PhD Thesis
Zhuofeng Yu

The wastewater plasmidome and its derived resistome: Insight into their dynamics in the urban water systems
THE WASTEWATER PLASMIDOME AND ITS DERIVED RESISTOME: INSIGHT INTO THEIR DYNAMICS IN THE URBAN WATER SYSTEMS
The wastewater plasmidome and its derived resistome: Insights into their dynamics in the urban water systems
The wastewater plasmidome and its derived resistome: Insight into their dynamics in the urban water systems

PhD Thesis
Zhuofeng Yu

Principal supervisor: Prof. Søren Johannes Sørensen
Project collaboration partners: Asst. Prof. Joseph Nesme, Asst. Prof. Jonas Stenløkke Madsen
Affiliation: Section of Microbiology, Department of Biology, Faculty of Science, University of Copenhagen
Submitted: 16 May 2022
Note: This thesis has been submitted to the PhD School of Faculty of Science, University of Copenhagen
The Covers

Cover page

Top figure: Schematic of the mechanisms for plasmids mobility.

1). The red-orange block: conjugation (illustrated by the yellow-purple coloured plasmid), conjugative mobilisation (illustrated by the blue coloured plasmid; specifically, the black cross stands for the mobilisation of the blue coloured mobilisable plasmid by the yellow-purple coloured conjugative plasmid), and conduction (illustrated by the dark green coloured plasmid).

2). The violet block: plasmid transduction (illustrated by the brown-beige coloured plasmid and grey coloured bacteriophage).

3). The cyan block: plasmid transformation (illustrated by the brown-beige coloured plasmid) and vesiduction (illustrated by the pink and light green coloured plasmids, together with the cyan extracellular vesicles).

Note: The concept of this figure was inspired by Rodriguez-Beltrán J. et al., Nat Rev Microbiol., 2021.

Bottom figure: Flow of the wastewater-derived microbiome and associated mobilome (plasmidome).

Note: The khaki-coloured water flow represents wastewater flow. Several bacteria that are active in horizontal gene transfer of plasmids are amplified on the top layer of the wastewater flow for visualization. This illustration doesn’t represent the actual sizes of the displayed microbial species.

Back cover

Diagram of construction for plasmid sequences by collecting retrieved reads from plasmid DNA shotgun metagenomic sequencing data.

Note: This figure was created by using licenced vectors on Adobe Stock (https://stock.adobe.com/) and drawn with Adobe Illustrator v26.3 (Adobe Inc., San Jose, CA, USA).
Everything will be okay in the end. If it is not okay, it's not the end.

John Lennon
Preface

It has been four years since I first came to MME (Section of Microbiology, Department of Biology, Faculty of Science, University of Copenhagen) and worked as a PhD student. I can still recall many unforgettable scenes here and there in the offices, corridor, kitchen, and labs. I almost learned as a beginner here since I majored quite differently (Environment Engineering) for my master’s degree and had little knowledge in traditional microbiology. The reason why I suddenly jumped into this field was all by coincidence: I did my master’s thesis on municipal solid waste landfill leachates and found little was known of the occurrence and prevalence of antibiotic resistance genes in such an environment at that time, and I became interested in the topic of antibiotic resistance since then. I thought I was fairly lucky since when I was searching for PhD positions, my primary supervisor Søren J. Sørensen just had a collaborative international project on antibiotic resistance called DARWIN (see Appendix 1.1) and he found me as a potential PhD candidate for the project after several times of email exchanges. The interview with him was enjoyable and he enlightened me on many aspects I wasn’t clear about during my master’s study.

Upon my arrival in MME, I found the people here were from all over the world but got along like a big family. And they are very kind to me and patient to guide me on research and getting adapted to the science culture here. I really learnt a lot from them both related to working abilities and personality during these years regardless of whether they are newly arrived or have already left MME. This thesis would not be completed without those people around me, supporting me and caring for me, especially for the days I felt alone in the darkness, and they said you should never give up. I must appreciate my principal supervisor Søren J. Sørensen for the first and foremost, who not only gives me guidance on the project but creates the working atmosphere for me to thrive and develop critical thinking and novel ideas. I can recall I listed a large number of targets and aims in the first draft of my PhD plan, which seems quite impractical. At that point, Søren kindly talked to me and patiently guided me on how to make a realistic PhD plan, which benefited me a lot. With his accompany, I presented a comprehensive progress report of the DARWIN project in London during my first year of PhD.

In addition, Søren offered me the opportunity to go to the Royal Microscopical Society (United Kingdom) for learning how to use flow cytometry, which benefited my studies to a great extent. Secondly, tusinde tak to my ’everyday supervisor’ Joseph Nesme, we had countless discussions on my way to finishing this thesis, and he is always there to support and encourage me. I learned many things from him: having assured protocols and materials ready before any experiments, keeping notes for everything done in the lab, and the spirit of doubt. And I could still remember he and Asmus helped me modify the slides and gave suggestions and tips for the presentation till midnight before the day of my oral talk for the BAGECO conference. Also, many thanks to my co-supervisor Jonas Stenløkke Madsen, who puts up lots of key points in my PhD projects and experiments in the regular meetings and offers helpful advice on my work at all times. I am also grateful for Prof. Barth Smets at the Department of Environmental Engineering of the Technical University of Denmark (DTU), who gave me the chance for around half a year’s stay at his lab (METLab) as a change of environment during my
PhD study. At METLab, Barth and Dr. Arnaud Dechesne have helped, mentored and collaborated closely with me on top of my PhD project and the SandBAR project (see Appendix 1.2). Meanwhile, I want to appreciate Rafael Pinilla Redondo and Qin Qin Wang for their timely help during my wet lab time. They both provide me with numerous hints and tips on the experiments I had difficulties with. I also would like to express my particular thanks to my key collaborator Wanli He for his tremendous assistance in the bioinformatic analysis part of my research. He is pretty generous and ready to help when I call him for help. Besides, I need to thank Anette Hørdum Løth, Ayoe Lüchau and Jannie Christensen for their professional technician techniques whenever I ask for advice. In addition, I am grateful to Lorrie Maccario, Rodrigo Chávez, Xuanji Li, Bruce Wu, Mads Frederik Hansen, Nan Yang, Sarah Camara, Shuang Peng, Tim Evison, Trine Bøgeskov Madsen, Zhuang Gong, Shaodong Wei, Dingrong Kang, Shashank Guptanp, Xiaofen Wu, Franziska Klincke and many other MME members (no matter they have already left the section or are currently active) I couldn’t remember on my mind at the moment. I thank you all for having regardless of long or short time with me on scientific, study, leisure or just a casual chat during these years. You all have made my PhD time colourful and meaningful.

I would like to say a special thank you to my parents. They are always there and support me without any reason. I really appreciate their mountains of encouragement, advice, and love, which have kept me afloat through the final stages. At last, a big thank you to ‘me’. Thank you for making it possible all along the hard way to the final point of my PhD although it might not be the ideal one when I thought at the beginning of the study. Thank you for not giving up when you felt disappointed at the experiment results so many times. Thank you for finally being able to say out I have completed this thesis and PhD study. I think one will eventually succeed to some extent if he really pursues, seeks possible approaches and help, walks on the right track, makes attempts to the utmost, and “if you fall, don’t lose faith”. I will remember these to heart and always be motivated in the future. Of course, finishing PhD doesn’t mean an end to study, there are still enormous questions and unknown phenomena behind my research. I will make my efforts to close those doors and hopefully, bring about some novel discoveries.

Zhuofeng Yu
Copenhagen, Denmark, May 2022

Photo of the lovely MME family members
Summary

Wastewater treatment plants (WWTPs) have been recognised as reservoirs for antibiotic resistance and environments in which human pathogens commonly dwell all the time given that WWTPs constantly collect residual pharmaceutical compounds and massive amounts of human-gut associated bacteria. As recurrent antibiotic selective pressure and high cell densities favour horizontal gene transfer for dissemination of antibiotic resistance genes (ARGs) and considering the possibility of ARGs transmission from wastewater microbial communities to human pathogens, it is highly important to reveal and assess the potential barrier to prevent potential human health risks. However, the bulk of our current knowledge on the prevalence and abundance of the main driver of the horizontal gene transfer – mobile genetic elements (MGEs) is limited to those in databases and previous studies. Therefore, we have little overall information about the extent of participation and niche preference of the MGEs in a complex environment such as the wastewater ecosystem.

In this thesis, we targeted three urban water systems located in Denmark, Spain, and the United Kingdom to decipher the dynamics and fate of the total pools of MGEs, namely the mobilome. This thesis is an attempt to ascertain the general features and distinct patterns of the wastewater mobilome and its associated antibiotic resistome with the dynamic changes of wastewater treatment stages – from the sewer sources to the interior biological treatment processes in the WWTPs. Three European countries with different antibiotic use practices were considered, and sampling campaigns were constructed in two seasons (summer and winter, 2018). 16S-rRNA amplicons, high-throughput qPCR array, exogenous plasmid isolation, permissiveness test, mobilome and metagenome were produced from the recruited sewer and wastewater samples for the different types of data to shed light on the interaction of microbiome, mobilome and antibiotic resistome. Furthermore, metadata containing physicochemical, hydrology and other parameters of the related sampling campaigns was utilized for linking and exploring different aspects of the dataset. The thesis is principally divided into two chapters. The Introductory chapter explains the detailed background, aims and hypothesis of the research targets. And in the Manuscript chapter, the main subjects, results, and opinions are thoroughly covered. This thesis has resulted in three manuscripts, along with several collaborations that have not been listed here, which hopefully will lead to further publications and new investigations.

Preliminarily, in a collaborative study using the same constructed samples “Extended-spectrum beta-lactamase and carbapenemase genes are substantially and sequentially reduced during conveyance and treatment of urban sewage” (DOI: 10.1021/acs.est.0c08548, Appendix 1.1), we have systematically quantified 70 extended-spectrum beta-lactamase (ESBL) and carbapenemase genes with spatiotemporal variations. This study gives us the message that the urban water systems harbour persistent clinically important ARGs, and they were correlated with the fate of distinct taxonomic groups and MGEs. In order to understand the interplay of ARGs and MGEs dynamics across the treatment compartments, we primarily paid attention to the ARGs associated MGEs, particularly plasmids. That is because plasmids are able to shape bacterial evolution as key HGT
molecular vehicles that have efficient and rapid transmission capacity, and plasmids can be involved in various HGT mechanisms such as conjugation, transduction and transformation. In Manuscript 1 “Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater”, we used a fluorescent-reporter-gene based exogenous isolation approach to capture ESBL encoding mobile determinants from sewer microbiome samples that enter an urban water system in Denmark. After screening and sequencing, we isolated one complete 73,132 Kbp circular contig and it was subsequently characterized as an IncN plasmid carrying a complete set of functional conjugative genes. Furthermore, this plasmid showed multiple antibiotic resistance with corresponding ARGs. This plasmid was then gfp tagged and tested for permissiveness retransferring to the wastewater communities. As a result, the potential host range of this plasmid was revealed. In addition, a sequence similarity search across curated plasmid repositories revealed that this IncN plasmid derives from an IncN backbone harboured by environmental and nosocomial Enterobacterial isolates. Finally, the abundance of this IncN plasmid was measured using a short-reads mapping from shotgun metagenomic datasets from four locations in three European countries. This plasmid displayed the highest relative abundance in the hospital sewers than in any other compartments across all countries.

Based on the findings of the occurrence and variance of an IncN plasmid encoding multiple ARGs in the urban water systems, we were curious about the fate of other plasmids and moreover, the plasmidome (total pool of plasmids) and its derived ARGs along with the treatment processing. In Manuscript 2 “Insights into the circular: the cryptic plasmidome and its derived antibiotic resistome in the urban water systems”, we employed a direct plasmid DNA metagenome sequencing strategy skipping the widely used transposon-aided capture and multiple displacement amplification and aiming to uncover larger-sized plasmids. We processed the raw data using our Plaspline pipeline (version: v1.1, https://github.com/Wanli-HE/Plaspline) in the circular module to generate ascertained circular-topology plasmid contigs. Through this approach, we recovered 10,942 non-redundant putative plasmid contigs representing a broad diversity of mobility (MOB) types and incompatibility (Inc) groups, of which 87% were novel plasmids compared to the PLSDB database. Plasmid size-diversity distribution benchmarked the feasibility and reliability of the methodology. We presented that the highest prevalent plasmids were the non-mobilisable ones and the plasmidome richness generally decreased with the treatment processing. Plasmid communities could be clearly clustered based on the treatment compartments, while the seasonal variations seemed not influential. Plasmids encoding ARGs were dominantly mobilisable, and the predominant plasmid-borne ARGs targeted resistance to aminoglycosides (ant, aph), beta-lactams (blaOXA, blaSCO and blaTEM) and tetracycline (tet). To our surprise, there was a common plasmid pool shared by the three countries, and a few plasmids were persistent across the different treatment stages. Finally, we showed some persistent plasmids carried ARGs, implying persistent plasmid-mediated antibiotic resistance emerged in the pan-European urban water systems.

Since the focus of Manuscript 2 was on circular-topology plasmidome and consequently, the derived resistome analysis was made exclusively on the circular plasmid contigs, the image of those small linear plasmid contigs
was lacking. Regarding Manuscript 3 “Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems”, we applied the same sequencing output from Manuscript 2 considering that our rigorous experimental methodology has recovered almost pure plasmid DNA from the sewer/wastewater samples. Accordingly, we considered the assembled circular and linear contigs from this pool were all plasmids (at least predominantly, there may be other MGEs included), while the linear contigs were not circularized due to sequencing depth or intrinsic unique genetic contexts. We removed the chromosome noise and accordingly, generated the target plasmidome in this study. We identified 225 ARGs belonging to 180 groups of ARG families in this ‘full version’ plasmidome. The plasmid resistome richness and relative abundance detected in Spanish UWS samples were significantly \( p < 0.05 \) higher than the other studied countries, which mirrors the remarkable domestic antibiotic use in Spain. And we only detected significant \( p < 0.05 \) differences in ARG risk scores between the hospital and residential sewers in Spain. Meanwhile, different sewer compartments showed a partitioning role for the resistome richness and abundance distributions. Intriguingly, we perceived a group of shared ARGs among the three countries regardless of treatment stages. Further, \( \geq 80\% \) of ARG types in the wastewater treatment plants could be found in the sewer sources, implying these ARGs were persistent in the urban water systems. Correlations analyses were performed for linking the plasmidome derived resistome, microbiome, and the total resistome (documented by a high-throughput qPCR array). Overall, Manuscript 3 demonstrates the important role of plasmidome derived resistome in the total antibiotic resistome, and such plasmidome resistome was shaped by geographic-regional and treatment-sectional variations in the urban water systems.
Dansk Resumé

Spildevandsrensningsanlæg er blevet erkendt som reservoir for antibiotikaresistens og som et miljø, hvor menneskelige patogener normalt altid findes, da rensningsanlæg konstant opsamler farmaceutiske produkter og enorme mængder af bakterier fra den humane tarm. Da det selektive antibiotika tryk og høje celletæthed favoriserer horisontal genoverførsel til spredning af antibiotika resistensgener (ARG'er), og i betragtning af muligheden for ARG'er-overførsel fra det mikrobielle spildevandssamfund til menneskelige patogener, er det yderst vigtigt at finde en måde til at forhindre disse potentielle sundhedsrisici for mennesker. Imidlertid er hovedparten af vores nuværende viden om udbredelsen og hyppigheden af “the main driver” af den horisontale genoverførsel - mobile genetiske elementer (MGE'er) begrænset til dem i databaser og i tidligere undersøgelser. Derfor har vi kun få overordnede oplysninger om omfanget og nichepræferencen for MGE'erne i et komplekt miljø som spildevandsøkosystemet.

I denne afhandling målrettede vi undersøgelserne til tre bynære rensningsanlæg i Danmark, Spanien og Storbritannien for at kunne afdække dynamikken og skæbnen af de samlede puljer af MGE'er, nemlig mobilomen i disse anlæg. Det er et forsøg på at undersøge de generelle træk og distinkte mønstre af spildevands mobilomet og dets associerede antibiotika resistom med de dynamiske ændringer i de forskellige stadier i et rensningsanlæg- fra kloak til de interne biologiske behandlingsprocesser i renseanlæggene. Tre europæiske lande med forskellige antibiotikapraksis blev valgt, og prøveudtagningskampagner blev foretaget i to forskellige årstider (sommer og vinter, 2018). 16S-rRNA-amplikoner, high-throughput qPCR-array, exogen plasmidisolering, permissivitetstest, mobilom og metagenom blev udført med kloak- og spildevandsprøverne for at opnå forskellige typer data, der kan belyse interaktionen mellem mikrobiom, mobilom og antibiotikaresistom. Desuden blev metadata indeholdende fysisk-kemiske, hydrologi og andre parametre fra de relaterede prøvetagningskampagner undersøgt for at opnå og udforske forskellige aspekter af datasættet. Afhandlingen er hovedsageligt opdelt i to kapitler. Det indledende kapitel forklarer den detaljerede baggrund, mål og hypotser for forskningen. Og i manuskriptkapitlet er hovedemnerne, resultaterne og diskussionen grundigt dækket. Dette arbejde har resulteret i tre manuskripters samt flere samarbejder, der ikke er nævnt her, og som forhåbentlig vil føre til yderligere publikationer og nye undersøgelser.

Foreløbig har vi i en samarbejdsundersøgelse, der anvender de samme prøver “Udvidet-spektrum beta-lactamase og carbapenemase gener er væsentligt og sekventielt reduceret under transport og behandling af byspildevand” (DOI: 10.1021/acs.est.0c08548), systematisk kvantificeret 70 “extended-spektrum” beta-lactamase (ESBL) og carbapenemase gener med spatiotemporale variationer. Denne undersøgelse viser, at de urbane rensningsanlæg rummer persistente, klinisk
Vigtige ARG'er, som var korreleret med forskellige taksonomiske grupper og MGE'ers skæbne. For at forstå samspillet mellem ARG'er og MGE's dynamik på tværs af rensningsanlæggets forskellige afdelinger var vi primært opmærksomme på de ARG'er associerede MGE'er, især plasmider. Det er fordi plasmider er i stand til at skabe bakteriel evolution med HGT som nøglefunktion, der har effektiv og hurtig transmissionskapacitet, og være involveret i forskellige HGT mekanismer såsom konjugation, transduktion og transformation. I Manuskript 1 “Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater” brugte vi en Escherichia coli-recipientstamme og anvendte et eksogent plasmid isolations assay til at fange ESBL-plasmider i kloakprøverne.


Udbredte plasmider var de ikke-mobiliserbare, og plasmidomet normalt faldt med rensningsprocessen. Plasmidsamfund kunne tydeligt grupperes baseret på stadier i processerne, mens sæsonvariationerne ikke syntes at have indflydelse. Plasmider, der koder for ARG’er, var fortrinsvis mobiliserbare, og havde målrettede resistens over for aminoglykosider (ANT, APH), beta-lactamer (OXA, SCO og TEM) og tetracyclin (TET). Til vores overraskelse var der den samme spildevands plasmidpulje, i de tre lande, og nogle få plasmider var vedholdende til stede i de forskellige behandlingsstadier. Yderligere viste vi, at nogle persistente plasmider bar ARG’er, hvilket antyder, at vedvarende plasmid-mediator antibiotikaresistens opstod i de paneuropæiske, urbane rensningsanlæg. Endelig viste vi interaktionen mellem spildevands plasmidomet, brugen af antibiotika i hjemmet, beta-lactamase-gener og det mikrobielle samfund.

Da fokus i Manuscript 2 var på cirkulær-topologi-plasmidom, og den afledte resistom analyse derfor udelukkende blev lavet på de cirkulære plasmid contings, manglede “the image” af disse små lineære plasmid contings. I Manuskript 3 “Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems”, brugte vi den samme sekventering som i Manuscript 2, behandlede dataene gennem både de cirkulære og lineære moduler i Plasplinen (version: v1.1) for at opnå henholdsvis cirkulære og lineære plasmid contigs (de lineære blev ikke cirkulariseret på grund af sekventerings dybde eller iboende unikke genetiske kontekst. Begge plasmid contings udgjorde target plasmidomet i manuskript 3. Vi identificerede 225 ARG’er, der tilhørte 180 grupper i spildevandets ’totale’ plasmidom. Den plasmidom afledte resistom hyppighed i Spanien var signifikant (p < 0,05) højere end i de to andre lande. Den blandede kloak plasmidom afledte resistom hyppighed var mere ens i urbant kloakvand i alle tre lande, hvilket kunne skyldes, at de var den primære kilde til ARG’er nedstrøms, og dette var i overensstemmelse med fundene præsenteret i Manuskript 2. Vi observerede, at 18 enkelt resistens-ARG’er bidrog med >60% til den samlede ARG hyppighed. Specifikt bidrog aminoglycosid-, tetracyclin-, makrolid- og phenicol antibiotica med ca. 50 %. Yderligere fandt vi en signifikant negativ korrelation (p < 0,0001) mellem den plasmidom afledte resistom rigdom og mikrobiom hyppighed og en signifikant positiv korrelation (p < 0,01) mellem den relative plasmidom afledte resistom rigdom og ARGs relative hyppighed (data fra qPCR-array). Endelig koblede vi det plasmidom afledte resistom og det mikrobielle samfund med meta data for at find nøglerollene i samspillet. Vi konkluderede, at det spildevands plasmidom afledte resistom blev udviklet af landets antibiotika forbrugsvaner, behandlingsstadier og mikrobielle samfund.
List of Manuscripts

I. Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater

Zhuofeng Yu, a,∗ Joseph Nesme, a,∗ Qinqin Wang, a Jonas Stenløkke Madsen, a Rafael Pinilla-Redondo, a Kamille Anna Dam Clasen, a Hanadi Ananbeh, b,c Asmus Kalckar Olesen, a Zhuang Gong, a Nan Yang, a Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen, a,*

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
c Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

II. Insights into the circular: the cryptic plasmidome and its derived antibiotic resistome in the urban water systems

Zhuofeng Yu, a,∗ Wanli He, a,∗ Joseph Nesme, a,* Franziska Klincke, a Jonas Stenløkke Madsen, a Witold Kot, b,c Lars Hestbjerg Hansen, b,c Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen, a,*

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
c Department of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
III. Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems

Zhuofeng Yu,⁎⁎, Wanli He,⁎⁎, Joseph Nesme,⁎⁎, Franziska Klincke,⁎ Jonas Stenløkke Madsen,⁎ Witold Kot,⁎⁎ Lars Hestbjerg Hansen,⁎⁎ Arnaud Dechesne,⁎⁎⁎ Barth Smets,⁎⁎⁎ Søren Johannes Sørensen⁎⁎

⁎Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
⁎⁎Department of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
⁎⁎ Department of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
⁎⁎⁎ Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
Table of Contents

Part 1: Introduction ........................................................................................................... 1

1. Urban water systems .................................................................................................. 2
   1.1 The concept of urban water systems .................................................................. 2
   1.2 Why urban water systems .................................................................................. 3

2. Antibiotic resistance .................................................................................................. 3
   2.1 Antibiotic resistance and ARGs ......................................................................... 3
       2.1.1 What is antibiotic resistance ...................................................................... 3
       2.1.2 Mechanisms of antibiotic resistance .......................................................... 4
   2.2 Antibiotic resistome ............................................................................................. 5
   2.3 Antibiotic resistance in the wastewater environment ......................................... 5

3. Mobile genetic elements ............................................................................................ 6
   3.1 Categories of MGEs ............................................................................................... 6
       3.1.1 Plasmids ....................................................................................................... 6
           3.1.1.1 Plasmid typing ....................................................................................... 7
           1) MOB typing ................................................................................................. 7
           2) Replicon typing ............................................................................................ 8
           3) Inc groups and antibiotic resistance .............................................................. 10
           3.1.1.2 Plasmid mobility .................................................................................... 12
           3.1.1.3 Plasmid maintenance, persistence, and fitness cost ................................. 13
       3.1.2 Insertion sequence elements and transposons .................................................. 14
       3.1.3 Gene cassettes and integrons ......................................................................... 14
       3.1.4 Genomic islands ............................................................................................ 15
       3.1.5 Viruses .......................................................................................................... 16
           3.1.5.1 Phages/Bacteriophages ......................................................................... 16
   3.2 Horizontal gene transfer ....................................................................................... 16
       3.2.1 Canonical HGT mechanisms ....................................................................... 16
           3.2.1.1 Conjugation ............................................................................................ 16
           3.2.1.2 Transformation ..................................................................................... 17
           3.2.1.3 Transduction ......................................................................................... 17
       3.2.2 Emerging HGT mechanisms .......................................................................... 18
           3.2.2.1 Vesiduction (outer membrane vesicles) .................................................. 18
           3.2.2.2 Gene transfer agents (DNA packaging) .................................................. 18
4. Plasmidome and mobilome ................................................................. 18
   4.1 Plasmidome .................................................................................. 18
   4.1.1 Plasmidome applications and limitations ................................. 18
   4.1.2 Plasmidome experimental methodology .................................. 19
      4.1.2.1 Traditional plasmidome approaches .................................. 19
         1) Cultural-dependent methods ............................................... 19
            (i) Selective screening and isolation .................................... 19
         2) Cultural-independent methods ............................................. 19
            (i) Exogenous plasmid isolation ......................................... 19
            (ii) Total community and plasmid DNA extraction .................. 20
            (iii) Single-cell genomics ............................................... 21
            (iv) epicPCR .................................................................... 21
            (v) Magnetic bead adsorption ........................................... 21
            (vi) Other approaches ..................................................... 21
      4.1.2.2 Modern plasmidome approaches ...................................... 22
         1) Long-reads-sequencing-based plasmidome metagenomics ........ 22
         2) Hi-C/methylomes .......................................................... 22
         3) Optical mapping ............................................................. 23
   4.2 Mobilome .................................................................................... 23
      4.2.1 Mobilome experimental methodology .................................. 23
      4.2.2 Mobilome applications and prospects ................................. 24
   4.3 Outlook for probing plasmidome/mobilome .............................. 25
5. Aims and hypothesis .................................................................... 28
   5.1 Aims ......................................................................................... 28
   5.2 Hypothesis ............................................................................... 28
6. Results and discussion ............................................................... 30
   6.1 Transfer and retransfer potential of ESBL plasmids .................... 30
   6.2 Plasmidome and its derived antibiotic resistome in the UWSs ...... 30
      6.2.1 The UWS plasmidome ....................................................... 30
      6.2.2 The UWS plasmidome derived antibiotic resistome ............... 31
7. Conclusions ................................................................................. 32
8. Perspectives .................................................................................. 33
9. References .................................................................................... 35
Part 3: Conference Presentations (abstracts and posters)........... 158

1. Conference abstracts ................................................................. 159
   1.1 DMS 2018.................................................................................. 159
   1.2 BAGECO 15.............................................................................. 160
   1.3 EDAR 5..................................................................................... 162
   1.4 DMS 2019................................................................................ 164

2. Conference posters...................................................................... 166
   2.1 DMS 2018................................................................................ 166
   2.2 EDAR 5..................................................................................... 167

Part 4: Appendix............................................................................. 168

Appendix 1: PhD project briefings..................................................... 169
Appendix 2: Sampling information of the UWSs ......................... 173
Appendix 3: List of collaborations.................................................... 174
Part 1: Introduction
1. Urban water systems

1.1 The concept of urban water systems

The treatment of urban wastewater from the residential areas and workplaces (industries, hospitals, etc.) is fundamental to ensuring public health and maintaining environmental quality. Generally, wastewater treatment plants (WWTPs) are considered the central part of the urban water systems (UWSs) with the main objective is to protect the surface waters from the adverse effects of direct wastewater discharges (oxygen-consuming organic pollution, microbiological contamination with pathogens). A complete UWS comprises the set-up of a sewage collecting and conveying system from the sewer sources (residential areas and workplaces), then the sewers with their specific pipes are transported to the WWTPs with various facilities to treat the collected wastewater and finally treated wastewaters are sanitarily discharged to the downstream rivers. In the WWTPs, wastewaters normally flow through screens, settling tanks, and biological treatment basins (tertiary treatment systems can be recruited to eliminate nitrogen and phosphorus contaminations). Afterwards, wastewaters are pumped to a secondary settler, and diverse advanced treatment processes may be used to reach the discharge standards. Within the biological treatment basins, aerobic and/or anaerobic biological processes are applied to reduce organic matters, and numerous studies have been made focusing on this compartment since it is a key step for wastewater treatment.

Figure 1 Diagram of a representative urban water system studied in this PhD project.

Some facilities and operations steps in this diagram varied in countries (e.g., biological treatment process and tertiary filter). Sampling sites were constructed at all sites (Appendix 2).
1.2 Why urban water systems

As sewers receive excreted antibiotics, antibiotic-resistant bacteria (ARB), and antibiotic resistance genes (ARGs), the UWS are potential central conduits of antibiotic resistance to and from pathogens and environmental strains, posing severe human health risks. On the other hand, WWTPs are documented as hotspots for promoting the dissemination of antibiotic resistance, which is because WWTPs maintain high microbial densities and the co-mingling of different wastes. The co-occurrence of antibiotics, biocides, metals, and microbes accelerates the horizontal gene transfer (HGT) of ARGs and facilitates the blossom of multidrug resistance (MDR). Myriad studies have made efforts to reveal the diversity and abundance of ARGs and ARB within the WWTPs, however, investigating the antibiotic resistance issue from a macroscopic perspective of the UWS (including the full life cycle of the sewer flow) appears to be a new trend and worthy of attention. Studying the fate of antibiotic resistance in UWS will give us a better understanding of the dynamics of ARGs across different treatment compartments, and possibly propose strategies and barriers to reduce or avoid the spread of ARGs during wastewater treatment.

2. Antibiotic resistance

2.1 Antibiotic resistance and ARGs

2.1.1 What is antibiotic resistance

Antibiotics are indisputably one of the most successful forms of chemotherapy developed in the 20th century. However, with the overuse and misuse of antibiotics, the antibiotic resistance crisis has been listed as one of the emerging problems by the World Health Organization. Studies have shown antibiotic resistance and ARGs are ubiquitous in the environment, while the definition of antibiotic resistance can be addressed differently based on the subject and purpose.

In the clinical field, antibiotic resistance is normally assessed by minimum inhibitory concentration (MIC) breakpoints. The likelihood of therapeutic failure in human patients is often linked with antibiotic resistance. In the epidemiological field, the European Committee on Antimicrobial Susceptibility Testing (EUCAST) has established common epidemiological cut-off values for resistance (ECOFFs). And the ECOFFs are defined as the MIC value that corresponds to the upper limit of the wild-type population of a particular species. For a given bacterial species, the determination of ECOFF breakpoints demands analysing large numbers of independent isolates to establish the normal distribution of MICs. On the other hand, the operational definition of antibiotic
resistance is species-centric, focusing on the genetic context. It is based on the pairwise comparison of a parental strain with a mutant strain or with a strain containing an ARG that has been acquired via HGT \(^{12}\). Accordingly, this strain is considered antibiotic-resistant if it has a higher MIC value for the studied antibiotic than its parental wildtype strain \(^{16,17,18,19,20}\). And this definition is the most relevant definition for genomic studies of antibiotic resistance because it enables researchers to define a specific gene that encodes the corresponding resistance phenotype \(^{16,21}\).

Thus, from the operational perspective, a gene that confers resistance to antibiotics when it is present or when it increases susceptibility to antibiotics when it is absent is called ARG \(^{22}\). But one should also pay attention to the fact that transporter genes with mutations that impair the uptake of antibiotics are not ARGs because the absence of these genes makes bacteria more resistant \(^{12}\). Similarly, target genes with mutations that impair the mechanism of action of an antibiotic are not ARGs unless they can be transferred and dominantly expressed over the wild-type susceptible allele in a recipient host. Other specific situations such as involving the global gene regulator MarA also require notice \(^{23}\).

