.69	57.0	JAGTON000000000
CEKP-KP3A	Drain basin from hospital	83	50x	5.74	56.7	JAGTOL000000000
CEKP-KP4A	Hospital drain	60	57x	5.52	57.2	JAGTOK000000000
EKP-SCC5	Raw wastewater	92	63x	5.51	57.0	JAGSZT000000000

Table S3 (cont.) - 3rd generation cephalosporin-resistant *K. pneumoniae* isolates genomes used for comparative genomic analysis.

Isolate name	Isolate origin	Contigs number	Genome coverage	Genome size (Mbp)	G+C content (%)	Accession number
EKP-SCC17	Secondary Treatment Wastewater	87	50x	5.34	57.3	JAGSZU000000000
EKP-SCC100	Terciary Treatment Wastewater	87	53x	5.49	57.2	JAGSZV000000000
CKP-927	Expectoration	77	64x	5.43	57.2	JAGSZW000000000

The level of contamination was below 5% and the completeness above 90%, for all genomes.

Table S4 - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins. + indicates detection of the trait, - indicates not detection of the trait, S, I and R indicate susceptible, intermediate and resistant to the antibiotic, and NA indicates not assessed because the trait was not detected in the donor cell.

		ARGs					Antimicrobial Susceptibility												
	Isolate	blaCTX	blaOXA	blaSHV	blaTEM	blaKPC	AMC	AML	ATM	MEM	CTX	CAZ	CN	AK	CIP	SXT	RL	W	TE
Donor	**EKP-SCC3**	+	-	+	-	-	S	R	R	S	R	R	S	S	R	R	R	R	R
Transconjugant	3B	+	NA	+	NA	NA	NA	R	R	NA	R	R	NA	NA	S	R	R	R	S
Donor	**EKP-SCC10**	+	+	+	+	-	S	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	10A	+	+	-	+	NA	NA	R	R	NA	R	R	R	NA	S	R	R	R	NA
	10C	+	+	-	+	NA	NA	R	R	NA	R	R	R	NA	S	R	R	R	NA
	10D	+	+	-	+	NA	NA	R	R	NA	R	R	R	NA	S	R	R	R	NA
	10E	+	+	-	+	NA	NA	R	R	NA	R	R	R	NA	S	R	R	R	NA
Donor	**EKP-SCC11**	+	+	+	+	-	R	R	R	S	R	R	R	S	R	R	R	R	R

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	blaCTX	blaOXA	blaSHV	blaTEM	blaKPC	AMC	AML	ATM	MEM	CTX	CAZ	CN	AK	CIP	SXT	RL	W	TE
Transconjugants	11-7	+	+	-	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	R
	11-8	+	+	-	+	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
	11-9	+	+	-	+	NA	I	R	R	NA	R	I	R	NA	S	R	R	R	R
	11-10	+	+	-	+	NA	I	R	R	NA	R	I	R	NA	S	R	R	R	R
Donor	**EKP-SCC15**	+	+	+	-	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	15C	+	+	-	NA	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	15D	+	+	-	NA	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	15G	+	+	-	NA	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	15-1	+	+	-	NA	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
Donor	**EKP-SCC16**	+	+	+	+	-	I	R	R	S	R	R	S	S	R	S	I	S	S
Transconjugants	16A	+	+	-	+	NA	I	R	R	NA	R	R	NA	NA	S	NA	S	NA	NA
	16B	+	+	-	+	NA	S	R	R	NA	R	R	NA	NA	S	NA	S	NA	NA
	16C	+	+	-	+	NA	S	R	R	NA	R	I	NA	NA	S	NA	S	NA	NA
	16D	+	+	-	+	NA	S	R	R	NA	R	R	NA	NA	S	NA	S	NA	NA
Donor	**EKP-SCC44**	+	-	+	+	-	S	R	R	S	R	R	S	S	S	R	R	R	S
Transconjugants	44A	+	NA	-	+	NA	NA	R	R	NA	R	I	NA	NA	NA	R	R	R	NA
	44B	+	NA	-	+	NA	NA	R	R	NA	R	R	NA	NA	NA	R	R	R	NA
	44D	+	NA	-	+	NA	NA	R	R	NA	R	R	NA	NA	NA	R	R	R	NA
	44E	+	NA	-	+	NA	NA	R	R	NA	R	R	NA	NA	NA	R	R	R	NA
Donor	**EKP-SCC75**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	R

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	**blaCTX**	**blaOXA**	**blaSHV**	**blaTEM**	**blaKPC**	**AMC**	**AML**	**ATM**	**MEM**	**CTX**	**CAZ**	**CN**	**AK**	**CIP**	**SXT**	**RL**	**W**	**TE**
Transconjugants	75A	+	+	-	+	NA	I	R	R	NA	R	I	R	NA	I	R	R	R	R
	75B	+	+	-	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	R
	75C	+	+	-	+	NA	I	R	R	NA	R	R	R	NA	I	R	R	R	R
	75E	+	+	-	+	NA	I	R	R	NA	R	R	R	NA	I	R	R	R	R
Donor	**EKP-H1FC25**	+	-	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	H1FC25A	-	NA	+	+	NA	I	R	R	NA	R	R	S	NA	S	S	S	S	NA
	H1FC25B	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1FC25C	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1FC25D	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
Donor	**EKP-H1PC18**	+	-	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	H1PC18-1	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1PC18-2	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1PC18-3	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1PC18-4	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
	H1PC18-5	-	NA	+	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	NA
Donor	**EKP-H1PC38**	+	+	-	+	-	R	R	R	R	R	R	R	S	R	R	R	R	R
Transconjugants	H1PC38-25	+	+	-	+	NA	I	R	R	S	R	R	S	R	S	S	S	S	S
	H1PC38-5	-	-	NA	-	NA	R	R	S	S	R	R	S	S	S	S	S	S	S
	H1PC38-5	NA	NA	NA	NA	NA	R	R	S	S	I	R	S	S	S	S	S	S	S
Donor	**CKP-KP2-448**	+	+	+	-	+	R	R	R	R	R	R	R	I	R	R	R	R	S

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	**blaCTX**	**blaOXA**	**blaSHV**	**blaTEM**	**blaKPC**	**AMC**	**AML**	**ATM**	**MEM**	**CTX**	**CAZ**	**CN**	**AK**	**CIP**	**SXT**	**RL**	**W**	**TE**
Transconjugants	448-1	-	-	-	NA	+	R	R	R	I	R	R	S	S	S	S	S	S	NA
	448-1	-	-	-	NA	+	R	R	R	I	R	R	S	S	S	S	S	S	NA
Donor	**CKP-608**	+	-	+	-	-	I	R	R	S	R	R	R	S	R	R	R	R	R
Transconjugants	608-1	-	NA	-	+	+	R	R	R	NA	R	R	S	NA	S	R	R	R	S
	608-2	-	NA	-	+	+	R	R	R	NA	R	R	S	NA	S	R	R	R	S
	608-3	-	NA	-	+	+	R	R	R	NA	R	R	S	NA	S	R	R	R	S
	608-5	-	NA	-	+	+	R	R	R	NA	R	R	I	NA	S	R	R	R	S
Donor	**CKP-179**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	R
Transconjugants	179-1	+	+	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	R
	179-6	+	+	-	+fraco	NA	R	R	R	NA	R	R	R	NA	S	R	R	R	R
Donor	**CKP-847**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	R
Transconjugants	847-1	+	+	-	+	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
	847-4	+	+	-	+	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
	847-5	+	+	-	+	NA	I	R	R	NA	R	I	R	NA	S	R	R	R	I
	847-6	+	+	-	+	NA	S	R	R	NA	R	I	R	NA	R	R	R	R	R
	847-7	+	+	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	R
	847-8	+	+	-	+	NA	I	R	R	NA	R	I	R	NA	R	R	R	R	R
Donor	**CKP-018**	+	-	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugant	018-5	+	NA	-	+	NA	R	R	R	NA	R	R	R	NA	S	R	R	R	NA

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	blaCTX	blaOXA	blaSHV	blaTEM	blaKPC	AMC	AML	ATM	MEM	CTX	CAZ	CN	AK	CIP	SXT	RL	W	TE
Donor	**CKP-038**	+	-	+	+	-	I	R	R	S	R	R	S	S	S	S	R	S	S
Transconjugants	038-1	+	NA	-	+	NA	R	R	R	NA	R	I	NA	NA	NA	NA	S	NA	NA
	038-2	+	NA	-	+	NA	R	R	R	NA	R	I	NA	NA	NA	NA	S	NA	NA
Donor	**CKP-792**	+	+	+	+	-	R	R	R	S	R	R	R	I	R	R	R	R	R
Transconjugants	792-10	+	+	-	+	NA	S	R	R	NA	R	R	R	S	S	R	R	R	R
	792-22	+	+	-	+	NA	S	R	R	NA	R	R	S	S	S	R	R	R	S
	792-15	+	+	-	+	NA	S	R	R	NA	R	I	R	S	S	R	R	R	R
	792-1	+	+	-	+	NA	I	R	R	NA	R	R	R	S	I	R	R	R	R
Donor	**CKP-813**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	R
Transconjugants	813-1	-	-	-	+	NA	R	R	R	NA	R	R	I	NA	I	R	R	R	S
	813-2	-	-	-	+	NA	R	R	R	NA	R	R	S	NA	S	R	R	R	S
	813-3	-	-	-	+	NA	R	R	R	NA	R	R	S	NA	I	R	R	R	S
Donor	**CKP-KP1-080**	+	-	+	+	+	R	R	R	R	R	R	R	I	R	R	R	R	S
Transconjugants	8080-1	-	NA	-	+	+	R	R	R	I	R	R	I	I	S	R	R	R	NA
	8080-5	-	NA	-	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA
	8080-10	-	NA	-	+	+	R	R	R	I	R	R	I	S	S	R	R	R	NA
	8080-15	-	NA	-	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA
Donor	**CKP-KP1-349**	+	-	+	+	+	R	R	R	R	R	R	R	R	R	R	R	R	S
Transconjugants	349-1	-	NA	-	+	+	R	R	R	I	R	R	I	S	S	R	R	R	NA
	349-5	-	NA	-	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	**blaCTX**	**blaOXA**	**blaSHV**	**blaTEM**	**blaKPC**	**AMC**	**AML**	**ATM**	**MEM**	**CTX**	**CAZ**	**CN**	**AK**	**CIP**	**SXT**	**RL**	**W**	**TE**
Transconjugants	349-10	-	NA	-	+	+	R	R	R	I	R	R	I	S	S	R	R	R	NA
	349-15	-	NA	-	+	+	R	R	R	I	R	R	I	I	S	R	R	R	NA
Donor	**CKP-379**	+	+	+	-	-	R	R	R	S	R	R	R	S	R	R	R	R	R
Transconjugants	379-3	+	+	-	NA	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
	379-10	+	+	-	NA	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
Donor	**CKP-562**	+	+	+	+	-	I	R	R	S	R	R	R	S	I	R	R	R	R
Transconjugants	562-1	+	+	-	+	NA	I	R	R	NA	R	R	R	NA	S	R	R	R	R
	562-23	+	+	-	+	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	R
	562-18	+	+	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	R
	562-10	+	+	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	R
Donor	**CKP-927**	+	-	+	-	-	R	R	R	S	R	R	S	S	R	R	R	R	S
Transconjugants	927-1	+	NA	-	NA	NA	S	R	R	NA	R	I	NA	NA	S	S	S	R	NA
	927-3	+	NA	-	NA	NA	S	R	R	NA	R	I	NA	NA	S	S	S	R	NA
Donor	**CKP-048**	+	-	-	+	-	R	R	R	S	R	R	R	S	S	R	R	R	R
Transconjugant	048-5	-	NA	NA	+	NA	R	R	R	NA	R	R	I	NA	NA	R	R	R	S
Donor	**CKP-124**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	124-3	+	-	-	+	NA	S	R	R	NA	R	R	R	NA	S	R	R	R	NA
	124-4	+	-	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	NA

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	blaCTX	blaOXA	blaSHV	blaTEM	blaKPC	AMC	AML	ATM	MEM	CTX	CAZ	CN	AK	CIP	SXT	RL	W	TE
Donor	**CKP-125**	+	+	+	+	-	I	R	R	S	R	R	R	S	R	R	R	R	S
Transconjugants	125-2	+	-	-	+	NA	S	R	R	NA	R	I	R	NA	S	R	R	R	NA
	125-3	+	-	-	+	NA	R	R	R	NA	R	I	R	NA	S	R	R	R	NA
Donor	**CKP-KP1-388**	+	-	+	+	+	R	R	R	R	R	R	R	I	R	R	R	R	S
Transconjugants	388-1	-	NA	+	+	+	R	R	R	I	R	R	S	I	S	R	R	R	NA
	388-2	-	NA	-	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA
	388-3	-	NA	+	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA
	388-4	-	NA	+	+	+	R	R	R	I	R	R	S	S	S	R	R	R	NA
	388-5	-	NA	+	+	-	R	R	R	I	R	R	S	S	S	R	R	R	NA
Donor	**CKP-KP2-465**	+	+	+	-	+	R	R	R	R	R	R	S	S	R	R	R	R	S
Transconjugants	465-5	-	-	-	NA	+	R	R	R	I	R	R	NA	NA	S	S	S	S	NA
	465-10	-	-	-	NA	+	R	R	R	I	R	R	NA	NA	S	S	S	S	NA
	465-15	-	-	-	NA	+	R	R	R	R	R	R	NA	NA	S	S	S	S	NA
Donor	**CEKP-KP4A**	+	+	+	+	-	R	R	R	S	R	R	S	I	R	R	R	R	R
Transconjugant	KP4A-1	+	+	-	+	NA	R	R	R	NA	R	R	NA	S	R	R	R	R	R
Donor	**CKP-KP13**	+	+	+	+	-	R	R	R	S	R	R	S	R	R	R	R	R	R
Transconjugants	KP13-1	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
	KP13-2	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
Donor	**CKP-KP30**	+	+	+	+	-	R	R	R	S	R	R	S	R	R	R	R	R	R

Table S4 (cont.) - Antimicrobial resistance phenotype and genotype transferred to transconjugants by the conjugative isolates from the 59 isolates of *K. pneumoniae* resistant to 3rd generation cephalosporins.

		ARGs					Antimicrobial Susceptibility												
	Isolate	**blaCTX**	**blaOXA**	**blaSHV**	**blaTEM**	**blaKPC**	**AMC**	**AML**	**ATM**	**MEM**	**CTX**	**CAZ**	**CN**	**AK**	**CIP**	**SXT**	**RL**	**W**	**TE**
Transconjugants	KP30-3	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
	KP30-4	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
Donor	**CEKP-KP55A**	+	+	+	+	-	R	R	R	S	R	R	S	R	R	R	R	R	R
Transconjugants	KP55-1	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
	KP55-4	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
Donor	**CEKP-KP3A**	+	+	+	+	-	R	R	R	S	R	R	S	S	R	R	R	R	R
Transconjugants	KP3A-1	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	R
	KP3A-2	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	I
Donor	**CEKP-KP1A**	+	+	+	+	-	R	R	R	S	R	R	S	R	R	R	R	R	R
Transconjugants	KP1A-2	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
	KP1A-5	+	+	-	+	NA	R	R	R	NA	R	R	NA	R	S	S	S	S	S
Donor	**CKP-KP3**	+	+	+	+	-	R	R	R	S	R	R	S	S	R	R	R	R	R
Transconjugants	KP3-2	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	R
	KP3-7	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	R
Donor	**CKP-KP40**	+	+	+	+	-	R	R	R	S	R	R	S	S	R	R	R	R	R
Transconjugants	KP40-2	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	R
	KP40-4	+	+	-	+	NA	R	R	R	NA	R	R	NA	NA	S	R	R	R	R

Table S5 - Clusters obtained based on the pheno- and genotypic characteristics of the 3rd generation cephalosporin-resistant *K. pneumoniae* isolates and dominant characteristics observed in each cluster.

Cluster	No. of isolates	Wastewater isolates	Clinical isolates	Dominant characteristics
A	24	5	19	Presence of β-lactamase encoding genes (*bla*$_{CTX}$, *bla*$_{OXA}$, *bla*$_{SHV}$ and *bla*$_{TEM}$) Resistance to tetracycline High molecular weight plasmids Presence of plasmid replicon type FIIK Capacity to transfer ARGs and MDR to *E. coli* J53 Moderate/strong biofilm forming capacity *G. mellonella* infection indices varied between 0 to 9
B	14	9	5	Tetracycline susceptibility Incapacity to conjugate with *E. coli* J53 *G. mellonella* infection indices varied between 1 to 9
C	6	5	1	Tetracycline resistance Presence of 3 plasmids Absence of the plasmid replicon type FIIK Capacity to conjugate observed only in a hospital effluent isolate *G. mellonella* infection indices varied between 2 and 8
D	5	2	3	Absence of the *bla*$_{CTX}$ and *bla*$_{OXA}$ genes Only the clinical isolates were meropenem resistant, due to the presence of the gene *bla*$_{KPC}$ Able to transfer MDR phenotypes to *E. coli* J53 *G. mellonella* infection indices varied between 1 and 5
E	3	0	3	Able to conjugate with *E. coli* J53 MDR transfer was not observed *G. mellonela* infection indices of 1-2
F	3	3	0	Resistant to amoxicillin with clavulanic acid and to sulfonamides Detection of a single plasmid larger than 150 Kbp Not able to conjugate with *E. coli* J53 Moderate/strong capacity to form biofilms *G. mellonela* infection indices of 2-9

Table S6 - Genes searched related to clinically-relevant properties on selected 3rd generation cephalosporin-resistant *K. pneumoniae* genomes. 1 indicates the detection of the gene and 0 indicates the not detection of the gene. For each gene it is indicated the allele that was detected, the number of genomes in which it was detected, the length of the gene and of the query, the number of different nucleotides in the query sequence and the number of gaps. The accession number and the database from which each gene allele was downloaded is also indicated.

Due to high dimension this table is provided in the following link:

https://docs.google.com/spreadsheets/d/1pniXrw-a-EfcLBfWB-ipgI_YNSLQqitl/edit?usp=sharing&ouid=109610140440067111385&rtpof=true&sd=true

CHAPTER 6

ENVIRONMENTAL AND CLINICAL *KLEBSIELLA PNEUMONIAE*: TWO SIDES OF THE SAME COIN

The work presented in this chapter is part of a manuscript submitted:

Rocha, J., Henriques, I., Gomila, M., Manaia, C.M. (submitted). Environmental and clinical *Klebsiella pneumoniae*: two sides of the same coin.

Contribution:

Contributor Role	**JR**	**IH**	**MG**	**CM**
Conceptualisation				x
Methodology	x		x	x
Validation			x	x
Investigation	x		x	x
Data Curation	x		x	
Writing – Original Draft Preparation	x			x
Writing – Review and Editing		x	x	x
Supervision			x	x
Project Administration				x
Funding			x	x

Abstract

Members of *Klebsiella pneumoniae* complex are ubiquitous bacteria that can be found in different environmental compartments, such as soil, plants and water and as humans' opportunistic pathogens. This study aimed at exploring common and distinctive features in *K. pneumoniae* of clinical and environmental origin inferred from their genomes. Whole genome sequences of clinical (n=78) and environmental (n=61) belonging to the *K. pneumoniae* complex, available in the GenBank, obtained from 21 countries, were compared based on phylogenetic, pangenome and fitness-related traits. Multi-locus sequence typing divided the isolates into 56 sequence types, eight of which were common to clinical and environmental origins. For the genetic categories evaluated were observed more antibiotic resistance gene/plasmids exclusive to clinical than to environmental isolates, while latter presented higher genetic diversity of genes related with functions such as efflux or oxidative stress. These observations suggest the sharing and adaptation of *K. pneumoniae* between distinct niches.