### 2.1.2 Mechanisms of antibiotic resistance

Antibiotics can be classified into five main groups according to their mode of action \(^{6,24}\):

1. cell wall synthesis inhibition (e.g., vancomycin, cephalosporins, beta-lactams, bacitracin);
2. protein synthesis inhibition (e.g., aminoglycosides, chloramphenicol, tetracycline, linezolid);
3. nucleic acid synthesis inhibition (e.g., rifampin, metronidazole, quinolones, fluoroquinolones);
4. antimetabolites (e.g., trimethoprim, dapsone, sulphonamide);
5. cell membrane disintegration (e.g., polymyxin, daptomycin).

And bacteria have developed four major types of resistance mechanisms against antibiotics:

1. efflux pumps that effectively excrete antibiotics from the cell \(^{25}\);
2. inactivation of antibiotics occurs when the activity of the antibiotic substance is directly hindered by hydrolysis, or by conversion of functional groups etc \(^{26,27}\);
3. target by-pass: e.g., creating new pathways to circumvent the originally targeted enzyme, overproduction of the target compound \(^{28}\), structural changes in the cell wall \(^{29}\), and prevention of the antibiotic to bind to its target \(^{24}\);
4. target modification: occurs through modification of the antibiotic targets themselves \(^{24}\).
Remarkably, multiple antibiotic resistance mechanisms can confer resistance against one antibiotic at the same time 6, and one type of antibiotic resistance mechanism can confer resistance against more than one type of antibiotic.

2.2 Antibiotic resistome

Antibiotic resistome or in short, resistome is defined as a compilation of all the ARGs and their precursors in pathogenic and non-pathogenic bacteria communities 8.

2.3 Antibiotic resistance in the wastewater environment

In terms of engineering missions, WWTPs are exclusively designed to remove the majority of solids, organic matter and nutrients in wastewaters and ensure sanitary effluent discharging to the downstream rivers, yet not intended to reduce antibiotics or ARGs though it is arguable that WWTPs are able to remove those hazards to some degree 30, 31. Di Cesare et al. have reported the co-occurrence of antibiotics, ARGs, mobile genetic elements (MGEs) and heavy metal resistance genes (HMRGs) in the WWTPs 32. And numerous studies have shown the prevalence and dissemination of ARGs and ARB in the wastewater environment 33, 34, 35, 36, 37, 38. Furthermore, a number of correlation analyses have been made for the wastewater microbial communities and the antibiotic resistome either detected by high-throughput qPCR analysis 39, 40, 41, or combined 16S-rRNA profiling and metagenomics 42, 43, 44, 45, 46, 47. This popularity in this area is because linking the taxonomic and functional resistome profile can help probe the putative ARG carriers and the ‘superbugs’ 48. On the other side, some ‘hardcore’ researchers are interested in deciphering the core resistome by characterising the shared ARGs found in all assessed matrices, for example, resistome, plasmidome, mobilome and virulome 49. In some cases, a fraction of the wastewater resistome was solely extracted and analysed for the sake of research aims and/or significance in HGT, such as the wastewater plasmidome derived resistome 50, 51, 52, and the wastewater mobilome derived resistome 53. However, only activated sludge in the biological treatment processes of the WWTPs were studied in these studies given that the activated sludges normally shelter high biomass and potential accumulation of ARGs 45, 54.

Alternatively, studying the ARGs transfer potential, risk potential, pathway and origin in the wastewater environment can be important given that the current wastewater antibiotic resistance risk assessment and management are defective, and a comprehensive understanding of these contents is still missing 49, 55. Meanwhile, inspecting the resistome dynamics across different sections of the
UWSs uncovers the proliferation and dissemination of specific groups of ARGs, thus the transfer potential and ‘origin’ of these ARGs may be revealed.

3. Mobile genetic elements

In general, mobile genetic elements (MGEs) are freely genetic elements, moving around within a genome. Meanwhile, MGEs can be transferred from one species or replicon to another. Furthermore, MGEs can rearrange genes in the host genome, which are believed to be the agents of ‘open-source’ evolution through HGT.

3.1 Categories of MGEs

3.1.1 Plasmids

Plasmids are ubiquitous in nature, they are present in prokaryotes (bacteria and archaea) and sometimes in eukaryotic organisms such as yeast. Plasmids are a collection of functional genetic modules that are organized into a stable, self-replicating entity or ‘replicon’. Plasmids are smaller than the cellular chromosome and usually do not contain genes required for essential cellular functions. Basically, plasmids are circular extrachromosomal DNA molecules that replicate and are transmitted independently of chromosomal DNA, while linear plasmids are also known for their presence in some species, such as Escherichia coli. A plasmid normally includes essential backbones for replicative functions and a variable assortment of accessory genes (some researchers call them cargo genes). These accessory traits can be accumulated in the cell without altering the gene content of the bacterial chromosome. Generally, these accessory genes contain ARGs, HMRGs, virulence factors, and biocide resistance genes, which can be gained or lost from time to time as a trade-off of the fitness cost of the plasmid to the host even though compensatory mutations can occur.

San Millan et al. reported that plasmid encoding blaTEM-1 accelerated resistance evolution by increasing the rate of the emergence of novel TEM-1 mutations than the chromosome-carrying blaTEM-1 in Escherichia coli. This is noteworthy as the plasmid they constructed was a non-transmissible multicopy plasmid (5,369 bp). Meanwhile, Rodriguez-Beltran et al. doubted multicopy plasmids could allow bacteria to escape from fitness trade-offs during evolutionary innovation. However, the general mechanism for plasmid persistence remains an evolutionary paradox. Considering plasmids are autonomously evolving entities, researchers have worked hard on uncovering the long-term stability, determinants for persistence, changes in mobility and gene
repertoires in the plasmid communities. In this section, we primarily introduce the general classification of plasmids. Then the mobility of plasmids is briefly discussed. And lastly, we elaborate on the plasmid maintenance, persistence, and fitness cost to the host.

**Figure 2** Distribution of the different MOB families among clades.
‘F, G, I, T’ (clades above Proteobacteria) are the four types of MPF classifications found in four different monophyletic Proteobacteria groups. ‘FATA’ stands for Firmicutes, Actinobacteria, Tenericutes and Archaea. ‘FA’ stands for Firmicutes and Actinobacteria. This figure is adapted from the work of Guglielmini et al. and redrawn.

### 3.1.1.1 Plasmid typing
Plasmids can be classified in various ways (fertility F-plasmids, resistance plasmids, Col plasmids, degradative plasmids, virulence plasmids), while the genetic-traits-based grouping methods such as MOB typing, and replicon typing are more universally applied in the research filed.

**1) MOB typing**
MOB typing exploits the loci encoding plasmid mobility functions (i.e., relaxases). As a result, MOB typing only types relaxase-encoding plasmids. Relaxases are thought to be universally present
amongst plasmids that mobilise via the relaxase-\textit{in-cis} mechanism \cite{71}, whist Coluzzi \textit{et al.} found MOB\textsubscript{less} plasmids are also abundant in the database. These MOB\textsubscript{less} plasmids do not encode relaxases, and only possess \textit{oriT} for mobility or basic plasmid-associated molecular machineries for replication/recombination/repair, stability/transfer/defence, integration/excision and phage related functions \cite{65}. These MOB\textsubscript{less} plasmids are often ignored in observation as a lack of reports for their mobilisation and/or conjugation and neglected \textit{in silico} analysis since they are difficult to classify using contemporary bioinformatic tools.

Relaxase homology can be distant, even amongst plasmids of the same MOB type \cite{70}. Currently, MOB types are usually assigned using profile-based searches such as PSI-BLAST. In addition, \textit{in silico} analysis of MOB types is also widely used, which probes each relaxase family \cite{67,70,72}. Six probes (MOBV, MOBO, MOBP, MOBH, MOBF, MOBC) have been used to detect relaxases of Gammaproteobacterial plasmids, while another two probes (MOBB and MOBT) have been used to detect relaxases in other taxa \cite{67}. PSI-BLAST uncovers distant homology. As a result, MOB typing normally provides a lower resolution, but potentially more inclusive classification compared to replicon typing. Researchers have investigated and re-investigated the performance of MOB typing for publicly available complete plasmid sequences from time to time in order to understand the timely retrieved plasmid distributions \cite{73}. And Enterobacteriaceae plasmids have often been comprehensively assessed \cite{71,73,74}.

\textbf{2) Replicon typing}

Replicon typing utilizes the loci encoding plasmid replication (replicons) \cite{75,76}. Plasmid replicons include various different loci, none of which are universal across plasmids \cite{77}. Replicon types can be assigned by querying plasmids against various replicon sequences using BLASTN \cite{76}. Within the replicon typing framework, plasmids can be classified at hierarchical resolutions: for common replicon types, plasmid multilocus sequence typing (pMLST) schemes have been devised for sub-typing \cite{78,79}, whilst replicon types also belong to broader replicon families.

Most replicon families were originally defined according to plasmid incompatibility (which refers to the inability of two plasmids to coexist stably over a number of generations in the same bacterial cell line), and represent the genetic relatedness of replicons \cite{80}. An \textit{in silico} analysis of replication initiation protein genes has indicated that based on sequence similarity, the traditional incompatibility (Inc) groups represent phylogenetically-coherent groups \cite{76}. And the classification by replicon typing is mostly consistent with the conjugation-based scheme \cite{81}. Consequently, replicon typing is also called
rep typing or Inc grouping in some references \cite{74, 82}. However, a key limitation of replicon typing is that individual plasmids can contain multiple replicons, complicating classification; in contrast, just one relaxase is encoded on one plasmid \cite{83}. On the other hand, replicon typing provides higher resolution and detailed information on plasmid relatedness, particularly if a pMLST scheme is available \cite{84}.

**Figure 3** Distribution of different plasmid Inc types isolated from human, animal and environment across Europe, Asia and the Americas.

*Group ‘other’ includes ColE, IncB/O, IncK, IncL/M, IncN, IncP, IncR, IncT, IncU, IncW, IncX, IncY and IncZ. This figure is adapted from the work of Rozwandowicz et al. \cite{74} and redrawn.*
Enterobacteriaceae-associated plasmids are well documented with replicon typing \(^{74,76,85}\). In 2014, all publicly available clinically relevant Enterobacteriaceae plasmids were *in silico* replicon typed \(^{76}\). In 2015, Shintani *et al.* reported that 75% of publicly available Enterobacteriaceae plasmids could be replicon typed, and 44% of Gammaproteobacterial plasmids could be MOB typed \(^{85}\). Nevertheless, the ‘typeability’ of plasmids using either replicon typing or MOB typing is likely to be influenced by biases in the sequencing datasets, as well as the computational workflow. And replicon typing can also fail to identify divergent or novel replicons \(^{86}\). Thus, the most accurate typing method is to determine and characterize the full-length plasmid DNA sequence. In Enterobacteriaceae only, 28 different Inc groups were discovered \(^{85}\). With the development of the PCR-based replicon typing (PBRT), researchers speed up the typing procedure \(^{75}\). This method targets different plasmid regions that are specific for each plasmid group, such as replicon genes, iterons and RNAI. Meanwhile, bioinformatic tools simultaneously providing MOB and replicon typing information developed rapidly \(^{87}\).

3) *Inc* groups and antibiotic resistance

Since replicon typing provides a prediction of conjugation potential and given that antibiotic resistance will be potentially widely spread once harboured by the conjugative plasmids \(^{88}\), researchers found it remarkable to record different Inc groups of plasmids with their typical types of antibiotic resistance onboard \(^{74,82}\).

In general, IncF, IncI, IncA/C, IncL (previously designated IncL/M), IncN and IncH plasmids harbour the most diverse types of ARGs \(^{74}\). In addition, IncP plasmids were reported to carry ARGs conferring resistance to ESBLs \(^{89}\), sulphonamides \(^{90}\), aminoglycosides, tetracyclines and colistin \(^{91,92}\), which alarms for attention.
Figure 4 Distribution of genes encoding resistance to different antimicrobial classes carried by different plasmid Inc types.

Plasmid group ‘other’ includes ColE, IncB/O, IncK, IncL/M, IncN, IncP, IncR, IncT, IncU, IncW, IncX, IncY and IncZ. Genes group ‘other’ includes genes encoding resistance to: trimethoprim, chloramphenicol, florfenicol, colistin, fosfomycin. This figure is adapted from the work of Rozwandowicz et al. 74 and redrawn.
3.1.1.2 Plasmid mobility

Plasmids are the key vectors of HGT and essential tools in genetic engineering\textsuperscript{93}. To acknowledge the prevalence and evolution of the various bacterial traits (e.g., antibiotic resistance, virulence and detoxication), investigating the plasmid mobility is necessary.

**Figure 5** (A) Schematic view of the genetic constitution of transmissible plasmids. (B) Scheme of some essential interactions in the process of conjugation. (C) Distribution of conjugative, mobilisable, and non-transmissible plasmids according to plasmid size.  
*This figure is adapted from the work of Smillie et al.\textsuperscript{93} and redrawn*. ori\textsubscript{T} = an origin of transfer, T\textsubscript{4}CP = type IV coupling protein, T\textsubscript{4}SS = type IV secretion system, R = relaxase, MPF = mating pair formation.

Essentially, an origin of transfer (ori\textsubscript{T}), a relaxase and the type IV coupling protein (T\textsubscript{4}CP) are regarded as the set of mobility genes for mobilisation. Together with a membrane-associated mating
pair formation (MPF) complex providing the mating channel, one plasmid, therefore, possesses the full set of conjugation systems. The MPF is a form of a type IV secretion system (T4SS), which is among those I to IX bacterial secretion systems \(^{94, 95, 96}\). A plasmid that codes for its own set of MPF genes is considered conjugal (self-transmissible). If it requires an MPF of another genetic element present in the cell, it is regarded as mobilizable. Some other plasmids are called non-mobilisable or non-transmissible because they are neither conjugal nor mobilizable. Accordingly, conjugal, mobilisable, and non-mobilisable plasmids are classified \(^{93}\). As for their sizes, mobilisable plasmids are often < 30 Kbp, while conjugal plasmids are generally larger than mobilisable plasmids (Fig. 4). The classification of plasmids in terms of mobility shows that conjugal plasmids distribute around an average of 100 Kbp in the database, while mobilisable plasmids have a mean peak at 5 Kbp and a broad, flat, secondary peak at around 150 Kbp. Non-mobilizable plasmids show three distinctive peaks for small (~ 5 Kbp), average (~ 35 Kbp), and large (> 300 Kbp) plasmids \(^{93}\).

3.1.1.3 Plasmid maintenance, persistence, and fitness cost

During plasmid replication, plasmids must be distributed between daughter cells when division takes place. Small plasmids are usually kept at a high copy number, and they are efficiently maintained by random segregation to the daughter cells. With regard to the larger-sized plasmids, which are normally retained at a low copy number to minimize the burden on their hosts. In a non-selection pressure environment, larger-sized plasmids generally risk being outcompeted by plasmid-free counterparts. As a result, large plasmids encode functional modules such as segregational stability genes for being maintained \(^{97}\). Plasmid maintenance usually involves multimer resolution, partitioning, and post-segregational killing systems \(^{98}\).

Plasmid maintenance is important for the host vertically retains the plasmid traits, while HGT mediated by broad-host-range (particular, conjugal) plasmids is an important mechanism of antibiotic resistance dissemination in the environment \(^{88}\). Plasmids are maintained at different levels (copy numbers) in different bacteria. Plasmid can be persistent or eliminated due to is fitness cost to the host \(^{99}\). On the one side, plasmids provide bacteria with beneficial/adaptive traits. On the other side, plasmids pose a metabolic burden in the non-selection for plasmid associated traits \(^{64}\). As a consequence, some plasmids can be disappeared in the nonelection environment or remain at a low level of abundance due to competition \(^{100, 101}\). One can argue that the initial cost of the plasmid carriage limits its transmission \(^{102}\) despite plasmid fitness cost can be alleviated by compensatory mutation \(^{103}\). Accordingly, it is widely accepted that the fitness effects of plasmids play a crucial role on their interaction with the host and mobility for horizontal transmission of antibiotic resistance.
When the fitness effect of a plasmid turns into an advantage outcompeting those plasmid-free cells, plasmids can be persistent. The observation of an IncP plasmid in multiple phylotypes of Aeromonas, Enterobacteriaceae, and Pseudomonas in the serial growth experiments demonstrates some plasmids could be community-wide beneficial rather than only favouring and contributing to a single taxon. With recognising the benefit of plasmid-encoded traits, microbes may tend to maintain such plasmids in the long term. Meanwhile, a simulation study paved the way for modelling evaluation of plasmid persistence by different plasmid interactions, and the results showed the interactions between two different plasmids can favour their persistence in the microbial communities. Further, the work from Jordt et al. suggests two distinct plasmids in communities could coevolve and emerge novel MDR plasmids under antibiotic selection. Several other hypotheses have also been proposed to explain plasmid persistence, such as host-plasmid co-adaptation, plasmid hitchhiking, cross-ecotype transfer, and high plasmid transfer rates. Nonetheless, a common plasmid persistence mechanism remains mysterious.

### 3.1.2 Insertion sequence elements and transposons

Insertion sequence (IS) elements consist of one (sometimes two) transposase (\(tnp\)) gene flanked by short inverted repeats. Transposons are generally DNA elements that can move between genomic locations. The term transposon generally refers to elements that carry accessory genes, distinguishing them from IS elements.

Basically, IS don’t harbour passenger genes, whereas IS are able to transfer ARGs as part of a composite (or compound) transposon. That is a genetic context which is bounded by two copies of the same or related IS. These IS possess the capacity of movement as a single unit. On the other hand, unit transposons are historically regarded as MGEs larger than IS, enclosed in the boundary of inverted repeats, yet not a pair of IS. Unit transposons own a transposase gene and an internal passenger gene that have the possibility of carrying ARGs.

### 3.1.3 Gene cassettes and integrons

A gene cassette is a small mobile element, ~ 0.5 to 1 Kbp, consisting of a single gene (occasionally two). Gene cassettes can present in a non-replicative free circular form, which is usually discovered inserted into an integron.

In many research reports, integron is detected by the \(IntI\) gene. Nonetheless, integron also includes an \(attI\) recombination site and a promoter distinguishing from gene cassettes. Integrons can capture,
rearrange, and express gene cassettes \(^{110}\). Basically, many gene cassettes carry ARGs and the corresponding inserted integrons can be found on transposons and/or plasmids.

![Figure 6 MGEs and their possible roles in HGT.](image)

*This figure is adapted from the work of Frost et al. \(^{56}\) and redrawn.*

### 3.1.4 Genomic islands

A genomic island is a distinct region of a bacterial chromosome that has been acquired via HGT. Taking an example, integrative and conjugative elements (ICEs) are conjugative DNAs that insert into host chromosomes. On the one hand, ICEs can be inserted into the chromosome of the recipient cell after their conjugational transfer from a donor cell. On the other hand, the ICEs originally in the donor cell can be reintegrated into the chromosome thereafter \(^{109}\).

Backboned by phage-like integration and/or excision functioning system, plasmid-like conjugation and/or maintenance components, and a regulation module, ICEs can also harbour cargo genes, where ARGs may be found.
3.1.5 Viruses

3.1.5.1 Phages/Bacteriophages

Bacteriophages were widely exploited for use in molecular biology and genomics. Phages are normally characterized by single/double-stranded DNA/RNA, with a few to several 100 Kbp lengths. They are mainly constituted by some specific replication genes, hijacking machinery and packaging genes (capsid). It is well acknowledged that the phage genome can be integrated into the bacterial chromosome and replicate with it as a prophage in the process of lysogeny. However, phages can also replicate independently as circular/linear plasmids in some reports 56. A new phenotype may occur due to the carriage of a prophage 111, and further recombination with other prophages and MGEs lead to the mosaic structure of phages 112.

3.2 Horizontal gene transfer

Horizontal gene transfer (HGT) is a driving force for the innovation of genetic diversity and versatility in bacterial species. Although horizontally acquired genes tend to be more frequently shared between closely related bacterial species 113, HGT has caused an ever-increasing magnitude of ARGs to propagate from commensal and environmental species to pathogens 114, 115.

3.2.1 Canonical HGT mechanisms

3.2.1.1 Conjugation

Conjugation requires a cell to cell contact via cell surface pili or adhesins, and the DNA is transferred through the ‘conduit’ to the recipient 93. Conjugation is facilitated by the functional conjugative machinery by a set of genes on plasmids or by ICEs in the chromosome 116. Conjugation can occur among closely and distantly related bacterial species or even in dissimilar niches. Several researchers highlight this mechanism as a key factor contributing to the plasticity and evolvability of bacterial genomes and predominantly responsible for spreading ARGs in the environment 60, 82, 117.
As known, the conjugative plasmid can freely transfer between bacteria. Conjugative mobilisation helps the mobilisable plasmids to be transferred to the recipient cell by offering the conjugation machinery for the mobilisable plasmids. Non-mobilisable plasmids can be transferred by plasmid conduction (co-integration via recombination of non-mobilisable and mobilisable/conjugative plasmids, then the co-integrated plasmid is resolved after the transfer). This figure is adapted from the work of Rodríguez-Beltrán et al. and redrawn.

### 3.2.1.2 Transformation

Natural transformation happens in certain bacteria. And transformation includes the steps of uptake, integration, and functional expression of naked fragments of extracellular DNA. Specifically, plasmid transformation occurs when the microbe uptakes the free plasmid DNA from the cell outside. However, due to limited free plasmids in the environment and the physiology of the host, nature transformation is detected at low rates.

### 3.2.1.3 Transduction

Bacteriophages can transfer genes (chromosomal DNA and/or MGEs) which are beneficial to the hosts to improve their maintenance and dissemination through transduction. Specifically, plasmid transduction happens when a bacteriophage (phage) packages plasmid DNA in viral particles and infects another bacterial cell. In the recipient cell, the plasmid can be excised and re-circularized again.
3.2.2 Emerging HGT mechanisms

3.2.2.1 Vesiduction (outer membrane vesicles)
In recent years, some researchers have showed a certain group of plasmids employed specialized vesicles for spreading themselves to other cells \(^{122}\), which was detected at high rates in biofilms \(^{123}\). And vesiduction was therefore unearthed and proposed as the fourth way of HGT \(^{124}\). Meanwhile, plasmid transmission via intracellular connections mediated by nanotubes has also been suggested \(^{125}\).

Extracellular vesicles (EVs) are secreted by bacteria from the cell membrane (or outer membrane). Extracellular vesicles offer the cell a variety of physiological functions and also can harbour DNA and thus, potentially play a role in HGT \(^{124}\). Vesiduction is different from transduction given that it is not viruses related. However, the regulator and functioning factors, as well as the detailed mechanism for vesiduction call for further studies.

3.2.2.2 Gene transfer agents (DNA packaging)
Identified in *Rhodobacter capsulatus* in 1974 \(^{126}\), gene transfer agents are host-cell produced particles. These particles seemingly possess phage structures and can transfer genetic materials. von Wintersonff *et al.* have reviewed the ability of gene transfer agents to exchange genes between phyla, and they further suggested that gene transfer agents could act as ARGs’ drivers in the environment \(^{116}\).

4. Plasmidome and mobilome

4.1 Plasmidome
Plasmidome refers to the total plasmid populations from a given environment or a particular microbiome \(^{60, 127}\).

4.1.1 Plasmidome applications and limitations
Plasmidome have been studied in diverse and complex environments, such as cow rumen \(^{128}\), rat cecum \(^{129}\), swine gut \(^{130}\), human gut \(^{131, 132, 133, 134}\), brewery and insect isolates \(^{135}\), pork production system \(^{136}\), sediments \(^{137}\), surface water \(^{138}\), groundwater \(^{139}\), and extensively in wastewater or activated sludge at WWTPs \(^{48, 50, 51, 52, 140, 141, 142}\). In addition, researchers have also been intrigued of the specific plasmidome of *Acinetobacter lwofii* \(^{143}\), *Bacillus cytotoxicus* \(^{144}\), *Escherichia coli* \(^{145, 146}\), Firmicutes \(^{147}\) and *Salmonella enterica* \(^{148}\). However, the methodologies in these reports were
compromised for biased targets and a lack of needed bioinformatic tools. Many plasmidome studies only isolated a fraction of plasmidome \(^{149, 150, 151}\). Current direct shotgun metagenomics uses the community DNA, large quantities of sequencing-yield reads are discarded due to chromosomal ‘contamination’ in those untargeted sequencing approaches, which tends to be not cost-effective. One then may argue to use plasmid DNA extracts for sequencing. However, to date, the state-to-art short-reads-based plasmidome sequencing inherently brings difficulties in assembling the repetitive genomic context \(^{150}\). Consequently, modern plasmidome approaches, such as long-read sequencing, and Hi-C, have been recently utilized and benchmarked \(^{52, 53}\).

4.1.2 Plasmidome experimental methodology

4.1.2.1 Traditional plasmidome approaches

1) Cultural-dependent methods

(i) Selective screening and isolation

Cultural-dependent methods mainly use a specific antibiotic (or a selective reagent) to isolate antibiotic-resistant (targeted) microorganisms on an appropriate medium \(^{152, 153}\). Afterwards, antimicrobial susceptibility testing (AST) and MIC test may be applied to these isolates to determine their resistance profiles. Isolates then can be PBRT typed, shotgun metagenomic or whole genome sequenced \(^{152, 154, 155}\). The direct screening and isolation plasmidome method do not necessarily represent the real plasmid diversity in a specific environment since only a few microorganisms can be cultured.

2) Cultural-independent methods

(i) Exogenous plasmid isolation

Exogenous plasmid isolation employs a bacterial conjugation assay for capturing the plasmids directly from the complex samples (e.g., soil, wastewater) using a recipient bacterium (usually tagged with an ARG marker and/or a fluorescent reporter for screening) \(^{156}\). The resultant strain (i.e., transconjugant) after the conjugation assay is selected by an appropriate medium (usually supplemented with antibiotics or targeted selective reagents) \(^{145}\). The transconjugant is then AST and MIC tested and processed for plasmid isolation and sequencing (or other plasmid typing methods such as PBRT). Exogenous plasmid isolation benefits the researchers by offering the host information of the studied plasmid(s). In addition, exogenous plasmid isolation covers plasmid mobility compared to the direct screening and isolation method. However, non-mobilisable plasmids cannot be studied through this method.
(ii) Total community and plasmid DNA extraction

a. Shotgun metagenomics (short-read-based sequencing)

Shotgun metagenomics directly isolates the total community DNA for sequencing. Afterwards, it normally recruits metagenomics-assembled genome binning (MAG) for plasmid analysis in silico. This method results in relatively low sequencing depth for plasmid analysis or at a high price for reaching the required sequencing depth and inherently meets difficulties in assembling and identifying some MGEs due to repetitive sequences. Nevertheless, shotgun metagenomics is the most applied methodology for plasmidome studies.

![Plasmidome extraction diagram](image)

**Plasmidome extraction:**
- Microbial community
- DNA extraction and shearing
- Metagenomics
- TRACA
- MDA
- Sheared DNA removal

**Transposon-aided capture (TRACA):**
- Plasmidome
- Transposon insertion
- E. coli library preparation

**Multiple displacement amplification (MDA):**
- Plasmidome
- Binding of random primers
- Amplification with phi29 polymerase
- Strand displacement + primer binding

Figure 8 Workflow of the canonical plasmidome shotgun metagenomics.

*TRACA and MDA require common pre-processed metagenomic/plasmidome DNA. This figure is adapted from the work of Saak et al. 151 and redrawn.*

b. Plasmidome metagenomics (short-read-based sequencing)

In plasmidome metagenomics, plasmid DNA was primarily extracted and purified, with the option of aiding with transposon-aided capture (TRACA) or multiple displacement amplification (MDA) to
obtain sufficient plasmid DNA for sequencing\textsuperscript{159, 160}. Plasmidome shotgun metagenomics facilitates the assembly of plasmids and viruses compared to metagenomics by decreasing the complexity of the sample. However, this method still has the ordinary drawbacks in PCR amplification coverage bias, lacking host information, and biased amplification towards small plasmids once using TRAC or MDA.

(iii) Single-cell genomics
Through single-cell genomics, plasmids and host cells are feasibly linked, and it enables to distinguish of strains from the same species in a high resolution\textsuperscript{161}. The shortcoming of single-cell genomics is the running price, which is relatively high compared to other methods.

(iv) epicPCR
epicPCR, short for ‘emulsion, paired isolation and concatenation PCR’, is a method based on the premise of encapsulating individual cells in polyacrylamide beads before lysis\textsuperscript{162}. Target DNA is attached to the beads and amplified by fusion PCR. As an approach to exploring the plasmid’s host, epicPCR can generate a physical linkage of 16S-rRNA gene with plasmid associated genes\textsuperscript{151}, and this technique has been used to assign ARGs and a class I integron to their hosts in wastewater treatment plant communities\textsuperscript{163}.

(v) Magnetic bead adsorption
The concept of this methodology is originally built on the probe-and-capture strategy\textsuperscript{164, 165}, where multiple probes are designed for each ARG of interest to capture plasmid DNA from the samples that contain sequences similar to the probes by hybridization. Hybridized sequences and their flanking regions will be immobilized on Streptavidin-coated magnetic beads, retrieved with a magnet and then sequenced\textsuperscript{166}. This targeted technique provides the genetic context of the ARGs, from which we can infer a potential association between ARGs and MGEs or hosts.

(vi) Other approaches
Calderón-Franco \textit{et al.} have recently reported a method that applied chromatography to isolate and enrich extracellular DNA without causing cell lysis\textsuperscript{167}. The isolated free-floating extracellular DNA was then characterized using metagenomics. This method returns higher DNA yields compared to other plasmid isolation methods, yet a higher sample volume was required to obtain high-quality DNA (~ 1,000 mL)\textsuperscript{167}.
Ott et al. suggested linked reads technology could help the assembly of the repetitive genome features\textsuperscript{168}, and this would possibly provide a new framework for analysing the metagenomic datasets.

### 4.1.2.2 Modern plasmidome approaches

1) **Long-reads-sequencing-based plasmidome metagenomics**

Long-read sequencing method by Oxford Nanopore and PacBio single-molecule real-time (SMRT) overcomes the assembly issue of repetitive sequence structure in short-reads sequencing. These technologies generate sequencing reads of many kilobases, which inherently preserve genomic context and improve assemblies of large plasmids\textsuperscript{151}.

![Diagram A](image1.png)  
**Figure 9 (A)** Chromosome conformation capture (3C)-based metagenomics for assignment of MGEs to their hosts in microbial communities.  
**Figure 9 (B)** Methylation profile-based assignment of MGEs to their hosts in microbial communities using SMRT sequencing. 

*This figure is adapted from the work of Carr et al.\textsuperscript{169} and redrawn.*

2) **Hi-C/methylomes**

Chromosome conformation capture (3C) technologies and their based technologies, for instance, Hi-C, physically cross-links chromosomal and extrachromosomal genetic elements within a cell. This opens a brand-new avenue to identify the plasmidome and plasmid associated ARGs, together with their host associations\textsuperscript{52,170,171}. Unfortunately, current data from Hi-C based sequencing trails contain considerable noise because of spurious intercellular contacts and intracellular sharing of genetic
elements, and its requirement for sequencing depth is significantly higher than conventional metagenomics.

An alternative approach to link MGEs of interest to their host is based on analysing methylation profiles. In short, prokaryotes often encode restriction-modification defence systems, which employ restriction endonucleases that target specific DNA sequence motifs. These defence systems also encode methyltransferases to prevent self-targeting. Meanwhile, methylation profiles are often associated with MGEs and differ even in closely related strains. Accordingly, one can recognize the distinct combinations of methylation motifs and MGEs since all DNA molecules are exposed to the same restriction-modification systems within a given host and thus, receive the same methylation marks by the host. However, few plasmidome studies have been reported using methylomes.

3) Optical mapping
By adding long-range information, optical mapping eases the assembly of plasmids and MGEs repetitive genomic regions. This methodology labels MGEs’ AT- and GC-rich regions using different fluorescent probes, and then visualize individual DNA molecules in nanofluidic devices with fluorescence microscopy. It appears promising for further applications of this method in the plasmidome field since only one study was found.

4.2 Mobilome
Historically, we and others have rendered the term plasmidome to mobilome and vice versa, thus creating an ambiguity between them. Herein, to make the message clear, mobilome is defined as all MGEs within a given microbiome as suggested by Carr et al.

4.2.1 Mobilome experimental methodology
In brief, the experimental methodology for mobilome studies is simply the same as plasmidome with a focus on the total pool of MGEs rather than plasmids exclusively. With the popularization of the emerging long-read sequencing technology, mobilome metagenomic studies have gained more attention and revealed a high complexity and diversity of MGEs in human gut and wastewaters. The long-read metagenomics improves the de novo assembly and generates longer contigs and bins than those of short-reads-based assembly by distinguishing among homologous sequences. Therefore, more and more studies tend to apply Oxford Nanopore and/or SMRT in detecting those large MGEs.
with complex and repetitive regions, such as plasmids and mobilisable genetic elements (Fig. 10). And accordingly, covering a whole gene with interest is not an issue.

![Diagram showing the distribution of MGEs](image)

**Figure 10** Length of MGEs.

*Data were recorded the review from Carr et al. and plotted to show the approximate length distribution of some MGEs.*

### 4.2.2 Mobilome applications and prospects

Mobilome has been studied in various environments, for example, dairy cattle rumen, human gut, soils and wastewaters. Furthermore, great efforts have been paid in studying the mobilome of specific microbial communities/taxa of organisms: probiotics, acidophilic prokaryotes, *Acinetobacter baumannii*, *Bifidobacterium*, *Dichotomius (Luederwaldtinia) schiffleri* (Coleoptera: Scarabaeidae), ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter sp*), *Passiflora*, *Pseudomonas stutzeri*, *Saccharolobus shibatae*, *Salmonella*, *Shewanella algae* and *Vibrio spp*.

Given that multiple ARGs and HMRGs co-occurred in wastewater plasmidome studies, nowadays researchers paid much attention to the antibiotic resistance issues in mobilome. Calderón-Franco et al. reported ARGs co-localized with MGEs on genetic contigs of free-floating extracellular DNA in wastewaters. de Nies et al. showed a core group of 15 AMR categories disseminated across more than 50 sampling time points in a 1.5 years sampling campaign from a WWTP, in which bacitracin was exclusively associated with plasmids; aminoglycoside, MLS (macrolide, lincosamide and streptogramin) and sulfonamide resistance genes were primarily disseminated via plasmids;
whereas fosfomycin and peptide resistance were linked with phages 195. Since more and more mobile resistome studies arise, it is critical to notice that a comparable/universal ARG database for comparison of antibiotic resistance mobilome studies is required 196.

Linking the host for the mobilome and mobile resistome (i.e., antibiotic resistance mobilome) can be important to dissecting the dynamics of the community-wide HGT and resistance evolution 42, 47, 52, 197. With the abundance-trending correlation analysis of the mobilome and the corresponding microbial community profiling dataset, one can decipher the association between detected MGEs and their potential hosts 42, 46. However, spurious correlations might occur between these independent variables due to one causal correlation to a common divisor used during the normalisation of different datasets 170. Thus, we can hardly confirm the true lineage of a microbe carrying a specific MGE by performing such calculations. To address this issue, proximity ligation methods such as Hi-C and methylomes analysis turn out to be solutions for more trustable linkages 52. Comparative genomics, fluorescent reporter system, and single-cell fusion PCR (e.g., epicPCR) can also be alternatives based on the research interests 198.

4.3 Outlook for probing plasmidome/mobilome

For a long time, plasmidome/mobilome studies were considered exceptional approaches to explore plasmid/MGEs pools, given that metagenomics has been believed to be an inclusive and standardised method for universal studies 199. However, researchers have realized the cons brought by the untargeted metagenomics using barely short reads: the inherent prone errors in assembling some MGEs repeated regions, the insufficient depth and the poor recovery of MGEs by using MAGs 157. Consequently, direct plasmid DNA metagenome sequencing is gaining more and more fame and applications 151, 169, 199. For example, wastewater plasmidome/mobilome studies updated and modified the methodology itself over time, from metagenomics to plasmidome metagenomics 50, 51, 52, 140, 141, 142, 159, 200.

In most reports, researchers currently only target a fraction of the entire plasmidome/mobilome and are not able to comprehensively capture all HGT in microbial communities even using the most updated technologies 198, which is due to the sample processing, sequencing approaches and bioinformatic tools. Efficient approaches to probe the plasmidome/mobilome, its associated resistome and microbiome profile are under development. As we have shown here (Fig. 11), researchers have documented a number of experimental methodologies in order to classify known MGEs, investigate novel MGEs, and assign them to their host microbes. Meanwhile, we suggest applying a combined
sequencing strategy for the purpose of high accuracy and more coverage that performs short-reads-based and long-reads-based sequencing at the same time, which will promote uncovering longer unknown MGEs, for instance, plasmids and phages.

![Diagram of different plasmidome/mobilome approaches]  

**Figure 11** Different plasmidome/mobilome approaches to detecting plasmids/MGEs and/or identifying their original or potential host(s) in microbial communities. Approaches were compared and plotted based on reviews from Rios Miguel et al. 150 and Saak et al. 151. Some approaches may be assigned to an arguably wrong cluster as a lack of research documentation. However, they would be systematically studied and designated to a proper category in the future. The coloured rounded rectangle represents different degrees of the ability to measure the host of MGEs by the corresponding approach. The font colours stand for the target range of MGEs by the corresponding approach. Exog. isol. = exogenous plasmid isolation, TRACA = transposon-aided capture, MDA = multiple displacement amplification, WGS = whole-genome sequencing, Sel. screening and isol. = selective screening and isolation, Linked reads technol. = linked reads technology, n.d. = not determined.

Nowadays, many researchers use the publicly available pipelines for plasmidome/mobilome analysis 201, 202, 203. Meanwhile, diverse bioinformatic tools are currently available and exploited for assembly, prediction/identification/verification, classification and evaluation of the sequencing data 169, 204, 205.
Although the dedicated bioinformatic tools for plasmidome/mobilome analysis are not systematically dissected in this Introduction chapter, reviews and tools’ developers have evaluated and benchmarked their performance\textsuperscript{158, 169, 210, 211, 212}. With the nowadays fast-paced development environment, the tools we use may be superseded soon\textsuperscript{213}. However, machine learning seems intriguing over time\textsuperscript{214, 215, 216}, and some novel machine learning models have identified more MGE contigs and unearthed the dataset with higher resolution than common tools\textsuperscript{133}. One should be aware of there is no such golden law for using a specific tool, and a tool which displays high accuracy and coverage in one report may not perform as well in another since datasets vary from each other\textsuperscript{212, 217, 218}. Therefore, it is recommended to critically evaluate and compare tools’ performance (feasibility and reliability) for each plasmidome/mobilome study\textsuperscript{212}.

Characterising the plasmidome/mobilome with high-quality reads is always welcomed by the bioinformatics tools. Contemporary shotgun metagenomics generates highly accurate short reads and has been comprehensively employed for investigating the host-MGEs associations using MAGs\textsuperscript{45}. Given that MAGs are not omnipotent\textsuperscript{157, 170}, we consider a hybrid assembly of reads for generating complete and accurate contiguous sequences from SMRT long-reads-based sequencing and proximity ligation on short-reads-based sequencing. This complementary approach will certainly provide us with high-resolution genetic contexts and a better understanding of the integrated host-MGE-ARG associations. Nonetheless, one may not be able to perform such an approach due to the high cost, we then suggest further applications can address the following points within budget: (1) possibly construct the whole plasmid/phage contigs of larger sizes; (2) evade assembly errors; (3) analyse ARGs and MGEs with higher resolution in terms of accuracy and completeness; (4) provide insights into some underestimated genetic features\textsuperscript{53, 219}. 
5. Aims and hypothesis

The globalization of ARGs in the natural environment and nosocomial circumstances has become one of the biggest health concerns in the current century. Particularly, WWTPs are indisputable reservoirs of ARGs and are considered unique conduits between humans and environments. Given that plasmid-mediated HGT plays a critical role in disseminating ARGs in the environment, it is important to inspect the transfer potential of transmissible plasmids and their dynamics to have a better understanding of whether these mobile ARGs can be hosted by opportunistic pathogens, their fate during the wastewater treatment and should be included in One Health’s considerations.

5.1 Aims

My PhD work has been part of a European collaborative research project – DARWIN (Appendix 1.1) that aims at uncovering the ARG hosts across the entire UWSs and identifying where key HGT events occur with a surveillance goal to assess the risk of environmental exposure of ARGs. As an integrated part of this project, the overall aim of my PhD work was to investigate the pool of plasmids in the UWSs and the associated antibiotic resistance encoded by these mobile genetic elements.

Specifically, I aim to:

1. Evaluate the transfer and retransfer potential of wastewater derived ESBL encoding plasmids to an opportunistic pathogen (the enteric bacterium Escherichia coli) and the wastewater communities, respectively.

2. Reveal the pool of plasmids (plasmidome) and its derived antibiotic resistome in different stages of UWSs based on a direct plasmidome sequencing strategy and a rigorous bioinformatic method to prevent false positives.

5.2 Hypothesis

In recent years, many researchers have made attempts to elaborate on the wastewater plasmidome/mobilome and the associated resistome as indicated in the Introduction chapter. However, the methodologies in these previous reports were compromised in terms of biased targets and lack of needed bioinformatic tools. And notably, numerous studies on the association between ARGs and MGEs were made by observing the nuanced correlations in abundances. I propose using
exogenous plasmid isolation primarily to check the potential of capturing resistance plasmids in the wastewater environment and then applying a direct plasmidome approach to decipher the plasmidome and the associated resistome in high resolution in different treatment compartments of comparable UWSs in different countries.

(1) ESBL encoding plasmids enter, persist, and can propagate in the wastewater environment by transfer to environmental strains. And these strains can retransfer the plasmids into opportunistic pathogens.

(2) Plasmidome in the UWSs is predominantly constituted by relatively small-sized, non-mobilisable plasmids. Some plasmids are internationally spread, and some are persistent across treatment compartments in the UWSs. Given that three UWSs in three different countries were studied and sampling was performed in two seasons, I postulate that country, seasonal and treatment compartment variations play important roles in shaping the richness and abundance of UWS plasmidome and the associated resistome considering different selection pressure (caused by different antibiotic usage) and distinct microbial communities.
6. Results and discussion

6.1 Transfer and retransfer potential of ESBL plasmids

By applying a fluorescent-reporter-gene based exogenous isolation approach to capture ESBL encoding mobile determinants from sewer microbiome samples that enter an UWS in Denmark, I and my collaborators isolated a ~73 Kbp IncN plasmid (pDK_DARWIN) that harboured and expressed multiple ARGs after screening and sequencing. Using a dual fluorescent reporter gene system, we then showed that this plasmid can transfer into native urban water communities. We demonstrated the transfer of pDK_DARWIN to microbiome members of both the sewer (in the upstream UWS compartment) and wastewater treatment (in the downstream UWS compartment) microbiomes.

Sequence similarity search across curated plasmid repositories revealed that pDK_DARWIN derives from an IncN backbone harboured by environmental and nosocomial Enterobacterial isolates, suggesting the constantly evolving accessory region and genes rearrangement on these IncN plasmids. Furthermore, we searched for pDK_DARWIN sequence matches in wastewater metagenomes from three countries, revealing that this plasmid can be detected in all of them, with a higher relative abundance in hospital sewers compared to residential sewers. Considering the relatively low detection limits of bulk shotgun sequencing, confident detection of pDK_DARWIN sequence based on reads mapping coverage in these samples means that this plasmid is present in non-negligible amounts in these samples.

Above all, “Aim 1” is thereby achieved by this study. We obtained several ESBL encoding plasmids isolates from the hospital sewer of a UWS, which were captured by an Escherichia coli strain. Then we showed that one of these ESBL encoding plasmids was MDR encoding and could be transferred back to the wastewater communities. Through metagenome searching, this ESBL plasmid was relatively more abundant in the hospital sewer than in any other UWS compartments. “Hypothesis 1” was partially verified, further DNA extraction and 16S-rRNA sequencing for the re-transformants was in need to potentially infer the host range of this plasmid in situ.

6.2 Plasmidome and its derived antibiotic resistome in the UWSs

6.2.1 The UWS plasmidome

With a direct plasmidome wet-lab methodology and focused on the circular compilation of the plasmidome in silico, we investigated the plasmidome and the ARGs carrying plasmids in different compartments of UWSs in three European countries representing different antibiotic usage regimes.
We identified 9,538 novel sequences in a total of 10,942 recovered plasmids. Of these, 66 were typed as conjugative, 1,906 mobilisable, 8,970 non-mobilisable plasmids. The UWS plasmidome was dominated by small plasmids (≤10 Kbp) representing a broad diversity of MOB types and Inc groups. Seasonal variation was not influential. A shared collection of plasmids from different countries was detected in all treatment compartments, and plasmids could be source-tracked in the UWSs. These patterns showed plasmid persistence was a common feature in the UWSs. In addition, we discovered the abundance and richness of ARGs harbouring plasmids differed notably among different countries and treatment compartments.

6.2.2 The UWS plasmidome derived antibiotic resistome
Considering the target plasmidome in section 6.2.1 was exclusively circular-topology, we potentially lost the message from those linear contigs in the dataset. We, therefore, recruited the assembled putative plasmid contigs from both circular and linear modules through our Plaspline pipeline. After removing chromosome associated contigs, we accordingly generated the ‘full version’ plasmidome. In this plasmidome, we identified 225 ARGs belonging to 180 groups of ARG families. ARGs from aminoglycoside, tetracycline, macrolide and phenicol drug classes contributed more than half of the plasmidome resistome. The plasmid resistome richness and relative abundance detected in Spanish UWS samples were significantly (p < 0.05) higher than the other studied countries, which mirrors the remarkable domestic antibiotic use in Spain. And we only detected significant (p < 0.05) differences in ARG risk scores between the hospital and residential sewers in Spain. Meanwhile, different sewer compartments showed a partitioning role for the resistome richness and abundance distributions. We thus concluded that the plasmidome derived resistome was shaped by geographic-regional and UWS-sectional variations in the UWS environment. Intriguingly, we perceived a group of shared ARGs among the three countries regardless of treatment stages. Further, ≥80% of ARG types in the wastewater treatment plants could be found in the sewer sources, which implies these ARGs were persistent in the UWSs. Finally, correlation analysis revealed that plasmidome derived resistome was an important part of the total resistome.

Taken together, “Aim 2” is thereby realized by these two studies. We comprehensively demonstrated the general patterns (compositions, richness, abundance) of plasmidome and plasmidome derived resistome, their dynamics and influential factors along with the UWS processing by using a direct plasmidome methodology. We benchmarked both the experimental and computational analysis for the UWS plasmidome and plasmidome derived resistome. “Hypothesis 2” was partially correct, seasonal variations did not significantly affect the plasmidome and plasmidome derived resistome.
patterns. However, geographic-regional (country’s antibiotic use) and treatment compartment variations indeed took part in shaping the UWS plasmidome and plasmidome derived resistome. We were able to correlate the plasmidome derived resistome to the microbiome and total resistome with significant associations, as well as the plasmidome derived beta-lactam resistome and the total beta-lactam resistome. These indicate the plasmidome derived resistome was critical to be considered for studying the resistome in metagenomics in the future.

7. Conclusions

Through the Introduction chapter, I primarily describe the general background and importance of the research target (i.e., urban water systems, UWS) in this PhD project. Then I present the literature reviews on the research topics in this PhD project, which include antibiotic resistance, mobile genetic elements (particularly plasmids), horizontal gene transfer as well as plasmidome and mobilome. With the research questions of ‘Do UWS derived ESBL plasmids possess the horizontal transmission potential?’ and ‘What are the patterns and dynamics of the UWS plasmidome and plasmidome associated antibiotic resistome?’, I have participated in the experiment design, implemented the related laboratory experiments, and partially analyzed the sequencing data. I have learnt many techniques and expertise throughout my PhD project, from the basic bacteria culture, AST, MIC, various sample pre-treatment methods, plasmid and community DNA isolation, library preparation for the next- and third-generation sequencing approaches to the sequencing data analysis and some plotting skills. It’s an indisputably fruitful studying process and I indeed enjoy my PhD time. I will elaborate on all that I have learned and applied during this PhD project in the next Manuscript chapter. Herein, I will mainly list the conclusions I found from my PhD work.

(1) IncN plasmid is prevalent across Europe and an efficient vector capable of disseminating multiple antibiotic resistance genes (including ESBL) in the UWSs.

(2) The UWS plasmidome is dominated by cryptic plasmids, which are non-mobilizable, not ‘typeable’ under the current MOB and Inc typing scheme and previously unknown (i.e., novel) plasmids. Considering that some of these plasmids carry ARGs, are prevalent across three countries and can persist throughout the UWS compartments, we call for attention to these ‘priority’ plasmids when perform further regular surveillance and risk management in the relevant environment.
Plasmidome derived resistome plays an important part in the total resistome, and this resistome is shaped by geographic-regional and UWS-sectional variations in the UWS environment.

8. Perspectives

In this PhD work, I have aimed to shed light on the horizontal transmission of multiple ARGs encoding plasmids in the wastewater communities and the fate of the total pool of plasmids (plasmidome) together with the plasmidome derived resistome across different treatment compartments in comparable UWSs located in different European countries. This work was part of the European collaborative research project – DARWIN (Appendix 1.1), and all studies focused on the same UWSs from different sampling campaigns (Appendix 2).

I have successfully executed and realized the aims as shown above, and most of the hypotheses were verified. In addition, I have also carried out some other studies during this PhD project. For example, I have performed a direct selective screening and isolation from the wastewater activated sludge community (section 1.2 in Part3). I confirmed the presence of colistin-resistant strains in the activated sludge, and most were identified as *Pseudochrobactrum* sp. and *Ochrobactrum* sp. using full-length 16S rRNA gene sequencing. Interestingly, a *Pseudochrobactrum* isolate hosted a similar-sized IncN plasmid as pDK_DARWIN (section 6.1) also encoding complete conjugative transfer systems and multiple ARGs. Further filter mating experiments of this plasmid with a colistin-sensitive recipient strain confirmed the mobility of this IncN plasmid and its associated colistin resistance. However, the responsible determinant could not be readily identified by sequence homology as one of the already characterized mobilized colistin resistance *mcr* gene variants. Together with section 6.1, I propose that IncN plasmids have already been widely disseminated in the wastewater environment, rearranging the cargo genes, constantly evolving and spreading ARGs among different microbial communities (section 2.2 in Part3). Unfortunately, this *Pseudochrobactrum* isolate was lost in the laboratory when I was about to carry out studying the corresponding transmissible colistin resistance. This taught me a hard lesson on strain storage.

With respect to the underlying research questions in my PhD work, I suggest investing in other plasmids that have been exogenously isolated and directly isolated with screening since many of them haven’t been sequenced and studied. I expect that there might be pools of other multiple ARGs encoding plasmids and/or similar IncN plasmids in the wastewater environment.
Meanwhile, we could:

(1) Perform deep sequencing for several other UWS treatment stages, e.g., the primary settler, the secondary settler, the tertiary filter, and upstream and downstream rivers in the same constructed UWS of this work (Fig. 1), to obtain a full-frame image of the UWS plasmidome. Alternatively, I propose a combined Hi-C and methylation analysis that applies both short-reads-based and long-reads-based sequencing technologies in combination with other molecular approaches (transcriptomics, proteomics, and metabolomics) to reveal the in situ links, lifestyle, evolution and interactions of the resistome, mobilome and microbiome in the UWSs 52, 151, 171.

(2) Investigate the genetic context of the plasmids from the retrieved plasmidome dataset.
   (i) Update the database on MOB typing and replicon typing tethering with Inc groups and host range given that we observed large numbers of miscellaneous rep clusters and MOB_unknown that couldn’t be classified using the contemporary database.
   (ii) Inspect the genetic background of the plasmidome, sequence similarity with the database and ARG risk scores of the mobilisable and conjugative plasmids as well as those persistent non-mobilisable ARGs encoding plasmids.
   (iii) Explore the co-occurrence of ARGs, HMRGs, virulence factors, biocide resistance genes, and/or the emerging antifungal resistance genes 220.
   (iv) Censor the presence of other types of MGE loci on the plasmids – their open reading frames genetic distance between each other and ARGs 221.
   (v) Probe the pan-immune system, such as toxin-antitoxin, restriction enzymes and the adaptive CRISPR/Cas systems 222.

(3) Systematically benchmark the plasmidome/mobilome experimental methodologies (Fig. 11) and bioinformatic tools/workflow to evaluate their performance, pros, and cons.

Of course, some of the outlooks appear not realistic in the recent future, whilst with the rapid development of novel technologies and in silico tools, I envisage the plasmidome/mobilome studies will be routine tasks in the resistome research field and a necessity in surveillance and risk management for antibiotic resistance environmental dissemination and exposure.
9. References


Li L, Dechesne A, Madsen JS, Nesme J, Sørensen SJ, Smets BF. Plasmids persist in a microbial community by providing fitness benefit to multiple phylotypes. *ISME J* 2020; 14: 1354-1363.


Bolotin E, Hershberg R. Horizontally acquired genes are often shared between closely related bacterial species. *Front Microbiol* 2017; 8: 1536.


42.


188. Botelho J, Cazares A, Schulenburg H. The ESKAPE mobilome contributes to the spread of antimicrobial resistance and CRISPR-mediated conflict between mobile genetic elements. bioRxiv 2022; e-pub ahead of print 04 January 2022; doi:0.1101/2022.01.03.474784.


221. Razavi M, Kristiansson E, Flach CF, Larsson DGJ. The association between insertion sequences and antibiotic resistance genes. mSphere 2020; 5.

Part 2: Manuscripts
1. Manuscript 1

Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater

Zhuofeng Yu, a, * Joseph Nesme, a, * Qin Qin Wang, a Jonas Stenløkke Madsen, a Rafael Pinilla-Redondo, a Kamille Anna Dam Clasen, a Hanadi Ananbeh, b,c Asmus Kalekar Olesen, a Zhuang Gong, a Nan Yang, a Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen a, *

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
c Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
Title of the paper
Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater

Authors and affiliations
Zhuofeng Yu, a, * Joseph Nesme, a, * Qinquin Wang, a Jonas Stenløkke Madsen, a Rafael Pinilla-Redondo, a Kamille Anna Dam Clasen, a Hanadi Ananbeh, b, c Asmus Kalekar Olesen, a Zhuang Gong, a Nan Yang, a Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen a, #

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
c Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

* These authors contributed equally to this work.
# Corresponding Author.

Competing Interests
The authors declare no other competing financial interests.
1.1 Abstract

Wastewater treatment plants (WWTPs) are considered reservoirs of antibiotic resistance genes (ARGs). Given that plasmid-mediated horizontal gene transfer plays a critical role in disseminating ARGs in the environment, it is important to inspect the transfer potential of transmissible plasmids to have a better understanding of whether these mobile ARGs can be hosted by opportunistic pathogens and should be included in One Health’s considerations. In this study, we used a fluorescent-reporter-gene based exogenous isolation approach to capture extended-spectrum beta-lactamases encoding mobile determinants from sewer microbiome samples that enter an urban water system (UWS) in Denmark. After screening and sequencing, we isolated a ~73 Kbp IncN plasmid (pDK_DARWIN) that harboured and expressed multiple ARGs. Using a dual fluorescent reporter gene system, we show that this plasmid can transfer into resident urban water communities. We demonstrated the transfer of pDK_DARWIN to microbiome members of both the sewer (in the upstream UWS compartment) and wastewater treatment (in the downstream UWS compartment) microbiomes. Sequence similarity search across curated plasmid repositories revealed that pDK_DARWIN derives from an IncN backbone harbourd by environmental and nosocomial Enterobacterial isolates. Furthermore, we searched for pDK_DARWIN sequence matches in UWS metagenomes from three countries, revealing that this plasmid can be detected in all of them, with a higher relative abundance in hospital sewers compared to residential sewers. Overall, this study demonstrates that this IncN plasmid is prevalent across Europe and an efficient vector capable of disseminating multiple ARGs in the urban water systems.

1.2 Keywords

IncN
Plasmids
Antibiotic resistance
Horizontal gene transfer
Urban wastewater
1.3 Introduction

Plasmids are extrachromosomal mobile genetic elements (MGEs) capable of self-replication. Bacteria can gain and lose plasmids for the sake of survival under selection pressure and reducing fitness costs in the nonselective environment, respectively. In some cases, plasmids persist due to inherency or adaptation to environmental changes, which are attributed to the accompanying accessory genes to help the hosts gain competitive advantages over the others. These non-core accessory genes promote phenotypic plasticity and significantly increase the diversity of the pan-genome of a given species and can constitute up to 90% of it. Moreover, the accessory genes can include various resistance genes (e.g., towards antibiotics, biocides and heavy metals), virulence factors, and even the recent emerging CRISPR-Cas systems. Plasmids also can possess other traits, such as stability and mobility systems, which improve plasmid maintenance in an active cell population and facilitate the plasmid transfer to new hosts, respectively. Natural plasmid transfer as an important mechanism of horizontal gene transfer (HGT), normally requires a dedicated cadre of gene systems: the mobility (MOB) genes including oriT, a relaxase and a type IV coupling protein (T4CP), and the mating pair formation genes (the type IV secretion system, T4SS). Plasmid HGT makes it possible for plasmid-borne genes to be shared in both closely and distantly related bacterial species or even in dissimilar ecological niches. And these events have led to bacterial antibiotic resistance becoming prevalent and ubiquitous in diverse environmental settings, posing potential risks to human health.

Plasmids can be classified into various incompatibility (Inc) groups, based on replication typing. This classification is broad, including 28 Inc groups documented for Enterobacteriaceae alone. Plasmids that belong to the same Inc group cannot coexist stably in the same bacterial cell because they share, and thus compete for, components of the replication and partitioning machinery. Plasmids of the same Inc group generally share backbones with similar organizations and functions. IncF, IncI, IncA/C, IncH plasmids are the most predominant Inc groups among the human, animal and environment samples. Together with IncL and IncN, these plasmid groups are believed to carry the greatest diversity of antibiotic resistance genes (ARGs).

Bacterial resistance to beta-lactams can be caused by efflux pumps, the modification and/or reduced production of outer membrane porins, alterations of the beta-lactams targeted molecular penicillin-binding proteins, and enzymatic inactivation (i.e., beta-lactamases). Among these, the production of beta-lactamases represents the most relevant mechanism of beta-lactam resistance in Gram-negative pathogens, thus arguably constituting the biggest challenge to the use of beta-lactams. Considering that beta-lactamases are often carried by plasmids and frequently co-occur with other types of ARGs, understanding their dissemination routes and potential hosts is crucial to the call
of the “One Health” strategy by the World Health Organization. Long et al. demonstrated the co-transfer of mcr-1, mcr-3.5, blaNDM-5 and rmtB mediated by a dual-replicon IncHI2(ST3)/IncN plasmid pMCR1_025943 in Escherichia coli strains 15. And Kayama et al. suggested IncN plasmid pKPI-6 favoured dissemination of blaIMP-6 and blaCTX-M-2 in some Enterobacteriaceae members (Klebsiella pneumoniae, Klebsiella oxytoca and Escherichia coli) 16. Moreover, a ca. 50 Kbp IncN plasmid backbone with blaIMP-6 has been widely shared among Enterobacterales from 22 hospital isolates 17. These facts imply some Inc groups of plasmids can be readily harbour and spread beta-lactamases genes in the nosocomial environment.

Wastewater treatment plants (WWTPs) are regarded as a permissive facilitator in the dissemination of clinically relevant blaAmpC genes into other aquatic ecosystems, thus raising public health concerns 18. Indeed, WWTPs have been recognised to act as reservoirs of various ARGs and hotspots of HGT 19, 20, 21, 22. Even after wastewater treatment, the effluents of WWTP still retain diverse and abundant ARGs 23, 24. One wastewater metagenomic study shows that the detection frequency of extended-spectrum beta-lactamase (ESBL) genes (blaCTX-M) and carbapenemases genes (blaNDM, blaVIM, blaGES and blaOXA-48) increased during the wastewater treatment 25. And ESBL genes (blaTEM, blashV-12, blaCTX-M-1 and blaCTX-M-15) carrying and carbapenem-resistant isolates from the family of Enterobacteriaceae could be discovered in the WWTP effluent 26. Furthermore, wastewater irrigation practices have been proven to introduce ESBL genes (blaCARB-3, blaOXA-347, blaOXA-5 and blaRm3) to the agriculture field soils 27. Wastewater flow normally brings together broad-host-range plasmids (e.g., IncN, IncP, IncQ and IncW) 28. And IncA/C, IncN, IncP and IncW plasmids could be detected in the WWTP effluent and receiving streams 29.

Taken together, WWTPs effectively comprise genetic exchange public parks, where beta-lactamase genes are often linked to efficient MGE vehicles, such as plasmids. Therefore, it is crucial to investigate the potential genetic drivers of beta-lactamase spread and their transfer host range in the wastewater environment. We aimed to uncover the beta-lactamase mobile genetic determinants in wastewaters and these elements’ transfer and retransfer potential in horizontal transmission. In the present study, we applied an exogenous plasmid isolation method to capture ESBL plasmids in the residential and hospital sewers from an urban water system in Odense, Denmark. We used a mCherry red-fluorescent Escherichia coli CSH26 recipient strain in metaparental mating assays with sewer microbiomes and screened for transconjugants exhibiting ampicillin resistance. We identified that the transconjugants showed multiple-drug resistance (MDR) in addition to ampicillin resistance. After sequencing the captured plasmids both by Illumina MiSeq and Oxford Nanopore followed hybrid assembly, we verified the MDR-responsible plasmid belonged to the IncN group and encoded ARGs against aminoglycosides, beta-lactams, fluoroquinolones and sulfonamides. Consequently, it is
critical to investigate the potential host range of such MDR IncN plasmids. In this study, we performed a permissiveness test to explore its potential transfer to indigenous sewer/wastewater bacteria and applied metagenome and PLSDB similarity search for this plasmid to reveal its favoured habitat (treatment compartment) and putative ancestors (epidemic sources and lineages). This study contributes to the understanding of IncN plasmid’s evolution and its relevant transmissible antibiotic resistance.

1.4 Materials and Methods

The overall methodology applied in this study is illustrated in Fig. 1.

Sampling and pre-treatment

Hospital and residential sewer samples from a Danish urban water system (Odense, Denmark) were taken using an ISCO automatic sampler in 24-hour time proportional sampling (50 mL per 5 minutes) in May 2017. Three replicates from two hospital sewer pipelines and one residential sewer pipeline were taken. All samples (n = 9) were immediately cooled with ice before transport. For each sample, 100 mL was centrifuged at 10,000 g for 8 minutes at 4 °C (Eppendorf, Hamburg, Germany) in the laboratory. Supernatants were then removed, while the pellets were resuspended in 20% of glycerol (Sigma-Aldrich, St. Louis, MO, USA) to reach a final volume of 10 mL. After homogenizing for 30 minutes, samples were kept at -80 °C for storage.

Before any experiments, 2 mL of each thawed sample was vortexed with sterile metal beads for 3 minutes, then transferred to a 2 mL tube (Eppendorf) and spun down shortly to settle large debris. The supernatant was transferred to a new Eppendorf tube and filtered using sterile 10 μm filters (Frisenette Aps, Knebel, Denmark). Afterwards, the sample buffer was changed to 0.9% of NaCl (Sigma-Aldrich).
Figure 1 Workflow of this study.

a) *E. coli* CSH26::mCherry was the *Escherichia coli* CSH26::lacIq-Plpp-mCherry-Kan² strain.

b) *E. coli* MG1655::mCherry/pDK_DARWIN::pA1O4O3-gfpmut3 was the *Escherichia coli* MG1655::lacIq-Plpp-mcherry-Kan²//pDK_DARWIN::pA1O4O3-gfpmut3 strain.

¹) During the screening (primarily using ampicillin) of the transconjugants obtained from the mCherry channel of FACS, several ampicillin-resistant transconjugants were discovered as shown on the Sierra blue plate.
'Interesting transconjugant(s)’ means transconjugants showed multiple antibiotic resistance in the AST. One transconjugant (DK_DARWIN) appeared to be one candidate.