1. Introduction

The species *Klebsiella pneumoniae*, within the family *Enterobacteriaceae*, is considered a major pathogen, associated with urinary, respiratory, gastrointestinal and skin infections (Podschun & Ullmann, 1998; Wyres & Holt, 2016). Despite its ubiquitous character, *K. pneumoniae* are severe opportunistic pathogens, whose control is impaired by the frequent multidrug resistance phenotype, representing a major threat for neonates, elderly, and immunocompromised patients (Pendleton *et al*., 2013). Previous reports indicate that about one-third of the Gram-negative bacterial infections observed in hospitals are attributed to members of this species (Navon-Venezia *et al*., 2017). The genome of *K. pneumoniae* is ~5.5 Mbp in size, with ~5500 genes, of which ~3500 are accessory, suggesting its dynamic and plastic character (Bialek-Davenet *et al*., 2014; Holt *et al*., 2015). The accessory genome comprises acquired antimicrobial and metal resistance and virulence genes, as well as plasmids and other elements associated with horizontal gene transfer, evidenced by domains of varying guanine/cytosine content (Navon-Venezia *et al*., 2017; Wyres & Holt, 2018, 2016). A wide range of virulence factors involved in iron uptake, capsule production, biofilm production, and others, facilitate the evasion of host defence mechanisms (Gomez-Simmonds & Uhlemann, 2017; Paczosa & Mecsas, 2016). In addition, antibiotic resistance genes acquired by members of this species cover all classes of antibiotics, being *K. pneumoniae* sometimes among the first to spread emerging resistance genes (e.g., bla_{CTX}, bla_{KPC}, bla_{NDM}) (Liu *et al*., 2019; Yu *et al*., 2019; Zhou *et al*., 2020; Navon-Venezia *et al*., 2017).

As other opportunistic pathogens, *K. pneumoniae* are ubiquitous bacteria able to thrive in environmental compartments (e.g., soil, plants, and waterways) (Bagley, 1985; Podschun & Ullmann, 1998). While some *K. pneumoniae* are likely emitted from human and animal sources, and hence can be considered environmental contaminants, others are truly environmental strains thriving in their natural habitat. Indeed, water, vegetation and soil are described as the native environment of *K. pneumoniae* (Brisse *et al*., 2006). Given its wide distribution (Wyres & Holt, 2018), it is important to investigate the common and distinctive features of environmental and human-associated strains of *K. pneumoniae*. Although distinguishing between environmental and human bacteria can be challenging, the characterization of populations from both origins is crucial to assess the major drivers of evolution, the stability of acquired genetic traits, and ultimately to infer about privileged paths of transmission from the environment to humans.

This study aimed to test the hypothesis that human-associated and environmental *K. pneumoniae* may belong to distinct genetic lineages and display different features, particularly regarding the pattern of mobile genetic elements, acquired genes, allelic genetic diversity,

among others. It is also hypothesized that genes acquired by human-associated *K. pneumoniae* may be stable in these bacteria, even when they thrive in the environment. The study was designed to cover the broadest geographic distribution of clinical and environmental isolates. Therefore, whole genome sequences of *K. pneumoniae* were retrieved from NCBI database, which were complemented with whole genome sequences available in the in-house collection.

2. Material and Methods

2.1. Whole genome sequences selection

Genomes were searched and retrieved from NCBI, between May 3 and October 31, 2018, using the keyword *Klebsiella pneumoniae*, filtered for assembled genomes (Figure S1). This search resulted in 231 *K. pneumoniae* complete genomes, 174 of clinical origin, 10 environmental and 47 with unreported origin, according to the information available in the NCBI database. For this study, 56 complete clinical genomes were downloaded based on the criterion that they were obtained from different clinical samples (e.g., blood, wound, urine, respiratory tract, among others). Given the low number of complete genomes of environmental origin, it was necessary to also include 43 draft-genomes available in the same database. Among those of environmental origin (e.g. sewage, river, soil, food) these genomes were selected because they contained fewer contigs or scaffolds (less than 74 and 99, respectively). This collection was complemented with 30 draft genomes (22 clinical and 8 environmental) under study by the same authors (Rocha *et al*., in revision; Gomila *et al*., in preparation), making a total of 139 genomes, 78 clinical and 61 environmental (Table S1). The inclusion of genomes from potentially clonal strains was avoided. For this, when the genomes shared 100% of average nucleotide identity based on BLAST algorithm (ANIb) they were included only when originating from different samples or if they harboured different genes/pangenome.

2.2. Phylogenetic inference

The partial sequences of the housekeeping genes *gapA*, *infB*, *mdh*, *pgi*, *phoE*, *rpoB* and *tonB* were extracted from the genome to determine the Multi-Locus Sequence Type (Diancourt *et al*., 2005). Sequence types were determined using the BIGSdb (Jolley & Maiden, 2010). Gene fragments were concatenated and a total of 3,012 bp was used for multilocus sequence analysis. Alignment was performed using MEGA7 and phylogenetic trees were constructed using Neighbor-Joining, Maximum Likelihood and Minimum Parsimony methods, with

bootstraps of 1000 replicates. Average nucleotide identity (ANIb) values (9,661 pairwise comparisons) were determined using the online service JSpeciesWS (http://jspecies.ribohost.com/jspeciesws/#analyse). PRIMER6 software (Clarke & Gorley, 2006) was used to calculate Euclidean distance and to construct a dendrogram using the Unweighted Pair Group Method using Averages (UPGMA).

2.3. Comparative genomics analysis

Genomes were annotated using PROKKA version 1.12 (Seemann *et al*., 2014) in order to standardize the annotation for all the genomes. PROKKA software predicts open reading frames and performs the annotation of deduced amino acid sequences. Pangenome analysis of clinical and environmental genomes was determined for each group separately using the GET_HOMOLOGUES software and the criteria 70% similarity and 50% of coverage (Contreras-Moreira & Vinuesa, 2013). To ensure that all core genes were included in the analysis and to avoid the bias caused by truncated genes in the draft genomes, the soft core genes (present in >95% of the genomes) of the clinical and of the environmental pangenomes were selected for further comparisons. To detect sequences exclusive of each group or common to both, the deduced amino acid sequences belonging to the clinical and to the environmental genomes soft cores were compared using CD_HIT website (Li & Godzik, 2006) (70% similarity over 50% of coverage). The exclusive gene sequences were extracted and validated manually based on BLASTn searches against the opposite group (clinical *vs*. environmental). The functional categories of cloud, shell, soft core and core genes for all the approaches were determined using the KOALA (KEGG Orthology And Links Annotation) database. Genes encoding putatively clinically- and fitness-relevant properties, such as antibiotic and metal resistance, virulence, *quorum sensing*, and oxidative stress, and sequences of different plasmid replicon types were screened in the whole genome sequences. BLASTn was used to screen 237 genes downloaded from specialized databases: metal resistance (n=38), virulence (n=87); efflux systems (n=17) (from Institute Pasteur database, https://bigsdb.pasteur.fr/cgibin/bigsdb/bigsdb.pl?db=pubmlst_klebsiella_seqdef&page=downloadAlleles), antibiotic resistance (n=44) (from ResFinder database (Zankari *et al*., 2012) and cross-checked in CARD database, https://card.mcmaster.ca/), plasmids replicon type (n=35) (PlasmidFinder 2.1 tool from Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/PlasmidFinder/), *quorum sensing* (n=7) and oxidative stress (n=9) (NCBI). These two last groups were searched in the Uniprot database using the terms "*Klebsiella pneumoniae*" and "*quorum sensing*" or "*Klebsiella pneumoniae*" and "oxidative stress" and downloaded from NCBI database. Based on this information, a presence/absence matrix was constructed and similarity between genomes was calculated using the Jaccard

index. Results were represented in a UPGMA clustering dendrogram. Variants of each gene were determined based on nucleotide differences (>1 nucleotide difference) between sequences.

2.4. Statistical analysis

Fisher's exact tests were applied to assess the statistically significant differences between the proportion of clinical and environmental isolates that harbored the screened genes and the respective alleles ($p<0.05$). The diversity of the alleles observed for each gene was compared based on the Shannon diversity index, which reflects the relative abundance of the gene's alleles ($p<0.05$).

3. Results

3.1. Genome's diversity

The study examined 139 genomes (Table S1, Figure S1) with origin in 21 countries: USA (23/139, 17%), UK, Portugal and Spain (each 15/139, 33%), China (14/139, 10%), Germany (13/139, 9%), Thailand (11/139, 8%) and other countries (each <8 isolates, 33/139, 24%) and belonged to 56 MLST. Among the identified sequence types (STs), 23 were identified exclusively in clinical isolates, 25 STs exclusively in environmental isolates, and eight STs (ST11, ST14, ST15, ST37, ST45, ST147, ST348 and ST437) included isolates classified in both categories (Table 1 and Table S1). The predominant STs among clinical isolates were ST147 (18%), ST11, ST23, and ST258 (each 8%) and among environmental isolates were ST14 (8%), ST895 and ST3128 (each 7%) (Table 1 and Table S1). The number of unique STs, meaning an ST detected once among the isolates, was similar in clinical (n=16) and in environmental (n=19) *K. pneumoniae* genomes ($p>0.05$).

Table 1 - Summary of the *Klebisella* spp. genomes features used in this study.

	Clinical	Environmental
Number of isolates	78	61
Number of countries	14	14
Sequencing technologies	Illumina PacBio Nanopore	Illumina PacBio 454
Number of STs	31	33
Unidentified STs*	no	yes (n=1)
Most abundant STs (identified in at least 3% of the 139 genomes):		
ST11	6	1
ST14	2	5
ST15	4	2
ST23	6	0
ST37	2	2
ST147	14	2
ST258	6	0
ST307	4	0
ST326	4	0
ST392	4	0
ST895	0	4
ST3128	0	4

* Unidentified ST refers to a ST which has not been determined as it has not been possible to obtain the housekeeping gene sequences required for this determination.

The occurrence of clinical and environmental isolates identified with the same ST was not circumstantial, as it was observed in different countries. For example, it was observed for ST11 (n=7) in Japan, Germany, China, USA and Spain, ST14 (n=7) in USA and Algeria, ST15 (n=6) in Portugal, Nepal, USA and China, ST37 (n=4) in Thailand, USA and China, ST45 (n=3) in Thailand and UK and ST147 (n=16) in Portugal, Switzerland, Germany, United Arab Emirates, Thailand, Pakistan and Spain. Also, some STs represented by more than one isolate were reported in a single country (USA, ST16, n=2 clinical; ST941, n=2 clinical; Portugal, ST348, n=2 environmental, n=1 clinical; Spain, ST392, n=4 clinical; ST326, n=4 clinical; ST405, n=2 clinical; Germany, ST3128, n=4 clinical), although most of the times were reported by the same authors. This latter situation was observed for ST258 in USA (n=6 clinical) and ST437, (n=2 environmental, n=1 clinical) in Brazil (Figure 1).

The ANIb values varied between 93% and 100% (Table S2). Values below 95% corresponded to 13 isolates that at the time of the genomes download were identified as *K. pneumoniae* and were later revised to *K. variicola* (n=6, 1clinical and 5 environmental) and to *K. quasipneumoniae* (n=7, 1 clinical and 6 environmental) (Rodrigues *et al*., 2018) that belong to the *K. pneumoniae* complex. The integration of *K. pneumoniae* and other species of the complex in distinct clusters was also indicated by the analysis based on the concatenated sequences of the MLST genes (Neighbor-Joining, Maximum Likelihood and Maximum Parsimony) (Figure 1 and Figure S2). Those 13 genomes were affiliated to different STs (ST3870, ST3013, ST2355, ST138, ST477, ST2045, ST3011, ST3851, ST146, ST906, ST3940, ST208, and ST355) (Figure 1 and Table S1). The option to maintain these genomes in the study was justified by the fact that they belong to the same complex and their inclusion avoided the disproportion on the number of clinical and environmental isolates that might bias the results. Possible biases in the results due to the inclusion of these isolates were also critically assessed. The analysis based on the ANIb values (Figure S3) clustered the genomes according to the ST although, in some cases, such as the ST11, ST23, ST37, ST258 and ST392, the ST was divided in different groups (Figure 1 and Figure S3). In all cases except for ST258 and ST392, this separation was associated with distinct countries (Figure S3).

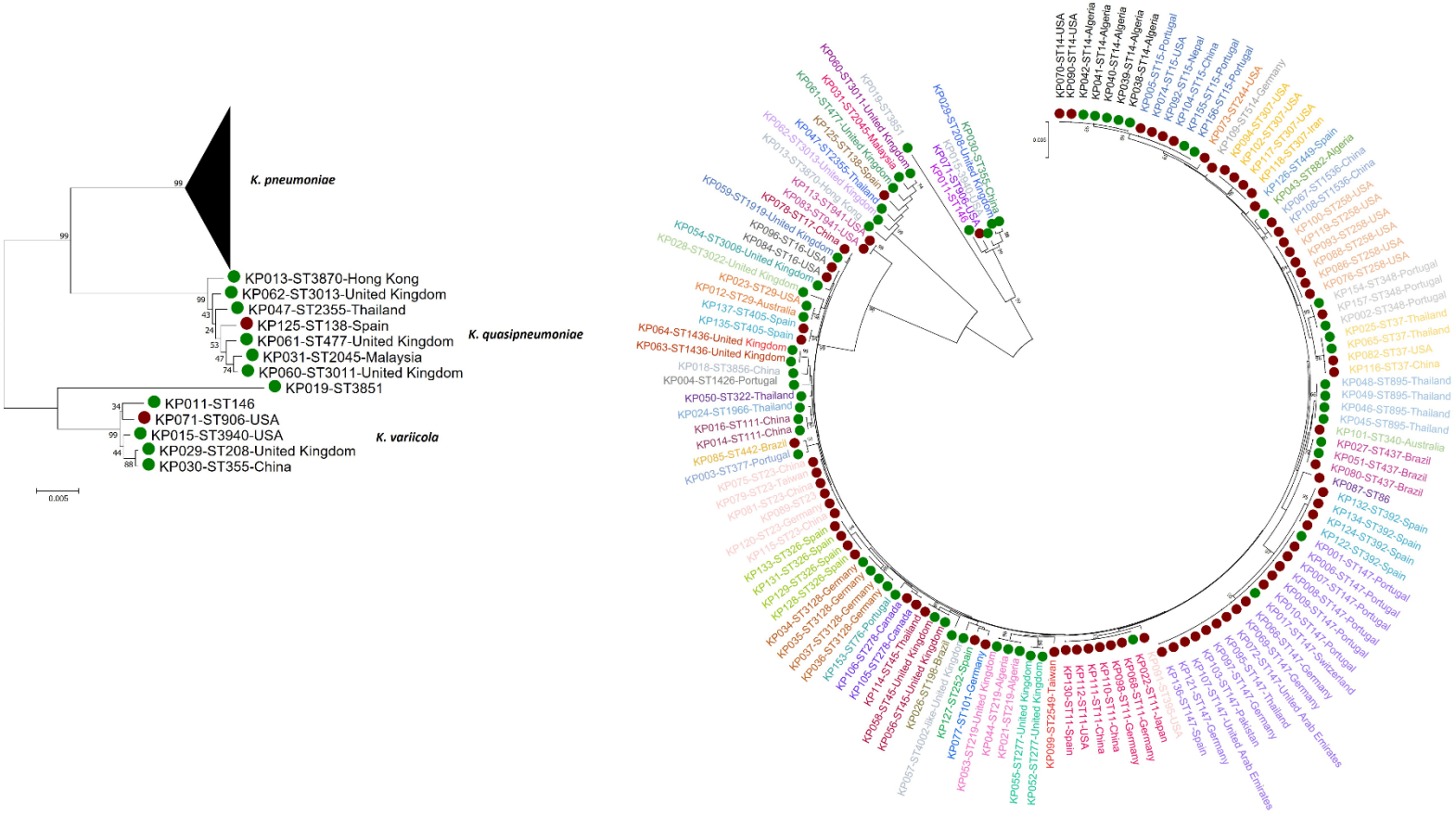


Figure 1 - Phylogenetic tree obtained by concatenated MLST gene sequences (*gapA*; *infB*; *mdh*; *pgi*; *phoE*; *rpoB*; *tonB*) and using the Neighbor-Joining method and a bootstrap of 1000 replicates – condensed (left) and extended (right) representation. Similar results were obtained with Maximum Likelihood and Maximum Parsimony methods. Bootstrap values (>60%) are indicated at the nodes. Clinical isolates are indicated in red and environmental in green.

3.2. Genome analysis: pangenome and core genome approach

The number of coding sequences that were detected in clinical and environmental pangenomes was similar (10 210 in clinical and 10 288 in environmental) (Figure S4). The number of deduced amino acid sequences identified in the core, soft core and shell was higher in clinical genomes than in environmental genomes, while the number of cloud genes (present in 1 or 2 genomes) was higher in environmental isolates (Figure S4). The functional categories of core amino acid sequences were mostly related with genetic information processing (587/2713, 21% clinical, 435/2319, 19% environmental), environmental information processing (298/2713, 11% clinical, 261/2319 11% environmental), signalling and cellular processes (298 /2713 11% clinical, 281/2319 12% environmental) and carbohydrate metabolism (288/2713 11% clinical, 249/2319 11% environmental) and not different in both groups under analysis (Figure S5). The analysis of the soft core genes for each group did not reveal any exclusive gene in either group. Cloud amino acid sequences were mostly related to the functional categories previously mentioned (Figure S5). Clinically relevant genes (e.g., antibiotic resistance and *quorum sensing*) were detected in similar values in clinical and environmental isolates, and were distributed across the core, soft core, shell and cloud dimensions of the pangenomes.

This result motivated a deeper investigation of target genes associated with clinically-relevant traits in clinical and environmental genomes. Therefore, the genomes were further compared based on presence/absence and number of alleles of genes (herein represented by sequences that differ in at least one nucleotide) related with antibiotic and metal resistance, plasmid replicon type, virulence, efflux systems, oxidative stress and *quorum sensing*. This analysis revealed that 3 metal resistance (*pbrA, pbrBC* and *pbrR* genes), 37 virulence genes (*iro*, *clb*, *iuc*, among other genes), 13 antibiotic resistance genes (*rmt*, *mef*, *mcr*, among other genes) and 14 plasmid replicon types (*IncHI2*, *IncQ1*, *IncU*, *IncX3*, *psL483*, among others) were exclusive of clinical isolates (Figure S6). Only one of these isolates harbouring these exclusive genes, specifically the plasmid replicon type *psL483* was affiliated to *K. quasipneumoniae* (Table S3). In contrast, one antibiotic resistance gene and six plasmid replicon types were detected exclusively in environmental isolates (Figure S6). Only one of these isolates was affiliated *K. quasipneumoniae*, the only environmental isolate (1/61 isolates) harbouring the replicon types *Col(IMG531)* and *Col(IRGK)* (Table S3). These differences apart, 20 genes in these categories were detected in all genomes (antibiotic resistance n=1, virulence n=1, efflux systems n=8, oxidative stress n=5, *quorum sensing* n=5) and most genes of the others (163/237) were detected in both groups (Table S3 and Table S4). However, some of the genes common to both groups presented significantly different prevalence values ($p<0.05$). The genes related with antibiotic (*bla*$_{KPC}$, *bla*$_{OXA}$, *bla*$_{SHV}$, *bla*$_{TEM}$

and *aac(6')-Ib-cr*) or metal resistance (*mer*, *ter*) and virulence (yersiniabactin - *fyuA*, *irp*, *ybt*) were significantly more prevalent among clinical genomes, irrespective of the inclusion in the analysis of *K. variicola* strain KP071 (*bla*$_{OXA}$, *bla*$_{TEM}$, *aac(6')-Ib-cr* and *ter) and K. quasipneumoniae* strain KP125 (*bla*$_{TEM}$). In turn, the resistance gene *bla*$_{OKP-A}$ and replicon types *Col(MGD2)* and *Col(BS512)* were more significantly more frequent in environmental genomes (Figure 2-A and Table S4). However, this result was attributed to *K. quasipneumoniae* isolates KP013, K031, KP047, KP060, KP061 and KP062 that harboured the gene *bla*$_{OKP-A,}$ eventually intrinsic in this species here represented by distinct sequence types and geographic origins. A dendrogram based on the presence/absence of the 237 genes clustered the genomes mainly in agreement with the sequence types and/or geography (Figure S7). This analysis showed that a single ST could be subdivided according to geography – e.g. ST14 isolates were divided in Algeria and USA subgroups, ST147 isolates were split into groups from Germany, Portugal, Thailand or United Arab Emirates (Table S4, Figure S7).