Exogenous plasmid isolation

A modified exogenous plasmid isolation protocol was applied in the present study using fluorescence-activated cell sorting (FACS). The hospital and residential sewer samples were used as the donors in the exogenous plasmid isolation experiment. And a chromosomally lacIq-Plpp-mCherry-Kan<sup>8</sup> tagged <i>Escherichia coli</i> CSH26 strain (intrinsically resistant to rifampicin and nalidixic acid) was used as the recipient. The recipient strain was thus resistant to antibiotics kanamycin (Kan), rifampicin (Rif) and nalidixic acid (Nal), while sensitive to ampicillin (Amp).

The live cell numbers of the donors and recipients were detected to calculate the live cells ratios in the recipient and donor cultures, respectively. Briefly, cell cultures were diluted in 1xPBS (Sigma-Aldrich) to obtain ≈ 3000 events/sec when loaded on the FACS Aria III instrument (BD, Franklin Lakes, NJ, USA) at the flow rate of 1, each sample was recorded for 60 sec and accordingly, cell numbers were counted. The number of live bacterial cells in the recipient overnight cultures was detected in the red fluorescence channel. The live donor cell numbers were determined by using a FilmTrace™ LIVE/DEAD™ Biofilm Viability kit (Invitrogen, Waltham, MA, USA) for the pretreated hospital and residential sewage samples (controls were applied). Conjugation assay was performed by adding live donor and live recipient cells at a ratio of 1:1 and a density of ca. 30,000 bacteria/mm<sup>2</sup> onto an MCE 0.22 µm membrane (Fisher Scientific, Pittsburgh, PA, USA). This step ensured an equal number of 3.0 × 10<sup>7</sup> donor and recipient cells on the membrane. The membrane was on a synthetic sewage agar plate (Table S1), which was supplemented with 50 µg/mL of cycloheximide to inhibit fungal growth. All plates were incubated at 25°C for 48 hours. Negative controls were performed for all samples by adding only donor or only recipient onto the membranes at a density of 6.0 × 10<sup>7</sup> bacteria per membrane. After incubation, the membrane was transferred to a 15 mL Falcon tube (Corning, New York, NY, USA) that contains 2 mL of 1xPBS (Sigma-Aldrich). The tubes were then vortexed for elution of cells for 5 minutes. Cell eluent was stored at 4 °C for 24 hours to increase the fluorescent intensity. Afterwards, the recipient and transconjugant cells were sorted out based on their red fluorescence, while donor cells were excluded. Sorted cells were serial diluted 10<sup>1</sup> to 10<sup>7</sup> times. 100 µL of the diluted cells were plated on selective plates. Recipients and transconjugants were selected on Rif/Nal Luria Broth (LB) agar plates, and transconjugants were selected on Amp/Rif/Nal LB agar plates. Three replicates and one control were applied. All plates were incubated at 37 °C for 24 hours. Afterwards, CFUs were counted, and the transfer ratios were calculated by the equation shown below.
Transfer ratio = \frac{\text{Transconjugant CFUs}}{\text{Recipient CFUs}} \times 100\%

Transconjugant colonies were then re-plated and finally harvested from the Amp/Rif/Nal LB agar plates for three times. Stocks of the verified transconjugants were prepared using 15% of glycerol (Sigma-Aldrich) and stored at -80 °C.

**Antibiotic susceptibility testing**

Antibiotic susceptibility testing (AST) was performed according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) broth microdilution method for aminoglycosides (kanamycin, gentamicin, streptomycin), beta-lactams (cephalexin, penicillin, ampicillin, cefotaxime, carbenicillin, imipenem, meropenem), chloramphenicol, colistin, fluoroquinolones (ciprofloxacin, nalidixic acid, olaquindox), macrolides (erythromycin, tylosin), rifampicin, sulfamethoxazole, tetracycline, thiostrepton, trimethoprim, vancomycin\(^{32}\). Meanwhile, the EUCAST disk diffusion method was applied for the rapid test of beta-lactams (ampicillin, cefotaxime, imipenem, meropenem), chloramphenicol, ciprofloxacin, colistin, erythromycin, gentamicin, rifampicin, sulfamethoxazole, tetracycline and trimethoprim\(^{32}\).

**Plasmid sequencing and data analysis**

Fresh transconjugant(s) grown on Amp/Rif/Nal LB agar plates were cultured in Amp/Rif/Nal LB broth overnight. 2 mL of the overnight culture was used for plasmid DNA isolation. A Plasmid Mini AX kit (A&A Biotechnology, Poland) was employed. Thereafter, a digestion with Plasmid-Safe DNase (Lucigen, Madison, WI, USA) was performed to purify the extracted plasmid DNA. After DNA quantity and quality check, sequencing was carried out on a MinION device (Oxford Nanopore Technologies, Oxford, United Kingdom) and a MiSeq machine (Illumina, San Diego, CA, USA) with Rapid Barcoding Kit SQK-RBK004 (Oxford Nanopore Technologies) and MiSeq Reagent Kit v3 (Illumina), respectively.

Raw plasmid reads signals from the MiniON sequencer were primarily sorted by barcode using *deepBinner* 0.2.0\(^{33}\) and base-called using *Albacore* 2.2.1 (https://nanoporetech.com/), and then were assembled using *Unicyler* (minasm2 + Racoon polishing)\(^{34}\). Errors were corrected by Illumina MiSeq short reads polish. Plasmid circular contigs were visually inspected and extracted from the *de novo* De Bruijn assembly graphs using *Bandage* 0.81 software\(^{35}\). Circular sequences were primarily submitted for annotation on the RAST server with RASTTk scheme v2.0\(^{36}\). Afterwards, more features of the open reading frames were added by searching the sequences against databases of *PlasmidFinder* 2.1\(^{37,38}\), RGI 5.1.1/CARD 3.1.0\(^{39}\), *ResFinder* 4.0\(^{40}\), *VirulenceFinder* 2.0\(^{41}\), and *oriTfinder* 1.0\(^{42}\) to identify plasmid replicon, antibiotic resistance, virulence factors and mobility regions on each plasmid. The unannotated open reading frames (typically known as hypothetical
protein CDS after RAST annotations) were translated into protein sequences and blasted against the non-redundant protein sequences database using the blastp PSI-BLAST tool. Further analysis using function-prediction HHpred tool with COG_KOG_V1.0 and Pfam-A_v33.1 databases were applied when no results from blastp. Annotations obtained from PSI-BLAST and HHpred tools were labelled beginning with “Putative”. All detections were run in tools’ built-in default parameters.

Plasmid maps were created and visualized by Geneious Prime 2022.1 (Biomatters, Auckland, New Zealand). Similar plasmids were searched by mash dist strategy (tool version: 2.1.1) with default setting (max. p-value=0.1, max. distance=0.1) against the PLSDB database (database version: v.2021_06_23_v2). One clinical and one environmental relevant plasmid were randomly picked from those similar plasmids and selected as candidates for plasmid sequence comparison analysis.

The comparison analysis was done by Mauve Alignment (progressiveMauve algorithm, match seed weight = 15 minimum LCB score = 30,000, gapped aligner: MUSCLE 3.6) and Geneious Alignment (code matrix = 65% similarity [5.0/-4.0]) in Geneious Prime 2022.1 (Biomatters).

Tagging the plasmid with a fluorescent marker

pDK_DARWIN captured from the exogenous isolated was selected for GFP tagging in the permissiveness test. The tagging for pDK_DARWIN was performed by complementing the plasmid sequence with the fragment pA1O4O3-gfpmut3 using the helper plasmid pSPIN-GFP following a previous protocol. In brief, pSPIN::pA1O4O3-gfpmut3 was initially constructed using the USER® Cloning Kit (New England Biolabs Inc., Ipswich, MA, USA), and the fragment gfpmut3 and plasmid pSPIN (plasmid #160730 from Addgene, Watertown, MA, USA) were assembled according to the manufacturer's protocol. 32-nt spacer was then cloned into the constructed vector pSPIN::pA1O4O3-gfpmut3 at the BsaI restriction site following the manufacturer’s protocol (BsaI-HF, New England Biolabs Inc.; T4 DNA ligase, Thermo Scientific, Waltham, MA, USA). The ligated product was transformed into NEB® 5α Competent Escherichia coli following the manufacturer's instructions (New England Biolabs Inc.). Selection for Escherichia coli DH5α/pSPIN::pA1O4O3-gfpmut3-space was done overnight at 30°C on LB agar medium supplemented with Kan (50 µg/ml). Colonies were then harvested and screened by colony PCR using the pFB50 and pFB51 primers. Details for the strains and primers were documented in Table S2 and Table S3, respectively.

100-150 ng of the plasmid pDK_DARWIN was transformed by electroporation (1.8 KV for ~ 5 ms) into the competent Escherichia coli DH5α/pSPIN::pA1O4O3-gfpmut3-space. Transformed cells were cultured for 30 hours at 30°C in LB agar medium containing cefotaxime (2 µg/ml) and Kan (50 µg/ml). Plasmid pSPIN::pA1O4O3-gfpmut3-space was then removed and verified by colony PCR with primers pFB50 and pFB51. Afterwards, the inserted fragment on pDK_DARWIN::pA1O4O3-gfpmut3 was amplified using primers Sanger_seq_Primer-F and Sanger_seq_Primer-R (insertion site
was indicated on the plasmid map, Fig. 2A). The specific PCR products were purified and confirmed by Sanger sequencing (Eurofins Genomics, Ebersberg, Germany). Candidates with correct insertion of the pA1O4O3-gfpmut3 fragment were transferred to Electrocomp™ GeneHogs Escherichia coli (Invitrogen) with electroporation (1.8 KV for ca. 5 ms) and stored at -80°C. Subsequently, pDK_DARWIN::pA1O4O3-gfpmut3 was transformed to a plasmid-free Escherichia coli MG1655::lacIq-Plpp-mCherry-KanR strain by a conjugation assay. Potential transformants were characterized and sorted out by FACS (BD) at the mCherry channel and plated on Amp LB agar medium for 24 hours at 37°C. Colonies on the plate were checked for fluorescence by a confocal laser scanning microscope Zeiss LSM 800 (Carl Zeiss AG, Oberkochen, Germany). Transformants with correct fluorescence phenotype were replated and checked on the confocal laser scanning microscopy (Fig. S1). Accordingly, the Escherichia coli MG1655::lacIq-Plpp-mCherry-KanR/pDK_DARWIN::pA1O4O3-gfpmut3 strain was constructed.

Permissiveness test of the captured IncN plasmid

The Escherichia coli MG1655::lacIq-Plpp-mCherry-KanR/pDK_DARWIN::pA1O4O3-gfpmut3 strain was applied as the donor strain in the permissiveness test. Hospital and residential sewer samples, mixed sewer samples, and wastewater samples in the biological treatment basin (BTP) of a WWTP were used as recipient communities. All these sewer and wastewater samples were taken from the same urban water system as aforementioned in the “Exogenous plasmid isolation”. The related sampling campaign was conducted in 2018 and has been documented for microbial and beta-lactam ARG profiles together with two other countries’ comparable urban water systems in a previous study 49.

The live donor cells and WWTP recipient community were mixed at 1:1 cell ratio and placed onto an MCE 0.22 μm sterilized membrane (Fisher Scientific) in a LB medium agar plate and an OECD medium agar plate (Table S1) as the same procedures in the conjugation assay of “Exogenous plasmid isolation”. Three replicates and quality controls were applied. After 24 hours of incubation at 30 °C and 48 hours of GFP maturation at 4°C, transfer events were detected by a BD FACS machine and transfer ratio was documented as the ratio of transconjugants (only green) to the original WWTP recipient (nonfluorescent) cell number. Sorting was performed in the Single Cell Precision mode targeting to sort ca. 10,500 cells as suggested by Olesen et al. 30 with parameters set as shown in Table S4. A counting step for 10 to 20 minutes was performed to record events in different gates and calculate the transfer ratio for each sample. All data were stored and analyzed in the BD FACSDiva software version 8.0.3. The detailed permissiveness test method refers to a previous paper 30.
Abundance of the IncN plasmid pDK_DARWIN in the pan-European urban water systems

Previously, we have performed two sampling campaigns (summer and winter, 2018) at different locations in three urban water systems from three European cities of similar size (Odense, Denmark; Santiago de Compostela, Spain; Durham, United Kingdom)\textsuperscript{49}. Samples from the hospital and residential sewers, the end of sewer (i.e., mixed sewer, WWTP inlet) and BTP in the WWTP (n = 78 samples; 3 replicates per location in two seasons) were sequenced using Illumina NovaSeq sequencing generating 2.3E9 reads in total (mean = 3.1E7 per sample). Reads were mapped against the sequence of pDK_DARWIN using \textit{bowtie2} \textsuperscript{50} in \textit{anvi’o} \textsuperscript{51} using metagenomic-reference workflow. A detection threshold of >50\% of pDK_DARWIN sequence coverage by mapped reads was set to determine if the plasmid is detected in the sample and calculate relative abundance. Relative abundance was calculated by dividing the number of reads mapping to pDK_DARWIN in each sample by the total number of reads in that sample.

1.5 Results

Isolation of antibiotic-resistance plasmids from wastewater

Through the exogenous plasmid isolation, we successfully captured several highly transmissible plasmids [transfer ratio (transconjugant CFUs/recipient CFUs): 3.34\% ± 1.15\%] encoding ampicillin resistance from the sewer communities\textsuperscript{52}. Then we did antibiotic susceptibility testing, and we chose to focus on a transconjugant “DK_DARWIN” (i.e., \textit{Escherichia coli} CSH26::\textit{lacIq-Plpp-mCherry-KanR}/pDK_DARWIN) that showed a multiple resistance profile, including resistance to aminoglycosides (gentamicin, streptomycin), beta-lactams (ampicillin, cefotaxime, carbenicillin), sulfonamides and its combinations (sulfamethoxazole and trimethoprim) as indicated in Table 1.

Table 1 Antibiotic susceptibility testing for the strains involved in the exogenous plasmid isolation*.

| Recipient (Plasmid-free) | KAN | GEM | STR | LEX | PEN | AMP | CTX | CAR | IMP | MEM | CIP | NAL | OLA | VAN | COL | ERY | TYL | THI | CHL | RIF | SMX | TMP | TET |
|-------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Transconjugant (Plasmid-harboured) | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT | BMDT | CDT |

* KAN = kanamycin, GEM = gentamicin, STR = streptomycin, LEX = cephalaxin, PEN = penicillin, AMP = ampicillin, CTX = cefotaxime, CAR = carbenicillin, IMP = imipenem, MEM = meropenem, CIP = ciprofloxacin, NAL = nalidixic acid, OLA = olaquindox, VAN = vancomycin, COL = colistin, ERY = erythromycin, TYL = tylosin, THI = thiostrepton, CHL = chloramphenicol, RIF = rifampicin,
SMX = sulfamethoxazole, TMP = trimethoprim, TET = tetracycline. BMDT = broth microdilution testing method, DDT = disk diffusion testing method. Resis. = resistant, Sens. = sensitive. Recipient = Escherichia coli CSH26::lacIq-Plpp-mCherry-KanR, Transconjugant = DK_DARWIN (i.e., Escherichia coli CSH26::lacIq-Plpp-mCherry-KanR/pDK_DARWIN).

The BMDT and DDT testing approaches were applied according to EUCAST recommendations (version 10.0) 32. The combined results of the two testing methods were shown in the table. Contradictory results were found using the two testing methods for measuring the susceptibility of MEM and CIP for the transconjugant and recipient, respectively. Therefore, both results (MEM and CIP) were not considered acquired resistance.

Next, by plasmid sequencing, we identified the captured plasmid. pDK_DARWIN is a 73,132 bp IncN plasmid carrying a complete set of functional conjugative genes and multiple ARGs (Fig. 2A). Specifically, these ARGs target antibiotic classes of aminoglycosides (aac(3)-lla, aph(3’’)-lb, aph(6)-ld), beta-lactams (blaCTX-M-15, blaOXA-1, blaTEM-1B), chloramphenicol (catB3), fluoroquinolones (aac(6’)-lb-cr, qnrB1), sulfonamides (sul2), trimethoprim (dfrA14) and tunicamycin (tmrB). Meanwhile, we observed antirestriction protein ArdA, Class 1 integron integrase IntI1, diverse transposases and integrases, and miscellaneous putative functional genes on this plasmid.

**Plasmid similarity search and analysis**

We then performed sequence similarity searches with pDK_DARWIN across curated plasmid sequence repositories (PLSDB) 45, 46. This resulted in 1,169 hits to previously isolated plasmids (derived from the associated BioSample) from diverse sites around the world, particularly in East Asia, Europe and North America. Notably, these plasmids were predominantly documented in the last five years (Figs. 2B and 2C) and harboured by members of Enterobacteriaceae (ca. 97%, 1,130 counts out from the 1,169 hits), particularly from the *Escherichia* and *Klebsiella* genera (ca. 79%), Fig. 2D. The median length of pDK_DARWIN’s similar plasmids was 89 Kbp (Fig. 2E), and the median GC content was 52% (Fig. 2F). Detailed results of the similar plasmids were recorded in Table S5.
Figure 2 Plasmid map and PLSDB similarity search of pDK_DARWIN.

(A) Plasmid map of pDK_DARWIN. (B) Global BioSample location distribution of similar plasmids of pDK_DARWIN in PLSDB database. (C) BioSample creation date of similar plasmids of pDK_DARWIN in PLSDB database. (D) Taxonomic composition of similar plasmids of pDK_DARWIN in PLSDB database (species, genus, and family taxa). (E) Sequence length of similar plasmids of pDK_DARWIN in PLSDB database. (F) GC content of similar plasmids of pDK_DARWIN in PLSDB database.

A plasmid comparison analysis was performed to detect whether this plasmid shared some common regions with other plasmids and possible evolutionary relations. One clinically and one environmentally relevant plasmid were selected from the PLSDB similarity search result. The results showed that pDK_DARWIN shared the genetic regions encoding conjugation module (relaxase, T4CP, T4SS), ARGs (dfrA14, blaCTX-M-15) and MGEs (IntI1, IS1380), stability (StdA, StdB and StdC) and antirestriction proteins with other IncN plasmids (Fig. 3A). Alignment of these three plasmids showed 67.1% pairwise identity and 58.8% identical sites (Fig. 3B). The main distinct area was the
accessory genes region (harbouring a high content of ARGs and MGEs) where the three plasmids showed the lowest identity.

**Figure 3** Plasmid sequence alignment of pDK_DARWIN with its similar plasmids.  
(A) Mauve alignment of pDK_DARWIN and its similar plasmids. (B) Geneious Alignment of pDK_DARWIN and its similar plasmids. ‘AP019402’ stands for the 51 Kbp IncN plasmid isolated from a patient with accession no. AP019402, and ‘CP056317’ stands for the 56 Kbp IncN plasmid isolated from freshwater downstream of a WWTP with accession no. CP056317.

In Fig. 3A, each contiguously coloured region is a locally collinear block (LCB), a region without rearrangement of homologous backbone sequence (match seed weight: 15, minimum LCB score: 30000). LCBs below a sequence’s centre line are in the reverse complement orientation relative to the reference sequence. Lines between plasmid sequences trace orthologous LCBs between plasmids. White regions within an LCB signify the presence of lineage-specific sequence at that site. Specifically, the large light blue block stands for the region of the T4SS gene cluster; the large light green block stands for the region of relaxase, T4CP, and several stability and antirestriction proteins; the red block stands for the region of ARGs towards tetracycline resistance which pDK_DARWIN didn’t carry.
In Fig. 3B, in the identity bar, the green-coloured region stands for 100% mean pairwise identity over all pairs in the column, the khaki-coloured region stands for at least 30% and under 100% identity, and the red-coloured region stands for below 30% identity.

Retransfer of pDK_DARWIN to the sewer and wastewater communities

Given that pDK_DARWIN was derived from the hospital sewer, we tested its retransfer potential to the sewer/wastewater environment by performing permissiveness tests in different communities from the hospital, residential and mixed sewers, and in the BTP of the receiving WWTP. We observed transfer events of pDK_DARWIN, displayed as green events in the GFP channel (Fig. S2). Transfer ratios were not significantly different (Turkey’s and Wilcoxon’s tests) among the different recipient communities using LB medium, yet significantly different ($p < 0.05$) between HS and RS, HS and MS, HS and BTP, RS and MS, as well as RS and BTP using OECD medium (Turkey’s test) as shown in Fig. 4. In general, we showed that pDK_DARWIN can have high transfer efficiencies (mean transfer ratio: $0.50\% \pm 0.28\%$ on LB medium and $0.11\% \pm 0.08\%$ on OECD medium) into the sewer and wastewater samples communities.

Figure 4 Transfer ratio (T/R) of pDK_DARWIN to the sewer or wastewater communities.

$DK_{1a}$, $DK_2$, $DK_4$, and $DK_6$ represent the recipient communities from hospital sewer, residential sewer, mixed sewers, and wastewater in the BTP of the receiving WWTP, respectively. The conjugation assays were done in triplicates, shown individually as coloured dots, the black dots are the mean of the triplicates and error bars show the standard deviation. The logarithmic y-axis shows
transconjugants per recipient (T/R, i.e., transfer ratio). The numbers between different recipient communities in Fig. 4B were the “p adj” values in Turkey’s test.

Prevalence and abundance of pDK_DARWIN in the sewers and wastewaters

The abundance of pDK_DARWIN was measured using short-reads mapping from shotgun metagenomic datasets from 4 locations in 3 European countries (manuscript under preparation). In all three countries (Denmark, Spain, and the United Kingdom), pDK_DARWIN mapped with the highest relative abundance in hospital sewers, indicating this IncN plasmid was more prevalent in hospital sewers than in any other compartments across all countries (Fig. 5). Significant differences in abundance were observed in Spain and the United Kingdom between hospital and residential sewers samples, and between hospital and end of sewer in Spain. In Denmark, the plasmid was not detected in the BTP of the WWTP, while it was detected in one sample out of 6 both in Spanish and British samples.

Figure 5 Plasmid pDK_DARWIN relative abundance in three comparable urban water systems from three different countries.

DK = Odense, Denmark; SP = Santiago de Compostela, Spain; UK = Durham, United Kingdom. Relative abundance is only displayed for samples where pDK_DARWIN sequence was above the detection threshold (>50% coverage), as described in the methods section. Relative abundance is
compared between hospital sewer, residential sewer and end of sewer (wastewater treatment plant inlet) locations in each country using Wilcoxon’s test, and the p values were shown between different compartments in the figure. Biological treatment basin and biofilter were the different BTPs used in Spain and the United Kingdom, respectively.

1.6 Discussion

IncN plasmids encode MDR

IncN plasmids can be multilocus sequence typed by the repA, traJ and korA loci. They are frequently typed in clinical samples and analysed due to harbouring and disseminating sulphonamide, quinolone, aminoglycoside, tetracycline and streptomycin resistance genes. IncN plasmids were among the most frequently detected ESBL gene carriers (IncFII, IncA/C, IncL/M, IncN and IncI1) in Enterobacteriaceae, which are responsible for the spread of blaCTX-M variants, blaKPC, blavim, and blanDM-1. From a phylogenetic perspective, IncN plasmids are broad-host-range and conjugative plasmids, and they have been frequently detected in various human and food-animal samples. These indications are alarming since beta-lactam antibiotics are currently the most widely used antibiotics in the hospital settings around the world, the IncN plasmids are prone to facilitate the dissemination of beta-lactam resistance from nosocomial-relevant systems to adjacent or remote environmental settings via possible pathways such as the urban wastewater network.

Here, we show that the IncN plasmid pDK_DARWIN captured in the hospital sewer promoted MDR when transferred to a previously sensitive host. The plasmid sequence evidenced genes known to confer the observed resistance phenotypes except catB3 (Fig. 2A, Table 1). Plasmids evolve over time and offer the host beneficial traits to be stably maintained and represent a HGT toolbox for exchanging ARGs among bacteria of different origins and sources. Two main antibiotic resistance genetic regions were discovered on pDK_DARWIN: one contained [aac(6')-lb-cr, blaOXA-1, aac(3)-lla and tmr], the other with [aph(6)-ld], aph(3')-lb and sul2]. A sequence query view for the top 100 blastn results of pDK_DARWIN indicated that these two resistance genetic blocks might be horizontally acquired from plasmids associated with the genera Citrobacter, Enterobacter, Escherichia, Klebsiella and Salmonella (Fig. S3, Table S6). Meanwhile, we observed the left hit region including Citrobacter, Enterobacter, Escherichia, Klebsiella, Klyvera, Morganella, Pantoea, Salmonella and Yersinia sourced plasmid sequences overlapped the T4SS gene clusters and groups of plasmid maintenance and stability genes with pDK_DARWIN. And the right hit region exhibited an even more diverse and larger number of hosts, incorporating Leclercia and Raoultella. And this region mainly contains relaxase, T4CP, multiple stability and putative functional genes. Interestingly, we found one single sequence could hit both left and right regions, but exclusively cover the middle
main ARG region we described above for once (Table S6, Fig. S4). This indicates the IncN backbones do not necessarily or often carry the set of ARGs for horizontal transmission as shown in this study. However, considering the constantly evolving accessory region (middle hit region) and gene rearrangement, we assumed pDK_DARWIN is among those groups of highly infective plasmids.

Li et al. isolated ESBLs producing Escherichia coli strains in wastewaters of a Chinese WWTP and found that 70.0% (35/50) of these isolates showed conjugation ability and MDR. Further, they identified that one ESBL plasmid with an IncN replicon carried multiple ARGs (blaCTX-M-15, blaTEM-1, qnrS1). Genes rmtB and blaCTX-M-1 encoded on IncN plasmids could be commonly captured from pig farms, farmworkers, and the adjacent soil environment. Various ESBL genes and qnrS could be found on IncN plasmids harboured by Escherichia coli isolates from waterbirds. All these investigations imply that IncN-facilitated MDR has occurred and spread in diverse environments. Eikmeyer et al. reported that four IncN plasmids harbouring a common blaTEM-1 gene emerged in the effluent of a German WWTP. And here we displayed that the MDR IncN plasmid pDK_DARWIN possessed a high transfer ratio in the exogenous isolation and preserved an efficient transmission ability transferring back to the sewer and wastewater communities in this study, illustrating that IncN plasmids are potent vehicles for ARG dissemination in sewage water microbial communities. Moreover, pDK_DARWIN shared backbones with both surface water and clinical settings sourcing plasmids, which demonstrates it’s been actively involved in horizontal transmission, constantly evolving its accessory genetic context to be stably maintained in different microbial communities in different environment, arguably bridging the barrier of the ecological niches for HGT.

**Enterobacteriaceae are potential hosts of the sewer derived IncN plasmid**

IncN plasmids belong to the MOB family, size ranging from 30 to 70 Kbp, and are a group of low-copy-number, broad-host-range plasmids, which are usually colocalized with IncF plasmids. IncN plasmids can be found in most Enterobacteriaceae. IncN plasmids carrying ARG have already spread throughout Europe and often isolated from Escherichia coli of animal sources. Herein, we attempted to investigate the host range of the exogenously isolated urban sewer derived IncN plasmid pDK_DARWIN in the related wastewater environment. We evidenced conjugation transfer by using a genetically modified pDK_DARWIN reporter plasmid and FACS (Figs. 4 and S2C – 2E). However, due to low sorting speed, we failed to harvest sufficient cells for DNA extraction and subsequent sequencing. Olesen et al. suggested ca. 3000 events/sec was ideal for FACS sorting, whereas we were far away from that number. Further concentration of the mating matrix and modification for the permissiveness test protocol can be applied in order to reach a satisfying sorting speed and investigate recipients’ taxonomy by means of 16S-rRNA gene sequencing for example. Despite not identifying
putative hosts of IncN pDK_DARWIN in our samples, we documented the transfer ratios of different sampling sites and transfer conditions. Transfer ratios were generally higher using LB medium than OECD medium (Fig. 4). Nonetheless, the OECD medium would probably offer a less biased window to explore those non-fast-growing and/or rare microorganisms since it has alleviated bias. Meanwhile, a lower transfer ratio was detected in hospital sewer compared to other treatment compartments. This was possibly because the hospital sewer communities had already carried other IncN plasmids, resulting in the occurrence of incompatibility with pDK_DARWIN.

Based on the PLSDB similarity search for this IncN plasmid, we retrieved 1,169 plasmids from the database. And these similar plasmids were dominantly hosted by Enterobacteriaceae (97%), but also could be hosted by the family of Aeromonadaceae, Erwiniaaceae, Moraxellaceae, Morganellaceae, Pseudomonadaceae, Shewanellaceae, Vibrionaceae and Yersiniaceae. Carattoli et al. showed the MDR IncN plasmid pKOX105 played a crucial role in the propagation of diverse carbapenemase genes in Klebsiella spp. And another MDR IncN plasmid pKP96 that encodes qnra1, aac(6')-ib-cr and blaCTX-M-24 could be determined from clinical Klebsiella pneumoniae isolates. Sequence alignment of pDK_DARWIN with clinically- and environmentally- relevant plasmids indicates the backbones of this IncN plasmid as well as the possibility of pDK_DARWIN acquiring various cargo genes from remote sources (Fig. 3B).

**Transfer potentials of the IncN plasmid in the urban water systems**

Carattoli, A. earlier suggested that IncN plasmids were preferably habited in the nosocomial and animal husbandry surroundings, whilst IncN plasmids can also emerge in the environmental settings, such as soil, piggery manure, surface water and wastewater.

In the present study, we monitored the occurrence and abundance of pDK_DARWIN in a metagenome dataset containing sewer and wastewater samples from different treatment stages of three urban water systems located in three different countries (manuscript in preparation). We found that pDK_DARWIN was detected of a higher relative abundance in hospital sewer samples than the other treatment stages in Denmark (Fig. 5). Importantly, we showed that pDK_DARWIN can persist up to the mixed sewer (wastewater inlet), but also inside the BTPs of Spanish and British WWTPs.

The higher relative abundance of reads mapping to pDK_DARWIN sequence from the British BTP sample can be explained by the aerobic process used in that plant that favours the growth of Enterobacteriaceae (identified as probable hosts of this plasmid), whereas WWTPs from Denmark and Spain use an anaerobic treatment process less suitable for the persistence of these bacteria. Considering the relatively low detection limits of bulk shotgun sequencing, confident detection of pDK_DARWIN sequence, based on reads mapping coverage, in these samples means that this plasmid is present in non-negligible amounts in these samples. Accordingly, we detected a high
transfer ratio of pDK_DARWIN to the wastewater communities in the BTP in the permissiveness test (Fig. 4). This implies the transfer potential of pDK_DARWIN was still vigorous in vitro although the fitness cost in the indigenous circumstances was overwhelming. It might be also interesting to search those similar plasmids of pDK_DARWIN in the metagenome dataset as compensatory evolution can occur, and those plasmids with pDK_DARWIN backbones and lowered fitness costs may replicate and maintain themselves in new hosts expanding their host range in the wastewaters. However, the plasmid fitness impact cannot be inferred from results obtained with other host-plasmid combinations, even if these are closely related. Further studies on the evolutionary potential of this IncN plasmid for the purpose of deciphering its adaptive trajectories can be important considering its harboured MDR and high transmission capability in HGT.
1.7 Acknowledgements

This research was funded by a Joint Programming Initiative-Antimicrobial Resistance grant (JPI-AMR; DARWIN project #7044-00004B) to BFS, and the DFF-Research Project 2 Grants from the Danish Council for Independent Research | Technology and Production (7017-00210A) to BFS. 
ZFY wants to thank Xiao Peng (Institute of Microbiology, Chinese Academy of Sciences, China) for assistance in re-streaking the transconjugant strains.