In the group of the 237 genes analysed were observed 2661 gene variants (1601 in clinical and 1648 in environmental isolates), indicated by at least one nucleotide substitution (Table S3). The rationale of this analysis was to assess genetic variation, irrespective of the implications on the phenotype. The highest number of gene variants was observed for virulence, especially to capsular related genes (*wzc* n=55, *wzi* n=64), *quorum sensing* (e.g. *lsrB* n=74, *tqsA* n=83) and oxidative stress (e.g. *msrQ* n=37, *oxyR* n=37), also observed in *K. quasipneumoniae* and *K. variicola* isolates (Table S3 and Table S5). In 26 out of the 163 common genes (4 antibiotic resistance, 2 metal resistance, 11 virulence, 3 efflux systems, 2 oxidative stress and 4 *quorum sensing*) significant differences ($p<0.05$) in the prevalence of alleles detected in clinical and environmental isolates were observed (Figure S6). The prevalence of 16 out of these 26 was significantly different, irrespective of the inclusion of *K. quasipneumoniae* and *K. variicola* in the analysis (e.g. *aadA*, *bla*$_{OXA}$, *ibpB*, *lsrC*, among others). Indeed, 17 out those 26 genes were among the three *Klebsiella* species, and in the case of 10 of those genes, *K. quasipneumoniae* and *K. variicola* were responsible for the differences observed between clinical and environmental isolates. Moreover, in 23 genes common to both clinical and environmental isolates, it was observed a significantly higher Shannon diversity index in environmental than in clinical genomes ($p<0.05$) for genes of the categories efflux systems (n=9), oxidative stress (n=8) and *quorum sensing* (n=6 (Table S6). Again, these differences were due to *K. quasipneumoniae* and *K. variicola* genomes, responsible for the observed significant differences in 14 of those 23 genes (5 efflux systems, 5 oxidative stress, 4 *quorum sensing*). The Shannon diversity of the alleles related to copper resistance (*pco*), adhesion (*mrk*), fosfomycin resistance (*fosA*), among others, was also higher in environmental isolates ($p<0.05$) (Table S6 and Figure 2-B). Only for copper and fosfomycin resistance the

exclusion of *K. quasipneumoniae* and *K. variicola* genomes did not alter the significant differences ($p<0.05$). In contrast, the allelic Shannon diversity in genes related to antibiotic resistance (n=8), metal resistance (n=2), virulence (n=12) and the plasmid replicon types (n=3) was higher in clinical genomes ($p<0.05$) (Table S6 and Figure 2-B). These differences in Shannon diversity were maintained when *K. quasipneumoniae* and *K. variicola* genomes were excluded from the analysis ($p<0.05$), except for 2 replicon types (*IncFII (pKP91)* and *IncN*) ($p>0.05$).

Genomes in which was detected the lowest number of the screened genes (<30% of 237) corresponded to 21 clinical (of 13 sequence types) and 25 environmental (to 17 sequence types) isolates. These genomes typically contained metal resistance genes (*pco*, *sil*, *ars*) in environmental isolates, or antibiotic resistance (bla_{KPC}) and virulence genes (yersiniabactin) in clinical isolates. Genomes with the highest number of the screened genes (>50% of the 237) corresponded to 3 clinical isolates (ST14, USA and two ST23, China) (Table S4).

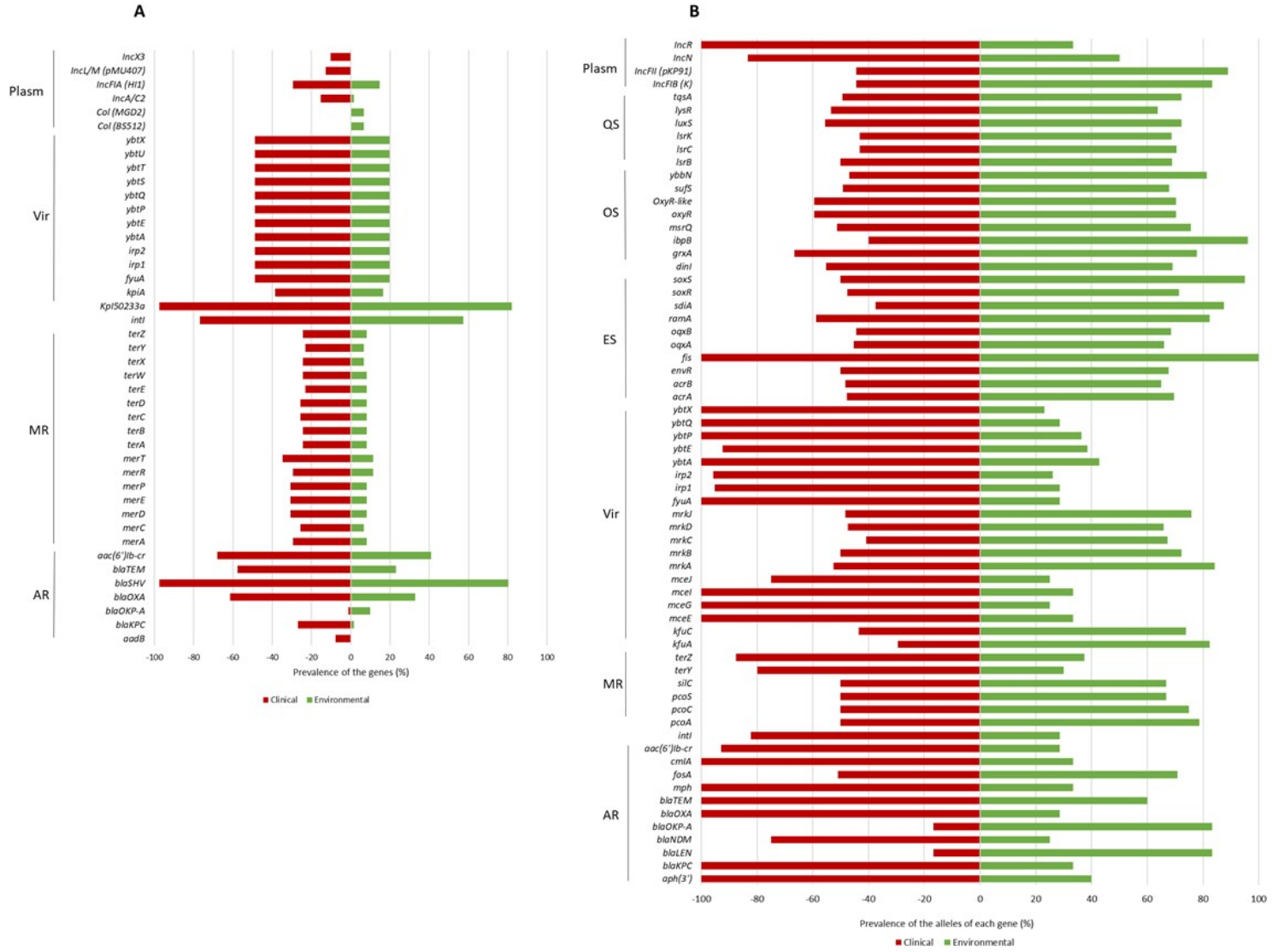

A
Plasm
IncX3
IncL/M (pMU407)
IncFIA (HI1)
IncA/C2
Col (MGD2)
Col (BS512)
Vir
ybtX
ybtU
ybtT
ybtS
ybtQ
ybtP
ybtE
ybtA
irp2
irp1
fyuA
kpiA
Kpl50233a
intI
MR
terZ
terY
terX
terW
terE
terD
terC
terB
terA
merT
merR
merP
merE
merD
merC
merA
AR
aac(6')Ib-cr
blaTEM
blaSHV
blaOXA
blaOKP-A
blaKPC
aadB
-100 -80 -60 -40 -20 0 20 40 60 80 100
Prevalence of the genes (%)
Clinical
Environmental
B
Plasm
IncR
IncN
IncFII (pKP91)
IncFIB (K)
QS
tqsA
lysR
luxS
lsrK
lsrC
lsrB
OS
ybbN
sufS
OxyR-like
oxyR
msrQ
ibpB
grxA
dinI
ES
soxS
soxR
sdiA
ramA
oqxB
oqxA
fis
envR
acrB
acrA
Vir
ybtX
ybtQ
ybtP
ybtE
ybtA
irp2
irp1
fyuA
mrkJ
mrkD
mrkC
mrkB
mrkA
mceJ
mceI
mceG
mceE
kfuC
kfuA
MR
terZ
terY
silC
pcoS
pcoC
pcoA
intI
AR
aac(6')Ib-cr
cmlA
fosA
mph
blaTEM
blaOXA
blaOKP-A
blaNDM
blaLEN
blaKPC
aph(3')
-100 -80 -60 -40 -20 0 20 40 60 80 100
Prevalence of the alleles of each gene (%)
Clinical
Environmental

Figure 2 - Statistically significant differences observed between clinical and environmental *K. pneumoniae* isolates. A) Prevalence (%) of genes (Fisher's exact test and p-value < 0.05); B) Prevalence (%) of gene-alleles, meaning variants of a single gene that differ in at least one nucleotide (Shannon diversity index (Table S6) and p-value < 0.05). The prevalence of the genes or of the gene-alleles was determined based on the following formula: Prevalence (%) = 100 x (Number of clinical or environmental genomes containing the gene A/ total number of clinical or environmental genomes) or 100 x (Number of observed variants of gene A in clinical or environmental genomes/ total number of observations of the gene A in clinical or environmental genomes). Some genes such as *bla*$_{LEN}$ and *bla*$_{OKP-A}$ were only observed in the *K. variicola* and *K. quasipneumoniae* species, respectively. AR – antibiotic resistance; MR – metal resistance; Vir – virulence; Plasm – plasmids; ES – efflux systems; OS – oxidative stress; QS – quorum sensing.

4. Discussion

The hypothesis of the study was that human-associated and environmental *K. pneumoniae* may belong to distinct genetic lineages and yield distinct genome features, mainly those that were associated with increased fitness, such as antibiotic and metal resistance, virulence, oxidative stress or *quorum sensing* (Runcharoen *et al*., 2017). To test this hypothesis, genomes from clinical or environmental sources and wide geographic distribution were compared. A first shortcoming was the limited number of whole genome sequences available for environmental isolates. This suggests the bias existing in public databases towards clinical isolates, which may represent a limitation to investigate the interface between humans and the environment. A solution adopted to overcome this shortcoming was the inclusion of high-quality draft genomes of environmental isolates, which we assumed would support a reliable comparison between clinical and environmental isolates. Also concerning the environmental genomes, the supporting information is sometimes insufficient or inaccurate to support robust ecology studies. The supply of accurate and reliable data relative to environmental isolate genomes should be encouraged among the scientific community.

The group of isolates that at the time of genomes download for this study were identified as *K. pneumoniae*, later reclassified as *K. quasipneumoniae* and *K. variicola,* were observed to represent a distinct group based on ANIb and MLST. While the whole genome sequence analysis revealed the differentiation of the species, *K. pneumoniae* and *K. quasipneumoniae* and *K. variicola* have been considered phenotypically and phylogenetically close and difficult

to distinguish (Long *et al*., 2017). Nevertheless, it was reported that genes such as *bla*$_{LEN}$, *bla*$_{OKP}$, *gyrA*, *parC*, among others, may contribute to the reliable distinction of these species (Rodrigues *et al*., 2018). In fact, the genes *bla*$_{OKP-A}$ and *bla*$_{LEN}$ were only detected in isolates affiliated to *K. quasipneumoniae* and *K. variicola*, respectively. The situations where the inclusion of these 13 isolates (2 clinical and 11 environmental) might have influenced the results were identified. Nevertheless, this situation highlights the importance of the correct species affiliation in public databases, as the information may be used without the verification that was made in this study and that revealed the misidentification.

The majority of the STs (n=53/61) were only observed either in clinical or environmental isolates, and in most of these cases each ST was represented by one isolate. The ST11, ST14, ST15, ST37, ST45, ST147, ST348, ST437 comprised clinical and environmental isolates of this study. These STs, widely distributed worldwide, have been associated to outbreaks (Wyres & Holt, 2016). Particularly, the STs with the widest geographic distribution, ST11, ST23 and ST147, which comprised mostly or exclusively clinical isolates, may suggest a certain degree of specialization. Among the environmental isolates, the same ST generally did not share the same origin, except for ST14 corresponding to five isolates from crops in Algeria. This may suggest a wider geographic distribution of the genetic lineages in the environment than in patients, where selection is known to occur and the diversity of habitats is incomparably narrower, as has been noted before (Wyres *et al*., 2020).

Among the 237 screened genes, it was mostly among clinical isolates that exclusive antibiotic and metal resistance, virulence and plasmid replicon types were observed. This observation was not unexpected given the strong selection pressures that bacteria are subjected to in an infection episode (Baishya & Wakeman, 2019; Martin & Bachman, 2018). It may also explain a significantly higher diversity in virulence and antibiotic resistance genes and gene alleles in these isolates. In contrast, efflux, oxidative stress and *quorum sensing* related genes were common to both groups and in some cases significantly more diverse in environmental isolates. Although these results may be explained in part by the influence of the isolates affiliated with K. *quasipneumoniae* and *K. variicola* (specifically for *envR*, *oxyR*, *luxS*, among others) the absence of selection pressure may also suggest the important advantage that these functions may confer to adapt to natural environments where external conditions are supposed to vary more than in the human body (Bambeke *et al*., 2000; Guan *et al*., 2017; Williams & Cámara, 2009). The continuum between clinical and environmental habitats is known (Teixeira *et al*., 2020; Rozman *et al*., 2020; Zagui *et al*., 2020) and explains why genes associated to clinical origins (e.g. *bla*$_{CTX-M}$, *bla*$_{KPC}$, *bla*$_{NDM}$) were detected in environmental isolates, although in lower prevalence. In addition, different mobilomes might be associated to clinical or environmental isolates, considering that plasmid replicon types such as *IncA/C2*, *IncFIA(H1)*, *IncL/M (pMU407)*, *IncX3* were more associated to clinical origins and the replicon

types *Col(BS512)* and *Col(MGD2)* to the environmental origins. The hypothesis that clinical and environmental isolates yield distinct genome features was confirmed.

An important conclusion of the study was that distinct lineages prevail in clinical and environmental settings, although some lineages may occur in both contexts. Another conclusion was that the natural environment may offer favourable conditions for the diversification/evolution of fitness-relevant genetic determinants, such as those related with efflux, oxidative stress and *quorum sensing*, while the clinical context drives the acquisition and evolution of virulence and antibiotic and metal resistance genes as well as some plasmid replicon types, which may be selected during colonization or infection. Comparative genomics studies with clinical and environmental isolates will be determinant to better understand the implications and risks of the spread of pathogens in the environment. However, it is essential that the public databases are gradually enriched with complete sequences of environmental isolates, which are still scarce when compared with the data available for clinical isolates.

5. References

Bagley, S. T. (1985). Habitat association of *Klebsiella* species. Infection Control, *6*(2), 52–58. https://doi.org/10.1017/S0195941700062603

Baishya, J., & Wakeman, C. A. (2019). Selective pressures during chronic infection drive microbial competition and cooperation. *Npj Biofilms and Microbiomes*, *5*(16), 1–16. https://doi.org/10.1038/s41522-019-0089-2

Bambeke, F. Van, Balzi, E., & Tulkens, P. M. (2000). Antibiotic efflux pumps. *Biochemical Pharmacology*, *60*, 457–470.

Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Nicolas-Chanoine, M. H., Decré, D., & Brisse, S. (2014). Development of a multiplex PCR assay for identification of *Klebsiella pneumoniae* hypervirulent clones of capsular serotype K2. *Journal of Medical Microbiology*, *63*, 1608–1614. https://doi.org/10.1099/jmm.0.081448-0

Brisse, S., Grimont, F., & Grimont, P. A. D. (2006). The genus *Klebsiella*. p159-196. In: Prokaryotes. (Eds Dworkin, M., Falkow, S., Rosenberg, E., Schleifer,K-H- & Stackebrand, E.), Springer, Singapore.

Clarke, K. R., & Gorley, R. N. (2006). *PRIMER v6: User Manual/Tutorial (Plymouth Routines in Multivariate Ecological Research)* (PRIMER-E (ed.)).

Contreras-Moreira, B., & Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Applied and Environmental Microbiology*, *79*(24), 7696–7701. https://doi.org/10.1128/AEM.02411-13

Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. D., & Brisse, S. (2005). Multilocus sequence typing of Klebsiella pneumoniae nosocomial isolates. *Journal of Clinical Microbiology*, *43*(8), 4178–4182. https://doi.org/10.1128/JCM.43.8.4178-4182.2005

Gomez-Simmonds, A., & Uhlemann, A.C. (2017). Clinical implications of genomic adaptation and evolution of carbapenem-resistant *Klebsiella pneumoniae*. *Journal of Infectious Diseases*, *215*(1), S18–S27. https://doi.org/10.1093/infdis/jiw378

Guan, N., Li, J., Shin, H.-d, Du, G., Chen, J., & Liu, L. (2017). Microbial response to environmental stresses: from fundamental mechanisms to practical applications. *Applied Microbiology and Biotechnology*, *101*(10), 3991–4008. https://doi.org/10.1007/s00253-017-8264-y

Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., Jenney, A., Connor, T. R., Hsu, L. Y., Severin, J., Brisse, S., Cao, H., Wilksch, J., Gorrie, C., Schultz, M. B., Edwards, D. J., Van Nguyen, K., Nguyen, T. V., Dao, T. T., … 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 of the United States of America*, *112*(27), E3574–E3581. https://doi.org/10.1073/pnas.1501049112

Jolley, K. A., & Maiden, M. C. J. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*, *11*, 1–11. https://doi.org/10.1186/1471-2105-11-595

Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, *22*(13), 1658–1659. https://doi.org/10.1093/bioinformatics/btl158

Liu, X., Zhang, J., Li, Y., Shen, Q., Jiang, W., Zhao, K., He, Y., Dai, P., Nie, Z., Xu, X., & Zhou, Y. (2019). Diversity and frequency of resistance and virulence genes in *bla*$_{KPC}$ and *bla*$_{NDM}$ co-producing *Klebsiella pneumoniae* strains from China. *Infection and Drug Resistance*, *12*, 2819–2826. https://doi.org/10.2147/IDR.S214960

Long, S. W., Lindson, S. E., Saavedra, M. O., Cantu, C., Davis, J. J., Brettin, T., & Olsen, R. J. (2017). Whole-genome sequencing of human clinical *Klebsiella pneumoniae* isolates reveals misidentification and misunderstandings of *Klebsiella pneumoniae*, *Klebsiella variicola*, and *Klebsiella quasipneumoniae*. *MSphere*, *2*(4), 1–15.

Martin, R. M., & Bachman, M. A. (2018). Colonization, infection, and the accessory genome of *Klebsiella pneumoniae*. *Frontiers in Cellular and Infection Microbiology*, *8*, 1–15. https://doi.org/10.3389/fcimb.2018.00004

Navon-Venezia, S., Kondratyeva, K., & Carattoli, A. (2017). *Klebsiella pneumoniae*: A major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiology Reviews*, *41*(3), 252–275. https://doi.org/10.1093/femsre/fux013

Paczosa, M. K., & Mecsas, J. (2016). *Klebsiella pneumoniae*: Going on the offense with a strong defense. *Microbiology and Molecular Biology Reviews*, *80*(3), 629–661. https://doi.org/10.1128/mmbr.00078-15

Pendleton, J. N., Gorman, S. P., & Gilmore, B. F. (2013). Clinical relevance of the ESKAPE pathogens. *Expert Review of Anti-Infective Therapy*, *11*(3), 297–308. https://doi.org/10.1586/eri.13.12

Podschun, R., & Ullmann, U. (1998). *Klebsiella* spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clinical Microbiology Reviews*, *11*(4), 589–603. https://doi.org/10.1128/cmr.11.4.589

Rodrigues, C., Passet, V., Rakotondrasoa, A., Brisse, S. (2018). Identification of *Klebsiella pneumoniae*, *Klebsiella quasipneumoniae*, *Klebsiella variicola* and related phylogroups by MALDI-TOF mass spectrometry. *Frontiers in Microbiology*, *9*, 1–7. https://doi.org/10.3389/fmicb.2018.03000

Rozman, U., Duh, D., Cimerman, M., & Turk, S. Š. (2020). Hospital wastewater effluent: hot spot for antibiotic resistant bacteria. *Journal of Water Sanitation and Hygiene for Development*, *10*(2), 171–178. https://doi.org/10.2166/washdev.2020.086

Runcharoen, C., Moradigaravand, D., Blane, B., Paksanont, S., Thammachote, J., Anun, S., Parkhill, J., Chantratita, N., & Peacock, S. J. (2017). Whole genome sequencing reveals high-resolution epidemiological links between clinical and environmental *Klebsiella pneumoniae*. *Genome Medicine*, *9*(1), 1–10. https://doi.org/10.1186/s13073-017-0397-1

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

Teixeira, P., Tacão, M., Baraúna, R. A., Silva, A., & Henriques, I. (2020). Genomic analysis of *Chromobacterium haemolyticum*: insights into the species resistome, virulence determinants and genome plasticity. *Molecular Genetics and Genomics*, *295*(4), 1001–1012. https://doi.org/10.1007/s00438-020-01676-8

Williams, P., & Cámara, M. (2009). Quorum sensing and environmental adaptation in *Pseudomonas aeruginosa*: a tale of regulatory networks and multifunctional signal molecules. *Current Opinion in Microbiology*, *12*(2), 182–191. https://doi.org/10.1016/j.mib.2009.01.005

Wyres, K. L., & Holt, K. E. (2016). *Klebsiella pneumoniae* population genomics and antimicrobial-resistant clones. *Trends in Microbiology*, *24*(12), 944–956. https://doi.org/10.1016/j.tim.2016.09.007

Wyres, K. L., & Holt, K. E. (2018). *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Current Opinion in Microbiology*, *45*, 131–139. https://doi.org/10.1016/j.mib.2018.04.004

Wyres, K. L., Lam, M. M. C., & Holt, K. E. (2020). Population genomics of *Klebsiella pneumoniae*. *Nature Reviews Microbiology*, *18*(6), 344–359. https://doi.org/10.1038/s41579-019-0315-1

Yu, X., Zhang, W., Zhao, Z., Ye, C., Zhou, S., Wu, S., Han, L., Han, Z., & Ye, H. (2019). Molecular characterization of carbapenem-resistant *Klebsiella pneumoniae* isolates with focus on antimicrobial resistance. *BMC Genomics*, *20*(1). https://doi.org/10.1186/s12864-019-6225-9

Zagui, G. S., de Andrade, L. N., Moreira, N. C., Silva, T. V., Machado, G. P., da Costa Darini, A. L., & Segura-Muñoz, S. I. (2020). Gram-negative bacteria carrying β-lactamase encoding genes in hospital and urban wastewater in Brazil. *Environmental Monitoring and Assessment*, *192*(376), 1–11. https://doi.org/10.1007/s10661-020-08319-w

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

Zhou, H., Zhang, K., Chen, W., Chen, J., Zheng, J., Liu, C., Cheng, L., Zhou, W., Shen, H., & Cao, X. (2020). Epidemiological characteristics of carbapenem-resistant *Enterobacteriaceae* collected from 17 hospitals in Nanjing district of China. *Antimicrobial Resistance and Infection Control*, *9*(1), 1–10. https://doi.org/10.1186/s13756-019-0674-4

6. Supplementary Material

6.1. Supplementary Figures

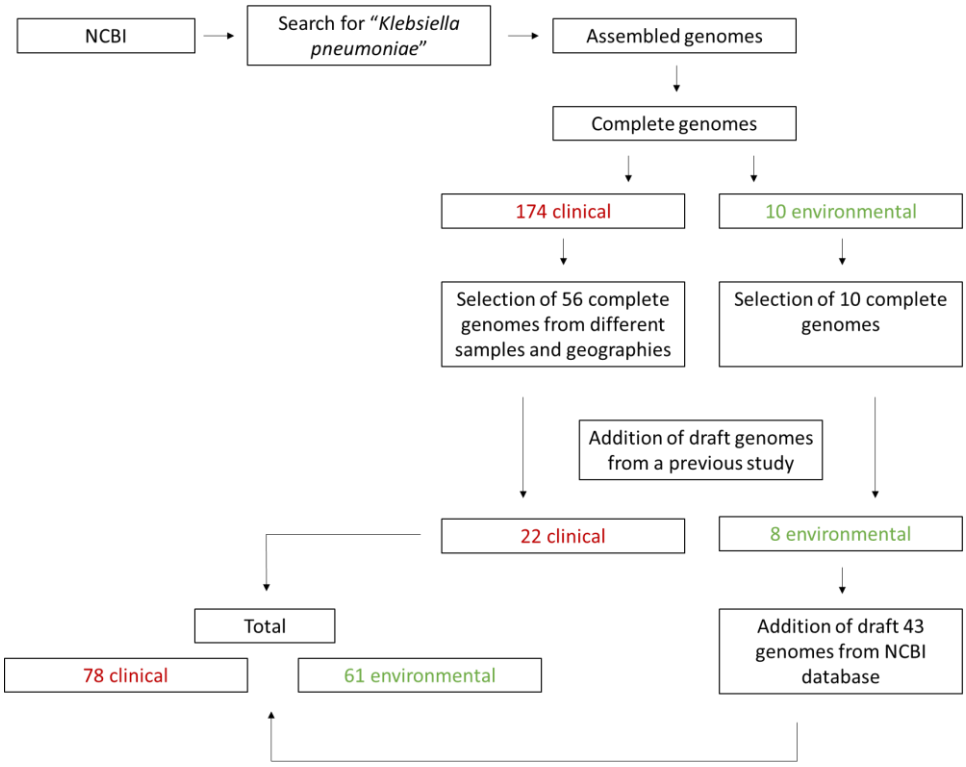


Figure S1 - Workflow followed to establish the *K. pneumoniae* collection of genomes used in this study.

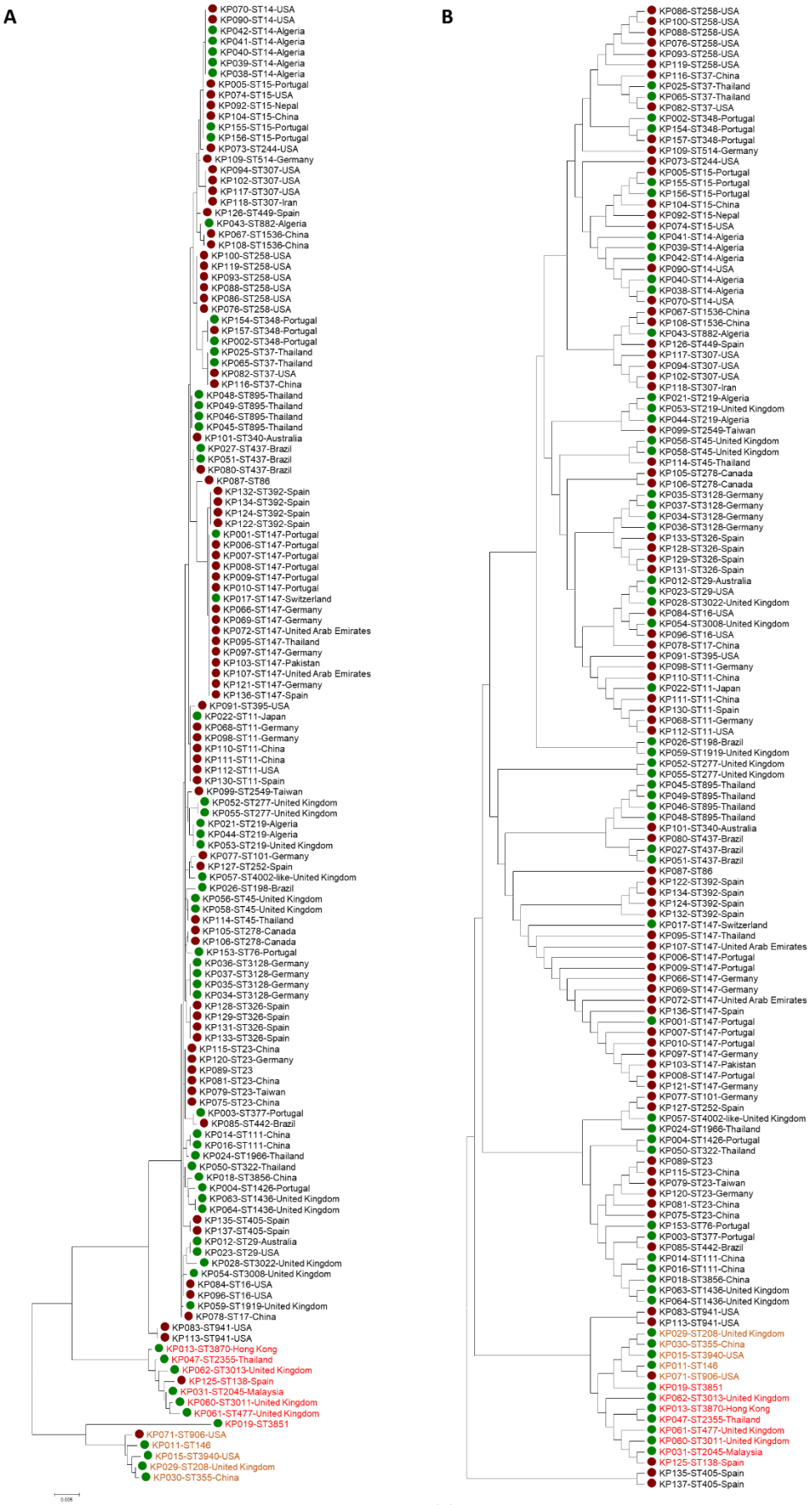

A
KP070-ST14-USA
KP090-ST14-USA
KP042-ST14-Algeria
KP041-ST14-Algeria
KP040-ST14-Algeria
KP039-ST14-Algeria
KP038-ST14-Algeria
KP005-ST15-Portugal
KP074-ST15-USA
KP092-ST15-Nepal
KP104-ST15-China
KP155-ST15-Portugal
KP156-ST15-Portugal
KP073-ST244-USA
KP109-ST514-Germany
KP094-ST307-USA
KP102-ST307-USA
KP117-ST307-USA
KP118-ST307-Iran
KP126-ST449-Spain
KP043-ST882-Algeria
KP067-ST1536-China
KP108-ST1536-China
KP100-ST258-USA
KP119-ST258-USA
KP093-ST258-USA
KP088-ST258-USA
KP086-ST258-USA
KP076-ST258-USA
KP154-ST348-Portugal
KP157-ST348-Portugal
KP002-ST348-Portugal
KP025-ST37-Thailand
KP065-ST37-Thailand
KP082-ST37-USA
KP116-ST37-China
KP048-ST895-Thailand
KP049-ST895-Thailand
KP046-ST895-Thailand
KP045-ST895-Thailand
KP101-ST340-Australia
KP027-ST437-Brazil
KP051-ST437-Brazil
KP080-ST437-Brazil
KP087-ST86
KP132-ST392-Spain
KP134-ST392-Spain
KP124-ST392-Spain
KP122-ST392-Spain
KP001-ST147-Portugal
KP006-ST147-Portugal
KP007-ST147-Portugal
KP008-ST147-Portugal
KP009-ST147-Portugal
KP010-ST147-Portugal
KP017-ST147-Switzerland
KP066-ST147-Germany
KP069-ST147-Germany
KP072-ST147-United Arab Emirates
KP095-ST147-Thailand
KP097-ST147-Germany
KP103-ST147-Pakistan
KP107-ST147-United Arab Emirates
KP121-ST147-Germany
KP136-ST147-Spain
KP091-ST395-USA
KP022-ST11-Japan
KP068-ST11-Germany
KP098-ST11-Germany
KP110-ST11-China
KP111-ST11-China
KP112-ST11-USA
KP130-ST11-Spain
KP099-ST2549-Taiwan
KP052-ST277-United Kingdom
KP055-ST277-United Kingdom
KP021-ST219-Algeria
KP044-ST219-Algeria
KP053-ST219-United Kingdom
KP077-ST101-Germany
KP127-ST252-Spain
KP057-ST4002-like-United Kingdom
KP026-ST198-Brazil
KP056-ST45-United Kingdom
KP058-ST45-United Kingdom
KP114-ST45-Thailand
KP105-ST278-Canada
KP106-ST278-Canada
KP153-ST76-Portugal
KP036-ST3128-Germany
KP037-ST3128-Germany
KP035-ST3128-Germany
KP034-ST3128-Germany
KP128-ST326-Spain
KP129-ST326-Spain
KP131-ST326-Spain
KP133-ST326-Spain
KP115-ST23-China
KP120-ST23-Germany
KP089-ST23
KP081-ST23-China
KP079-ST23-Taiwan
KP075-ST23-China
KP003-ST377-Portugal
KP085-ST442-Brazil
KP014-ST111-China
KP016-ST111-China
KP024-ST1966-Thailand
KP050-ST322-Thailand
KP018-ST3856-China
KP004-ST1426-Portugal
KP063-ST1436-United Kingdom
KP064-ST1436-United Kingdom
KP135-ST405-Spain
KP137-ST405-Spain
KP012-ST29-Australia
KP023-ST29-USA
KP028-ST3022-United Kingdom
KP054-ST3008-United Kingdom
KP084-ST16-USA
KP096-ST16-USA
KP059-ST1919-United Kingdom
KP078-ST17-China
KP083-ST941-USA
KP113-ST941-USA
KP013-ST3870-Hong Kong
KP047-ST2355-Thailand
KP062-ST3013-United Kingdom
KP125-ST138-Spain
KP031-ST2045-Malaysia
KP060-ST3011-United Kingdom
KP061-ST477-United Kingdom
KP019-ST3851
KP071-ST906-USA
KP011-ST146
KP015-ST3940-USA
KP029-ST208-United Kingdom
KP030-ST355-China
0.005
B
KP086-ST258-USA
KP100-ST258-USA
KP088-ST258-USA
KP076-ST258-USA
KP093-ST258-USA
KP119-ST258-USA
KP116-ST37-China
KP025-ST37-Thailand
KP065-ST37-Thailand
KP082-ST37-USA
KP002-ST348-Portugal
KP154-ST348-Portugal
KP157-ST348-Portugal
KP109-ST514-Germany
KP073-ST244-USA
KP005-ST15-Portugal
KP155-ST15-Portugal
KP156-ST15-Portugal
KP104-ST15-China
KP092-ST15-Nepal
KP074-ST15-USA
KP041-ST14-Algeria
KP039-ST14-Algeria
KP042-ST14-Algeria
KP090-ST14-USA
KP040-ST14-Algeria
KP038-ST14-Algeria
KP070-ST14-USA
KP067-ST1536-China
KP108-ST1536-China
KP043-ST882-Algeria
KP126-ST449-Spain
KP117-ST307-USA
KP094-ST307-USA
KP102-ST307-USA
KP118-ST307-Iran
KP021-ST219-Algeria
KP053-ST219-United Kingdom
KP044-ST219-Algeria
KP099-ST2549-Taiwan
KP056-ST45-United Kingdom
KP058-ST45-United Kingdom
KP114-ST45-Thailand
KP105-ST278-Canada
KP106-ST278-Canada
KP035-ST3128-Germany
KP037-ST3128-Germany
KP034-ST3128-Germany
KP036-ST3128-Germany
KP133-ST326-Spain
KP128-ST326-Spain
KP129-ST326-Spain
KP131-ST326-Spain
KP012-ST29-Australia
KP023-ST29-USA
KP028-ST3022-United Kingdom
KP084-ST16-USA
KP054-ST3008-United Kingdom
KP096-ST16-USA
KP078-ST17-China
KP091-ST395-USA
KP098-ST11-Germany
KP110-ST11-China
KP022-ST11-Japan
KP111-ST11-China
KP130-ST11-Spain
KP068-ST11-Germany
KP112-ST11-USA
KP026-ST198-Brazil
KP059-ST1919-United Kingdom
KP052-ST277-United Kingdom
KP055-ST277-United Kingdom
KP045-ST895-Thailand
KP049-ST895-Thailand
KP046-ST895-Thailand
KP048-ST895-Thailand
KP101-ST340-Australia
KP080-ST437-Brazil
KP027-ST437-Brazil
KP051-ST437-Brazil
KP087-ST86
KP122-ST392-Spain
KP134-ST392-Spain
KP124-ST392-Spain
KP132-ST392-Spain
KP017-ST147-Switzerland
KP095-ST147-Thailand
KP107-ST147-United Arab Emirates
KP006-ST147-Portugal
KP009-ST147-Portugal
KP066-ST147-Germany
KP069-ST147-Germany
KP072-ST147-United Arab Emirates
KP136-ST147-Spain
KP001-ST147-Portugal
KP007-ST147-Portugal
KP010-ST147-Portugal
KP097-ST147-Germany
KP103-ST147-Pakistan
KP008-ST147-Portugal
KP121-ST147-Germany
KP077-ST101-Germany
KP127-ST252-Spain
KP057-ST4002-like-United Kingdom
KP024-ST1966-Thailand
KP004-ST1426-Portugal
KP050-ST322-Thailand
KP089-ST23
KP115-ST23-China
KP079-ST23-Taiwan
KP120-ST23-Germany
KP081-ST23-China
KP075-ST23-China
KP153-ST76-Portugal
KP003-ST377-Portugal
KP085-ST442-Brazil
KP014-ST111-China
KP016-ST111-China
KP018-ST3856-China
KP063-ST1436-United Kingdom
KP064-ST1436-United Kingdom
KP083-ST941-USA
KP113-ST941-USA
KP029-ST208-United Kingdom
KP030-ST355-China
KP015-ST3940-USA
KP011-ST146
KP071-ST906-USA
KP019-ST3851
KP062-ST3013-United Kingdom
KP013-ST3870-Hong Kong
KP047-ST2355-Thailand
KP061-ST477-United Kingdom
KP060-ST3011-United Kingdom
KP031-ST2045-Malaysia
KP125-ST138-Spain
KP135-ST405-Spain
KP137-ST405-Spain

Figure S2 - Phylogenetic tree obtained concatenating MLST gene sequences (*gapA*; *infB*; *mdh*; *pgi*; *phoE*; *rpoB*; *tonB*) using Maximum Likelihood (A) and Maximum Parsimony (B) methods.