1.8 Contributions

ZFY performed the antibiotic susceptibility testing for the transconjugants from the exogenous plasmid isolation, plasmid sequencing and downstream analysis for pDK_DARWIN, as well as wrote the manuscript; JN processed the metagenome search for pDK_DARWIN, JN and RPR gave suggestions to the experimental work; QQQ has assisted in the construction of the GFP tagged pDK_DARWIN; NY has offered technical knowledge in the verification for the pDK_DARWIN::GFP construct; ZG has helped during the permissiveness test for pDK_DARWIN; JSM and RPR gave recommendations for the story scope and the laboratory work; KADC captured the pDK_DARWIN in the exogenous plasmid isolation; RPR and HA helped in the laboratory work; AKO gave the advice in using the FACS Aria III instrument and processed the FACS data analysis in the permissiveness test; JN, AD, BS and SJS made the initial proposal of this work; all the co-authors have reviewed and gave suggestions for this manuscript.

1.9 Strain Accessibility

Strains of *Escherichia coli* CSH26::lacIq-Plpp-mCherry-Kan<sup>8</sup> strain (SP1554), *Escherichia coli* CSH26::lacIq-Plpp-mCherry-Kan<sup>8</sup>/pDK_DARWIN strain (SP1555) and *Escherichia coli* MG1655::lacIq-Plpp-mcherry-Kan<sup>8</sup>/pDK_DARWIN::pA1O4O3-gfpmut3 (SP1557) are available upon request.
1.10 References


52. Clasen KAD. Investigating the transfer frequency of ESBL-resistance in an urban wastewater system utilising a quantitative approach combining exogenous plasmid isolation and FACS. Master Thesis 2018; University of Copenhagen, Copenhagen, Denmark.


1.11 Supporting Information

Supporting Information for
Horizontal transmission of a multidrug-resistant IncN plasmid isolated from urban wastewater

Authors and affiliations
Zhuofeng Yu, a, * Joseph Nesme, a, * Qinquin Wang, a Jonas Stenløkke Madsen, a Rafael Pinilla-Redondo, a Kamille Anna Dam Clasen, a Hanadi Ananbeh, b, c Asmus Kalckar Olesen, a Zhuang Gong, a Nan Yang, a Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen a, #

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
c Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

* These authors contributed equally to this work.
# Corresponding Author.
Table S1 Medium used for conjugation assay in this study.

<table>
<thead>
<tr>
<th>No.</th>
<th>Medium*</th>
<th>Reagents#</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LB agar medium</td>
<td>LB broth</td>
<td>20 g/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agar</td>
<td>15g/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Meat extract</td>
<td>12.8 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peptone</td>
<td>12.8 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Potato starch</td>
<td>42.7 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skim milk powder</td>
<td>50.8 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glycerol</td>
<td>34 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urea</td>
<td>75 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NH₄Cl</td>
<td>11 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sodium acetate</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K₂PO₄·3H₂O</td>
<td>23 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K₂HPO₄</td>
<td>28 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MgSO₄·7H₂O</td>
<td>2 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uric acid</td>
<td>10 mg/L</td>
</tr>
<tr>
<td>2</td>
<td>Synthetic sewage medium</td>
<td>Meat extract</td>
<td>110 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peptone</td>
<td>160 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urea</td>
<td>30 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K₂HPO₄</td>
<td>28 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MgSO₄·7H₂O (1000x)</td>
<td>2 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NaCl (1000x)</td>
<td>7 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CaCl₂·2H₂O (1000x)</td>
<td>5 mg/L</td>
</tr>
</tbody>
</table>

* LB agar medium and synthetic sewage medium were prepared with Purelab Flex ultrapure water (ELGA LabWater, Lane End, United Kingdom), while OECD medium was prepared with PBS.

# All reagents were bought from Sigma-Aldrich (Sigma-Aldrich, St. Louis, MO, USA).
<table>
<thead>
<tr>
<th>Strain</th>
<th>Growth temperature (°C)</th>
<th>Antibiotic resistance*</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em> DH5α</td>
<td>37</td>
<td>-</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td><em>Electrocomp™ GeneHogs Escherichia coli</em></td>
<td>37</td>
<td>-</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><em>Escherichia coli</em> CSH26::lacIq-Plpp-mCherry-KanR</td>
<td>37</td>
<td>KanR, RifR, NalR</td>
<td>Kindly shared by Gitte Charbon*</td>
</tr>
<tr>
<td><em>Escherichia coli</em> MG1655::lacIq-Plpp-mCherry-KanR</td>
<td>37</td>
<td>KanR</td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> MG1655::lacIq-Plpp-mcherry-KanR/pDK_DARWIN::pA1O4O3-gfpmut3</td>
<td>37</td>
<td>KanR, AmpR, CtxR</td>
<td>This study</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Growth temperature (°C)</th>
<th>Antibiotic resistance</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSPIN</td>
<td>37</td>
<td>KanR</td>
<td>Addgene plasmid #160730</td>
</tr>
<tr>
<td>pSPIN-GFP</td>
<td>37</td>
<td>KanR</td>
<td>This study</td>
</tr>
<tr>
<td>pDK_DARWIN::pA1O4O3-gfpmut3</td>
<td>37</td>
<td>AmpR, CtxR</td>
<td>This study</td>
</tr>
</tbody>
</table>

# The growth temperature stands for the ideal temperature for the strains growing.

* Kan = kanamycin, Rif = rifampicin, Nal = nalidixic acid, Amp = ampicillin, Ctx = cefotaxime.

* Contact information: [https://linkedin.com/in/gitte-ebersbach-charbon-5ba4633](https://linkedin.com/in/gitte-ebersbach-charbon-5ba4633)
Table S3 Primers used in this study.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pFB50</td>
<td>5’- CCGATTCATTAATGCAGCTGG -3’</td>
</tr>
<tr>
<td>pFB51</td>
<td>5’- CGATGGACATCACCCTAGC -3’</td>
</tr>
<tr>
<td>Sanger_seq_Primer-F</td>
<td>5’- AATGCGTCCGCGGATC -3’</td>
</tr>
<tr>
<td>Sanger_seq_Primer-R</td>
<td>5’- TTGTTCAGCGGGCG -3’</td>
</tr>
</tbody>
</table>
**Table S4** Parameters applied in the FACS sorting.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Voltage (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCS</td>
<td>500</td>
</tr>
<tr>
<td>SSC</td>
<td>312</td>
</tr>
<tr>
<td>FITC</td>
<td>555</td>
</tr>
<tr>
<td>PETxRed</td>
<td>676</td>
</tr>
<tr>
<td>PerCP-Cy5-5</td>
<td>747</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Threshold value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSC</td>
<td>200</td>
</tr>
<tr>
<td>SSC</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laser (nm)</th>
<th>Delay</th>
<th>Area scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>488</td>
<td>0.00</td>
<td>1.05</td>
</tr>
<tr>
<td>633</td>
<td>-83.26</td>
<td>0.71</td>
</tr>
<tr>
<td>561</td>
<td>-39.00</td>
<td>0.88</td>
</tr>
<tr>
<td>405</td>
<td>40.47</td>
<td>0.90</td>
</tr>
</tbody>
</table>
Table S5 PLSDB similarity search result of pDK_DARWIN*.
* This table was produced by PLSDB including the information of similar plasmids’ Accession no., Distance, p value, BioSample location, BioSample creation, BioSample isolation source, BioSample host, PlasmidFinder, pMLST, GC contents, Length, Taxon_info, etc. This table was uploaded on Google Sheets for sharing.

Link: https://drive.google.com/file/d/1akEw_JSHfyHJSGqlsH-Wce_sN1rGu0wE/view?usp=sharing
### Table S6 Hit sequence information of the blastn results of pDK_DARWIN*.

#### Left hit region

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence List Name</th>
<th>Description</th>
<th>Genus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CP008901</td>
<td>Enterobacter cloaceae ECNIH3 plasmid pKPC-47e, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>2</td>
<td>CP008962</td>
<td>Escherichia coli strain ECON1H3 plasmid pKPC-629, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>3</td>
<td>CP008964</td>
<td>Klebsiella pneumoniae subsp. pneumoniae strain KPNH1792 plasmid pKPC-6e4, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>4</td>
<td>CP008981</td>
<td>Pantoea spp. PSN1H1 plasmid pKPC-1e5, complete sequence</td>
<td>Pantoea</td>
</tr>
<tr>
<td>5</td>
<td>CP014524</td>
<td>Escherichia coli strain ZH063 plasmid pZ063 2, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>6</td>
<td>CP020049</td>
<td>Escherichia coli strain AR_0061 plasmid unitig_1, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>7</td>
<td>CP020066</td>
<td>Klebsiella pneumoniae strain AR_0177 plasmid unitig_5, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>8</td>
<td>CP023420</td>
<td>Klebsiella pneumoniae strain 1050 plasmid pK1050-4, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>9</td>
<td>CP025758</td>
<td>Citrobacter freundii complex sp. CFN124 plasmid pKPC-349c, complete sequence</td>
<td>Citrobacter</td>
</tr>
<tr>
<td>10</td>
<td>CP025865</td>
<td>Klebsiella pneumoniae strain WCHKP34 plasmid pNDM1_1L_34, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>11</td>
<td>CP026274</td>
<td>Klebsiella oxytoca strain KONIH4 plasmid pKPC-4666, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>12</td>
<td>CP026496</td>
<td>Klebsiella pneumoniae strain 616 plasmid pK616 1, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>13</td>
<td>CP026590</td>
<td>Klebsiella pneumoniae strain NUHL3045 plasmid p4, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>14</td>
<td>CP030310</td>
<td>Klebsiella pneumoniae strain 205880 plasmid p205880-8, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>15</td>
<td>CP039821</td>
<td>Klebsiella pneumoniae strain C2414 plasmid pC2414-3-NDM, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>16</td>
<td>CP040178</td>
<td>Klebsiella pneumoniae strain 2e plasmid unnamed3, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>17</td>
<td>CP040187</td>
<td>Salmonella enterica subsp. enterica serovar Cuba strain AR-0402 plasmid pAR-0402</td>
<td>Salmonella</td>
</tr>
<tr>
<td>18</td>
<td>CP046118</td>
<td>Enterobacter cloaceae strain CBG15936 plasmid pNDM1-CBG, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>19</td>
<td>CP047737</td>
<td>Enterobacter hormaechei strain Eho-5 plasmid pEoh-5, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>20</td>
<td>CP047744</td>
<td>Enterobacter hormaechei strain Eho-4 plasmid pEoh-4, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>21</td>
<td>CP047752</td>
<td>Enterobacter hormaechei strain Eho-3 plasmid pEoh-3, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>22</td>
<td>CP050859</td>
<td>Klebsiella pneumoniae strain SCH6109 plasmid pSC6109-NDM, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>23</td>
<td>CP068288</td>
<td>Enterobacter hormaechei strain 1575 plasmid p1575-1, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>24</td>
<td>HF545433</td>
<td>Escherichia coli plasmid pE66A, strain E66A, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>25</td>
<td>HF555507</td>
<td>Klebsiella pneumoniae plasmid pK45-67VM complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>26</td>
<td>KP077034</td>
<td>Klebsiella pneumoniae subsp. pneumoniae strain OW16C2 plasmid pOW16C2, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>27</td>
<td>KJ933392</td>
<td>Escherichia coli strain YD626E plasmid pYD626E, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>28</td>
<td>KM660724</td>
<td>Morganella morganii strain MRSN22709 plasmid pMR3-OXA181, complete sequence</td>
<td>Morganella</td>
</tr>
<tr>
<td>29</td>
<td>KU260255</td>
<td>Escherichia coli strain BK32602 plasmid pK32602, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>30</td>
<td>KY129843</td>
<td>Klebsiella oxytoca strain 97.38 plasmid pK38, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>31</td>
<td>KY128484</td>
<td>Klebsiella mechinans strain 97.58 plasmid pK38, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>32</td>
<td>LS998787</td>
<td>Escherichia coli isolate EC-TO75 genome assembly, plasmid: 3</td>
<td>Escherichia</td>
</tr>
<tr>
<td>33</td>
<td>LV221007</td>
<td>Yersinia pseudotuberculosis strain Yps.R1, plasmid pYps.R1 complete sequence</td>
<td>Yersinia</td>
</tr>
<tr>
<td>34</td>
<td>L1985266</td>
<td>Escherichia coli strain 177 genome assembly, plasmid: RCS53_p</td>
<td>Escherichia</td>
</tr>
<tr>
<td>35</td>
<td>MG764553</td>
<td>Klebsiella pneumoniae strain A3295 plasmid pA3295-KPC, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>36</td>
<td>MK862125</td>
<td>Enterobacter hormaechei strain 1-RC-17-04040-5 plasmid p04040-5-KPC40, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>37</td>
<td>MN178639</td>
<td>Khuvora erycrebspin strain SCW213 plasmid pCW213-1, complete sequence</td>
<td>Khuvora</td>
</tr>
<tr>
<td>38</td>
<td>MN891676</td>
<td>Klebsiella pneumoniae strain 24169 plasmid p24169-KPC, complete sequence</td>
<td>Klebsiella</td>
</tr>
</tbody>
</table>

#### Middle hit region

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence List Name</th>
<th>Description</th>
<th>Genus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AP022378</td>
<td>Citrobacter portucaulensis 10MTU157 DNA, complete genome</td>
<td>Citrobacter</td>
</tr>
<tr>
<td>2</td>
<td>CP006602</td>
<td>Klebsiella pneumoniae strain ATCC BAA-2146 plasmid pK1g, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>3</td>
<td>CP018460</td>
<td>Klebsiella pneumoniae strain Kp_Goe 39795 plasmid pKp_Goe 791-5, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>4</td>
<td>CP018819</td>
<td>Klebsiella pneumoniae strain AR_0049 plasmid unitig_3, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>5</td>
<td>CP021463</td>
<td>Salmonella enterica subsp. enterica serovar Typhimurium strain UGA14 plasmid pUGA14_1, complete sequence</td>
<td>Salmonella</td>
</tr>
<tr>
<td>6</td>
<td>CP021749</td>
<td>Enterobacter cloaceae complex sp. strain AR_0163 chromosome, complete genome</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>7</td>
<td>CP028413</td>
<td>Enterobacter sp. CR ENT-193 plasmid pCR ENT-193_1, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>8</td>
<td>CP031311</td>
<td>Klebsiella quasipneumoniae strain G458 plasmid pG458, 136.4kb, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>9</td>
<td>CP031317</td>
<td>Klebsiella quasipneumoniae strain G457 plasmid pG457, 150.8kb, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>10</td>
<td>CP052538</td>
<td>Klebsiella pneumoniae strain B16K014 plasmid pB16K014, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>11</td>
<td>CP061495</td>
<td>Enterobacter hormaechei subsp. xiangfangensis strain GENC200 plasmid pGENC200, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>12</td>
<td>CP061496</td>
<td>Enterobacter hormaechei subsp. steigerwaltii strain GENC174 plasmid pGENC174, complete sequence</td>
<td>Enterobacter</td>
</tr>
<tr>
<td>13</td>
<td>JX422423</td>
<td>Klebsiella pneumoniae plasmid pKD1, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>14</td>
<td>KX636133</td>
<td>Klebsiella pneumoniae strain KJ119 plasmid pKJ119-NDM1, complete sequence</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>15</td>
<td>KX696646</td>
<td>Salmonella enterica subsp. enterica serovar Typhimurium plasmid pSTe-A54650, complete sequence</td>
<td>Salmonella</td>
</tr>
<tr>
<td>16</td>
<td>MG878867</td>
<td>Escherichia coli strain Fc21617 plasmid pFc21617-310, complete sequence</td>
<td>Escherichia</td>
</tr>
<tr>
<td>17</td>
<td>MN540571</td>
<td>Escherichia coli strain E.coli46g plasmid pH1c2H2 CTX_M 15, complete sequence</td>
<td>Escherichia</td>
</tr>
</tbody>
</table>
Three distinct hit regions were discovered in the sequence query view (Fig. S3) and shown in detail above. Yellow coloured blocks represented the sequences that were found to hit two regions of pDK_DARWIN (left and right).
**Figure S1** Verification for the constructed GFP tagged pDK_DARWIN *Escherichia coli* strain using Zeiss LSM 800 (Carl Zeiss AG, Oberkochen, Germany) confocal laser scanning microscopy (CLSM).

Figs. **S1A, S1D** and **S1G** were imaged under the mCherry channel of CLSM, Figs. **S1B, S1E** and **S1H** were imaged under the gfp channel of CLSM, Figs. **S1C, S1F** and **S1I** showed the merged signals from both mCherry and gfp channels. The scale on the view is 5 µm as indicated by the white bar in the figures.

Clones from an overnight culture of *Escherichia coli* MG1655::lacIq-Plpp-mCherry-KanR/pDK_DARWIN::pA1O4O3-gfpmut3 on LB agar medium supplemented with kanamycin (50 µg/mL) and ampicillin (100 µg/mL) were checked and one clone was shown as an example in Figs.
S1A to S1C. Clones from an overnight culture of Escherichia coli MG1655::lacIq-Plpp-mCherry-KanR//pDK_DARWIN::pA1O4O3-gfpmut3 on LB agar medium supplemented with kanamycin (50 µg/mL), ampicillin (100 µg/mL) and IPTG (0.2 mM) were checked and two clones were shown as two distinct examples in Figs. S1D to S1F and Figs. S1G to S1H, respectively. Controls of mCherry, gfp and nonfluorescent strains were applied. CLSM check has been performed four times by replating the checked clones and afterwards, confident clones (Escherichia coli MG1655::lacIq-Plpp-mCherry-KanR//pDK_DARWIN::pA1O4O3-gfpmut3 strain) were stock prepared and stored at -80 °C.
Figure S2 Gating and event summary of FACS sorting in the permissiveness test.

(A) Overview of the gating used in the sorting experiment. (B) Overview of gating results and statistics for quality control (LB broth). (C) Overview of gating results and statistics for the donor culture (i.e., Escherichia coli MG1655::lacIq-PlpP-mCherry-Kan^R/pDK_DARWIN::pA1O4O3-gfpmut3). (D) Overview of gating results and statistics for the recipient culture (hospital sewer was used as an example). (E) Detailed gating results and statistics for the filter mating sample (transconjugants obtained in the hospital sewer were used as an example).
**Figure S3** Sequence query view of blastn results of pDK_DARWIN*. Three distinct hit regions were discovered in the sequence query view (left, middle and right). The detailed hit sequence information was documented in Table S6.
* Default parameters of blastn (nr_nt) were applied, and the top 100 results were picked up and shown in the figure. Accession numbers of the plasmid sequences were indicated on the left of the numbering. Visualization of the results was performed by Geneious Prime 2022.1 (Biomatters, Auckland, New Zealand) and further integrated with Adobe Illustrator with version 26.3 (Adobe Inc., San Jose, CA, USA).
Figure S4 Taxonomy information for the hosts of the plasmids that hit pDK_DARWIN’s different regions (left, middle and right) as shown in Fig. S3.

This figure was created based on the taxonomy information from Table S6 and plotted for displaying the composition of different genera in each hit region (the size of the bubble in each region stands for the percentage of a specific genus accounting for all taxa in this hit region). The intersection in the Venn diagram represents the same plasmid sequence in two hit regions (left and right). The middle hit region does not have same plasmid sequence with the other two regions.
References


2. Manuscript 2

Insights into the circular: the cryptic plasmidome and its derived antibiotic resistome in the urban water systems

Zhuofeng Yu,* Wanli He,* Joseph Nesme,a,* Franziska Klincke,a Jonas Stenløkke Madsen,a Witold Kot,b,c Lars Hestbjerg Hansen,b,c Arnaud Dechesne,d Barth Smets,d Søren Johannes Sørensen*a,*

aSection of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
bDepartment of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
cDepartment of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
dDepartment of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
Title Page

Title of the paper
Insights into the circular: the cryptic plasmidome and its derived antibiotic resistome in the urban water systems

Authors and affiliations
Zhuofeng Yu, a,*, Wanli He, a,* Joseph Nesme, a,## Franziska Klincke, a Jonas Stenløkke Madsen, a Witold Kot, b,## Lars Hestbjerg Hansen, b,## Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen, a,##

a Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
b Department of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
\textdegree Department of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
d Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

* These authors contributed equally to this work.
## Corresponding Authors.

Competing Interests
The authors declare no other competing financial interests.
2.1 Abstract

Plasmids have been a concern in the dissemination and evolution of antibiotic resistance in the environment. In this study, we investigated the total pool of plasmids (plasmidome) and its derived antibiotic resistance genes (ARGs) in different compartments of urban water systems (UWSs) in three European countries representing different antibiotic usage regimes. We applied a direct plasmidome approach using wet-lab methods to enrich circular DNA in the samples followed by shotgun sequencing and in silico contigs circularisation. We identified 9,538 novel sequences in a total of 10,942 recovered circular plasmids. Of these, 66 were identified as conjugative, 1,896 mobilisable and 8,970 non-mobilisable plasmids. The UWSs’ plasmidome was dominated by small plasmids (≤10 Kbp) representing a broad diversity of mobility (MOB) types and incompatibility (Inc) groups. A shared collection of plasmids from different countries was detected in all treatment compartments, and plasmids could be source-tracked in the UWSs. More than half of the ARGs-encoding plasmids carried mobility genes for mobilisation/conjugation. The richness and abundance of ARGs-encoding plasmids generally decreased with the flow while we observed non-mobilisable ARGs-harbouring plasmids maintained their abundance in the Spanish wastewater treatment plant. Overall, our work unravels the UWS plasmidome is dominated by cryptic (i.e., non-mobilizable, non-typeable and previously unknown) plasmids. Considering that some of these plasmids carried ARGs, were prevalent across three countries and could persist throughout the UWSs compartments, these results should alarm and call for attention.

2.2 Keywords

Urban water systems (UWS)
Circular plasmidome
Source tracking
Antibiotic resistance genes (ARGs)
ARGs harbouring plasmids (AR plasmids)
2.3 Introduction

In the last decades, horizontal gene transfer (HGT) has gained more and more awareness. Reasons behind this are HGT not only potentiates the cargo genes such as antibiotic resistance genes (ARGs) spread among similar and distinct microbes in various environments but even accelerates the occurrence of several multidrug resistant (MDR) organisms in different ecological niches. Currently, plasmids are of particular concern as they provide efficient and rapid transmission capacity. Moreover, other MGEs such as insertion sequences, integrons and gene cassettes as well as transposable elements can constitute the accessory regions of the plasmids. Specifically, conjugative plasmids are of particular interest since they possess several comprehensive genetic characteristics, such as mobility (MOB) genes [including an origin of transfer (oriT), a relaxase, a type IV coupling protein (T4CP)] and mating pair formation genes (i.e. type IV secretion system, T4SS). These features enhance these well-reported plasmids maintenance and mobility within bacterial populations by conjugation. However, plasmid fitness costs to the hosts under different environmental conditions vary remarkably, and plasmid persistence remains mysterious. The total plasmid populations in a specific niche (microbiome or environment) are known as plasmidome. We and others have previously used the term mobilome, but given that mobilome can cause ambiguity for referring to all the MGEs, we use the term plasmidome in the following text.

Various plasmidome approaches have been made in diverse environments, such as cow rumen, rat cecum, human gut, soil, sediments, surface water, groundwater, and extensively in wastewater. However, the methodology in previous studies were compromised in terms of biased targets and lack of needed bioinformatic tools. Conventionally, environmental plasmidome studies were performed under strong selective pressures that alter plasmid dynamics and bias towards phylogenetic evolution of plasmids compared to their indigenous habitats. Traditional cultural-dependent methods (antibiotic screening for resistance plasmids and in situ matings with fluorescent reporters) and cultural-independent techniques [exogenous plasmid isolation, PCR-based replicon typing (PBRT) and relaxase/MOB typing, epicPCR] targets a relatively small part of the entire plasmidome. Sequencing-based methods such as shotgun metagenomics and whole-genome sequencing and single-cell genomics facilitate an unbiased characterization of known and novel plasmids in a sample. However, large quantities of sequencing-yield reads are discarded due to chromosomal contamination in in nontargeted sequencing approaches, which tends to be not cost-effective. Accordingly, there is no consensus on universal methodology for plasmidome study.

Most Europeans reside in urban areas (>70%), and urban water systems (UWSs) play a crucial role in today’s city infrastructure and the water cycle. Nevertheless, with the intensive use of antibiotics for domestic and hospital purposes, wastewater treatment plants (WWTPs) have been considered as
hotspots of antibiotic resistance (AR) and potential reservoirs for ARGs evolution and dissemination to the environment. This is mainly due to the actual goal of WWTPs, which was exclusively to remove the majority of solids, organic matter and nutrients in wastewaters and ensure sanitary effluent discharging to the downstream water bodies, yet not intended to reduce ARGs. Superabundant studies have been made on detecting antibiotics, ARGs, antibiotic-resistant bacteria and ARGs harbouring plasmids (AR plasmids) through PCRs or selective screenings in wastewaters. Meanwhile, researchers have been intrigued by the possible association between microbiome and antibiotic resistome in the WWTPs by using high-throughput qPCR analysis, or combined 16S profiling and metagenomics. Nonetheless, the dynamics of the plasmidome and the plasmid-mediated mobile antibiotic resistome in the UWSs are poorly understood. On the other hand, different countries have specific antibiotic consumption practices, which have been documented as an explanation for distinctive human gut resistomes. Considering the discharge of human excrements to the sewers, the wastewater antibiotic resistome can be influenced by the regional antibiotic usage. Herein, we proposed investigating the UWS plasmidome in a dynamic perspective concerning different compartments during the UWS processing. Denmark (DK), Spain (SP) and the United Kingdom (UK), located in northern, southern and western Europe with country-specific antibiotic use practices, were chosen in this study. Sewage and wastewater samples were collected from the sewer catchments to the WWTP biological treatment process basin (BTP) during summer and winter in 2018, and a direct plasmid DNA metagenomic sequencing strategy skipping the widely used transposon-aided capture (TRACA) and multiple displacement amplification (MDA) was applied. This was in light of methodology studies that suggested direct sequencing of plasmid DNA samples uncovered more larger-sized plasmids. Moreover, we introduced a plasmid purification using an exonuclease plasmid-safe digestion to shear and remove chromosomal DNA before building sequencing libraries. To our best understanding, we present the full UWS plasmidome for the first time, shedding light on how the plasmids and plasmid derived mobile antibiotic resistome migrate in different compartments in the UWSs.
2.4 Materials and Methods

Comparable UWSs from Denmark, Spain and the United Kingdom located in Odense, Santiago de Compostela and Durham County, respectively, were recruited in this study (Fig. S2). At each UWS; hospital sewer (HS), residential sewer (RS), mixed sewer (MS, i.e., the mixture of HS and RS, and was regarded as the WWTP influent), and treated wastewater after the BTP in the WWTP were sampled (Fig. 1A). Specifically, for the BTP facilities used in the WWTPs, biological treatment basins were applied in DK and SP, while biofilter was employed in the UK. Sampling campaigns at the four sites were performed in the winter and summer of 2018 using ISCO automatic samplers for 24-hour flow (50 mL per 5 minutes) in DK and the UK, while 24-hour-time proportional samples in SP (mixing hourly samples according to flow information). Two sewer lines were operated and sampled for HS in DK (sewer lines of two different wards of the same hospital), while a single HS line was sampled in the other two countries. Three replicates per site and season were collected on three consecutive days without rain events. In total (n = 78), 24 samples from the UK, 24 samples from SP, and 30 samples from DK were obtained. Detailed sampling information was recorded in our previous study 47.

All samples were initially cooled with ice on-site, then 100 mL of each sample was spun down by a centrifuge (Eppendorf, Hamburg, Germany) at 10,000 g for 8 minutes at 4 °C in the laboratory. The supernatants were removed, while the pellets were resuspended in 20% of glycerol stock to reach a final volume of 10 mL for storage at -80°C.

The direct plasmidome approach

The direct plasmidome approach applied here refers to our collaborators 9, 44, and was further modified in the lab (Fig. S3). Briefly, samples were pretreated by filtration, vortex and sonication and resuspended in TE buffer. Afterwards, a specific pre-lysis with enzymatic cocktails (lysozyme, mutanolysin, and lysostaphin) was performed to break the Gram-positive cell walls. Then alkaline lysis and plasmid-safe DNase digestion were performed to obtain ‘fine’ plasmid DNA. Quality-checked DNA was proceeded to the library preparation and finally sequenced on an Illumina NextSeq platform with a v2.5 sequencing kit (Illumina, San Diego, CA, USA). The detailed methodology was introduced in Text S1.

Bioinformatic analysis

Processing the raw data

An in-house bioinformatic pipeline Plaspline (version: 1.1, https://github.com/Wanli-HE/Plaspline) for plasmidome analysis was implemented in this study. 150 bp paired-end raw reads from the 78 samples were processed through quality control, assembly for circular contigs (regarded as putative plasmids from then on), identification and removal of redundant circular plasmids (> 90% sequence
identity and > 95% coverage), calculations of plasmid abundance based on reads mapping (normalization by 16S- rRNA gene copy per sample thereafter), and plasmid annotation and classification. Specifically, we consider the presence of a plasmid in one sample when >80% of the plasmid sequence was covered by at least one read. And this was applied in the plasmid source tracking analysis. Meanwhile, RGI (version: 5.1.1) / CARD (version: 3.0.9) was applied for ARG annotations on the plasmids (default parameters were applied, https://card.mcmaster.ca/analyze/rgi).

**PLSDB similarity search**

All the retrieved plasmid sequences were submitted to PLSDB (database version: 2021_06_23; tool version: 0.4.1-386-gd7e4b70b05) and searched by blast+ (version: 2.12.0+) to evaluate plasmid similarity to the input (identity ≥ 0.8 and coverage ≥ 0.7). Plasmid sequences without hits were considered novel plasmids.

**Statistical analysis**

Richness index of the plasmidome was estimated by the “skbio.diversity” function (version: 0.4.2, http://scikit-bio.org/docs/0.4.2/diversity.html) in Python. Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity was applied to evidence the different patterns of plasmid communities’ through found in the different UWS compartments. actors influencing the plasmidome composition in samples such as: countries of origin, season, and UWS compartments were compared by significance tests (Wilcoxon rank-sum test), and plasmidome similarities (ANOSIM analysis) were observed by using python packages “skbio” (version: 0.5.7, http://scikit-bio.org/).

### 2.5 Results

**UWS plasmidome in a nutshell**

After the direct plasmid DNA metagenome sequencing, the raw data were implemented in the circular module of *Plaspline* and passed through the plasmid verification and clustering. We eventually recovered 10,942 non-redundant plasmids 66 conjugative, 1,906 mobilisable and 8,970 non-mobilisable plasmids. Based on the PLSDB similarity search, we discovered 9,538 novel plasmids (ca. 87.2% of the present plasmidome), which expands our knowledge of the UWS plasmidome (Fig. 1B). Non-mobilisable plasmids accounted for the most part in every compartment in the UWSs, followed by mobilisable and conjugative plasmids (Fig. 1C). Meanwhile, we observed a similar pattern with the plasmid database of the association between plasmid size and mobility in Fig. 1D. We illustrated the mean peak rankings of the plasmid length: conjugative plasmids > mobilisable plasmids > non-mobilisable plasmids (only one peak was detected in this study).
Figure 1 General information of the UWS plasmidome.

(A) Schematic of the sewer and wastewater flow (from HS and RS to MS, then from MS to BTP). (B) Number of plasmids recovered in this study compared to PLSDB database v2021_06_23. The intersection in the Venn diagrams represents the hit plasmids of the present plasmidome with PLSDB. (C) Distribution of plasmids in different treatment compartments in the UWSs. (D) Distribution of plasmid size from each mobility group.

Relative abundances of known MOB types (MOB_Q, MOB_V, MOB_F and MOB_P) of plasmids showed a declining trend along with the flow, yet the relative abundances of MOB_unknown steadily grew during the UWS processing (Fig. S4A). On the other hand, the plasmid Inc groups could be arguably divided into three groups with distinct tendency patterns (Fig. S4B). ‘Inc Group 1’ plasmids including several rep_cluster_xxxx showed higher relative abundance in the BTPs. “xxxx” here stands for the miscellaneous numbering of the rep_clusters, which is due to the fact that contemporary in silico analysis hasn’t incorporated these clusters with the canonical established replicon classification system. ‘Inc Group 2’ including IncQ1, IncQ2, IncI1, IncFII possessed the most diverse
Inc categories, and their relative abundances mostly showed a decreasing propensity with the UWS processing. ‘Inc Group 3’ such as IncW showed superior relative abundance in the sewer lines, however, were likely to be removed in the BTPs.