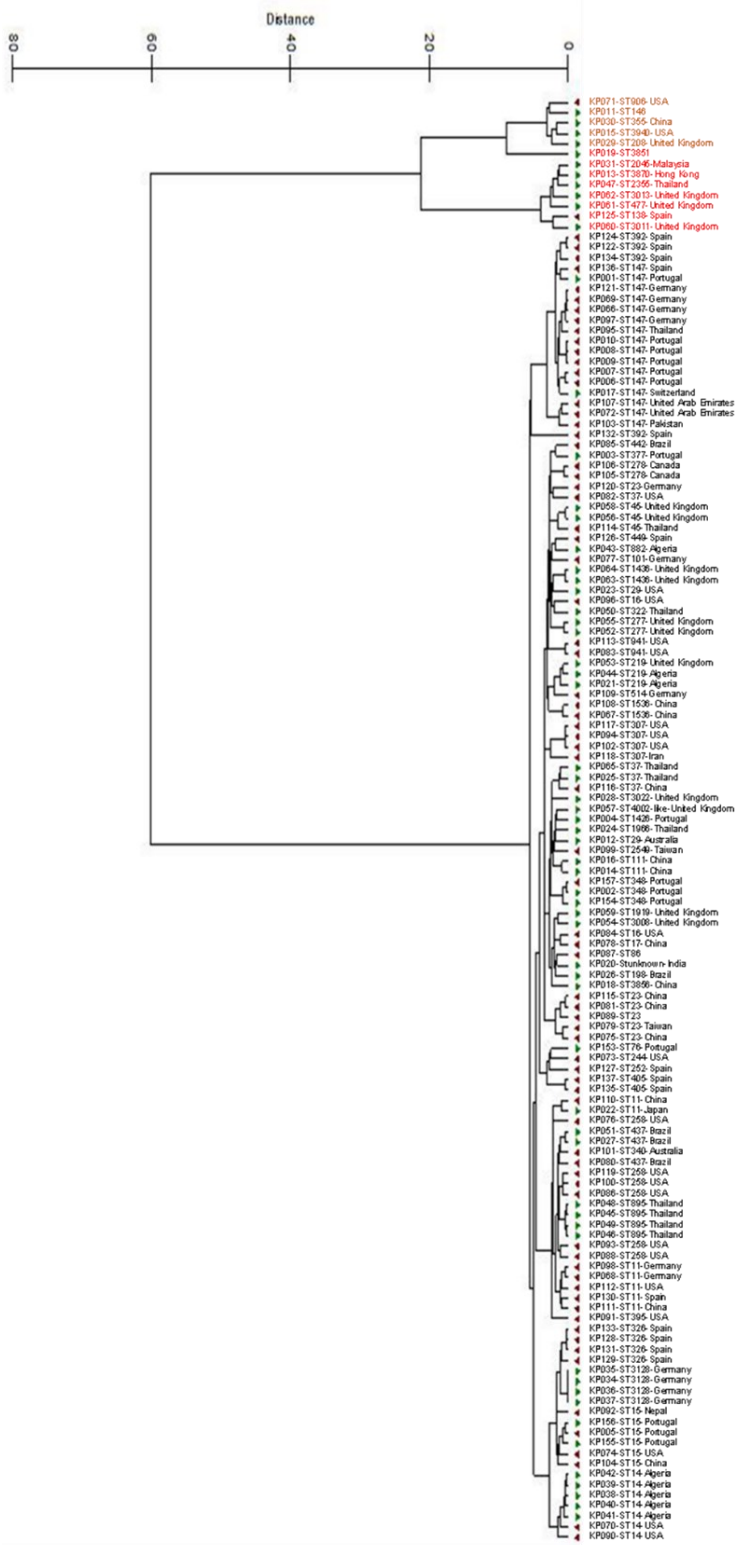


Figure S3 - Dendrogram obtained from a matrix based on pairwise ANIb comparisons among the 139 genomes. Red symbols represent clinical isolates (n=78) and green symbols represent environmental isolates (n=61).

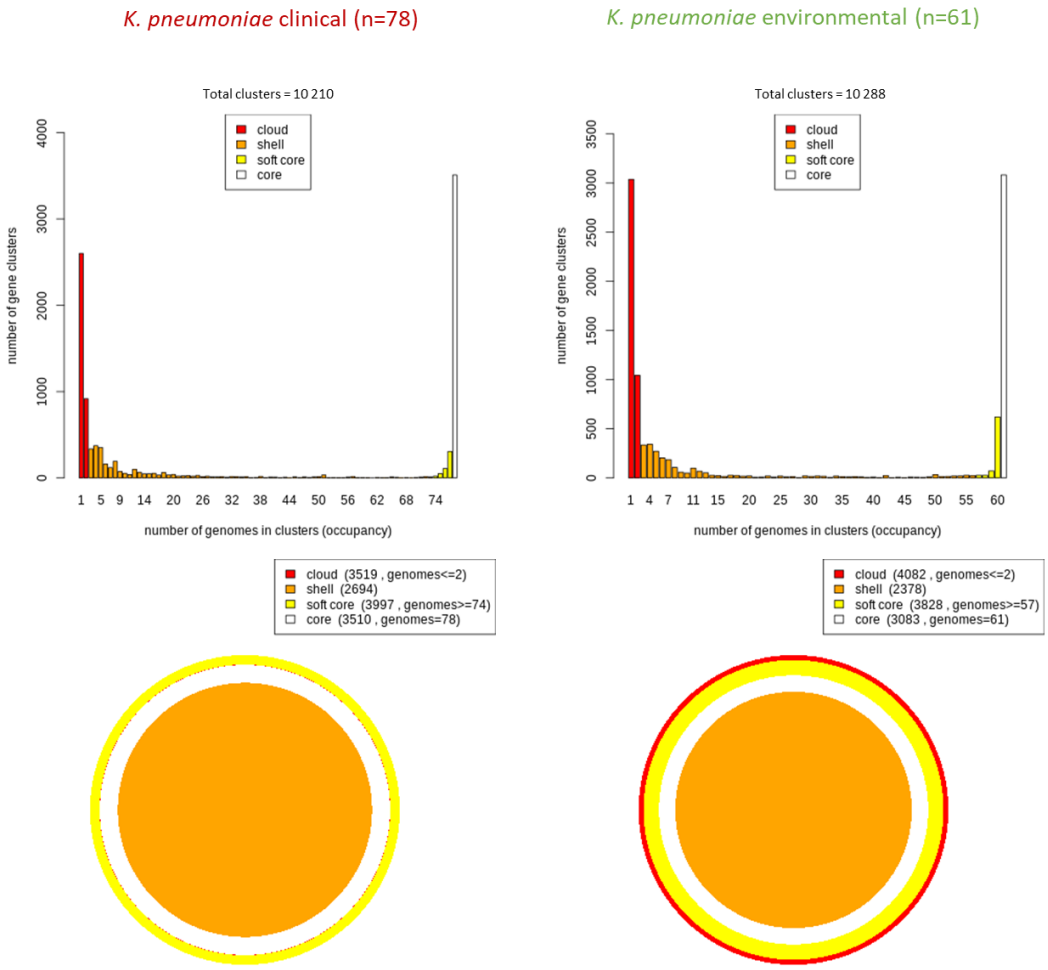


Figure S4 - Pangenome analysis of the individual clinical and environmental *K. pneumoniae* isolates based on the criteria of 50% of coverage and 70% identity between deduced amino acid sequences. Core genes – present in all genomes; soft core genes – present in 95% of the genomes, cloud genes – present only in less than 2 genomes and shell genes – present in more than 2 genomes and less than 95% of the genomes.

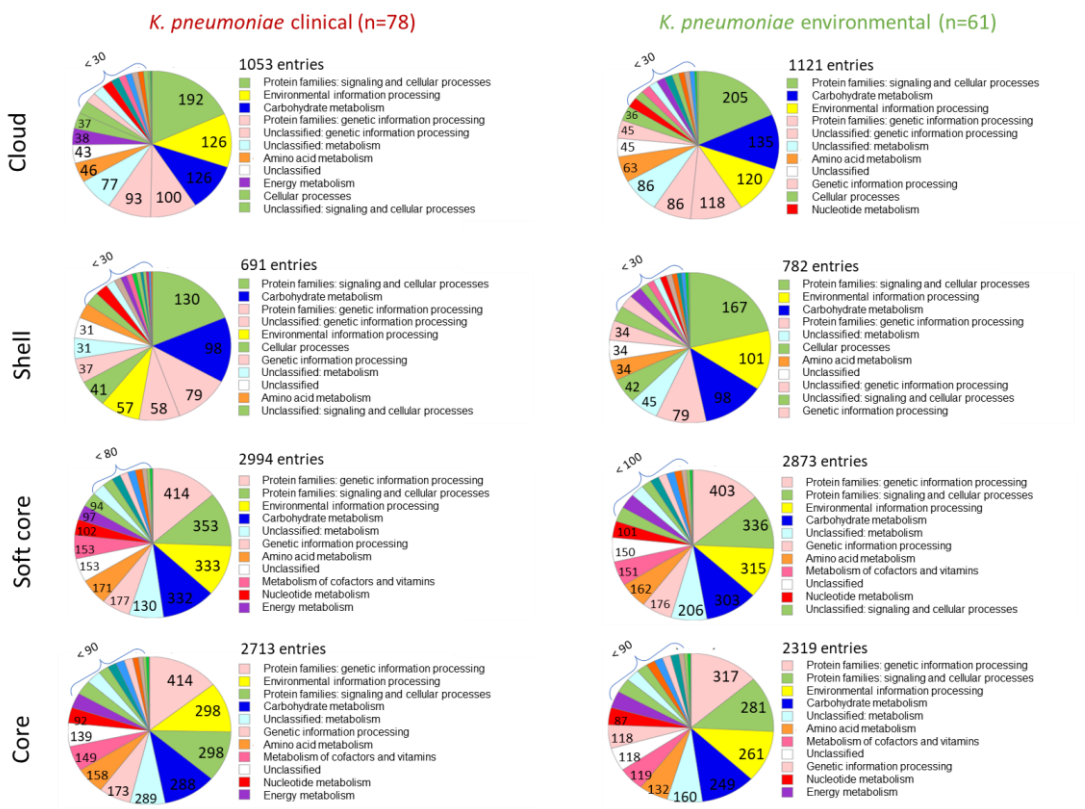


Figure S5 - Functional categories of the pangenome analysis of clinical and environmental *K. pneumoniae* isolates amino acid sequences. Numbers refer to distinct deduced amino acid sequences with an attributed functional category. Metabolism functional category include: carbohydrate, energy, lipid, nucleotide, amino acid metabolism, glycan biosynthesis, metabolism of cofactors and vitamins, metabolism of terpenoids and polyketides, biosynthesis of other secondary metabolites, and xenobiotics biodegradation and metabolism. Genetic information processing functional category include: translation, folding, sorting and degradation, and replication and repair. Environmental information processing functional category include: membrane transport, and signal transduction. Cellular processes functional category includes: transport and catabolism, cell growth and death, cellular community – prokaryotes, and cell motility.


Antibiotic Resistance
aac(3)
aac(6')
aadA
aadB
aph(3')
aph(4)
armA
rmtB
rmtC
rmtF
rmtG
strA
strB
blaCMY
blaCTX-M
blaDHA
blaIMP
blaKPC
blaLEN
blaNDM
blaOKP-A
blaOXA
blaSHV
blaTEM
ere
erm
mef
mph
msrE
sul
tet
mcr
dfrA
fosA
catA
catB
cmlA
flor
aac(6')Ib-cr
qnrA
qnrB
qnrS
arr
intI
Metal Resistance
arsA
arsB
arsC
arsD
arsR
merA
merC
merD
merE
merP
merR
merT
pbrA
pbrBC
pbrR
pcoA
pcoB
pcoC
pcoD
pcoE
pcoR
pcoS
silC
silE
silF
silG
silP
silR
silS
terA
terB
terC
terD
terE
terW
terX
terY
terZ
Virulence
amt
nikA2
nikE2
ureA2
ureC2
yut
allA
allB
allC
allD
allR
allS
arcC
clbJ
clbK
fdrA
gcl
glxK
hyi
kfuA
kfuB
kfuC
kvgA
kvgS
mceA
mceB
mceC
mceD
mceE
mceG
mceH
mceI
mceJ
mrkA
mrkB
mrkC
mrkD
mrkF
mrkH
mrkI
mrkJ
rmpA
rmpA2
ybbW
ybbY
ybtE
ybtF
iroB
iroC
iroD
iroN
clbA
clbB
clbC
clbD
clbE
clbF
clbG
clbI
clbL
clbM
clbN
clbO
clbP
clbQ
iucA
iucB
iucC
iucD
iutA
wzc
wzi
kpnA
KpvSO233a
kpiA
nikA2
fyuA
irp1
irp2
ybtA
ybtE
ybtP
ybtQ
ybtS
ybtT
ybtU
ybtX

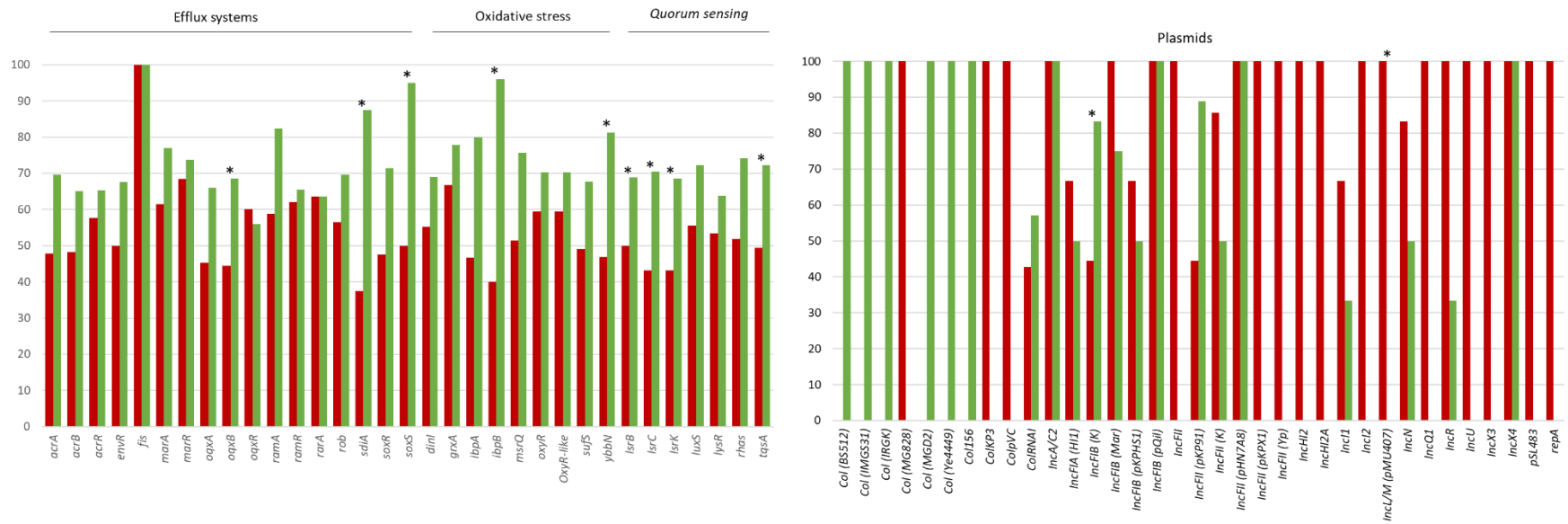


Figure S6 - Prevalence in percentage of the alleles found for each gene in clinical (n=78) and environmental (n=61) genomes analysed. The genes were organized by functional category. The asterisk indicates statistically significant differences between the prevalence of alleles found in the total of clinical and the total of environmental isolates compared to the total alleles found for the same gene, based on Fisher's exact test and p-value < 0.05.

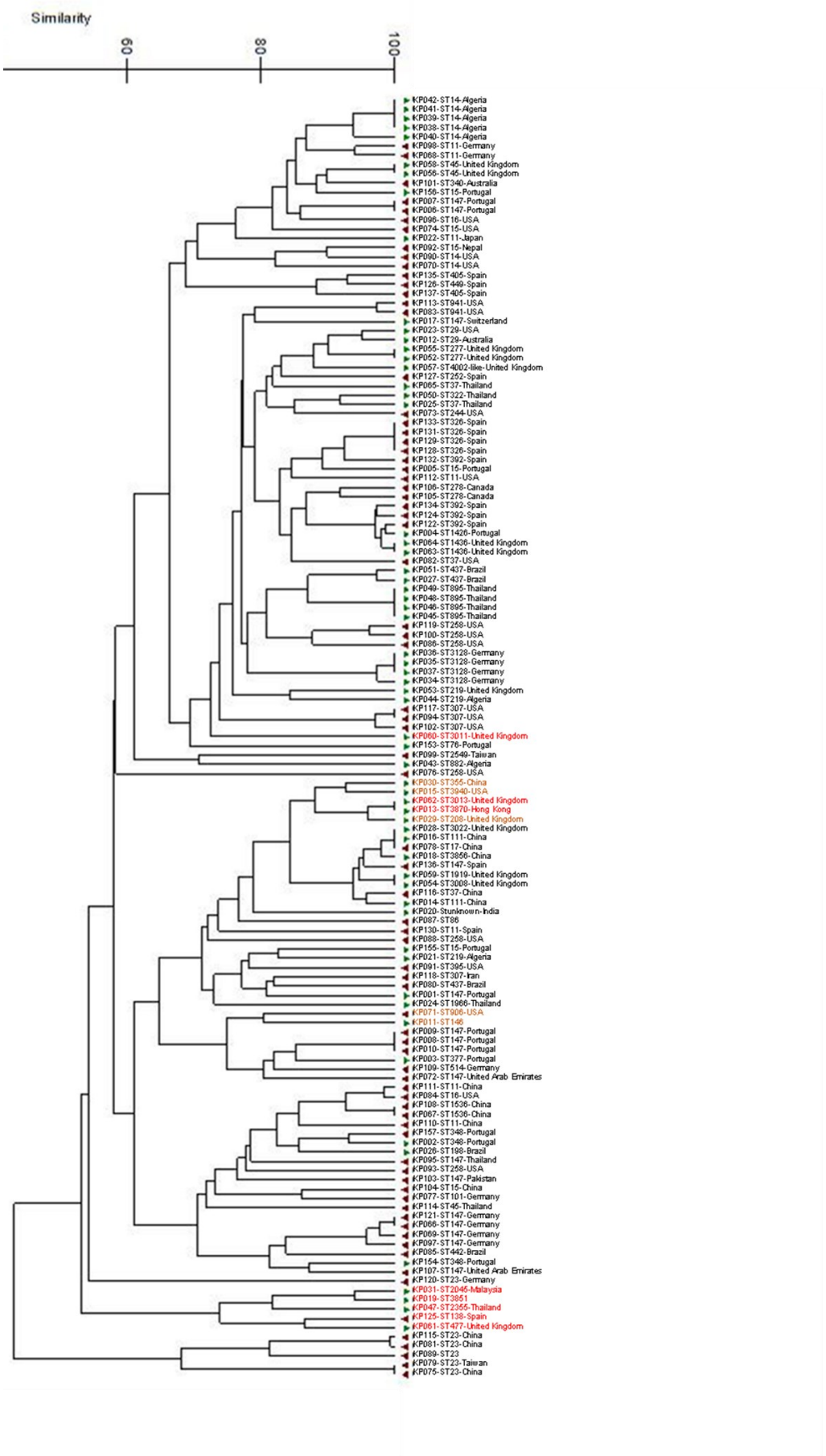

Similarity
60
80
100
KP042-ST14-Algeria
KP041-ST14-Algeria
KP039-ST14-Algeria
KP038-ST14-Algeria
KP040-ST14-Algeria
KP098-ST11-Germany
KP088-ST11-Germany
KP058-ST45-United Kingdom
KP056-ST45-United Kingdom
KP101-ST340-Australia
KP156-ST15-Portugal
KP007-ST147-Portugal
KP006-ST147-Portugal
KP096-ST16-USA
KP074-ST15-USA
KP022-ST11-Japan
KP092-ST15-Nepal
KP090-ST14-USA
KP070-ST14-USA
KP135-ST405-Spain
KP126-ST449-Spain
KP137-ST405-Spain
KP113-ST941-USA
KP083-ST941-USA
KP017-ST147-Switzerland
KP023-ST29-USA
KP012-ST29-Australia
KP055-ST277-United Kingdom
KP052-ST277-United Kingdom
KP057-ST4002-like-United Kingdom
KP127-ST252-Spain
KP065-ST37-Thailand
KP050-ST322-Thailand
KP025-ST37-Thailand
KP073-ST244-USA
KP133-ST326-Spain
KP131-ST326-Spain
KP129-ST326-Spain
KP128-ST326-Spain
KP132-ST392-Spain
KP005-ST15-Portugal
KP112-ST11-USA
KP106-ST278-Canada
KP105-ST278-Canada
KP134-ST392-Spain
KP124-ST392-Spain
KP122-ST392-Spain
KP004-ST1426-Portugal
KP064-ST1436-United Kingdom
KP063-ST1436-United Kingdom
KP082-ST37-USA
KP051-ST437-Brazil
KP027-ST437-Brazil
KP049-ST895-Thailand
KP048-ST895-Thailand
KP046-ST895-Thailand
KP045-ST895-Thailand
KP119-ST258-USA
KP100-ST258-USA
KP086-ST258-USA
KP036-ST3128-Germany
KP035-ST3128-Germany
KP037-ST3128-Germany
KP034-ST3128-Germany
KP053-ST219-United Kingdom
KP044-ST219-Algeria
KP117-ST307-USA
KP094-ST307-USA
KP102-ST307-USA
KP060-ST3011-United Kingdom
KP153-ST76-Portugal
KP099-ST2549-Taiwan
KP043-ST882-Algeria
KP076-ST258-USA
KP030-ST355-China
KP015-ST3940-USA
KP062-ST3013-United Kingdom
KP013-ST3870-Hong Kong
KP029-ST208-United Kingdom
KP028-ST3022-United Kingdom
KP016-ST111-China
KP078-ST17-China
KP018-ST3856-China
KP138-ST147-Spain
KP059-ST1919-United Kingdom
KP054-ST3008-United Kingdom
KP116-ST37-China
KP014-ST111-China
KP020-Stunknown-India
KP087-ST86
KP130-ST11-Spain
KP088-ST258-USA
KP155-ST15-Portugal
KP021-ST219-Algeria
KP091-ST395-USA
KP118-ST307-Iran
KP080-ST437-Brazil
KP001-ST147-Portugal
KP024-ST1966-Thailand
KP071-ST906-USA
KP011-ST146
KP009-ST147-Portugal
KP008-ST147-Portugal
KP010-ST147-Portugal
KP003-ST377-Portugal
KP109-ST514-Germany
KP072-ST147-United Arab Emirates
KP111-ST11-China
KP084-ST16-USA
KP108-ST1536-China
KP067-ST1536-China
KP110-ST11-China
KP157-ST348-Portugal
KP002-ST348-Portugal
KP026-ST198-Brazil
KP095-ST147-Thailand
KP093-ST258-USA
KP103-ST147-Pakistan
KP104-ST15-China
KP077-ST101-Germany
KP114-ST45-Thailand
KP121-ST147-Germany
KP066-ST147-Germany
KP069-ST147-Germany
KP097-ST147-Germany
KP085-ST442-Brazil
KP154-ST348-Portugal
KP107-ST147-United Arab Emirates
KP120-ST23-Germany
KP031-ST2045-Malaysia
KP019-ST3851
KP047-ST2355-Thailand
KP125-ST138-Spain
KP061-ST477-United Kingdom
KP115-ST23-China
KP081-ST23-China
KP089-ST23
KP079-ST23-Taiwan
KP075-ST23-China

Figure S7 – UPGMA dendrogram based on the presence/absence matrix of fitness-related genes in the 139 genomes. Red symbols represent clinical isolates (n=78) and green symbols represent environmental isolates (n=61).