Figure 2  Plasmidome diversity and distribution in the UWSs.

(A) The dynamic plasmidome richness along with the UWS processing. (B) PCoA analysis based on BrayCurtis distances for the UWS plasmidome from different UWS compartments in each country. HS_a and HS_b were the two sewer lines operated in Denmark. (C) PCoA analysis based on BrayCurtis distances for the UWS plasmidome from different countries in each UWS compartment.

Meanwhile, we perceived a general decreasing trend in plasmid richness with the flow regardless of country (Fig. 2A). There was consistently retained high plasmid richness in the sewers (HS, RS and MS) in Spain with plasmid richness in the Spanish RS being significantly ($p < 0.01$) higher than in the Danish and British RS. Overall, Denmark held the lowest plasmid richness throughout the UWS processing. It is noteworthy that the British BTP remained significantly ($p < 0.01$) higher plasmid...
richness compared to the other two countries. Basically, the plasmids in HS and RS were dissimilar except in Spain, MS was shaped by RS, and BTP plasmids were remarkably different from the sewers (Fig. 2B). Interestingly, the plasmid populations in the three countries appeared totally different in the separated compartments (Fig. 2C).

The common and persistent plasmids in the UWSs
Given the background of the globalization of ARGs, we assumed plasmids were also internationally spread and ‘shared’ by different countries. In this study, we discovered a shared pool of plasmids in every treatment compartment (Fig. 3A). The ‘biggest share’ was found in HS (n = 2,000), followed by MS (n = 1,857), RS (n = 1,776) and BTP (n = 67). Spain and the United Kingdom consistently shared the most plasmids in HS, MS and BTP, while Spain and Denmark shared more plasmids in RS. Notably, Spain possessed the highest number of country-distinct plasmids in all the sewers, while the United Kingdom owned the most distinct plasmids in BTP (Fig. 3A).

Based on the knowledge of the richness and abundance of most plasmids declined with the UWS processing (Fig. 2A and Fig. S4B), we wondered about the fate of a specific group (mobility) of plasmids. Therefore, we detected the plasmidome trajectory, i.e., the plasmid populations flowing from an upstream treatment stage to a downstream treatment stage in the UWSs. The plasmidome trajectories of different mobility groups from HS and RS to MS, and from MS to BTP were studied (Fig. 3B). RS predominantly contributed to the plasmid flow to MS in Spain and Denmark, however, HS played a bigger role in the United Kingdom for transporting plasmids to MS. And the majority (> 75%) of the plasmid flows were composed of non-mobilisable plasmids in the trajectories of HS to MS and RS to MS. Considering the striking plasmid richness in the British BTP (Fig. 2A), we were curious about where these plasmids came from and what could be the reasoning. The plasmidome trajectory showed these British BTP plasmids were dominantly occupied by non-mobilisable plasmids. Nevertheless, these BTP plasmids mainly originated from either RS or HS, or alternatively, the indigenous MS plasmid populations remained a question. Thereby, a source-tracking analysis was performed for the BTP plasmids. We found that 672, 131 and 1,612 BTP plasmids were trackable to the upstream sewer compartments in Spain, Denmark and the United Kingdom, respectively (Fig. 3C). 337, 87 and 488 plasmids were universally propagated in the sewer catchments in Spain, Denmark and the United Kingdom, respectively, and finally flowed to the BTP in each country. Intriguingly, MS was the key BTP plasmids contributor regardless of country (Fig. 3C).
Figure 3 The plasmidome persistence in the UWSs.

(A) Common plasmids shared by different countries in each UWS compartment. (B) Plasmid trajectories in the UWSs. (C) Source distributions of the source-trackable BTP plasmids in each country.

Numbers inside of the Venn diagram of Fig. 3A represent the percentage of the total number of shared plasmids. The numbers in Fig. 3B and the numbers inside the Venn diagram of Fig. 3C represent the absolute plasmid numbers. In Fig. 3B, the suffix ‘Mob’ and ‘Non’ after ‘HS’, ‘RS’, ‘MS’ and ‘BTP’ mean ‘plasmids with mobility (i.e., mobilisable or conjugative)’ and ‘plasmids without mobility (i.e., non-mobilisable)’, respectively.

In Fig. 3A, the intersection in the Venn diagrams represents the commonly shared plasmids between/among countries. In Fig. 3C, the intersection in the Venn diagrams represents the common plasmids found in both/all sewers that also showed occurrence in BTP.
The UWS plasmidome associated antibiotic resistome

AR plasmids

Non-mobilisable plasmids took a large part (82.1%) in the present UWS plasmidome, while in the identified 216 AR plasmids, we observed a high ratio (52.3%) of plasmids with mobility (Fig. 4A). Plasmid trajectory analysis reveals the persistence features of AR plasmids in the UWSs (Fig. 4B). The sewer sources (HS and RS) both contributed considerably to the end of sewer (MS) in the three countries, however, AR plasmids could hardly flow to the BTP in Spain and Denmark. In light of the tremendous plasmid populations transported from MS to BTP in the United Kingdom (Fig. 3B), we also detected remarkable AR plasmids that were delivered to the British BTP.

Figure 4 The general pattern of the AR plasmids in the UWSs.

(A) Mobility of the UWS plasmidome with comparison to the PLSDB database. (B) AR plasmid trajectories in the UWSs. The suffix ‘Mob’ and ‘Non’ after ‘HS’, ‘RS’, ‘MS’ and ‘BTP’ mean ‘AR plasmids with mobility (i.e., mobilisable or conjugative)’ and ‘AR plasmids without mobility (i.e., non-mobilisable)’, respectively. (C) The dynamic AR plasmid abundances along with the UWS processing.

By recruiting the microbiome dataset of the same sampling campaign 47, the plasmidome abundance was normalized per 16S-rRNA (Fig. 4C). We illustrated the mobilisable/conjugative AR plasmids
abundance generally dropped along with the flow in all countries. In Denmark and the United Kingdom, non-mobilisable AR plasmids consistently kept at a low level of abundance in RS, MS and BTP though it was comparable in HS with Spain. In contrast, the AR plasmid abundance was promoted in the Spanish MS from the sewer sources (HS and RS), and the BTP retained a similar level of abundance with MS in Spain.

**Figure 5** Distributions and abundances of the plasmid-borne ARGs in the UWSs.  
(A) PCoA analysis based on BrayCurtis distances for the plasmid-borne ARGs from different UWS compartments in each country. (B) PCoA analysis based on BrayCurtis distances for plasmid-borne ARGs from different countries in each UWS compartment. Specifically, in the figure of DK, the yellow and blue squares represent the two hospital sewer lines operated in Denmark. (C) The dynamic plasmid-borne ARG abundances along with the UWS processing. ARG family abbreviations: aph (aminoglycoside phosphotransferase), tet (tetracycline-resistant ribosomal protection protein), cat
(chloramphenicol acetyltransferase), bla (beta-lactamase), mph (macrolide phosphotransferase), ant (aminoglycoside nucleotidylyltransferase), abc-f (miscellaneous abc-f subfamily ATP-binding cassette ribosomal protection proteins), erm (erm 23s ribosomal RNA methyltransferase), aac (aminoglycoside acetyltransferase), lnu (lincomamide nucleotidylyltransferase), sul (sulfonamide resistant sul), gly (glycopeptide resistance gene cluster), qnr (quinolone resistance protein).

Plasmid-borne ARGs

Plasmid-borne ARGs were dragged and analyzed for exploring the plasmidome derived resistome. Basically, ARGs in RS and MS were similar in every country, and HS seemed also closely connected to other sewers in Spain (Fig. 5A). A seeming country coding pattern of ARGs was discovered in the sewer compartments, and it appeared the ARG groups from different countries were more distinct in the raw sewers (HS and RS) than in MS (Fig. 5B). We then examined the abundance of a specific family of the plasmid-borne ARGs (Fig. 5C). In Denmark and the United Kingdom, RS always harboured the lowest abundance of ARGs except for tet, lnu and qnr genes, whilst RS carried various and prosperous ARGs in Spain. Generally, Spain harboured the most diverse and abundant ARGs in the sewer catchments. United Kingdom showed unique high ARG abundances of bla, mph, ant and abc-f families in BTP. Denmark steadily exhibited low (or under detection threshold) ARG abundances except for tet, bla and erm. Particularly, qnr genes were exclusively detected in the Danish raw sewers (HS and RS).

2.6 Discussion

The state-of-art plasmidome methodology

With the present plasmidome methodology and bioinformatic analysis, we didn’t detect any significant changes in the abundance of plasmid per sample (Wilcoxon rank-sum test, \( p = 0.12 \)) from the two sampling seasons, therefore we combined them as replicates. The plasmid functional transfer system generally requires at least 2 Kbp length for mobilisation and ca. 15 Kbp for conjugation. Plasmids with larger sizes are normally of more interest considering HGT possibilities, and capacity for harbouring various cargo genes such as ARGs. Smillie et al. elaborated the association between plasmid size and mobility with detection of the distribution of different plasmids in the database according to plasmid size. Following their findings, we displayed the UWS plasmidome exhibiting a similar pattern (Fig. 1D), which somehow supports the feasibility and reliability of this work.

Aiming to improve the detection of larger-sized circular elements (>10 Kbp), we applied the direct plasmid DNA metagenome sequencing approach excluding TRACA and MDA. However, we
harvested considerable small-sized (≤10 Kbp) plasmid contigs (Fig. 1D). This may be caused by the fact that the target UWS plasmidome was substantially rich in smaller genetic elements, while one can also argue the fact that not recovering large-sized plasmids is due to sequence depth and proper assembly tools. One groundwater plasmidome study showed the capture of extremely large plasmids (>1 Mbp) using MDA with a deep sequencing strategy \textsuperscript{15}. This urges a necessity to perform deeper sequencing and improve the contemporary bioinformatic tools for the sake of a higher resolution for the large-sized fraction of the plasmidome.

**Uncovering the cryptic UWS plasmidome**

Though plasmids can be classified in various ways, the genetic-traits-based grouping methods such as MOB typing and replicon typing are universally applied \textsuperscript{51}. We showed the UWS plasmidome was basically a blend of plasmids with MOB\textsubscript{unknown} (undivided MOB) types (Fig. S4A). This suggests the nature of UWS plasmidome and a recent report indicated that the contemporary plasmid database contained large quantities of MOB\textsubscript{less} plasmids \textsuperscript{52}. Apart from the MOB typing, traditional plasmid Inc grouping, defined by conjugation experiment data, is also used from time to time given that it is the most consistent typing method with the plasmid mobility scheme \textsuperscript{4}. With the expansion of the whole genome sequencing data, the Inc group database tethering with replicon information gets rapidly developed, and many *in silico* typing tools are publicly accessible \textsuperscript{4,51,53,54}. More than 28 Inc groups have been recognised so far, and many contain subgroups. Our study uncovered that plasmids belonging to the miscellaneous rep\_clusters were the most prevalent groups in the UWSs (Fig. S4B), which had not been reported by conventional PBRT or metagenome studies in the wastewater environment \textsuperscript{15,55,56}. Taken together, our work showed the UWS plasmidome were dominated by these cryptic plasmids, which were non-mobilizable, not typeable under the current MOB and Inc typing scheme and previously unknown (i.e., novel) plasmids.

In addition, we perceived that the UWS plasmidome was essentially a collection of non-mobilisable plasmids, whereas PLSDB showed a prevailing pattern of plasmids with mobility (mobilisable/conjugative) in the database (Fig. 4A). This can be attributed to the fact that mobilisable and conjugative plasmids normally carry more functional traits or specific features than non-mobilisable plasmids, and they take critical roles in HGT for spreading ARGs. These conspicuous characteristics lead to massive “mobilisable plasmids” and “conjugative plasmids” oriented projects and thus, their sequences constituted a major part of the PLSDB database.

**The dynamic UWS plasmidome**

As a critical step in urban water management, WWTPs have been of particular concern in environmental studies for years \textsuperscript{23,26}. However, there’s a lack of examining the plasmidome dynamics along with the UWS processing due to method biases (targeting only some MGE marker genes) or...
incomplete sampling construction\textsuperscript{27, 33, 38, 56, 57, 58}. This study targeted the sewer catchments (HS, RS and MS) and the BTPs in the WWTPs. It is arguably true that the samples from the front- and post-BTPs (e.g., the primary settler, the secondary settler, and the tertiary filter) and the upstream and downstream rivers were also worth surveys to obtain a full-frame image of the fate of UWS plasmidome. Unfortunately, we met difficulties building the plasmidome libraries for sequencing these sites. Within this sampling scope, MS was sampled after the converge of the raw sewer lines and was supposed to be the convergence of HS and RS. Consequently, it is reasonable to speculate that MS contained the highest plasmid richness and abundance. However, this was not the common feature in the present study (Figs. 2A and S5B). The plasmid richness in MS mainly maintained the level of RS, while lower than HS. With regard to the plasmid abundance, it was hard to conclude the general pattern for this sophisticated plasmidome, whilst we observed increased plasmid abundance in MS compared to HS and RS in a few “Inc Group 2” plasmids with miscellaneous rep\_cluster\_xxxx replicons. Apart from this, we discovered that all the sewer lines (HS, RS and MS) carried significantly ($p < 0.01$) higher richness of plasmids than BTPs (Fig. 2A). This could be explained by the dramatic environmental changes for the plasmid hosts flowing from the sewer lines to the WWTPs, while plasmid maintenance is closely related to its host$^6$, and BTPs can remove several microbes$^{59, 60, 61, 62, 63}$. Strikingly, the British BTP retained significantly ($p < 0.01$) higher plasmid richness than the other two countries, which was probably owing to the different BTP facilities employed in the United Kingdom (biofilter) compared to Denmark and Spain (biological treatment basin). The biofilter is an aerobic process, which favours the growth of specific groups of microbes (plasmid carriers), whereas the anaerobic treatment process was less suitable for the persistence of those plasmid harbouring bacteria.

HS and RS can be considered as separated niches given that distinct selection pressure is exerted in the corresponding environment. And they usually possess different microbial community compositions and plasmid profiles considering the different sewer physicochemical properties$^{32, 35}$. Herein, we showed the distinct HS and RS plasmid populations in Denmark and the United Kingdom (Fig. 2B). Nonetheless, we discovered a close relationship between the plasmid populations in the Spanish HS and RS, which could be attributed to the extensive domestic use of antibiotics in Spain (Fig. S1), and indeed the Spanish HS plasmidome appeared distinct compared to Denmark and the United Kingdom (Fig. 2C). Interestingly, we figured out that the MS plasmidome was closely related to that in RS regardless of country. This indicates the MS plasmid communities might be notably shaped by the injection of RS plasmids with the flow.

On the other hand, we illustrated a clear routine of ‘vanishment’ for the HS-distinct plasmids, whereas RS-distinct and MS-distinct plasmids were not totally removed in the BTP (Fig. S5A). Nonetheless,
the HS-distinct plasmids should not be overlooked as hospital wastewaters were detected as pools of plasmids with mobility\textsuperscript{64}. In view of the mobility features carried by these HS-distinct plasmids, it is reasonable to speculate that the harboured ARGs might be already widely disseminated in the sewer microbiome\textsuperscript{35, 64, 65}.

**The common UWS plasmidome in the pan-European countries**

Recently, a network analysis has revealed the shared plasmid clusters of gut microbiomes from distinct human groups\textsuperscript{66}. This implies plasmids could be spread and shared among niches in despite of spatial-temporal differences. In this study, we showed the constantly evolving plasmid community structure along with the UWS processing as well as the documentation of different UWSs in different European countries. We revealed that plasmids from different countries were shared in every UWS stage (Fig. 3A). And we noticed a ‘bigger’ share in the sewers than in the BTPs, which was probably due to the distinct treatment settings implemented in the United Kingdom. Predecessors reported country-specific antibiotic resistance profiles in seven different European WWTPs\textsuperscript{67}. Our finding of the common pool of pan-European wastewater-derived plasmids alarms the bell since their emergence implied they probably have already been widespread in the urban water cycles in the whole of Europe and would potentially acquire ARGs from those AR plasmids through HGT and create cross-national resistance issues\textsuperscript{68, 69}.

**The persistent plasmidome in the UWSs**

Plasmid persistence is constantly a hot topic in the field and is frequently reported to infer its potential mechanism in diverse environments\textsuperscript{6}. In the present UWS plasmidome, we investigated the emergence and prevalence of plasmid persistence using a plasmidome trajectory model (Fig. 3B). We observed the persistent plasmidome was mainly composed of non-mobilisable plasmids. As known, plasmids can lower their mobility, become less dependent on mobilization and conjugation to be maintained in the bacterial communities\textsuperscript{70}. Meanwhile, compensatory mutations can happen to alleviate plasmid fitness costs to hosts\textsuperscript{71}. In view of the variations of selection pressure in different compartments of the UWSs, we assumed the plasmid groups with mobility gradually lost competitive advantages with the flow, resulting in non-mobilisable plasmids outcompeting mobilisable and conjugative plasmids. Alternatively, another explanation can be the overwhelming number of non-mobilisable plasmids in the raw sewers, and the reduction during sewer transportation and WWTP processing to them was minor.

**Source tracking of the BTP plasmids**

Considering the blossom of persistent plasmids in the UWSs, we were aware of their occurrence in the BTPs. BTPs are of particular importance because they constitute a potential mobile genetic and antibiotic resistance reservoir to the WWTP effluent and downstream rivers\textsuperscript{28, 29, 72}. Accordingly,
tracking the origin of the BTP plasmids would help us understand its composition (intrinsic vs nomadic) and the gateway to prevent specific groups of plasmids from flowing to the WWTPs. In the present study, a multitude of BTP plasmids was trackable to the front treatment compartments (Fig. 3C). And the majority of these trackable BTP plasmids were from a common pool contributed by the three sewer lines. Further, we noticed that the trackable BTP plasmids were dominantly contributed by MS. This seems plausible when considering MS was the confluence of HS and RS, as well as the direct inflow to the WWTP.

To this point, we imagined these ubiquitous plasmids were common goods shipped from different origins (HS, RS) to the transit hub (MS) and finally deposited in the retention centre (BTP). Meanwhile, they resembled goodies and were preserved at each site. Given that the BTP plasmids were essentially (>70%) occupied by these nomadic plasmids, we anticipated they might offer their hosts beneficial traits to be retained in the UWSs and thus became prevalent. Alternatively, these plasmids might luckily taxi on those durable microbes which survived along the sewer pipelines till the BTPs. Another hypothesis can be that the backbones of these persistent plasmids were not changed while some genetic contexts alternation/rearrangement/movement happened (compensatory mutation), for instance, accessory genes (e.g. mobility genes, ARGs) could be discarded to lower plasmid fitness costs to the hosts. In other words, plasmids can systematically adapt the cargo genes for specific environments and the variability of plasmid fitness effects promotes plasmid persistence in bacterial communities. Our findings of such ubiquitous plasmids in the UWSs paved the way toward understanding the biological and ecological facts of the BTP plasmids. Further studies on these plasmids are necessary to explore their genetic repertoire, which could be pathogen-related and poses health risks to the downstream rivers.

**Fate of the AR plasmids in the UWSs**

WWTPs are well-characterized hotspots of antibiotic resistance though they arguably remove antibiotics, ARGs and antibiotic-resistant bacteria to some extent. A pan-European survey of the UWSs shows the WWTP effluent were responsible for the antibiotic resistome enrichment in the receiving water bodies. This suggests the potential health risks posed by WWTPs and the significance of monitoring the antibiotic resistance issues along with the UWS processing to track the crucial barrier (compartment) for prevention.

In general, we found a dissimilar pattern of AR plasmids in RS and HS, and a close relationship between AR plasmids in RS and MS (Fig. S6A). These could be due to the distinct nature of different sewers, and the confluence impacts from the upstreams, respectively. RS derives from the urban residential area and is consistently a key source of antibiotics, ARGs and antibiotic-resistant bacteria flowing to the WWTPs. HS originates in the hospitals, and it appears plausible that HS
harbours higher ARG abundance than RS given a more intense amount of antibiotics are used in the clinical environment. However, Buelow et al. reported that HS could be neglected compared to RS as not contributing significantly to the quantity and diversity of ARGs in the UWSs. This contradiction typically emerges in case studies and can be explained by differences in antibiotic usage, local wastewater effluent load, hydrological conditions, sampling strategies, etc. The message from these studies is that HS and RS are both essential contributors to the wastewater antibiotic resistome in the UWSs. However, this point has been buried in numerous studies.

Antibiotic resistance issues in different countries vary considering differing antibiotic usage and practices. Similarly, we were able to cluster the AR plasmid populations by country in each sewer catchment. Specifically, Spanish and British AR plasmid populations tended to be similar, yet dissimilar to those in Denmark. Regarding the dynamics of the AR plasmids, we discovered:

1. In Denmark, HS harboured a relatively high AR plasmid abundance compared to RS, and flowed comparable numbers of AR plasmids with RS to MS. However, all AR plasmids in Denmark were eliminated in the BTP, which could be associated with the strict discharge regulations in the local environment.

2. In Spain, we discovered similar numbers of AR plasmids with and without mobility flowed from sewer sources to MS and then to BTP. Plasmids can act as selfish parasites, while plasmid-encoded ARG traits, in turn, promote the stability, maintenance and transmission of themselves. And AR plasmids with mobility have continuously been headlines in wastewater studies. A remarkable number of AR plasmids in the Spanish UWS plasmidome carried mobility genes, whereas they seemingly got removed in the BTP in view of abundance. And a group of stubborn non-mobilisable AR plasmids in the Spanish BTP was found to retain a similar plasmid abundance with MS. This could be due to the fact that Spain consumes high amounts of antibiotics in the community and hospital sectors and higher concentrations or relative abundances of antibiotics, antibiotic-resistant bacteria and ARGs were detected in the Spanish wastewaters than in western and northern European countries. Therefore, the high selection pressure in the Spanish BTP favours plasmids encoding ARGs, and these non-mobilisable AR plasmids were preserved for the sake of beneficial traits for survival and lower fitness costs compared to mobilisable/conjugative plasmids.

3. In the United Kingdom, we discovered a similar removal pattern of AR plasmids with the UWS processing although larger numbers of AR plasmids flowed from MS to BTP than in the other two countries.
Overall, the non-mobilisable AR plasmids remained abundant along with the UWS processing in Spain. These AR plasmids were not strictly settled in fixed niches but could be disseminated to other environments with the WWTP processing, while their prevalence and dominance depend on the trade-off of plasmid fitness costs and benefits to the hosts and abiotic factors (e.g., selective pressure), which demands further investigations to understand their genetic evolution and possible infection to naive communities though they were without mobility 52.

**The plasmidome derived antibiotic resistome**

Plasmids play a key role in bacterial ecology and evolution, moreover, the synteny features of the ARG loci on the plasmid contigs bring the possibility of ARG horizontally transferring to pathogens 68, 69. Our high-throughput qPCR-based resistome network study for the same sampling campaign demonstrated a significant ($p < 0.05$) correlation among the extended-spectrum beta-lactamase and carbapenemase genes and MGEs marker genes 47. And researchers showed that the evolution of resistome was mainly associated with the interaction of microbial communities and MGEs 35, 36.

In the present work, we inspected the antibiotic resistome in this circular plasmidome dataset and discovered similar country and treatment compartment patterns with the AR plasmids (Figs. 5A and 5B). Sewer-distinct ARGs were all eliminated after the UWS processing (Fig. S5B). Aminoglycoside, beta-lactam, macrolide, chloramphenicol, and tetracycline resistance genes were ubiquitous in the sewer lines while they were almost removed in the BTP except in the United Kingdom (Fig. 5C). These ARG types have been also frequently reported in previous wastewater studies 24, 27, 34, 39, which reflects the plasmid derived resistome contributes to the total resistome. Generally, Spanish RS harboured relatively higher abundant ARGs than the other two countries, which was probably due to the excessive community usage of antibiotics (Fig. S1). And the occurrence of ARGs belonging to *bla* (including *bla*OXA, *bla*SCO and *bla*TEM variants), *mph*, *ant* and *abc*-f in the British BTP suggests the incapability of reducing these ARGs using biofilters. ERM genes showed surprisingly high abundance in the Danish HS, which would be associated with the local clinical practices. Generally, the wastewater ARG abundance mirrored the pattern of antibiotic use, environmental temperature and WWTP size 67. Our results implied the treatment facilities also played an important role in shaping the highly mobile resistome in the UWSs.

Further investigations into those persistent mobilisable/conjugative AR plasmids are urgent to be carried out given their potential horizontal transmission roles in the UWS downstream rivers. Out of research proity, we exclusively focused on the circular plasmidome in this study. Meanwhile, inspecting the dataset on the linear contigs would also potentially unearth their encoded ARGs and provide a full resistome view of the plasmidome.
2.7 Acknowledgements

This research was funded by a Joint Programming Initiative-Antimicrobial Resistance grant (JPI-AMR; DARWIN project #7044-00004B) to BFS, and the DFF-Research Project 2 Grants from the Danish Council for Independent Research | Technology and Production (7017-00210A; SandBAR project) to BFS.

We appreciate Dr. Marcos Quintela-Baluja (Newcastle University, UK) and Dr. Sabela Balboa Méndez (University of Santiago de Compostela, Spain) for the sampling campaigns and wastewater physical-chemical parameters detection in the United Kingdom and Spain, respectively. ZFY wants to thank Xiao Peng (Institute of Microbiology, Chinese Academy of Sciences, China), Dr. Hanadi Ananbeh (Mendel University in Brno, Czech Republic) and Asmus Kalckar Olesen (University of Copenhagen, Denmark) for assistance in sample pre-treatment.

2.8 Contributions

ZFY performed the samples pre-treatment and sequencing library preparation, as well as writing the manuscript and modifying the figures; WLH processed the bioinformatic and statistical analysis, and produced the incorporated figures; JN gave recommendations for the sampling campaigns and the experimental work; JN, FK, JSM, AD, BS and SJS gave suggestions for the bioinformatical analysis; FK built the MGEs database; WK and LHH shared experience in plasmidome methodology, performed the plasmidome sequencing and produced the raw reads; AD constructed the sampling campaign in Denmark, organised regular meetings and collected data for the metadata file; BS and SJS made the initial proposal of the sampling campaign; all the co-authors have reviewed and gave suggestions for this manuscript.
2.9 References

15. Kothari A, Wu YW, Chandaonia JM, Charrier M, Rajeev L, Rocha AM et al. Large circular plasmids from groundwater plasmidomes span multiple incompatibility groups and are enriched in multimetal resistance genes. *mBio* 2019; **10**: e02899-02818.


2.10 Supporting Information

Supporting Information for
Insights into the circular: the cryptic plasmidome and its derived antibiotic resistome in the urban water systems

Authors and affiliations
Zhuofeng Yu,a,* Wanli He,a,* Joseph Nesme,a,# Franziska Klincke,a Jonas Stenløkke Madsen,a Witold Kot,b,c Lars Hestbjerg Hansen,b,c Arnaud Dechesne,d Barth Smets,d Søren Johannes Sørensen,a,#

aSection of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
bDepartment of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
cDepartment of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
dDepartment of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

* These authors contributed equally to this work.
# Corresponding Authors.
The direct plasmidome approach.

Sewer/wastewater samples were primarily thawed and homogenised on ice for 30 minutes, then 2.5 mL of each sample was transferred to a new sterile 15 mL Falcon tube (Corning, New York, NY, USA). Subsequently, 7.5 mL of PBS (Sigma-Aldrich, St. Louis, MO, USA) and 10 sterile metallic beads were added. Vortex (5 minutes) and sonication (10 minutes) were followed. The pretreated samples were then filtered through a sterile paper filter (Frisenette Aps, Knebel, Denmark) with the help of a pump. 0.5 mL of the filtrate was used for cell counting by a FACS Aria™ III cell sorter (BD, Franklin Lakes, NJ, USA), while the rest filtrate was pellet down by an Eppendorf centrifuge (8000 rcf, 7 minutes, 4 ºC).

The resultant pellets were primarily resuspended in 400 μL of 1 × TE buffer (Sigma-Aldrich) and then lysed with an enzyme cocktail targeting Gram-positive cell walls with 50 μL of lysozyme (10 mg/mL, Sigma-Aldrich), 6 μL of mutanolysin (25 KU/mL, Sigma-Aldrich), and 3 μL of lysostaphin (4 KU/mL, Sigma-Aldrich). Lysis was executed for one hour at 37 ºC, and the lysed samples were then plasmid isolated using a Plasmid mini AX kit (A&A Biotechnology, Gdynia, Poland). Specifically, doubled volumes of the A&A mini AX kit-inclusive cell suspension solution, lysis solution and neutralising solution were applied. Harvested DNA pellets were dissolved in 50 µL of Elution Buffer (5 mM Tris/HCl, pH 8.5; MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany). Afterwards, exonuclease plasmid-safe digestion was performed to remove residual chromosomal DNA. A reaction system of 42 μL of extracted DNA, 2 μL of 25 mM ATP (Lucigen, Madison, WI, USA), 5 μL of 10 × Reaction Buffer (Lucigen) and 1 μL of Plasmid-Safe DNase (10 U, Lucigen) were made. The digestion was done by incubating the mixture at 37°C for 16 hours and then inactivated at 70°C for 30 minutes. Purified plasmid DNA were stored at -20°C before further experiments. Meanwhile, all the samples going through plasmid DNA extraction and purification steps were checked for the degree of chromosomal DNA contamination by qPCR targeting the v1-v9 region of the bacterial 16S-rRNA gene.

Digested plasmid DNA was library prepared with a Nextera XT DNA Library Preparation kit (Illumina, San Diego, CA, USA) and indexed with a Nextera XT Index Kit v2 of set A (Illumina). Cleaned libraries were quality checked with a Fragment Analyser (Advanced Analytical Technologies, now a part of Agilent Technologies, Santa Clara, CA, USA). Normalised libraries were pooled and loaded on an Illumina NextSeq platform with a v2.5 sequencing kit (Illumina).
Figure S1 Mean antibiotic consumption data of Spain (SP), Denmark (DK) and the United Kingdom (UK) from 2016 to 2018.

Antibiotics consumption for systemic use (ATC group J01) in the community (primary care sector) and the hospital sector were recorded and shown in DDD per 1000 inhabitants and per day. This reference data from the ESAC-Net interactive database for Spain, Denmark and the United Kingdom on antimicrobial consumption were retrieved from the European Centre for Disease Prevention and Control website¹. Data from 2016 to 2018 were downloaded and exploited because the sampling campaign was performed in 2018, and data was not available for Spanish hospital sector antibiotic usage before the year 2016.
**Figure S2** Diagram for the three comparable urban water systems and sampling locations*.

* Some vectors used in Figure S2 were downloaded from “https://stock.adobe.com” under licenses of Adobe Stock (Adobe Inc., San Jose, CA, USA).

Sampling sites are marked in orange circles. Hospital sewers were transported using a single sewer line system in Spain and the United Kingdom. Whereas in Denmark, two sewer lines in parallel were applied. Biological treatment basins (anaerobic digestion tanks) were involved in the Spanish and Danish wastewater treatment plants as the biological treatment process, while a biofilter was employed in the United Kingdom. After the secondary settler, British and Spanish wastewaters were directly discharged to the nearby rivers, and Danish wastewater went through a tertiary filter before entering the downstream river. Details about the urban water systems (UWSs) were recorded in a previous study².
Figure S3 Workflow of the direct plasmidome approach.

(A) Workflow for sample pre-treatment. (B) Workflow for plasmid DNA isolation and purification. (C) Workflow for the direct plasmidome sequencing.

1 “Sewer/wastewater samples” refer to the samples taken from the constructed sampling sites (orange-marked) in Fig. S2.

2 The enzyme cocktails contain 50 μL of lysozyme (10 mg/mL, Sigma-Aldrich), 6 μL of mutanolysin (25 KU/mL, Sigma-Aldrich), and 3 μL of lysostaphin (4 KU/mL, Sigma-Aldrich).

3 qPCR was performed targeting the v1-v9 region of bacterial 16S-rRNA gene using primers Uni27F and Uni1492R 

4 An Advanced Analytical Technologies Fragment Analyzer (now part of Agilent Technologies) was used for library checks.

* This figure was primarily created on “https://biorender.com” with credit to BioRender (BioRender, Toronto, ON, Canada) and modified in Adobe Illustrator with version 26.3 (Adobe Inc.).
Figure S4 Plasmid classifications of the UWS plasmidome.

(A) Relative abundance of plasmids of different MOB types along with the UWS processing. (B) Heatmap of different Inc groups of plasmids according to relative abundances clustering (plasmids of $\geq 0.5\%$ relative abundance were plotted).
**Figure S5** Fate of distinct plasmids and distinct plasmid-encoding ARGs in the UWSs.

(A) The dynamic abundance of sewer-distinct plasmids along with the UWS processing. (B) The dynamic abundance of sewer-distinct plasmid-encoding ARGs along with the UWS processing.

‘HS-distinct’ stands for the plasmids exclusively harbouring in HS originally (in terms of positioning in the UWS), couldn’t be found in RS, while could possibly flow to MS and/or BTP with the flow. ‘RS-distinct’ stands for the plasmids exclusively harbouring in RS originally (in terms of positioning in the UWS), couldn’t be found in HS, while could possibly flow to MS and/or BTP with the flow. ‘MS-
distinct’ stands for the plasmids exclusively harbouring in MS originally (in terms of positioning in the UWS), could neither be found in HS nor RS, while could possibly flow to BTP with the flow.
Figure S6 AR plasmid distribution in the UWSs.

(A) PCoA analysis based on BrayCurtis distances for the AR plasmids from different UWS compartments in each country. HS_a and HS_b were the two hospital sewer lines operated in Denmark. (B) PCoA analysis based on BrayCurtis distances for the AR plasmids from different countries in each UWS compartment.
References

3. Manuscript 3

Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems

Zhuofeng Yu,a,* Wanli He,a,* Joseph Nesme,a,* Franziska Klincke,a Jonas Stenløkke Madsen,a Witold Kot,b,c
Lars Hestbjerg Hansen,b,c Arnaud Dechesne,d Barth Smets,d Søren Johannes Sørensen,a,*

aSection of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
bDepartment of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
cDepartment of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
dDepartment of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
Title of the paper

Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems

Authors and affiliations

Zhuofeng Yu,⁎⁎ Wanli He,⁎⁎ Joseph Nesme,⁎⁎ Franziska Klincke,⁎ Jonas Stenløkke Madsen,⁎ Witold Kot,⁎⁎ Lars Hestbjerg Hansen,⁎⁎ Arnaud Dechesne,⁎⁎ Barth Smets,⁎⁎ Søren Johannes Sørensen⁎⁎

⁎Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
⁎Department of Environmental Science, Aarhus University, Frederiksbergvej 399, DK-4000 Roskilde, Denmark
⁎Department of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
⁎⁎Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

⁎These authors contributed equally to this work.
⁎⁎Corresponding Authors.

Competing Interests

The authors declare no other competing financial interests.
3.1 Abstract
Mobile genetic elements such as plasmids drive the dissemination and evolution of antibiotic resistance in the microbiome. And plasmid-mediated antibiotic resistance has been a growing concern in the wastewaters since urban water systems (UWSs) are known to breed and spread antibiotic resistance to the downstream rivers, which constitutes environmental exposure risks to human health. Here we investigated the plasmid derived antibiotic resistance at different treatment compartments in three comparable UWSs located in Spain, Denmark, and the United Kingdom using direct plasmid DNA metagenome sequencing. We assembled contigs (circular and linear) and removed chromosome ones using bioinformatic tools. Thereby, a plasmidome dataset was generated and analysed. We identified 225 different ARGs belonging to 180 groups of ARG families in the UWS plasmidome. Genes conferring resistance to aminoglycoside, tetracycline, macrolide and phenicol drug contributed to more than half of the plasmidome resistome abundance across samples. The plasmid resistome richness and relative abundance detected in Spanish UWS samples were significantly ($p < 0.05$) higher than the other studied countries, which mirrors the remarkable domestic antibiotic use in Spain. And we only detected significant ($p < 0.05$) differences in ARG risk scores between the hospital and residential sewers in Spain. Meanwhile, different sewer compartments showed a partitioning role for the resistome richness and abundance distributions. Intriguingly, we perceived a group of shared ARGs among the three countries regardless of treatment stages. Further, ≥80% of ARG types in the wastewater treatment plants could be found in the sewer sources, which implies these ARGs were persistent in the UWSs. Overall, this study shed light on the plasmidome derived resistome as an important part of the total resistome, and this resistome is shaped by geographic-regional and UWS-sectional variations in the UWS environment.

3.2 Keywords
Urban water systems (UWS)
Plasmidome
Plasmidome derived antibiotic resistome
Geographic regions
UWS compartments
3.3 Introduction

Unban water systems (UWSs) are designed to evacuate wastes from the human activity area (such as hospitals and residential places) to low human exposure areas and gradually reinstate them into natural watercourses. Meanwhile, abundant and diverse ARGs are also leaked and carried by the flow due to the globalization and widespread of antibiotic resistance genes (ARGs) in the natural environment and clinical settings, which has become one of the biggest health concerns in the current century [1]. UWSs as the unique conduit between humans and environments are indisputable hotspots of horizontal gene transfer (HGT) for the spread and evolution of ARGs [2, 3, 4, 5]. Mobile genetic elements (MGEs), especially plasmids and integrative conjugative elements (ICEs), are the primary vectors disseminating ARGs among close and/or remote phylogenetically related microbes, as well as causing the accompanying multidrug resistance (MDR) issues in the problematic nosocomial pathogens of the ESKAPEE group (Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter spp., and Escherichia coli) [6].

Extensive studies have been paid to uncover the antibiotic residues, ARGs and antibiotic resistome (i.e., the entire set of ARG populations) in the wastewater treatment plants (WWTPs) [5, 8, 9]. And some researchers forethoughtfully explored the linkage among the resistome, plasmidome (all plasmids within a given environment or microbial community) and the microbiome in activated sludges from the WWTPs [11, 12]. Much work has been done to determine the specific resistome features in a particular treatment chamber in the WWTPs, especially the biological treatment processes (BTPs), which provide helpful advice for the wastewater risk management and further operational modification [13, 14, 15]. A nine-year longitude temporal metagenomic study deciphers the consistent relationship between the ARGs and MGEs in the activated sludges [16], although long-term characterizations of the sludges reveal variable changes in the microbiome composition [17, 18]. Meanwhile, Yin et al. compared the resistome and mobilome (i.e., all MGEs in a studied microbiome) of the influent, activated sludge and effluent samples from different WWTPs with the purpose of assessing the spatial variations within the WWTPs [20]. In general, the aforementioned studies comprehensively investigate the composition and distribution of the antibiotic resistome and mobilome in the WWTPs or receiving rivers, whereas lacking a full picture of taking the whole UWS into consideration, for instance, the upstream sewer catchments.

In this study, we used our previously documented UWS plasmidome dataset, which applied a direct plasmid DNA metagenome sequencing approach to assess different treatment stages in three distinct UWSs of three European countries [21]. We benchmarked our rigorous experimental methodology for harvesting almost pure plasmid DNA from the sewer/wastewater samples [21]. Accordingly, we
believed the assembled circular and linear contigs were all plasmids (at least predominantly, there may be other MGEs included), while the linear contigs were not circularized due to sequencing depth or intrinsic unique genetic contexts. Then we removed the chromosome noise in the dataset, and the remaining pool of the contigs constituted the target plasmidome in this study. Through in silico analysis, we showed the dynamic plasmidome derived mobile resistome compositions, abundances, and distributions from sewer sources [hospital sewer (HS) and residential sewer (RS) lines], to the end of sewer lines (also known as the mixed sewer, i.e., MS, was often regarded as the WWTP influent), and finally in the BTP of the WWTP (Fig. 1A). Meanwhile, we detected the common ARGs shared by the three UWSs and inspected the epidemic origins of the ARGs emerging in BTPs by source tracking. Further, we recruited the microbiome and high-throughput qPCR datasets of the same sampling campaign. Accordingly, the unrevealing correlations among the plasmid-mediated antibiotic resistome, microbiome and the ‘total’ resistome were demonstrated. Overall, this study contributes to the monitoring of antibiotic resistome in the perspective of high mobility for transmission and evolution, and a manifest understanding of the mobile gene reservoir that potentially originates ARGs dissemination in the UWSs.

### 3.4 Materials and Methods

**Sampling construction, sample pre-treatment and plasmidome sequencing**

Detailed information regarding the sampling campaign and sample pre-treatment refer to a previous study. Plasmidome was probed using a direct plasmid DNA metagenomic sequencing strategy skipping transposon-aided capture and multiple displacement amplification. A dedicated plasmid purification step with an exonuclease Plasmid-Safe™ (Lucigen, Madison, WI, USA) digestion was performed to shear and remove chromosomal DNA before building the sequencing libraries. An Illumina NextSeq platform (Illumina, San Diego, CA, USA) with a v2.5 sequencing kit (Illumina) was utilized to perform the plasmidome sequencing. A full description of the employed plasmidome methodology has been introduced before.

**High-throughput qPCR array and microbiome documentation**

Archived datasets of the high-throughput qPCR array targeting 120 ARGs and MGEs together with 16S-rRNA amplicon sequencing regarding the studied UWSs were recruited from our collaborators’ work.

**Bioinformatic analysis**

The computational work for raw reads processing and assembly for the putative circular and linear plasmid contigs was performed through the Plaspline pipeline (https://github.com/Wanli-HE/Plaspline.git) and described as previously. Chromosomal contigs were removed in the assembly
by comparing to a database built of chromosomes of organisms from the kingdom Bacteria on the NCBI Refseq Genomes FTP (default parameters were applied, details were introduced in the workflow of Plaspline). Accordingly, a pool of the remaining contigs (containing both linear and circular plasmid contigs) was created and regarded as the target plasmidome in this study.

Gene calling for this ‘full version’ plasmidome dataset was performed by using Prodigal (version: 2.6.3) 24. Thereafter, the non-redundant gene set was generated by CD-hit (version: 4.6.2) 25 with a gene sequence identity of > 90% and coverage of > 95% in each gene cluster. RGI (version: 5.1.1) / CARD (version: 3.0.9) 26 was then applied for ARGs annotation (default parameters were applied, and some efflux pump genes were removed manually in the retrieved CARD database, details referred to the functions embedded in Plaspline). Briefly, the gene abundance was calculated using the reads mapping method. Samtools (version: 1.9) 27 was exploited for the observation of reads mapped to contigs. Mapped reads were filtered with a requirement of alignment length ≥ 90 bp and coverage ≥ 95%. Accordingly, a relative abundance table was generated from mSamtools (version: 0.9.6, https://github.com/arumugamlab/msamtools). A final step of normalization for the gene abundance was performed by the varying wastewater conditions (e.g., volatile suspended solids), theoretical cell volume, and carbon content per unit of cell volume according to the report by Li et al. 22.

The ARG risk score is an indicator of the possibility of wild environmental ARGs transferring and conferring resistance to human pathogens 28. In light of the importance of surveillance on ARGs with high mobility 20, we examined the risk scores of the plasmidome derived ARGs with MetaCompare 29.

The plasmidome derived resistome ARG richness index was evaluated by “skbio.diversity” (version: 0.4.2, http://scikit-bio.org/docs/0.4.2/diversity.html) in Python 30. Principal coordinates analysis (PCoA) based Bray-Curtis distance was applied to visualize the different patterns of ARG populations in different UWS compartments. Meanwhile, variations of geographic regions and UWS compartments were examined by Wilcoxon rank-sum test, and the similarities of the ARGs (ANOSIM analysis) were performed by “skbio” packages (version: 0.5.7, http://scikit-bio.org/) in python.

3.5 Results

Resistome profile of the UWS plasmidome

HS, RS, MS and BTP samples (78 in total) constructed at three reported comparable UWSs located in Denmark (DK), Spain (SP), and the United Kingdom (UK) were recruited for this study (Fig. 1A). We showed the UWS plasmidome derived antibiotic resistome with a diversity of 225 ARGs belonging to 180 groups of ARG families according to the ARO id in CARD database 26. In terms of
ARG richness of the entire UWS in a studied country, the plasmidome resistome in Spain was significantly \((p < 0.05)\) higher than in the other two countries, while Denmark and the United Kingdom showed no significant difference (Fig. 1B). We observed 15 ARGs that included \(\text{inu}, \text{ant}(6), \text{aph}(3'), \text{sul}, \text{aac}(6'), \text{dfr}, \text{abc-f}, \text{cat}, \text{qnr}, \text{mph}, \text{erm}, \text{tet}, \text{bla}_{OXA}, \text{ant}(3'')\) and \(\text{mfs}\), contributed >60\% to the total relative abundance of the plasmidome derived resistome at the gene family level (Fig. 1C). Meanwhile, we found aminoglycoside, tetracycline, macrolide and phenicol were the dominant (50\%) drug classes of the plasmidome resistome associated ARGs (Fig. 1D).

**Figure 1** General information of the plasmidome derived antibiotic resistome.

(A) Schematic of the sewer/wastewater flow in the three studied comparable UWSs. (B) ARG richness of the plasmidome derived antibiotic resistome in each country. (C) Dynamics of plasmidome derived...
antibiotic resistome relative abundances in the UWSs (gene family level). (D) Dynamics of plasmidome derived antibiotic resistome relative abundances in the UWSs (drug class level).

Specifically, we discovered that ARG richness in HS was considerably ($p < 0.05$) higher than RS, and RS showed no significant difference with MS in all three countries (Fig. 2A). Remarkably, ARG richness in MS was significantly ($p < 0.001$) higher than BTP in Spain and Denmark. Further, we illustrated the ARG richness in the three sewer lines (HS, RS and MS) were dissimilar from each other, yet MS showed high similarity to BTP in Spain and the United Kingdom (Fig. 2B). Intriguingly, ARGs in the two operated HS lines (HS_a and HS_b) in Denmark exhibited significantly ($p < 0.001$) different similarities. ARGs in HS_b were surprisingly similar to RS in Denmark, whereas HS_a was dissimilar from those two.

**Figure 2** The dynamic plasmidome derived antibiotic resistome in the UWSs.

(A) The dynamic ARG richness in the UWSs. (B) PCoA analysis based on Bray-Curtis distances for the plasmidome derived antibiotic resistome richness in different treatment compartments of each country.

HS_a and HS_b were the two operated hospital sewer lines in Denmark. ‘NS’ in Fig. 2A stands for ‘no significant difference’ in Wilcoxon rank-sum test.
The common UWS plasmidome resistome

Since ARGs have been globally disseminated through the MGEs mediated HGT, we anticipated a shared ARG pool from different countries. Herein, we showed a number of ARGs that could be shared among different countries in every treatment compartment of the UWSs (Fig. 3). The ‘biggest share’ of common ARGs was found between Spain and Denmark in the sewer lines, while Spain and the United Kingdom shared more ARGs in the BTP. Overall, HS harboured the largest number of shared ARGs, and BTP possessed the least.

Figure 3 Common ARGs shared by different countries in each treatment compartment. The intersection in the Venn diagrams represents the commonly shared plasmids between/among countries. The background yellowish water flow shows the sewer/wastewater flow from the sewer sources (HS and RS) to MS, then from MS to BTP.

Risk scoring for the plasmidome resistome in the UWSs

Risk scores were calculated and compared among different treatment compartments in each country (Fig. 4A). No significant differences in risk scores were detected among the different treatment compartments in Denmark and the United Kingdom. We exclusively discovered a remarkable ($p < 0.05$) difference between HS and RS in Spain.
Plasmid-borne ARGs source tracking in the UWSs

Previously, we showed source-trackable antibiotic resistance encoding plasmids in the same dataset 21. In this study, we employed a similar trajectory model for inspecting the ARG flow from sewer sources HS (cyan-coloured) and RS (green-coloured) to MS (pink-coloured), then to the BTP (orange-coloured) as shown in Fig. 4B. RS and HS resembled playing equivalent roles in transporting ARGs to MS in terms of ARG diversity. Approximately 53%, 36% and 63% of the ARGs in MS were retained in the BTP of Spain, Denmark and the United Kingdom, respectively. As for the BTP emerging ARGs, we discovered comparable numbers of ARG types were also found in both sewer sources, i.e., HS and RS, in Spain and Denmark, whereas a higher number of ARG types was found in RS than HS in the United Kingdom (Fig. 4B).

Figure 4 The plasmidome derived ARG risk scores and ARG persistence.

(A) ARG risk scores of the plasmidome derived resistome in different treatment compartments of each country. (B) ARG trajectory analysis of the plasmidome derived resistome in the UWSs and source tracking analysis of the BTP retained plasmidome derived ARGs in each country.

"*" means “p < 0.05” and ‘NS’ stands for ‘no significant difference’ in the Wilcoxon rank-sum test. ARGs discovered in the BTPs were source tracked to the sewer sources by manually checking their
existence in the HS and/or RS. The number close to the sewer lines (HS, RS and MS) represents the ARG diversity (absolute number of ARG types found at this site), and the percentage sitting in the middle of sewer sources and MS represents the percentage of ARGs flowed from sewer sources to MS. The arrows represent the flow direction of the corresponding (same coloured) sewer line. The percentage in the pie figures of BTPs represents the percentage of ARGs retained or removed from MS to BTP. The percentage in the pie figures in the source tracking analysis represents the percentage of the BTP retained ARGs that could be found in the sewer sources (HS or RS). The source tracking analysis was performed by manually checking the abundance of BTP retained ARGs in the HS and RS. We consider a BTP retained ARG was source-trackable when we found the abundance of this ARG in the tested sewer line > 0.

Linking the UWS plasmidome resistome to the microbiome and total resistome
Correlation analyses were performed for investigating the interplay of plasmidome derived antibiotic resistome with the wastewater microbiome and the total resistome (data from qPCR array 22) in the studied UWSs. We found a significant \( p < 0.0001 \) negative correlation between the plasmidome derived resistome richness and the microbiome richness (Fig. 5A), and a significant \( p < 0.01 \) positive correlation between the plasmidome derived resistome relative abundance and ARG relative abundance detected by qPCR (Fig. 5B). Meanwhile, we extracted the beta-lactam resistome by recruiting all detected beta-lactamase genes and pooling their relative abundances (Fig. 5C). We detected all four classes of beta-lactamase genes in the plasmidome (Class A: 59%, Class B: 23%, Class C: 9%, Class D: 9% accounting for the total beta-lactam resistome diversity). \( \text{bla}_{\text{OXA}} \) was the most prevalent and abundant beta-lactamase gene across the UWSs. Further, we perceived the relative abundance of plasmidome derived beta-lactam resistome notably \( p < 0.0001 \) positively correlated with the relative abundance of beta-lactamase genes detected by qPCR (Fig. 5D).
Figure 5 Correlation analysis for the plasmidome derived resistome.

(A) Correlation analysis of the plasmidome derived resistome richness with the corresponding microbiome richness in the same UWS. (B) Correlation analysis of the plasmidome derived resistome relative abundance with the ARG relative abundance (documented by qPCR in the same UWS). (C) Relative abundance (log10) of all beta-lactamase genes in the plasmidome. (D) Correlation analysis of the plasmidome derived beta-lactam resistome relative abundance with the beta-lactamase genes relative abundance documented by qPCR in the same UWS.
3.6 Discussion

The dynamic plasmidome derived resistome in the UWSs

Potential roles of antibiotic use in different countries

Our previous plasmidome study has exclusively focused on the circular contigs in the dataset \(^{21}\), while we might miss the messages from those linear contigs, such as ARGs. In the present study, we recalled all the assembled contigs and removed those that were chromosomally associated. Therefore, we generated the ‘full-version’ of the UWS plasmidome, in which we identified 1.7 times more ARG numbers compared to the circular version UWS plasmidome \(^{21}\). We showed the 15 ARGs [\(\text{Inu, ant}(6), \alpha\text{ph}(3'), \text{sul, aac}(6'), \text{dfrr, abc-f, cat, qnr, mph, erm, tet, bla}_{\text{OXA}}, \text{ant}(3'')\) and \(\text{mfs}\)] out of the 225 detected ARGs in total dominated the ‘full-version’ plasmidome derived resistome (Fig. 1C). We then focused on these 15 ARGs to censor their dynamics in the UWSs. By comparing these 15 ARGs' relative abundances between different countries in each UWS compartment, we found significant \((p < 0.05)\) differences mainly emerged in the sewer lines, namely HS, RS and MS (Tables S1 to S3), while the relative abundances were not significantly different in BTP between different countries for most ARGs. This could be explained by the varied selection pressure caused by the sewer conditions (e.g., antibiotic residues, temperature) among different countries \(^1\), and BTP in the WWTPs universally removed ARGs to some extent, which would result in low abundance comparisons \(^{31}\). Similarly, the comparison of the ARGs’ relative abundances in the drug class level showed the two sewer sources (HS and RS) were significantly \((p < 0.05)\) different among different countries (Tables S4 to S6).

Generally, we discovered the resistome richness and relative abundance in Spain were significantly different \((p < 0.05)\) from the other two countries considering the UWS compartments of HS and RS (Fig. 1B, Tables S1, S2, S4 and S5). This could be attributed to the fact that antibiotics have been extensively used in Spain in both community and hospital sectors \(^{32}\), and higher relative abundances of ARGs in the wastewaters have been more frequently detected in Spain than in western and northern European countries \(^9, 33\). Interestingly, ARGs belonging to the glycopeptide drug class exhibited exclusively relative high abundance in HS and also in the Spanish MS (Fig. 1D), which indicates this group of ARGs was more clinically relevant and its spread of contamination in the sewer lines of Spain.

The partitioning patterns of different UWS compartments

We used a comparative analysis for the ARG relative abundances of different UWS compartments in each country. Accordingly, we found that the HS resistome relative abundance was generally significantly different from other sewers (RS and MS) and the wastewater in BTP in all three countries no matter in the ARG gene family level (Tables S7 to S9) or drug class level (Tables S10 to S12). This demonstrates the relatively high ARG load in the HS, and other researchers have reported that
HS harboured more prevalent antibiotic-resistant bacteria than RS and more abundant ARGs were discovered in HS than RS since a more intense amount of antibiotics were used in the nosocomial environment. RS basically showed no significant difference from MS in view of ARG relative abundances whilst they arguable exhibited dissimilarity in ARG richness (Tables S7 to S12, Fig. 2B). This implies some ARGs were lost during the flow from RS to MS, while the abundances of the retained ARGs were not significantly affected by the sewer transportation. And Auguet et al. also showed the relative concentrations of ARGs (qnrS, sulI, sul2, blaTEM, blaKPC, ermB, tetM and tetW) didn’t significantly decrease due to the sewer pumping in the sewer pipe.

On the other hand, the ARG richness differed in the studied UWS treatment compartments, and we noticed an overall pattern of HS > RS, HS > MS and MS >BTP in all three countries (Fig. 2A). We were also aware of the close relations of ARG richness between MS and BTP in Spain and Denmark (Fig. 2B) and the significant differences in relative abundance between them (Tables S7 and S9). This suggests the ARG diversity was kept, while its abundance dropped remarkably after the WWTP processing in Spain and Denmark. And it could arguably indicate the BTP in the United Kingdom removed several types of ARGs and reduced ARG abundances in the meantime (Fig. 2B, Table S8). This difference emerged because a distinct BTP facility was employed in the United Kingdom (biofilter) compared to the systems used in Denmark and Spain (activated-sludge-based biological treatment basin). Although the ARG richness in the two constructed HS lines (HS_a and HS_b) was not significantly different, they demonstrated remarkable dissimilarity in the PCoA analysis based on BrayCurtis distances (Fig. 2B). ARG richness in HS_a was more ‘HS’-alike based on the observation of the ‘ordinary’ pattern of HS in the other two countries; whereas HS_a was more RS-alike. This was mainly due to the acceptance of this line was predominantly domestic sewage produced in the hospital surroundings.

MDR was prevalent along with the UWS processing (Fig. 1D). Their emergence and abundance (> 40% relatively) in the BTPs constituted the potential risk of being transferred to pathogenic bacteria considering this step usually involved high biomass and retention time. Above all, our findings of these dynamic variations of the plasmidome derived resistome in the UWSs displayed the specific partitioning patterns of the different treatment compartments and roles of antibiotic uses in different countries. This draws attention to the necessity of monitoring transmissible resistome in different stages of UWSs to improve risk management and determine barriers to antibiotic resistance prevention.
Plasmid-borne ARGs are shared and transported as common goods in the UWSs

WWTPs are playgrounds of various ARGs and MGEs. In light of our discovery of shared plasmids from different countries along with the UWS processing, we thereby detected the possible shared pools of ARGs. We illustrated that there was always a common reservoir of shared plasmid-borne ARGs in every UWS compartment (Fig. 3), suggesting that these shared ARGs have been internationally spread. Considering the potential risks brought by the detected ARGs in the local circumstance, we then applied MetaCompare to uncover the relatively different environmental resistome risks associated with each assessed plasmidome sample and matrix at different UWS compartments. Ekwanzala et al. reported the highest resistome risk score of a South African UWS was observed in the HS (mean = 46.34%), followed by the WWTP inlet (mean = 39.67%), WWTP’s BTP (mean = 33.53%), WWTP effluent (mean = 28.84%) and riverbed sediment (mean = 25.55%). Herein, we displayed the general pattern of high resistome risk scores in different UWSs (Fig. 4A). We didn’t detect significant differences among different treatment compartments except between the raw sewers (HS vs RS) in Spain. This, again, describes the high ARG contamination in the Spanish HS.

ARGs as cargo genes located on the plasmid manifest beneficial traits for the host’s survival during environmental selection. Therefore, the plasmid-borne ARGs’ dynamics are closely related to the plasmid fitness and persistence, and here we anticipated some ARGs would be constantly found in different UWS compartments since we have discovered persistent AR plasmids in the UWSs. We applied an ARG trajectory analysis and observed RS and HS both played important roles in spreading different types of ARGs to MS. More than half of the ARG types from the MS flow remained in the BTP of Spain and the United Kingdom, while only ca. 36% of the ARG types in MS were retained in Denmark (Fig. 4B). These maintained ARGs constituted a potential transmissible resistance pool to the post-WWTP facilities and the WWTP downstream surface waters. In an attempt to identify such a pool’s origin, we applied the so-called ARG source tracking analysis to check whether these ARGs were also discovered in the sewer sources. A comparably high number of BTP-retained ARGs were found in HS and RS in Spain and Denmark, while more BTP-retained ARGs could be found in RS than in HS in the United Kingdom. It is thus critical to acknowledge the ARGs that emerged in the BTPs have a genetic background in the sewer sources. However, further analyses are needed in order to understand these ARGs' genetic context, positioning on the plasmids (co-located and distance with other MGEs) and involvement in horizontal transmission.
Plasmidome derived resistome frames the total resistome

ARGs harbouring plasmids can functionally shape the microbiome community in some specific environment and contribute to the ever evolution resistome, resulting in MDR. Associations have been investigated between ARGs-encoding plasmids and bacterial communities using various target or nontarget experimental approaches. However, facts underlying the success of many of these associations remain unknown. One may apply the abundance-trending correlation analysis of the plasmidome and a corresponding microbial community profiling dataset to decipher the link between detected plasmids and their potential hosts. Although spurious correlations would occur throughout the calculations, such analysis has been constantly employed to explore the linkages of resistome, plasmidome and the microbiome.

In this study, the plasmidome resistome richness significantly \((p < 0.0001)\) negatively correlated with the corresponding microbiome richness (Fig. 5A), which illustrates the probable emergence of plasmid fitness cost in the populations when the microbiome is more diverse (under an alleviated selective environment), while the resistome develops and expands its map in specific groups of microbes (under possibly multiple selective environments). On the other side, the plasmidome resistome relative abundance significantly \((p < 0.01)\) positively correlated with the qPCR documented ARG relative abundance (Fig. 5B), this evidences the role of plasmidome resistome contributing to the total resistome.

Currently, beta-lactams are one of the most widely used groups of antibiotics in the world. Beta-lactamase is the major resistance determinant for beta-lactams in Gram-negative bacteria. Critically, beta-lactamases can often ride on plasmids and co-occur with other types of ARGs, creating difficulties in clinical therapies. Herein, we presented a highly prevalent and abundant pattern of Class D beta-lactamase gene \(\text{bla}_{\text{OXA}}\) in the UWS plasmidome regardless of compartment and country (Fig. 5C). In addition, \(\text{bla}_{\text{MOX}}\) and \(\text{bla}_{\text{IMP}}\) were also abundant in the sewer lines though both were eliminated in the BTPs in Spain and Denmark. Meanwhile, we showed a similar finding in Class A and Class B beta-lactamase genes accounting for the most part for the beta-lactam resistome as detected by high-throughput qPCR. Moreover, a significant \((p < 0.0001)\) positive correlation was drawn between the relative abundance of the plasmidome derived beta-lactam resistome and the relative abundance of beta-lactamase genes detected by qPCR (Fig. 5D). Consequently, for the other time, we proved the plasmidome resistome was an important part of the total resistome and plasmid facilitated beta-lactamase genes were core members of the beta-lactam resistome.

Overall, our discoveries highlight the important role of plasmidome derived resistome in the total antibiotic resistome. Regarding this case study, the UWS plasmidome derived resistome was dominantly shaped by the regional variation (i.e., country’s antibiotic use) and the treatment
compartments. The emergence of common and persistent plasmid harbouring ARGs during the UWS processing implies the prevalence and propagation of plasmid-mediated ARGs in the UWS environment, which emphasises the urgency to implement regular surveillance and risk management that to be appropriate for the different regions and treatment stages.

3.7 Acknowledgements

This research was funded by a Joint Programming Initiative-Antimicrobial Resistance grant (JPI-AMR; DARWIN project #7044-00004B) to BFS, and the DFF-Research Project 2 Grants from the Danish Council for Independent Research | Technology and Production (7017-00210A) to BFS.

3.8 Contributions

ZFY constructed the pre-treatment, sequencing library preparation for the samples, beautified the figures and wrote the manuscript; WLH performed the computing analysis and produced the figures; JN and FK gave recommendations for the bioinformatic pipelines and statistical analysis; WK and LHH processed the plasmidome sequencing and produced the raw reads; AD organised regular meetings and shared the metadata files regarding the sampling campaigns; BS and SJS made the initial proposal of this study; all the co-authors have reviewed and gave suggestions for this manuscript.
3.9 References


3.10 Supporting Information

Supporting Information for
Plasmidome derived antibiotic resistome reveals the partitioning of different geographic regions and treatment compartments in the urban water systems

Authors and affiliations
Zhuofeng Yu, a,* Wanli He, a,* Joseph Nesme, a,* Franziska Klincke, a Jonas Stenløkke Madsen, a Witold Kot, b,c Lars Hestbjerg Hansen, b,c Arnaud Dechesne, d Barth Smets, d Søren Johannes Sørensen a,*

aSection of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
bDepartment of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
cDepartment of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark
dDepartment of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

*These authors contributed equally to this work.
*Corresponding Authors.
**Table S1** *p* values in *Wilcoxon rank-sum* significance test of ARG (gene family level) relative abundance between SP and the UK in different treatment compartments. Green-coloured area means *p* < 0.01, red-coloured area means *p* < 0.05.