6.2. Supplementary Tables

Due to high dimension, supplementary tables are available in the following link:

https://docs.google.com/spreadsheets/d/1qLMIZ6S6cEVpdOlTru9osICZPjbJg2U8/edit?usp=sharing&ouid=109610140440067111385&rtpof=true&sd=true

Table S1 - List of *Klebsiella pneumoniae* isolates genomes used in this study.

Table S2 - Pairwise average nucleotide identity calculated for the list of *Klebsiella pneumoniae* isolates in study.

Table S3 - Clinically relevant genes screening and respective alleles on the *Klebsiella pneumoniae* isolates used in this study.

Table S4 - Clinically relevant genes presence (1) and absence (0).

Table S5 - Clinically relevant alleles presence (1) and absence (0).

Table S6 - Clinically relevant genes alleles diversity indices for genes detected in clinical and environmental *Klebsiella pneumoniae* isolates.

CHAPTER 7

GENERAL DISCUSSION AND MAIN CONCLUSIONS

Antibiotic resistance is a natural phenomenon in bacteria that after 80 decades of antibiotic use in human medicine, veterinary and food-animal production and in plant agriculture has been regarded as a major public health threat. Over the years, antibiotic residues, antibiotic resistant bacteria and antibiotic resistance genes have been disseminated in the environment, from pristine to highly impacted areas, in water, soil and wildlife (Dantas *et al*., 2008; D'Costa *et al*., 2011; Rizzo *et al*., 2013; Thaller *et al*., 2010; Vredenburg *et al*., 2014). In urban areas, wastewater treatment plants are major recipients, reservoirs and sources of these contaminants and therefore have been considered critical control points where antibiotic resistance can be monitored aiming the improvement of treatment processes and the attenuation of impacts of the discharge into the environment (Bürgmann *et al*., 2018; Manaia *et al*., 2016). Quantitative PCR (qPCR) has been adopted to monitor antibiotic resistance in wastewater and to study the effect of wastewater treatment processes (Cacace *et al*., 2019; Storteboom *et al*., 2010). Two major aims have been identified to be tackled through antibiotic resistance monitoring. The first, is the improvement and development of novel wastewater treatment processes. The second, is the comparison of the resistance loads entering and being emitted by wastewater treatment plants in regions with distinct climate and socio-economic contexts (Di Cesare *et al*., 2020; Pärnänen *et al*., 2019). However, the comparison of results may be challenging due to biases caused by different factors such as the DNA extraction kit used, the experience of the operator on molecular biology techniques, the samples matrix effects, qPCR equipment, protocols and reagents, among others (Djurhuus *et al*., 2017; Hinlo *et al*., 2017; Kim *et al*., 2013; Li *et al*., 2018; Rocha *et al*., 2020). One of the hypotheses of this thesis was that qPCR can be used in inter-laboratory assays for the quantification of antibiotic resistance genes in different types of wastewater to assess treatment efficiency and potential environmental impacts. This is a procedure currently used, however the conditions and most influencing variables are not properly identified, limiting the reliability of such comparisons. The monitoring of antibiotic resistance genes using qPCR can be made based on abundance (per volume) or prevalence (per 16S rRNA gene) determinations (Cacace *et al*., 2019; Gao *et al*., 2012; Narciso-da-Rocha *et al*., 2018; Rodriguez-Mozaz *et al*., 2015). However, the type of information retrieved is different and must be adjusted to the aims of the study. In this study it was demonstrated that treatment efficiency must be based on abundance values, rather than on prevalence, since total and antibiotic resistant bacteria are removed at approximately identical rates. It was also concluded that comparisons of determinations made in different laboratories may be reliable, mainly if targeting high abundance genes and if the procedures are harmonized. One of the factors recognized as crucial to the reliability of culture-independent microbiological studies is the quality of the DNA extract (Djurhuus *et al*., 2017; Hinlo *et al*., 2017; Li *et al*., 2018; Riediger *et al*., 2016). This can be a major obstacle for groups who have the expertise in wastewater

treatment and advanced treatment processes, but lack laboratory conditions and know-how on molecular biology. The second hypothesis of this thesis addressed this challenge and developed a method to control the DNA losses during extraction from wastewater and water samples, through the use of a cell-based internal standard. Several studies have assessed the effect of the use of standards for DNA extraction or qPCR experiments, however these were either based on cell-free DNA added to the samples prior to DNA extraction or added to the DNA extract prior to qPCR measurements (Cloud *et al*., 2003; Volkmann *et al*., 2007; Burggraf and Olgemöller, 2004). In both cases, the most important factor in DNA extraction, which is the cell lysis, was not considered. Therefore, the use of a cell-based internal standard is more reliable to assess cell losses during filtration, DNA extraction and in qPCR. The use of an internal standard that could be shared among different laboratories, mainly in those with limited experience in DNA extraction, was considered a valid contribution to promote the comparability of results worldwide.

A topic of intense debate is the selection of the best markers to assess the occurrence and spread of antibiotic resistance in wastewater and in the environment. The definition of priority markers is challenging and sometimes still based on variable criteria (Manaia *et al*., 2018). However, some criteria seem to be consensual, being the ubiquity in humans and wastewater, the high genomic plasticity and the history of acquired antibiotic resistance, leading arguments. *K. pneumoniae* meet these criteria, being, among others, identified by the World Health Organization and other identities (ECDC, 2019; WHO, 2014) as priority bacteria to tackle antibiotic resistance dissemination. Although members of this species can be found in human or environmental niches, the paths of transmission to humans are still not known. Environmental monitoring and assessment of risks of transmission to humans were major motivations to further investigate *K. pneumoniae.* The study was designed around the hypothesis that clinical isolates of *K. pneumoniae* may be lost or lose some antibiotic resistance and fitness features, once thriving in the environment. In fact, several comparative genomics studies have been discussing this question across different species (Youenou *et al*., 2015; Bodilis *et al*., 2018; Faoro *et al*., 2019; Ekwanzala *et al*., 2020). While in some studies clinical and environmental strains mostly shared common traits (Bodilis *et al*., 2018; Ekwanzala *et al*., 2020), in other studies, besides common traits, distinct features could be associated to clinical or environmental origins (Youenou *et al*., 2015; Faoro *et al*., 2019). This last observation is in line with the findings of this thesis where although it was not observed a specialization of *K. pneumoniae* genomes depending on the clinical or environmental origin, it was possible to identify traits more common in clinical isolates, such as antibiotic resistance genes, certain plasmid replicon types or virulence traits. Moreover, the higher level of diversity of efflux systems, oxidative stress and *quorum sensing* in environmental *K. pneumoniae* isolates might suggest that these traits are essential to their adaptation to different habitats.

These findings increased the knowledge on origin specific traits, which may be useful as genetic markers.

The main conclusions of this thesis can be summarized as:

- The assessment of the wastewater treatment efficiency should be expressed in abundance (per volume of sample) and not in prevalence (per 16S rRNA gene);
- The analysis of treatment efficiency and antibiotic resistance removal should be based on highly abundant genes, such as *sul1*;
- The DNA extraction, when performed by less experienced laboratories, might influence DNA extract quality, which is crucial to achieve reliable quantifications by qPCR;
- The simplification of DNA extraction procedures using cheaper or user-friendly DNA extraction kits might not create important bias on the comparison of antibiotic resistance gene quantification worldwide and might incentivize the participation of laboratories with lower resources in monitoring studies;
- The use of a cell-based internal standard is a procedure that will allow the control of losses occurred during DNA extraction procedure due to water samples matrix effect or during shipment of DNA extracts;
- The phenotypic and genomic studies did not evidence features that highlighted 3rd generation cephalosporins-resistant *K. pneumoniae* specialization to the isolation habitat, suggesting that clinical isolates once in wastewater may retain clinically relevant traits, even those that were acquired through horizontal gene transfer and were associated with transposons, insertion sequences or integrative and conjugative elements;
- The phylogeny, more than the isolates origin, was suggested to explain the profile of acquired traits, although genetic variation may occur within the same genetic lineage;
- The natural environment may offer favorable conditions for the diversification/evolution of housekeeping fitness-relevant features, such as efflux, oxidative stress and *quorum sensing* related, while the clinical context drives the acquisition and evolution of virulence and antibiotic and metal resistance genes as well as some plasmid replicon types.

CHAPTER 8

SUGGESTIONS OF FUTURE WORK

The present thesis revealed the influence of some factors on qPCR measurements and the potential clinical relevance of *K. pneumoniae* found in the environment. Some questions might be addressed as future work, such as:

- Study the application of the cell-based internal standard as a lyophilized product, more user-friendly and easier to share among different laboratories, as a way to allow the harmonization procedure among different laboratories;
- Study the effect of the water samples matrix effect and DNA extraction based on Gram-positive cell-based internal standard;
- Determine whether differences could be observed in the doses of *K. pneumoniae* needed to infect *G. mellonella*, among clinical and environmental isolates;
- Based on the previous suggestion, select the genes detected as more prevalent in clinical or environmental *K. pneumoniae* genomes in this thesis and investigate whether the gene expression is altered considering the different doses of clinical or environmental isolates tested;
- Perform comparative genome analysis similar to what was done in this thesis to the other ESKAPE microorganisms to increase the body of knowledge and allow the selection of target genes to be monitored in the surveillance studies.

REFERENCES

References cited in the Chapter 1, Chapter 2 and Chapter 7:

Aarestrup, F.M., & Woolhouse, M.E.J. (2020). Using sewage for surveillance of antimicrobial resistance. *Science 367*(6478), 630-632. https://doi.org/10.1126/science.aba3432

Akya, A., Chegenelorestani, R., Shahvaisi-Zadeh, J., & Bozorgomid, A. (2020). Antimicrobial resistance of *Staphylococcus aureus* isolated from hospital wastewater in Kermanshah, Iran. *Risk Management and Healthcare Policy, 13*, 1035–1042. https://doi.org/10.2147/RMHP.S261311

Alcalde-Sanz, L., & Gawlik, B. M. (2017). Minimum quality requirements for water reuse in agricultural irrigation and aquifer recharge - Towards a water reuse regulatory instrument at EU level. JRC Science for Policy Report. https://doi.org/10.2760/887727

Alexander, J., Bollmann, A., Seitz, W., & Schwartz, T. (2015). Microbiological characterization of aquatic microbiomes targeting taxonomical marker genes and antibiotic resistance genes of opportunistic bacteria. *Science of the Total Environment*, *512–513*, 316–325. https://doi.org/10.1016/j.scitotenv.2015.01.046

Alexander, J., Knopp, G., Dötsch, A., Wieland, A., &Schwartz, T. (2016). Ozone treatment of conditioned wastewater selects antibiotic resistance genes, opportunistic bacteria, and induce strong population shifts. *Science of the Total Environment*, *559*, 103–112. https://doi.org/10.1016/j.scitotenv.2016.03.154

Álvarez-Martínez, F.J., Barrajón-Catalán, E., & Micol, V. (2020). Tackling antibiotic resistance with compounds of natural origin: A comprehensive review. *Biomedicines 8*, 1–30. https://doi.org/10.3390/biomedicines8100405

Alves, M.S., Pereira, A., Araújo, S.M., Castro, B.B., Correia, A.C.M., & Henriques, I. (2014). Seawater is a reservoir of multi-resistant *Escherichia coli*, including strains hosting plasmid-mediated quinolones resistance and extended-spectrum beta-lactamases genes. *Frontiers in Microbiology*, *5*, 1–10. https://doi.org/10.3389/fmicb.2014.00426

Aminov, R.I. (2010). A brief history of the antibiotic era: Lessons learned and challenges for the future. *Frontiers in Microbiology*, *1*, 1–7. https://doi.org/10.3389/fmicb.2010.00134

An, X.L., Su, J.Q., Li, B., Ouyang, W.Y., Zhao, Y., Chen, Q.L., Cui, L., Chen, H., Gillings,

M.R., Zhang, T., & Zhu, Y.G. (2018). Tracking antibiotic resistome during wastewater treatment using high throughput quantitative PCR. *Environment International*, *117*, 146–153. https://doi.org/10.1016/j.envint.2018.05.011

Andersson, D.I., & Hughes, D. (2012). Evolution of antibiotic resistance at non-lethal drug concentrations. *Drug Resistance Updates, 15*, 162–172. https://doi.org/10.1016/j.drup.2012.03.005

Andersson, D.I., & Hughes, D. (2011). Persistence of antibiotic resistance in bacterial populations. *FEMS Microbiology Reviews*, *35*, 901–911. https://doi.org/10.1111/j.1574-6976.2011.00289.x

Baquero, F. (2012). Metagenomic epidemiology: A public health need for the control of antimicrobial resistance. *Clinical Microbiology and Infection*, *18*(4), 67–73. https://doi.org/10.1111/j.1469-0691.2012.03860.x

Barancheshme, F., & Munir, M. (2018). Strategies to combat antibiotic resistance in the wastewater treatment plants. *Frontiers in Microbiology*, *8*, 1-12. https://doi.org/10.3389/fmicb.2017.02603

Becerra-Castro, C., Macedo, G., Silva, A.M.T., Manaia, C.M., & Nunes, O.C. (2016). *Proteobacteria* become predominant during regrowth after water disinfection. *Science of the Total Environment*, *573*, 313–323. https://doi.org/10.1016/j.scitotenv.2016.08.054

Bengoechea, J.A., & Sa Pessoa, J. (2019). *Klebsiella pneumoniae* infection biology: Living to counteract host defences. *FEMS Microbiology Reviews, 43*(2), 123–144. https://doi.org/10.1093/femsre/fuy043

Bengtsson-Palme, J., Hammarén, R., Pal, C., Östman, M., Björlenius, B., Flach, C.F., Fick, J., Kristiansson, E., Tysklind, M., & Larsson, D.G.J. (2016). Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. *Science of the Total Environment*, *572*, 697–712. https://doi.org/10.1016/j.scitotenv.2016.06.228

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M.N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., & Martinez, J.L. (2015). Tackling antibiotic resistance: The environmental framework. *Nature Reviews Microbiology*, *13*(5), 310–317. https://doi.org/10.1038/nrmicro3439

Berman, J.J., (2012). Chapter 7: Gamma Proteobacteria. *In* Taxonomic Guide to Infectious Diseases. 37–47. https://doi.org/10.1016/b978-0-12-415895-5.00007-6

Biswal, B.K., Mazza, A., Masson, L., Gehr, R., & Frigon, D. (2014). Impact of wastewater treatment processes on antimicrobial resistance genes and their co-occurrence with virulence genes in *Escherichia coli*. *Water Research*, *50*, 245–253. https://doi.org/10.1016/j.watres.2013.11.047

Blair, J.M.A., Webber, M.A., Baylay, A.J., Ogbolu, D.O., & Piddock, L.J.V. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, *13*(1), 42–51. https://doi.org/10.1038/nrmicro3380

Böckelmann, U., Dörries, H.H., Ayuso-Gabella, M.N., De Marçay, M.S., Tandoi, V., Levantesi, C., Masciopinto, C., Houtte, E. Van, Szewzyk, U., Wintgens, T., Grohmann, E. (2009). Quantitative PCR monitoring of antibiotic resistance genes and bacterial pathogens in three european artificial groundwater recharge systems. *Applied and Environmental Microbiology*, *75*(1), 154–163. https://doi.org/10.1128/AEM.01649-08

Bodilis, J., Denet, E., Brothier, E., Graindorge, A., Favre-Bonté, S., & Nazaret, S. (2018). Comparative genomics of environmental and clinical *Burkholderia cenocepacia* strains closely related to the highly transmissible epidemic ET12 lineage. *Frontiers in Microbiology*, *9*, 1–12. https://doi.org/10.3389/fmicb.2018.00383

Boucher, H.W., Talbot, G.H., Bradley, J.S., Edwards, J.E., Gilbert, D., Rice, L.B., Scheld, M., Spellberg, B., & Bartlett, J. (2009). Bad bugs, no drugs: No ESKAPE! An update from the Infectious Diseases Society of America. *Clinical Infectious Diseases*, *48*(1), 1–12. https://doi.org/10.1086/595011

Brenner, D.J., & Farmer, J.J. (2015). *Enterobacteriaceae*. *In* Bergey's Manual of Systematics of Archaea and Bacteria 1–24. https://doi.org/10.1002/9781118960608.fbm00222

Burggraf, S., & Olgemöller, B. (2004). Simple technique for internal control of real-time amplification assays. *Clinical Chemistry*, *50*(5), 819–825. https://doi.org/10.1373/clinchem.2003.027961

Bürgmann, H., Frigon, D., Gaze, W. H., Manaia, C. M., Pruden, A., Singer, A. C., Smets, B. F., & Zhang, T. (2018). Water and sanitation: An essential battlefront in the war on antimicrobial resistance. *FEMS Microbiology Ecology*, *94*(9), 1–14. https://doi.org/10.1093/femsec/fiy101

Butler, M.S., & Buss, A.D. (2006). Natural products - The future scaffolds for novel antibiotics? *Biochemical Pharmacology*, *71*(7), 919–929. https://doi.org/10.1016/j.bcp.2005.10.012

Cacace, D., Fatta-Kassinos, D., Manaia, C.M., Cytryn, E., Kreuzinger, N., Rizzo, L., Karaolia, P., Schwartz, T., Alexander, J., Merlin, C., Garelick, H., Schmitt, H., de Vries, D., Schwermer, C.U., Meric, S., Ozkal, C.B., Pons, M.N., Kneis, D., & Berendonk, T.U. (2019). Antibiotic resistance genes in treated wastewater and in the receiving water bodies: A pan-European survey of urban settings. *Water Research, 162*, 320–330. https://doi.org/10.1016/j.watres.2019.06.039

Cangelosi, G.A., & Meschke, J.S. (2014). Dead or alive: Molecular assessment of microbial viability. *Applied and Environmental Microbiology*, *80*(19), 5884–5891. https://doi.org/10.1128/AEM.01763-14

Caputo, A., Fournier, P. E., & Raoult, D. (2019). Genome and pan-genome analysis to classify emerging bacteria. *Biology Direct*, *14*(1), 1–9. https://doi.org/10.1186/s13062-019-0234-0

Cloud, J. L., Hymas, W. C., Turlak, A., Croft, A., Reischl, U., Daly, J. A., & Carroll, K. C. (2003). Description of a multiplex *Bordetella pertussis* and *Bordetella parapertussis* LightCycler® PCR assay with inhibition control. *Diagnostic Microbiology and Infectious Disease*, *46*(3), 189–195. https://doi.org/10.1016/S0732-8893(03)00045-2

CLSI. (2016). *Clinical and Laboratory Standards Institute: Performance standards for antimicrobial susceptibility testing supplement M100S*.

Conlan, S., Park, M., Deming, C., Thomas, P.J., Young, A.C., Coleman, H., Sison, C., Weingarten, R.A., Lau, A.F., Dekker, J.P., Palmore, T.N., Frank, K.M., & Segre, J.A. (2016). Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. *mBio*, *7*(3), 1–9. https://doi.org/10.1128/mBio.00742-16

Council Directive (1991). Urban Waste Water Treatment Directive 91/271/EEC of the European Parliament and of the Council concerning urban waste-water treatment. Off. J. Eur. Parliam. 34, 1–18.

D´Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., & Wright, G.D. (2011). Antibiotic resistance is ancient. *Nature*, *477*(7365), 457–461.

https://doi.org/10.1038/nature10388

Dantas, G., Sommer, M.O.A., Oluwasegun, R.D., & Church, G.M. (2008). Bacteria subsisting on antibiotics. *Science*, *320*(5872), 100-103. https://doi.org/10.1126/science.1155157

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, *74*(3), 417–433. https://doi.org/10.1128/mmbr.00016-10

Davies, J., Spiegelman, G.B., & Yim, G. (2006). The world of subinhibitory antibiotic concentrations. *Current Opinion in Microbiology*, *9*(5), 445–453. https://doi.org/10.1016/j.mib.2006.08.006

Delgado-Blas, J. F., Ovejero, C. M., David, S., Montero, N., Calero-Caceres, W., Garcillan-Barcia, M. P., de la Cruz, F., Muniesa, M., Aanensen, D. M., & Gonzalez-Zorn, B. (2021). Population genomics and antimicrobial resistance dynamics of *Escherichia coli* in wastewater and river environments. *Communications Biology*, *4*(457), 1–13. https://doi.org/10.1038/s42003-021-01949-x

Di Cesare, A., Corno, G., Manaia, C.M., & Rizzo, L. (2020). Impact of disinfection processes on bacterial community in urban wastewater: Should we rethink microbial assessment methods? *Journal of Environmental Chemical Engineering*, *8*(5), 1-12. https://doi.org/10.1016/j.jece.2020.104393

Di Cesare, A., Fontaneto, D., Doppelbauer, J., & Corno, G. (2016). Fitness and recovery of bacterial communities and antibiotic resistance genes in urban wastewaters exposed to classical disinfection treatments. *Environmental Science and Technology*, *50*(18), 10153–10161. https://doi.org/10.1021/acs.est.6b02268

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., Goldsmith, D. B., Michisaki, R., Breitbart, M., Boehm, A. B., & Chavez, F. P. (2017). Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Frontiers in Marine Science*, *4*, 1–11. https://doi.org/10.3389/fmars.2017.00314

Drechsel, P., Scott, C., Raschid-Sally, L., Redwood, M., & Bahri, A. (2010). Wastewater irrigation and health: assessing and mitigating risk in low-income countries. Earthscan, IWMI, IDRC. London. ISBN 9781844077953

Du, J., Ren, H., Geng, J., Zhang, Y., Xu, K., & Ding, L. (2014). Occurrence and abundance of tetracycline, sulfonamide resistance genes, and class 1 integron in five wastewater treatment plants. *Environmental Science and Pollution Research*, *21*(12), 7276–7284. https://doi.org/10.1007/s11356-014-2613-5

Du, L., & Liu, W. (2012). Occurrence, fate, and ecotoxicity of antibiotics in agro-ecosystems. A review. *Agronomy for Sustainable Development*, *32*(2), 309–327. https://doi.org/10.1007/s13593-011-0062-9

ECDC. (2019). Antimicrobial resistance in the EU/EEA – Annual Epidemiological Report for 2019. Stockholm.

Ekwanzala, M. D., Dewar, J. B., Kamika, I., & Momba, M. N. B. (2020). Comparative genomics of vancomycin-resistant *Enterococcus* spp. revealed common resistome determinants from hospital wastewater to aquatic environments. *Science of the Total Environment*, *719*, 137275. https://doi.org/10.1016/j.scitotenv.2020.137275

EPA, (2012). Guidelines for Water Reuse. U.S. Agency for International Development. Whashington D.C.

EUCAST. (2021). Breakpoint tables for interpretation of MICs and zone diameters. Version 11.0, 2021.

Evans, K.C., Benomar, S., Camuy-Vélez, L.A., Nasseri, E.B., Wang, X., Neuenswander, B., & Chandler, J.R. (2018). Quorum-sensing control of antibiotic resistance stabilizes cooperation in *Chromobacterium violaceum*. *The ISME Journal*, *12*(5), 1263–1272. https://doi.org/10.1038/s41396-018-0047-7

FAO (2012). Coping with water scarcity - an action framework for agriculture and food security. Rome. ISBN 978-92-5-107304-9

Faoro, H., Oliveira, W. K., Weiss, V. A., Tadra-Sfeir, M. Z., Cardoso, R. L., Balsanelli, E., Brusamarello-Santos, L. C. C., Camilios-Neto, D., Cruz, L. M., Raittz, R. T., Marques, A. C. Q., Lipuma, J., Fadel-Picheth, C. M. T., Souza, E. M., & Pedrosa, F. O. (2019). Genome comparison between clinical and environmental strains of *Herbaspirillum seropedicae* reveals a potential new emerging bacterium adapted to human hosts. *BMC Genomics*, *20*(1), 1–15. https://doi.org/10.1186/s12864-019-5982-9

Finley, R.L., Collignon, P., Larsson, D.G.J., Mcewen, S.A., Li, X.Z., Gaze, W.H., Reid-Smith,

R., Timinouni, M., Graham, D.W., & Topp, E. (2013). The scourge of antibiotic resistance: The important role of the environment. *Clinical Infectious Diseases*, *57*(5), 704–710. https://doi.org/10.1093/cid/cit355

Fortunato, G., Vaz-Moreira, I., Becerra-Castro, C., Nunes, O.C., & Manaia, C.M. (2018). A rationale for the high limits of quantification of antibiotic resistance genes in soil. *Environmental Pollution*, *243*, 1696–1703. https://doi.org/10.1016/j.envpol.2018.09.128

Fuhrimann, S., Winkler, M.S., Stalder, M., Niwagaba, C.B., Babu, M., Kabatereine, N.B., Halage, A.A., Utzinger, J., Cissé, G., & Nauta, M. (2016). Disease burden due to gastrointestinal pathogens in a wastewater system in Kampala, Uganda. *Microbial Risk Analysis*, *4*, 16–28. https://doi.org/10.1016/j.mran.2016.11.003

Gao, P., Munir, M., & Xagoraraki, I. (2012). Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional municipal wastewater treatment plant. *Science of the Total Environment*, *421–422*, 173–183. https://doi.org/10.1016/j.scitotenv.2012.01.061

Gerner-Smidt, P., Besser, J., Concepción-Acevedo, J., Folster, J.P., Huffman, J., Joseph, L.A., Kucerova, Z., Nichols, M.C., Schwensohn, C.A., & Tolar, B. (2019). Whole genome sequencing: Bridging one-health surveillance of foodborne diseases. *Frontiers in Public Health*, *7*, 1–11. https://doi.org/10.3389/fpubh.2019.00172

Gorrie, C.L., Mirceta, M., Wick, R.R., Judd, L.M., Wyres, K.L., Thomson, N.R., Strugnell, R.A., Pratt, N.F., Garlick, J.S., Watson, K.M., Hunter, P.C., McGloughlin, S.A., Spelman, D.W., Jenney, A.W.J., & Holt, K.E. (2018). Antimicrobial-resistant *Klebsiella pneumoniae* carriage and infection in specialized geriatric care wards linked to acquisition in the referring hospital. *Clinical Infectious Diseases*, *67*(2), 161–170. https://doi.org/10.1093/cid/ciy027

Grady, C.P.L., Daigger, G.T., Love, N., & Filipe, C.D.M. (2011). Biological Wastewater Treatment. CRC press, NW.

Guevarra, R. B., Magez, S., Peeters, E., Chung, M. S., Kim, K. H., & Radwanska, M. (2021). Comprehensive genomic analysis reveals virulence factors and antibiotic resistance genes in Pantoea agglomerans KM1, a potential opportunistic pathogen. *PLoS ONE*, *16*, 1–27. https://doi.org/10.1371/journal.pone.0239792

Gupta, N., & Verma, V.K. (2019). Next-Generation Sequencing and Its Application:

Empowering in Public Health Beyond Reality *In* Microbial Technology for the Welfare of Society. Arora P. (eds). 313–341. https://doi.org/10.1007/978-981-13-8844-6_15

Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S.K., Kjeldgaard, J., Kaas, R.S., Clausen, P.T.L.C., Vogt, J.K., Leekitcharoenphon, P., van de Schans, M.G.M., Zuidema, T., de Roda Husman, A.M., Rasmussen, S., Petersen, B., … Aarestrup, F.M. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-08853-3

Henze, M., van Loosdrecht, M.C., Ekama, G.A., & Brdjanovic, D. (2008). Biological wastewater treatment. IWA publishing.

Higgins, P.G., Hrenovic, J., Seifert, H., & Dekic, S. (2018). Characterization of *Acinetobacter baumannii* from water and sludge line of secondary wastewater treatment plant. Water Research, *140*, 261–267. https://doi.org/10.1016/j.watres.2018.04.057

Hiller, C.X., Hübner, U., Fajnorova, S., Schwartz, T., & Drewes, J.E. (2019). Antibiotic microbial resistance (AMR) removal efficiencies by conventional and advanced wastewater treatment processes: A review. *Science of the Total Environment*, *685*, 596–608. https://doi.org/10.1016/j.scitotenv.2019.05.315

Hinlo, R., Gleeson, D., Lintermans, M., & Furlan, E. (2017). Methods to maximise recovery of environmental DNA from water samples. *PLoS ONE*, *12*(6), 1–22. https://doi.org/10.1371/journal.pone.0179251

Holt, K.E., Wertheim, H., Zadoks, R.N., Baker, S., Whitehouse, C.A., Dance, D., Jenney, A., Connor, T.R., Hsu, L.Y., Severin, J., Brisse, S., Cao, H., Wilksch, J., Gorrie, C., Schultz, M.B., Edwards, D.J., Van Nguyen, K., Nguyen, T.V., Dao, T.T., … 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 of the United States of America*, *112*(27), E3574–E3581. https://doi.org/10.1073/pnas.1501049112

Hutinel, M., Huijbers, P.M.C., Fick, J., Åhrén, C., Larsson, D.G.J., & Flach, C.F. (2019). Population-level surveillance of antibiotic resistance in *Escherichia coli* through sewage analysis. Eurosurveillance, *24*(37), 1–11. https://doi.org/10.2807/1560-7917.ES.2019.24.37.1800497

Iredell, J., Brown, J., & Tagg, K. (2016). Antibiotic resistance in *Enterobacteriaceae*: Mechanisms and clinical implications. *BMJ*, *352*, 1-19. https://doi.org/10.1136/bmj.h6420

ISO 7899. (2000). Water quality – Detection and enumeration of intestinal Enterococci. Geneva, Switzerland

ISO 9308. (2014). Water quality – Enumeration of *Escherichia coli* and coliform bacteria. Geneva, Switzerland

Kalavrouziotis, I.K., Kokkinos, P., Oron, G., Fatone, F., Bolzonella, D., Vatyliotou, M., Fatta-Kassinos, D., Koukoulakis, P.H., & Varnavas, S.P. (2015). Current status in wastewater treatment, reuse and research in some mediterranean countries. *Desalination and Water Treatment, 53*(8), 2015–2030. https://doi.org/10.1080/19443994.2013.860632

Kim, J., Lim, J., & Lee, C. (2013). Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. *Biotechnology Advances*, *31*(8), 1358–1373. https://doi.org/10.1016/j.biotechadv.2013.05.010

Klein, E.Y., Van Boeckel, T.P., Martinez, E.M., Pant, S., Gandra, S., Levin, S.A., Goossens, H., & Laxminarayan, R. (2018). Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(15), E3463–E3470. https://doi.org/10.1073/pnas.1717295115

Larsson, D.G.J., Andremont, A., Bengtsson-Palme, J., Brandt, K.K., de Roda Husman, A.M., Fagerstedt, P., Fick, J., Flach, C.F., Gaze, W.H., Kuroda, M., Kvint, K., Laxminarayan, R., Manaia, C.M., Nielsen, K.M., Plant, L., Ploy, M.C., Segovia, C., Simonet, P., Smalla, K., … Wernersson, A.S. (2018). Critical knowledge gaps and research needs related to the environmental dimensions of antibiotic resistance. *Environment International*, *117*, 132–138. https://doi.org/10.1016/j.envint.2018.04.041

Leite, J.P., Duarte, M., Paiva, A.M., Ferreira-Da-Silva, F., Matias, P.M., Nunes, O.C., & Gales, L. (2015). Structure-guided engineering of molinate hydrolase for the degradation of thiocarbamate pesticides. *PLoS One, 10*(4), 1–18. https://doi.org/10.1371/journal.pone.0123430

Li, A. D., Metch, J. W., Wang, Y., Garner, E., Zhang, A. N., Riquelme, M. V., Vikesland, P. J.,

Pruden, A., & Zhang, T. (2018). Effects of sample preservation and DNA extraction on enumeration of antibiotic resistance genes in wastewater. *FEMS Microbiology Ecology*, *94*(2), 1–11. https://doi.org/10.1093/femsec/fix189

Limayem, A., Wasson, S., Mehta, M., Pokhrel, A.R., Patil, S., Nguyen, M., Chen, J., & Nayak, B. (2019). High-throughput detection of bacterial community and its drug-resistance profiling from local reclaimed wastewater plants. *Frontiers in Cellular and Infection Microbiology*, *9*, 1-18. https://doi.org/10.3389/fcimb.2019.00303

Liotti, F.M., Posteraro, B., Mannu, F., Carta, F., Pantaleo, A., De Angelis, G., Menchinelli, G., Spanu, T., Fiori, P.L., Turrini, F., & Sanguinetti, M. (2019). Development of a multiplex PCR platform for the rapid detection of bacteria, antibiotic resistance, and *Candida* in Human blood samples. *Frontiers in Cellular and Infection Microbiology*, *9*, 1–12. https://doi.org/10.3389/fcimb.2019.00389

Lira, F., Vaz-Moreira, I., Tamames, J., Manaia, C.M., & Martínez, J.L. (2020). Metagenomic analysis of an urban resistome before and after wastewater treatment. *Scientific Reports, 10*(1), 1–9. https://doi.org/10.1038/s41598-020-65031-y

Long, S. W., Lindson, S. E., Saavedra, M. O., Cantu, C., Davis, J. J., Brettin, T., & Olsen, R. J. (2017). Whole-genome sequencing of human clinical *Klebsiella pneumoniae* isolates reveals misidentification and misunderstandings of *Klebsiella pneumoniae*, *Klebsiella variicola*, and *Klebsiella quasipneumoniae*. *mSphere*, *2*(4), 1–15.

Manaia, C.M. (2017). Assessing the risk of antibiotic resistance transmission from the environment to humans: Non-direct proportionality between abundance and risk. *Trends in Microbiology, 25*(3), 173–181. https://doi.org/10.1016/j.tim.2016.11.014

Manaia, C.M., Macedo, G., Fatta-Kassinos, D., & Nunes, O.C. (2016). Antibiotic resistance in urban aquatic environments: can it be controlled? *Applied Microbiology and Biotechnology*, *100*(4), 1543–1557. https://doi.org/10.1007/s00253-015-7202-0

Manaia, C.M., Rocha, J., Scaccia, N., Marano, R., Radu, E., Biancullo, F., Cerqueira, F., Fortunato, G., Iakovides, I.C., Zammit, I., Kampouris, I., Vaz-Moreira, I., & Nunes, O.C. (2018). Antibiotic resistance in wastewater treatment plants: Tackling the black box. *Environment International*, *115*, 312–324. https://doi.org/10.1016/j.envint.2018.03.044

Martin, R. M., Cao, J., Wu, W., Zhao, L., Manthei, D. M., Pirani, A., Snitkin, E., Malani, P. N., Rao, K., & Bachman, M. A. (2018). Identification of pathogenicity-associated loci in

Klebsiella pneumoniae from hospitalized patients. *mSystems*, *3*(3), 1–15. https://doi.org/10.1128/msystems.00015-18

Martinez, J.L. (2009). Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environmental Pollution*, *157*(11), 2893–2902. https://doi.org/10.1016/j.envpol.2009.05.051

Mbanga, J., Amoako, D.G., Abia, A.L.K., Allam, M., Ismail, A., & Essack, S.Y. (2021). Genomic analysis of *Enterococcus* spp. isolated from a wastewater treatment plant and its associated waters in Umgungundlovu district, South Africa. *Frontiers in Microbiology*, *12*, 1-13. https://doi.org/10.3389/fmicb.2021.648454

McEwen, S.A., & Collignon, P.J. (2018). Antimicrobial Resistance: A One Health Colloquium. *Microbiology Spectrum*, *6*(2), 1–26. https://doi.org/10.1128/microbiolspec.ARBA-0009-2017.Correspondence

McKinney, C.W., Dungan, R.S., Moore, A., & Leytem, A.B. (2018). Occurrence and abundance of antibiotic resistance genes in agricultural soil receiving dairy manure. *FEMS Microbiology Ecology*, *94*(3), 1–10. https://doi.org/10.1093/femsec/fiy010

McKinney, C.W., & Pruden, A. (2012). Ultraviolet disinfection of antibiotic resistant bacteria and their antibiotic resistance genes in water and wastewater. *Environmental Science and Technology*, *46*(24), 13393–13400. https://doi.org/10.1021/es303652q

Meredith, H., Srimani, J., Lee, A., Lopatkin, A., & You, L. (2015). Collective antibiotic tolerance: Mechanisms, dynamics, and intervention. *Nature Chemical Biology, 11*(3), 182–188. https://doi.org/10.1038/nchembio.1754.Collective

Morar, M., & Wright, G.D. (2010). The genomic enzymology of antibiotic resistance. *Annual Review of Genetics*, *44*, 25–51. https://doi.org/10.1146/annurev-genet-102209-163517

Munir, M., Wong, K., & Xagoraraki, I. (2011). Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Research*, *45*(2), 681–693. https://doi.org/10.1016/j.watres.2010.08.033

Murray, A.E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hugenholtz, P., Kämpfer, P., Konstantinidis, K.T., Lane, C.E., Papke, R.T., Parks, D.H., Rossello-Mora, R., Stott, M.B., Sutcliffe, I.C., Thrash, J.C., Venter, S.N., Whitman, W.B., Acinas, S.G., Amann, R.I., Anantharaman, … Reysenbach, A.L. (2020). Roadmap for naming uncultivated

Archaea and Bacteria. *Nature Microbiology*, 5(8), 987–994. https://doi.org/10.1038/s41564-020-0733-x

Nakano, S., Fujisawa, T., Ito, Y., Chang, B., Matsumura, Y., Yamamoto, M., Nagao, M., Suga, S., Ohnishi, M., & Ichiyama, S. (2018). Spread of meropenem-resistant *Streptococcus pneumonia* serotype 15A-ST63 clone in Japan, 2012-2014. *Emerging Infectious Diseases*, *24*(2), 275–283. https://doi.org/10.3201/eid2402.171268

Narciso-da-Rocha, C., & Manaia, C.M. (2017). The influence of the autochthonous wastewater microbiota and gene host on the fate of invasive antibiotic resistance genes. *Science of the Total Environment, 575*, 932–940. https://doi.org/10.1016/j.scitotenv.2016.09.157

Narciso-da-Rocha, C., Rocha, J., Vaz-Moreira, I., Lira, F., Tamames, J., Henriques, I., Martinez, J.L., & Manaia, C.M. (2018). Bacterial lineages putatively associated with the dissemination of antibiotic resistance genes in a full-scale urban wastewater treatment plant. *Environment International*, *118*, 179–188. https://doi.org/10.1016/j.envint.2018.05.040

Narciso-Da-Rocha, C., Varela, A.R., Schwartz, T., Nunes, O.C., & Manaia, C.M. (2014). *bla*$_{TEM}$ and *vanA* as indicator genes of antibiotic resistance contamination in a hospital-urban wastewater treatment plant system. *Journal of Global Antimicrobial Resistance*, *2*(4), 309–315. https://doi.org/10.1016/j.jgar.2014.10.001

Navon-Venezia, S., Kondratyeva, K., & Carattoli, A. (2017). *Klebsiella pneumoniae*: A major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiology Reviews*, *41*(3), 252–275. https://doi.org/10.1093/femsre/fux013

Ng, C., Tan, B., Jiang, X.T., Gu, X., Chen, H., Schmitz, B.W., Haller, L., Charles, F.R., Zhang, T., & Gin, K. (2019). Metagenomic and resistome analysis of a full-scale municipal wastewater treatment plant in Singapore containing membrane bioreactors. *Frontiers in Microbiology*, *10*, 1–13. https://doi.org/10.3389/fmicb.2019.00172

Nocker, A., Sossa-Fernandez, P., Burr, M.D., & Camper, A.K. (2007). Use of propidium monoazide for live/dead distinction in microbial ecology*. Applied and Environmental Microbiology*, *73*(16), 5111–5117. https://doi.org/10.1128/AEM.02987-06

Novo, A., André, S., Viana, P., Nunes, O.C., & Manaia, C.M. (2013). Antibiotic resistance, Antimicrobial residues and bacterial community composition in urban wastewater.

Water Research, *47*(5), 1875–1887. https://doi.org/10.1016/j.watres.2013.01.010

Novo, A., & Manaia, C.M. (2010). Factors influencing antibiotic resistance burden in municipal wastewater treatment plants. *Applied Microbiology and Biotechnology*, *87*(3), 1157–1166. https://doi.org/10.1007/s00253-010-2583-6

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G.A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C., & Iliopoulos, I. (2015). Metagenomics: Tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and Biology Insights*, *9*, 75–88. https://doi.org/10.4137/BBI.S12462

Paczosa, M.K., & Mecsas, J. (2016). *Klebsiella pneumoniae*: going on the offense with a strong defense. *Microbiology and Molecular Biology Reviews*, *80*(3), 629–661. https://doi.org/10.1128/mmbr.00078-15

Pallares-Vega, R., Blaak, H., van der Plaats, R., de Roda Husman, A.M., Hernandez Leal, L., van Loosdrecht, M.C.M., Weissbrodt, D.G., & Schmitt, H. (2019). Determinants of presence and removal of antibiotic resistance genes during WWTP treatment: A cross-sectional study. *Water Research*, *161*, 319–328. https://doi.org/10.1016/j.watres.2019.05.100

Pärnänen, K.M.M., Narciso-Da-Rocha, C., Kneis, D., Berendonk, T.U., Cacace, D., Do, T.T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J.L., Michael, S.G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., … Manaia, C.M. (2019). Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances*, *5*(eaau9124). https://doi.org/10.1126/sciadv.aau9124

Partridge, S.R., Pilato, V. Di, Doi, Y., Feldgarden, M., Haft, D.H., Klimke, W., Kumar-Singh, S., Liu, J.H., Malhotra-Kumar, S., Prasad, A., Rossolini, G.M., Schwarz, S., Shen, J., Walsh, T., Wang, Y., & Xavier, B.B. (2018). Proposal for assignment of allele numbers for mobile colistin resistance (*mcr*) genes. *Journal of Antimicrobial Chemotherapy*, *73*(10), 2625–2630. https://doi.org/10.1093/jac/dky262

Paul, D., Chakraborty, R., & Mandal, S.M. (2019). Biocides and health-care agents are more than just antibiotics: Inducing cross to co-resistance in microbes. *Ecotoxicology and Environmental Safety*, *174*, 601–610. https://doi.org/10.1016/j.ecoenv.2019.02.083

Pedrero, F., Kalavrouziotis, I., Alarcón, J.J., Koukoulakis, P., & Asano, T. (2010). Use of

treated municipal wastewater in irrigated agriculture-Review of some practices in Spain and Greece. *Agricultural Water Management, 97*(9), 1233–1241. https://doi.org/10.1016/j.agwat.2010.03.003

Pruden, A., Pei, R., Storteboom, H., & Carlson, K.H. (2006). Antibiotic resistance genes as emerging contaminants: Studies in northern Colorado. *Environmental Science and Technology*, *40*(23), 7445–7450. https://doi.org/10.1021/es060413l

Qiu, Z., Yu, Y., Chen, Z., Jin, M., Yang, D., Zhao, Z., Wang, J., Shen, Z., Wang, X., Qian, D., Huang, A., Zhang, B., & Li, J.W. (2012). Nanoalumina promotes the horizontal transfer of multiresistance genes mediated by plasmids across genera. *Proceedings of the National Academy of Sciences of the United States of America, 109*(13), 4944–4949. https://doi.org/10.1073/pnas.1107254109

Quainoo, S., Coolen, J.P.M., Sacha A. F. T. van Hijum, C., Martijn A. Huynen, c W.J.G.M., Willem van Schaik, E., & Wertheim, H.F.L. (2017). Whole-genome sequencing of bacterial pathogens : the future of nosocomial. *Clinical Microbiology Reviews*, *30*(4), 1015–1064.

Quintela-Baluja, M., Abouelnaga, M., Romalde, J., Su, J.Q., Yu, Y., Gomez-Lopez, M., Smets, B., Zhu, Y.G., & Graham, D.W. (2019). Spatial ecology of a wastewater network defines the antibiotic resistance genes in downstream receiving waters. *Water Research*, *162*, 347–357. https://doi.org/10.1016/j.watres.2019.06.075

Riediger, I. N., Hoffmaster, A. R., Casanovas-Massana, A., Biondo, A. W., Ko, A. I., & Stoddard, R. A. (2016). An optimized method for quantification of pathogenic Leptospira in environmental water samples. *PLoS ONE*, *11*(8), 1–12. https://doi.org/10.1371/journal.pone.0160523

Riquelme, M. V, Garner, E., Gupta, S., Metch, J., Zhu, N., Blair, M.F., Arango-Argoty, G., Maile-Moskowitz, A., Li, A., Flach, C.-F., Aga, D.S., Nambi, I., Larsson, D.G.J., Bürgmann, H., Zhang, T., Pruden, A., & Vikesland, P.J. (2021). Wastewater based epidemiology enabled surveillance of antibiotic resistance. *medRxiv*. https://doi.org/10.1101/2021.06.01.21258164

Rizzo, L., Gernjak, W., Krzeminski, P., Malato, S., McArdell, C.S., Perez, J.A.S., Schaar, H., & Fatta-Kassinos, D. (2020). Best available technologies and treatment trains to address current challenges in urban wastewater reuse for irrigation of crops in EU countries. *Science of the Total Environment*, *710*, 136312.

https://doi.org/10.1016/j.scitotenv.2019.136312

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., Michael, I., & Fatta-Kassinos, D. (2013). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of the Total Environment*, *447*, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Rocha, J., Cacace, D., Kampouris, I., Guilloteau, H., Jäger, T., Marano, R. B. M., Karaolia, P., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Berendonk, T. U., & Schwartz, T. (2020). Inter-laboratory calibration of quantitative analyses of antibiotic resistance genes. *Journal of Environmental Chemical Engineering*, *8*(1), 102214. https://doi.org/10.1016/j.jece.2018.02.022

Rodrigues, C., Passet, V., Rakotondrasoa, A., & Brisse, S. (2018). Identification of *Klebsiella pneumoniae*, *Klebsiella quasipneumoniae*, *Klebsiella variicola* and related phylogroups by MALDI-TOF mass spectrometry. *Frontiers in Microbiology*, *9*, 1–7. https://doi.org/10.3389/fmicb.2018.03000

Rodriguez-Mozaz, S., Chamorro, S., Marti, E., Huerta, B., Gros, M., Sànchez-Melsió, A., Borrego, C.M., Barceló, D., & Balcázar, J.L. (2015). Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Research*, *69*, 234–242. https://doi.org/10.1016/j.watres.2014.11.021

Rouli, L., Merhej, V., Fournier, P. E., & Raoult, D. (2015). The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbes and New Infections*, *7*, 72–85. https://doi.org/10.1016/j.nmni.2015.06.005

Runcharoen, C., Moradigaravand, D., Blane, B., Paksanont, S., Thammachote, J., Anun, S., Parkhill, J., Chantratita, N., & Peacock, S.J. (2017). Whole genome sequencing reveals high-resolution epidemiological links between clinical and environmental *Klebsiella pneumoniae*. *Genome Medicine*, *9*(1), 1–10. https://doi.org/10.1186/s13073-017-0397-1

Segerman, B. (2020). The most frequently used sequencing technologies and assembly methods in different time segments of the bacterial surveillance and RefSeq genome databases. *Frontiers in Cellular and Infection Microbiology*, *10*, 1–7. https://doi.org/10.3389/fcimb.2020.527102

Sharma, V.K., Johnson, N., Cizmas, L., McDonald, T.J., & Kim, H. (2016). A review of the influence of treatment strategies on antibiotic resistant bacteria and antibiotic resistance genes. *Chemosphere*, *150*, 702–714. https://doi.org/10.1016/j.chemosphere.2015.12.084

Sidstedt, M., Jansson, L., Nilsson, E., Noppa, L., Forsman, M., Rådström, P., & Hedman, J. (2015). Humic substances cause fluorescence inhibition in real-time polymerase chain reaction. *Analytical Biochemistry*, *487*, 30–37. https://doi.org/10.1016/j.ab.2015.07.002

Singer, A.C., Shaw, H., Rhodes, V., & Hart, A. (2016). Review of antimicrobial resistance in the environment and its relevance to environmental regulators. *Frontiers in Microbiology*, *7*, 1–22. https://doi.org/10.3389/fmicb.2016.01728

Smith, C.J., & Osborn, A.M. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiology Ecology*, *67*(1), 6–20. https://doi.org/10.1111/j.1574-6941.2008.00629.x

Soler, N., & Forterre, P. (2020). Vesiduction: the fourth way of HGT. *Environmental Microbiology*, *22*(7), 2457–2460. https://doi.org/10.1111/1462-2920.15056

Sorg, R.A., Lin, L., van Doorn, G.S., Sorg, M., Olson, J., Nizet, V., & Veening, J.W. (2016). Collective resistance in microbial communities by intracellular antibiotic deactivation. *PLoS Biology 14*(12), 1–19. https://doi.org/10.1371/journal.pbio.2000631

Sousa, J.M., Macedo, G., Pedrosa, M., Becerra-Castro, C., Castro-Silva, S., Pereira, M.F.R., Silva, A.M.T., Nunes, O.C., & Manaia, C.M. (2017). Ozonation and UV254nm radiation for the removal of microorganisms and antibiotic resistance genes from urban wastewater. *Journal of Hazardous Materials*, *323*, 434–441. https://doi.org/10.1016/j.jhazmat.2016.03.096

Sousa, P.S., Silva, I.N., Moreira, L.M., Veríssimo, A., & Costa, J. (2018). Differences in virulence between *Legionella pneumophila* isolates from human and non-human sources determined in *Galleria mellonella* infection model. *Frontiers in Cellular and Infection Microbiology*, *8*, 1–14. https://doi.org/10.3389/fcimb.2018.00097

Storteboom, H., Arabi, M., Davis, J.G., Crimi, B., & Pruden, A. (2010). Tracking antibiotic resistance genes in the south platte river basin using molecular signatures of urban, agricultural, and pristine sources. *Environmental Science and Technology*, *44*(19), 7397–7404. https://doi.org/10.1021/es101657s

Surleac, M., Barbu, I.C., Paraschiv, S., Popa, L.I., Gheorghe, I., Marutescu, L., Popa, M., Sarbu, I., Talapan, D., Nita, M., Iancu, A.V., Arbune, M., Manole, A., Nicolescu, S., Sandulescu, O., Streinu-Cercel, A., Otelea, D., & Chifiriuc, M.C. (2020). Whole genome sequencing snapshot of multidrug resistant *Klebsiella pneumoniae* strains from hospitals and receiving wastewater treatment plants in Southern Romania. *PLoS One 15*(1), 1–17. https://doi.org/10.1371/journal.pone.0228079

Szczepanowski, R., Linke, B., Krahn, I., Gartemann, K.H., Gützkow, T., Eichler, W., Pühler, A., & Schlüter, A. (2009). Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiology 155*(7), 2306–2319. https://doi.org/10.1099/mic.0.028233-0

Thaller, M.C., Migliore, L., Marquez, C., Tapia, W., Cedeño, V., Rossolini, G.M., & Gentile, G. (2010). Tracking acquired antibiotic resistance in commensal bacteria of Galápagos land iguanas: No man, no resistance. *PLoS One 5*(2), 3–6. https://doi.org/10.1371/journal.pone.0008989

The European Parliament and the Council, 2020. Regulation (EU) 2020/741, Minimum requirements for water reuse. Off. J. Eur. Union 177, 32–55.

United Nations. (2014). Water Scarcity. In *Water Scarcity*.

Van Dorp, L., Wang, Q., Shaw, L.P., Acman, M., Brynildsrud, O.B., Eldholm, V., Wang, R., Gao, H., Yin, Y., Chen, H., Ding, C., Farrer, R.A., Didelot, X., Balloux, F., & Wang, H. (2019). Rapid phenotypic evolution in multidrug-resistant *Klebsiella pneumoniae* hospital outbreak strains. *Microbial Genomics*, *5*(4), 1–11. https://doi.org/10.1099/mgen.0.000263

Van Goethem, M., Pierneef, R., Bezuidt, O., Van De Peer, Y., Cowan, D., & Makhalanyane, T. (2018). A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils. *Microbiome 6*(40), 1–12.

Varela, A.R., André, S., Nunes, O.C., & Manaia, C.M. (2014). Insights into the relationship between antimicrobial residues and bacterial populations ina hospital-urban wastewater treatment plant system. *Water Research*, *54*, 327–336. https://doi.org/10.1016/j.watres.2014.02.003

Varela, A.R., Macedo, G.N., Nunes, O.C., & Manaia, C.M. (2015a). Genetic characterization

of fluoroquinolone resistant *Escherichia coli* from urban streams and municipal and hospital effluents. *FEMS Microbiology Ecology*, *91*(5), 1–12. https://doi.org/10.1093/femsec/fiv015

Varela, A.R., Manageiro, V., Ferreira, E., Guimarães, M.A., Da Costa, P.M., Caniça, M., & Manaia, C.M. (2015b). Molecular evidence of the close relatedness of clinical, gull and wastewater isolates of quinolone-resistant *Escherichia coli*. *Journal of Global Antimicrobial Resistance*, *3*(4), 286–289. https://doi.org/10.1016/j.jgar.2015.07.008

Vartoukian, S.R., Palmer, R.M., & Wade, W.G. (2010). Strategies for culture of "unculturable" bacteria. *FEMS Microbiology Letters*, *309*(1), 1–7. https://doi.org/10.1111/j.1574-6968.2010.02000.x

Vaz-Moreira, I., Egas, C., Nunes, O.C., & Manaia, C.M. (2013). Bacterial diversity from the source to the tap: A comparative study based on 16S rRNA gene-DGGE and culture-dependent methods. *FEMS Microbiology Ecology*, *83*(2), 361–374. https://doi.org/10.1111/1574-6941.12002

Vaz-Moreira, I., Nunes, O.C., & Manaia, C.M. (2014). Bacterial diversity and antibiotic resistance in water habitats: Searching the links with the human microbiome. *FEMS Microbiology Reviews*, *38*(4), 761–778. https://doi.org/10.1111/1574-6976.12062

Vega, N.M., & Gore, J. (2014). Collective antibiotic resistance: Mechanisms and implications. *Current Opinion in Microbiology*, *21*, 28–34. https://doi.org/10.1016/j.mib.2014.09.003

Vernikos, G., Medini, D., Riley, D.R., & Tettelin, H. (2015). Ten years of pan-genome analyses. *Current Opinion in Microbiology*, *23*, 148–154. https://doi.org/10.1016/j.mib.2014.11.016

Volkmann, H., Schwartz, T., Kirchen, S., Stofer, C., & Obst, U. (2007). Evaluation of inhibition and cross-reaction effects on real-time PCR applied to the total DNA of wastewater samples for the quantification of bacterial antibiotic resistance genes and taxon-specific targets. *Molecular and Cellular Probes*, *21*(2), 125–133. https://doi.org/10.1016/j.mcp.2006.08.009

Vredenburg, J., Varela, A.R., Hasan, B., Bertilsson, S., Olsen, B., Narciso-da-Rocha, C., Bonnedahl, J., Stedt, J., Da Costa, P.M., & Manaia, C.M. (2014). Quinolone-resistant *Escherichia coli* isolated from birds of prey in Portugal are genetically distinct from those isolated from water environments and gulls in Portugal, Spain and Sweden.

Environmental Microbiology, *16*(4), 995–1004. https://doi.org/10.1111/1462-2920.12231

Whitman, R.L., Ge, Z., Nevers, M.B., Boehm, A.B., Chern, E.C., Haugland, R.A., Lukasik, A.M., Molina, M., Przybyla-Kelly, K., Shively, D.A., White, E.M., Zepp, R.G., & Byappanahalli, M.N. (2010). Relationship and variation of qPCR and culturable enterococci estimates in ambient surface waters are predictable. *Environmental Science and Technology*, *44*(13), 5049–5054. https://doi.org/10.1021/es9028974

WHO. (2015). Global Action Plan on Antimicrobial Resistance. Geneva, Switzerland. ISBN 978 92 4 150976 3

WHO. (2014). Antimicrobial resistance: Global report on surveillance. Geneva, Switzerland. ISBN 978 92 4 156474 8.

WHO (2020). Whole-genome sequencing for surveillance of antimicrobial resistance: Global Antimicrobial Resistance and Use Surveillance System (GLASS), Geneva, Switzerland. ISBN 978-92-4-001100-7

Woodford, N., & Ellington, M.J. (2007). The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection*, *13*(1), 5–18. https://doi.org/10.1111/j.1469-0691.2006.01492.x

Wu, Y., Zaiden, N., & Cao, B. (2018). The core- and pan-genomic analyses of the genus Comamonas: From environmental adaptation to potential virulence. *Frontiers in Microbiology*, *9*, 1–12. https://doi.org/10.3389/fmicb.2018.03096

Wyres, K.L., & Holt, K.E. (2018). *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Current Opinion in Microbiology*, *45*, 131–139. https://doi.org/10.1016/j.mib.2018.04.004

Wyres, K.L., Lam, M.M.C., & Holt, K.E. (2020). Population genomics of *Klebsiella pneumoniae*. *Nature Reviews Microbiology*, *18*(6), 344–359. https://doi.org/10.1038/s41579-019-0315-1

Youenou, B., Favre-Bonté, S., Bodilis, J., Brothier, E., Dubost, A., Muller, D., & Nazaret, S. (2015). Comparative genomics of environmental and clinical *Stenotrophomonas maltophilia* strains with different antibiotic resistance profiles. *Genome Biology and Evolution*, *7*(9), 2484–2505. https://doi.org/10.1093/gbe/evv161

Yuan, Q. Bin, Guo, M.T., & Yang, J. (2015). Fate of antibiotic resistant bacteria and genes during wastewater chlorination: Implication for antibiotic resistance control. *PLoS One 10*(3), 1–11. https://doi.org/10.1371/journal.pone.0119403

Zhang, T., Shao, M.F., & Ye, L. (2012). 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME Journal*, *6*, 1137–1147. https://doi.org/10.1038/ismej.2011.188

Zhuang, Y., Ren, H., Geng, J., Zhang, Yingying, Zhang, Yan, Ding, L., & Xu, K. (2015). Inactivation of antibiotic resistance genes in municipal wastewater by chlorination, ultraviolet, and ozonation disinfection. *Environmental Science and Pollution Research*, *22*(9), 7037–7044. https://doi.org/10.1007/s11356-014-3919-z