<table>
<thead>
<tr>
<th>Gene_ID</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.003691</td>
<td>0.42334</td>
<td>0.87278</td>
<td>0.078169</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.039352</td>
<td>0.33668</td>
<td>0.109315</td>
<td>0.748774</td>
</tr>
<tr>
<td>APH(3′)</td>
<td>0.000747</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.200185</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.399269</td>
<td><strong>0.037373</strong></td>
<td>0.006485</td>
<td>0.200185</td>
</tr>
<tr>
<td>AAC(6′)</td>
<td>0.133994</td>
<td>0.003948</td>
<td>0.149541</td>
<td>0.336668</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.001451</td>
<td>0.003948</td>
<td>0.42334</td>
<td>0.748774</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.001997</td>
<td><strong>0.010406</strong></td>
<td>0.024975</td>
<td>0.054664</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.778729</td>
<td>0.200185</td>
<td>0.87278</td>
<td>0.262332</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.091824</td>
<td>0.078169</td>
<td>0.054664</td>
<td>0.229766</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.133994</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.000747</td>
<td><strong>0.010406</strong></td>
<td><strong>0.037373</strong></td>
<td>0.109315</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.002726</td>
<td>0.87278</td>
<td>0.521839</td>
<td>0.688921</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.302895</td>
<td><strong>0.003948</strong></td>
<td>0.006485</td>
<td>0.078169</td>
</tr>
<tr>
<td>ANT(3′)</td>
<td>0.778729</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.006485</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td><strong>0.00873</strong></td>
<td>0.003948</td>
<td>0.003948</td>
<td><strong>0.010406</strong></td>
</tr>
</tbody>
</table>
Table S2 p values in Wilcoxon rank-sum significance test of ARG (gene family level) relative abundance between SP and the DK in different treatment compartments. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.078169</td>
<td>0.87278</td>
<td>0.200185</td>
<td>0.42334</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.078169</td>
<td>0.054664</td>
<td>1</td>
<td>0.262332</td>
</tr>
<tr>
<td>APH(3')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.006485</td>
<td>0.42334</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.109315</td>
<td>0.016309</td>
<td>0.006485</td>
<td>0.336668</td>
</tr>
<tr>
<td>AAC(6')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.054664</td>
<td>0.748774</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.262332</td>
<td>0.521839</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.037373</td>
<td>0.149541</td>
<td>0.262332</td>
<td>0.87278</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.003948</td>
<td>0.016309</td>
<td>0.521839</td>
<td>0.200185</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.262332</td>
<td>0.87278</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.037373</td>
<td>0.003948</td>
<td>0.024975</td>
<td>0.109315</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.336668</td>
<td>0.748774</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.003948</td>
<td>0.024975</td>
<td>0.003948</td>
<td>0.378478</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.748774</td>
</tr>
<tr>
<td>ANT(3&quot;)</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.006485</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td>0.521839</td>
<td>0.010406</td>
<td>0.200185</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S3  *p* values in Wilcoxon rank-sum significance test of ARG (gene family level) relative abundance between DK and the UK in different treatment compartments. Green-coloured area means *p* < 0.01, red-coloured area means *p* < 0.05.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.091824</td>
<td>1</td>
<td>0.149541</td>
<td>0.149541</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.707932</td>
<td>0.336668</td>
<td>0.054664</td>
<td>0.521839</td>
</tr>
<tr>
<td>APH(3')</td>
<td>0.006606</td>
<td>0.87278</td>
<td>0.149541</td>
<td>0.200185</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.302895</td>
<td>0.336668</td>
<td>0.748774</td>
<td>0.109315</td>
</tr>
<tr>
<td>AAC(6')</td>
<td>0.851412</td>
<td>0.936186</td>
<td>0.630954</td>
<td>0.054664</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.039352</td>
<td>0.262332</td>
<td>0.810181</td>
<td>0.200185</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.019208</td>
<td>0.024975</td>
<td>0.078169</td>
<td>0.229766</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.001997</td>
<td>0.078169</td>
<td>0.336668</td>
<td>0.262332</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.512075</td>
<td>0.054664</td>
<td>1</td>
<td>0.078169</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.189783</td>
<td>0.42334</td>
<td>0.024975</td>
<td>0.010406</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.348972</td>
<td>0.630954</td>
<td>0.109315</td>
<td>0.200185</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.133994</td>
<td>0.42334</td>
<td>0.003948</td>
<td>0.297953</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.001997</td>
<td>0.262332</td>
<td>0.748774</td>
<td>0.748774</td>
</tr>
<tr>
<td>ANT(3&quot;)</td>
<td>0.00873</td>
<td>0.336668</td>
<td>0.109315</td>
<td>0.336668</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td>0.002726</td>
<td>0.109315</td>
<td>0.024975</td>
<td>0.024975</td>
</tr>
</tbody>
</table>
Table S4 $p$ values in Wilcoxon rank-sum significance test of ARG (drug class level) relative abundance between SP and the UK in different treatment compartments. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>0.149541</td>
<td>0.630954</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>0.229766</td>
<td>0.109315</td>
<td>0.42334</td>
<td>1</td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.045328</td>
<td>0.262332</td>
<td>0.109315</td>
<td>0.748774</td>
</tr>
<tr>
<td>cephalexin</td>
<td>0.010406</td>
<td>0.630954</td>
<td>0.336668</td>
<td>0.87278</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.336668</td>
<td>0.149541</td>
<td>0.42334</td>
<td>0.521839</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td>0.262332</td>
<td>0.003948</td>
<td>0.336668</td>
<td>0.128205</td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td>0.149541</td>
<td>0.037373</td>
<td>1</td>
<td>0.336668</td>
</tr>
<tr>
<td>lincosamid antibiotic</td>
<td>0.87278</td>
<td>0.037373</td>
<td>0.262332</td>
<td>0.748774</td>
</tr>
<tr>
<td>penam</td>
<td>0.054664</td>
<td>0.037373</td>
<td>0.010406</td>
<td>0.87278</td>
</tr>
<tr>
<td>sulfonamid antibiotic</td>
<td>0.262332</td>
<td>0.016309</td>
<td>0.630954</td>
<td>0.87278</td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>1</td>
<td>0.936186</td>
<td>0.47117</td>
<td>1</td>
</tr>
<tr>
<td>diaminopyrimidine antibiotic</td>
<td>0.078169</td>
<td>0.016309</td>
<td>0.336668</td>
<td>0.521839</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.010406</td>
<td>0.109315</td>
<td>0.006485</td>
<td>0.336668</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.521839</td>
<td>0.200185</td>
<td>0.521839</td>
<td>0.87278</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td>0.336668</td>
<td>0.200185</td>
<td>0.078169</td>
<td>0.024975</td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.200185</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.003948</td>
<td>0.016309</td>
<td>0.016309</td>
<td>0.630954</td>
</tr>
</tbody>
</table>
**Table S5**  $p$ values in *Wilcoxon rank-sum* significance test of ARG (drug class level) relative abundance between SP and the DK in different treatment compartments. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05.$

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>0.091824</td>
<td>0.630954</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>0.399269</td>
<td>0.054664</td>
<td>0.149541</td>
<td>0.336668</td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.160057</td>
<td>0.336668</td>
<td>0.336668</td>
<td>1</td>
</tr>
<tr>
<td>cefamycin</td>
<td>0.000747</td>
<td><strong>0.045328</strong></td>
<td>0.109315</td>
<td>0.336668</td>
</tr>
<tr>
<td>carbapenem</td>
<td><strong>0.011446</strong></td>
<td>0.037373</td>
<td><strong>0.037373</strong></td>
<td>0.054664</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.92538</td>
<td>0.42334</td>
<td>0.748774</td>
<td>1</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td><strong>0.006606</strong></td>
<td>0.003948</td>
<td>0.054664</td>
<td><strong>0.016309</strong></td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td><strong>0.00873</strong></td>
<td>0.42334</td>
<td>0.078169</td>
<td><strong>0.003948</strong></td>
</tr>
<tr>
<td>lincosamide antibiotic</td>
<td><strong>0.004958</strong></td>
<td><strong>0.006485</strong></td>
<td>0.054664</td>
<td>0.149541</td>
</tr>
<tr>
<td>penam</td>
<td><strong>0.039325</strong></td>
<td>0.60954</td>
<td>0.810181</td>
<td>0.054664</td>
</tr>
<tr>
<td>sulfonamide antibiotic</td>
<td>0.061045</td>
<td>0.054664</td>
<td>0.200185</td>
<td><strong>0.006485</strong></td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>0.133994</td>
<td>0.47117</td>
<td>0.336668</td>
<td>1</td>
</tr>
<tr>
<td>dianinopyrimidine antibiotic</td>
<td><strong>0.004958</strong></td>
<td><strong>0.006485</strong></td>
<td>0.42334</td>
<td>0.748774</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.707932</td>
<td>0.630954</td>
<td>0.748774</td>
<td>0.054664</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.261054</td>
<td>0.42334</td>
<td>0.262332</td>
<td>0.092696</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td><strong>0.004958</strong></td>
<td><strong>0.0037373</strong></td>
<td>0.003948</td>
<td><strong>0.003948</strong></td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td><strong>0.000747</strong></td>
<td><strong>0.006485</strong></td>
<td><strong>0.010406</strong></td>
<td>0.42334</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.778729</td>
<td>0.200185</td>
<td>0.149541</td>
<td>0.748774</td>
</tr>
</tbody>
</table>
Table S6 $p$ values in Wilcoxon rank-sum significance test of ARG (drug class level) relative abundance between DK and the UK in different treatment compartments. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS</th>
<th>RS</th>
<th>MS</th>
<th>BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>0.399269</td>
<td>0.054664</td>
<td>0.149541</td>
<td>0.748774</td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.639575</td>
<td>0.87278</td>
<td>0.47117</td>
<td>1</td>
</tr>
<tr>
<td>cephalexin</td>
<td>0.039352</td>
<td>0.630954</td>
<td>0.575174</td>
<td>0.336668</td>
</tr>
<tr>
<td>carbapenem</td>
<td>0.189783</td>
<td>0.378478</td>
<td>0.336668</td>
<td>0.054664</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.031229</td>
<td>0.128205</td>
<td>0.42334</td>
<td>0.521839</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td>0.044045</td>
<td>0.630954</td>
<td>0.109315</td>
<td>0.016309</td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td>0.001451</td>
<td>0.42334</td>
<td>0.078169</td>
<td>0.037373</td>
</tr>
<tr>
<td>lincomycin antibiotic</td>
<td>0.049203</td>
<td>0.630954</td>
<td>0.262332</td>
<td>0.149541</td>
</tr>
<tr>
<td>penam</td>
<td>0.003691</td>
<td>0.109315</td>
<td>0.748774</td>
<td>0.054664</td>
</tr>
<tr>
<td>sulfonamide antibiotic</td>
<td>0.399269</td>
<td>0.748774</td>
<td>0.078169</td>
<td>0.006485</td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>0.189783</td>
<td>0.378478</td>
<td>0.748774</td>
<td>1</td>
</tr>
<tr>
<td>diaminopyrimidine antibiotic</td>
<td>0.019208</td>
<td>0.42334</td>
<td>0.936186</td>
<td>0.262332</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.019208</td>
<td>0.024975</td>
<td>0.054664</td>
<td>0.128205</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.014887</td>
<td>0.078169</td>
<td>0.630954</td>
<td>0.149541</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td>0.00873</td>
<td>0.149541</td>
<td>0.024975</td>
<td>0.003948</td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td>0.000747</td>
<td>0.078169</td>
<td>0.037373</td>
<td>0.109315</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.075156</td>
<td>0.054664</td>
<td>0.87278</td>
<td>0.748774</td>
</tr>
</tbody>
</table>
Table S7  $p$ values in Wilcoxon rank-sum significance test of ARG (gene family level) relative abundance between different treatment compartments in SP. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Gene_ID</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.262332</td>
<td>0.87278</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.003948</td>
<td>0.010406</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.149541</td>
<td>0.149541</td>
<td>0.024975</td>
<td>0.748774</td>
<td>0.054664</td>
<td>0.037373</td>
</tr>
<tr>
<td>APH(3')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.200185</td>
<td>0.006485</td>
<td>0.006485</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.016309</td>
<td>0.016309</td>
<td>0.003948</td>
<td>0.054664</td>
<td>0.078169</td>
<td>0.630954</td>
</tr>
<tr>
<td>AAC(6')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.024975</td>
<td>0.016309</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.521839</td>
<td>0.262332</td>
<td>0.003948</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.109315</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.262332</td>
<td>0.149541</td>
<td>0.003948</td>
<td>0.521839</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.42334</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.054664</td>
<td>0.003948</td>
<td>0.010406</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.109315</td>
<td>0.006485</td>
<td>0.010406</td>
<td>0.037373</td>
<td>0.024975</td>
<td>0.054664</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.003948</td>
<td>0.037373</td>
<td>0.054664</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.010406</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.109315</td>
<td>0.016309</td>
<td>0.003948</td>
<td>0.200185</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.748774</td>
<td>0.037373</td>
<td>0.054664</td>
<td>0.037373</td>
<td>0.054664</td>
<td>0.054664</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.010406</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.024975</td>
<td>0.024975</td>
</tr>
<tr>
<td>ANT(3'')</td>
<td>0.87278</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.006485</td>
<td>0.006485</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td>0.748774</td>
<td>0.521839</td>
<td>0.003948</td>
<td>0.262332</td>
<td>0.003948</td>
<td>0.006485</td>
</tr>
</tbody>
</table>
Table S8 $p$ values in Wilcoxon rank-sum significance test of ARG (gene family level) relative abundance between different treatment compartments in the UK. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.630954</td>
<td>1</td>
<td>0.010406</td>
<td>0.42334</td>
<td>0.010406</td>
<td>0.010406</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.003948</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.006485</td>
<td>0.016309</td>
<td>0.003948</td>
</tr>
<tr>
<td>APH(3')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.054664</td>
<td>0.024975</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.078169</td>
<td>0.024975</td>
<td>0.262332</td>
</tr>
<tr>
<td>AAC(6')</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.037373</td>
<td>0.748774</td>
<td>0.200185</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.006485</td>
<td>0.016309</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.010406</td>
<td>0.054664</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.87278</td>
<td>0.748774</td>
<td>0.003948</td>
<td>0.630954</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.521839</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.630954</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.003948</td>
<td>0.748774</td>
<td>0.003948</td>
<td>0.010406</td>
<td>0.748774</td>
<td>0.024975</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.748774</td>
<td>0.87278</td>
<td>0.109315</td>
<td>0.748774</td>
<td>0.109315</td>
<td>0.024975</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.262332</td>
<td>0.037373</td>
<td>0.006485</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.016309</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.024975</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.003948</td>
<td>0.010406</td>
<td>0.003948</td>
<td>0.054664</td>
<td>0.630954</td>
<td>0.521839</td>
</tr>
<tr>
<td>ANT(3'')</td>
<td>0.006485</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.87278</td>
<td>0.336668</td>
<td>0.42334</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td>0.003948</td>
<td>0.006485</td>
<td>0.003948</td>
<td>0.521839</td>
<td>0.078169</td>
<td>0.109315</td>
</tr>
</tbody>
</table>
Table S9 *p* values in *Wilcoxon rank-sum* significance test of ARG (gene family level) relative abundance between different treatment compartments in DK. Green-coloured area means *p* < 0.01, red-coloured area means *p* < 0.05.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincosamide nucleotidyltransferase (LNU)</td>
<td>0.024589</td>
<td>0.067799</td>
<td>0.021754</td>
<td>0.87278</td>
<td>0.003948</td>
<td>0.020241</td>
</tr>
<tr>
<td>ANT(6)</td>
<td>0.004958</td>
<td>0.001451</td>
<td>0.000747</td>
<td>0.262332</td>
<td>0.054664</td>
<td>0.200185</td>
</tr>
<tr>
<td>APH(3’)</td>
<td>0.006606</td>
<td>0.006606</td>
<td>0.001045</td>
<td>0.936186</td>
<td>0.045328</td>
<td>0.045328</td>
</tr>
<tr>
<td>sulfonamide resistant sul</td>
<td>0.011446</td>
<td>0.000747</td>
<td>0.001451</td>
<td>0.630954</td>
<td>0.521839</td>
<td>0.109315</td>
</tr>
<tr>
<td>AAC(6’)</td>
<td>0.001451</td>
<td>0.011446</td>
<td>0.000747</td>
<td>0.054664</td>
<td>0.149541</td>
<td>0.003948</td>
</tr>
<tr>
<td>trimethoprim resistant dihydrofolate reductase dfr</td>
<td>0.373597</td>
<td>0.54267</td>
<td>0.021754</td>
<td>0.936186</td>
<td>0.045328</td>
<td>0.045328</td>
</tr>
<tr>
<td>ABC-F ATP-binding cassette ribosomal protection protein</td>
<td>0.348972</td>
<td>0.261054</td>
<td>0.001451</td>
<td>0.87278</td>
<td>0.010406</td>
<td>0.006485</td>
</tr>
<tr>
<td>chloramphenicol acetyltransferase (CAT)</td>
<td>0.031229</td>
<td>0.003691</td>
<td>0.000747</td>
<td>0.521839</td>
<td>0.006485</td>
<td>0.010406</td>
</tr>
<tr>
<td>quinolone resistance protein (qnr)</td>
<td>0.302895</td>
<td>0.302895</td>
<td>0.000747</td>
<td>0.748774</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>macrolide phosphotransferase (MPH)</td>
<td>0.302895</td>
<td>0.002726</td>
<td>0.000747</td>
<td>0.42334</td>
<td>0.016309</td>
<td>0.045328</td>
</tr>
<tr>
<td>Erm 23S ribosomal RNA methyltransferase</td>
<td>0.851412</td>
<td>0.92538</td>
<td>0.000747</td>
<td>1</td>
<td>0.016309</td>
<td>0.037373</td>
</tr>
<tr>
<td>tetracycline-resistant ribosomal protection protein</td>
<td>0.019208</td>
<td>0.002726</td>
<td>0.000747</td>
<td>0.521839</td>
<td>0.024975</td>
<td>0.024975</td>
</tr>
<tr>
<td>OXA beta-lactamase</td>
<td>0.000747</td>
<td>0.000747</td>
<td>0.000747</td>
<td>0.336668</td>
<td>0.42334</td>
<td>0.037373</td>
</tr>
<tr>
<td>ANT(3”)</td>
<td>0.000747</td>
<td>0.000747</td>
<td>0.000747</td>
<td>0.42334</td>
<td>0.037373</td>
<td>0.010406</td>
</tr>
<tr>
<td>major facilitator superfamily (MFS) antibiotic efflux pump</td>
<td>0.061045</td>
<td>0.003691</td>
<td>0.000747</td>
<td>0.521839</td>
<td>0.006485</td>
<td>0.006485</td>
</tr>
</tbody>
</table>
**Table S10** p values in *Wilcoxon rank-sum* significance test of ARG (drug class level) relative abundance between different treatment compartments in SP. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>0.297953</td>
<td>0.149541</td>
<td>0.149541</td>
<td>0.630954</td>
<td>0.630954</td>
<td>1</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>0.521839</td>
<td>0.630954</td>
<td>0.630954</td>
<td>0.336668</td>
<td>0.336668</td>
<td></td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.092696</td>
<td>0.262332</td>
<td>0.521839</td>
<td>0.87278</td>
<td>0.521839</td>
<td>0.748774</td>
</tr>
<tr>
<td>cephemycin</td>
<td>0.003948</td>
<td>0.037373</td>
<td>0.87278</td>
<td>0.336668</td>
<td>0.200185</td>
<td>0.262332</td>
</tr>
<tr>
<td>carbapenem</td>
<td>0.748774</td>
<td>0.748774</td>
<td>0.262332</td>
<td>0.87278</td>
<td>0.748774</td>
<td>0.748774</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.037373</td>
<td>0.109315</td>
<td>0.630954</td>
<td>0.109315</td>
<td>0.521839</td>
<td>0.42334</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td>0.003948</td>
<td>0.200185</td>
<td>0.054664</td>
<td>0.024975</td>
<td>0.078169</td>
<td>0.045328</td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td>0.748774</td>
<td>0.42334</td>
<td>0.037373</td>
<td>0.87278</td>
<td>0.748774</td>
<td>0.748774</td>
</tr>
<tr>
<td>lincomamide antibiotic</td>
<td>0.037373</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.47117</td>
<td>0.630954</td>
<td>0.149541</td>
</tr>
<tr>
<td>penam</td>
<td>0.003948</td>
<td>0.262332</td>
<td>0.149541</td>
<td>0.003948</td>
<td>0.010406</td>
<td>0.010406</td>
</tr>
<tr>
<td>sulfonamide antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.109315</td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.47117</td>
<td>0.630954</td>
<td>0.149541</td>
</tr>
<tr>
<td>diaminopyrimidine antibiotic</td>
<td>0.054664</td>
<td>1</td>
<td>0.003948</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.109315</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.262332</td>
<td>0.149541</td>
<td>0.003948</td>
<td>0.109315</td>
<td>0.006485</td>
<td>0.010406</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.037373</td>
<td>0.262332</td>
<td>0.630954</td>
<td>0.87278</td>
<td>0.200185</td>
<td>0.200185</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.630954</td>
<td>0.006485</td>
<td>0.200185</td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.748774</td>
<td>1</td>
<td>0.42334</td>
<td>0.262332</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.037373</td>
<td>0.010406</td>
<td>0.336668</td>
<td>0.630954</td>
<td>0.200185</td>
<td>0.87278</td>
</tr>
</tbody>
</table>
**Table S11** $p$ values in *Wilcoxon rank-sum* significance test of ARG (drug class level) relative abundance between different treatment compartments in the UK. Green-coloured area means $p < 0.01$, red-coloured area means $p < 0.05$.

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>1</td>
<td>1</td>
<td>0.630954</td>
<td>1</td>
<td>0.630954</td>
<td>0.630954</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>1</td>
<td>0.630954</td>
<td>0.42334</td>
<td>0.054664</td>
<td>0.149541</td>
<td></td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.149541</td>
<td>0.688921</td>
<td>0.149541</td>
<td>0.42334</td>
<td>0.054664</td>
<td>0.149541</td>
</tr>
<tr>
<td>cephapomycin</td>
<td>0.262332</td>
<td>0.297953</td>
<td>0.521839</td>
<td>0.87278</td>
<td>0.262332</td>
<td>0.200185</td>
</tr>
<tr>
<td>carbapenem</td>
<td>0.87278</td>
<td>0.521839</td>
<td>0.200185</td>
<td>0.42334</td>
<td>0.200185</td>
<td>0.521839</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.87278</td>
<td>0.521839</td>
<td>0.016309</td>
<td>0.87278</td>
<td>0.030639</td>
<td>0.016309</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td>0.003948</td>
<td>0.010406</td>
<td>0.128205</td>
<td>0.037373</td>
<td>0.054664</td>
<td>0.054664</td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td>0.003948</td>
<td>0.037373</td>
<td>0.262332</td>
<td>0.024975</td>
<td>0.054664</td>
<td>0.109315</td>
</tr>
<tr>
<td>lincosamide antibiotic</td>
<td>0.010406</td>
<td>0.037373</td>
<td>0.748774</td>
<td>0.200185</td>
<td>0.037373</td>
<td>0.109315</td>
</tr>
<tr>
<td>penam</td>
<td>0.006485</td>
<td>0.037373</td>
<td>0.006485</td>
<td>0.024975</td>
<td>0.109315</td>
<td>0.016309</td>
</tr>
<tr>
<td>sulfonamide antibiotic</td>
<td>0.078169</td>
<td>0.262332</td>
<td>0.003948</td>
<td>0.024975</td>
<td>0.149541</td>
<td>0.003948</td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.378478</td>
<td>0.630954</td>
<td>0.149541</td>
</tr>
<tr>
<td>diaminopyrimidine antibiotic</td>
<td>0.149541</td>
<td>0.078169</td>
<td>0.003948</td>
<td>0.336668</td>
<td>0.010406</td>
<td>0.200185</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.748774</td>
<td>0.024975</td>
<td>0.054664</td>
<td>0.336668</td>
<td>0.109315</td>
<td>0.200185</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.200185</td>
<td>0.006485</td>
<td>0.87278</td>
<td>0.109315</td>
<td>0.521839</td>
<td>0.037373</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td>0.024975</td>
<td>0.109315</td>
<td>0.006485</td>
<td>0.42334</td>
<td>0.149541</td>
<td>0.054664</td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td>0.003948</td>
<td>0.010406</td>
<td>0.078169</td>
<td>0.109315</td>
<td>0.262332</td>
<td>0.748774</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.003948</td>
<td>0.003948</td>
<td>0.109315</td>
<td>0.037373</td>
<td>0.024975</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S12  *p* values in Wilcoxon rank-sum significance test of ARG (drug class level) relative abundance between different treatment compartments in the DK. Green-coloured area means *p* < 0.01, red-coloured area means *p* < 0.05.

<table>
<thead>
<tr>
<th>Drug class</th>
<th>HS vs RS</th>
<th>HS vs MS</th>
<th>HS vs BTP</th>
<th>RS vs MS</th>
<th>RS vs BTP</th>
<th>MS vs BTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>streptogramin antibiotic</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>peptide antibiotic</td>
<td>0.049203</td>
<td>0.20609</td>
<td>0.574149</td>
<td>0.262332</td>
<td>0.748774</td>
<td>1</td>
</tr>
<tr>
<td>fosfomycin</td>
<td>0.453695</td>
<td>0.399269</td>
<td>0.091824</td>
<td>0.810181</td>
<td>0.149541</td>
<td>0.149541</td>
</tr>
<tr>
<td>cephemycin</td>
<td>0.223391</td>
<td>0.302895</td>
<td>0.000747</td>
<td>0.688921</td>
<td>0.016309</td>
<td>0.016309</td>
</tr>
<tr>
<td>carbaopenem</td>
<td>0.574149</td>
<td>0.061045</td>
<td>0.000747</td>
<td>0.521839</td>
<td>0.016309</td>
<td>0.003948</td>
</tr>
<tr>
<td>nucleoside antibiotic</td>
<td>0.373597</td>
<td>0.639575</td>
<td>0.348972</td>
<td>0.521839</td>
<td>1</td>
<td>0.748774</td>
</tr>
<tr>
<td>elfamycin antibiotic</td>
<td>0.011446</td>
<td>0.024589</td>
<td>0.000747</td>
<td>0.010406</td>
<td>0.016309</td>
<td>0.016309</td>
</tr>
<tr>
<td>rifamycin antibiotic</td>
<td>0.399269</td>
<td>0.92538</td>
<td>0.001997</td>
<td>0.521839</td>
<td>0.006485</td>
<td>0.010406</td>
</tr>
<tr>
<td>lincosamide antibiotic</td>
<td>0.001045</td>
<td>0.016927</td>
<td>0.101208</td>
<td>0.748774</td>
<td>0.003948</td>
<td>0.020241</td>
</tr>
<tr>
<td>penam</td>
<td>0.054858</td>
<td>0.673417</td>
<td>0.574149</td>
<td>0.575174</td>
<td>0.336668</td>
<td>0.748774</td>
</tr>
<tr>
<td>sulfonamide antibiotic</td>
<td>0.160057</td>
<td>0.851412</td>
<td>0.000747</td>
<td>0.336668</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>glycopeptide antibiotic</td>
<td>0.001451</td>
<td>0.001045</td>
<td>0.000747</td>
<td>0.521839</td>
<td>0.149541</td>
<td>0.336668</td>
</tr>
<tr>
<td>dianinopyrimidine antibiotic</td>
<td>0.743059</td>
<td>0.673417</td>
<td>0.054858</td>
<td>0.936186</td>
<td>0.065552</td>
<td>0.065552</td>
</tr>
<tr>
<td>phenicol antibiotic</td>
<td>0.453695</td>
<td>0.049203</td>
<td>0.014887</td>
<td>0.037373</td>
<td>0.016309</td>
<td>0.054664</td>
</tr>
<tr>
<td>fluoroquinolone antibiotic</td>
<td>0.024589</td>
<td>0.075156</td>
<td>0.001045</td>
<td>0.521839</td>
<td>0.003948</td>
<td>0.003948</td>
</tr>
<tr>
<td>macrolide antibiotic</td>
<td>0.261054</td>
<td>0.001451</td>
<td>0.000747</td>
<td>0.200185</td>
<td>0.016309</td>
<td>0.065552</td>
</tr>
<tr>
<td>tetracycline antibiotic</td>
<td>0.001045</td>
<td>0.000747</td>
<td>0.261054</td>
<td>0.630954</td>
<td>0.87278</td>
<td>1</td>
</tr>
<tr>
<td>aminoglycoside antibiotic</td>
<td>0.004958</td>
<td>0.019208</td>
<td>0.261054</td>
<td>0.200185</td>
<td>0.037373</td>
<td>0.200185</td>
</tr>
</tbody>
</table>
Part 3: Conference Presentations (abstracts and posters)
1. Conference abstracts

1.1 DMS 2018

Conference full title: Annual Congress of the Danish Microbiological Society (Copenhagen, Denmark, 2018)

Abstract title: Danish wastewater harbours multiple mobilized colistin resistance (mcr) genes: a preliminary study on the environmental mcr reservoir

Zhuofeng Yu¹, Rafael Pinilla-Redondo¹, Joseph Nesme¹, Arnaud Deschene², Barth Smets² and Søren Sørensen¹

¹ Section of Microbiology, University of Copenhagen, Denmark
² Department of Environmental Engineering, Technical University of Denmark

Despite its neuro- and nephrotoxicity, colistin constitutes a last-resort antibiotic against multidrug-resistant Gram-negative pathogens as safer drugs gradually become obsolete. Thus, the fast-paced emergence and mobility of colistin resistance (mcr) genes pose a severe public health concern. To date, research on the prevalence of mcr genes has primarily focused on clinical and veterinary isolates, while the study of mcr genes in environmental reservoirs has been greatly overlooked. Incidentally, the recent emergence of mcr-1 and other mcr genes calls for epidemiological studies of the dynamics of plasmid-borne mcr. In this context, wastewater treatment plants (WWTP) are of high research interest since they comprise horizontal gene transfer hot spots for bacteria, and hence play a critical role in the dissemination of plasmid-borne antibiotic resistance. In this study, we investigated the activated sludge samples from the Odense municipality WWTP. We confirmed the presence of colistin resistant Enterobacteriaceae and the responsible mcr genes variants were identified by multiplex PCR targeting 5 different mcr genes (mcr-1; mcr-2; mcr-3; mcr-4; mcr-5). Plasmid DNA extracted from mcr PCR positive samples were sequenced and assembled thus revealing the array of plasmids vectors harbouring mcr genes in an urban WWTP and their host association.
1.2 BAGECO 15

Conference full title: 15th Symposium on Bacterial Genetics and Ecology (Lisbon, Portugal, 2019)

Abstract title: IncN plasmids are vectors disseminating colistin resistance in wastewater microbiota

Zhuofeng Yu¹, Rafael Pinilla-Redondo¹, Joseph Nesme³, Jonas Stenløkke Madsen¹, Nana McBean Andersen¹, Asmus Kalckar Olesen¹, Hanadi Ananbeh²,³, Arnaud Dechesne⁴, Barth Smets⁴ and Søren Sørensen¹

¹ Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
² Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
³ Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
⁴ Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

Wastewater treatment plants (WWTPs) collect residual pharmaceutical compounds and massive amounts of human-gut associated bacteria, some of which are closely related to human pathogens. As high cell densities and recurrent antibiotic selective pressure favour plasmid-mediated dissemination occur in WWTPs, those are regarded as particularly active hot spots for the spread of plasmid-borne antibiotic resistance genes (ARGs). Plasmids belonging to the IncN incompatibility group are often associated with the dissemination of clinically relevant ARGs. Their presence in WWTPs and effluent waters together with their broad host range, high transfer rate and ability to replicate in Enterobacteriaceae, a group comprising several critical pathogens makes them a critical concern to public health. However, little is known about their transfer dynamics in urban water systems.

In this work, we describe the capture and characterization of conjugative plasmids harbouring clusters of resistance determinants from a municipal WWTP activated sludge sample (Odense, DK). Plasmids were recovered by exogenous isolation using an *Escherichia coli* recipient strain. Nanopore sequencing revealed the presence of a 73Kbp IncN plasmid carrying a complete set of functional conjugative genes and multiple ARGs. Antibiotic susceptibility testing showed that this plasmid can transfer resistance of colistin, a last-resort antibiotic, to a sensitive *Escherichia coli* recipient. Direct
isolation from the WWTP community confirmed the presence of colistin-resistant strains in that environment. Most were identified as *Pseudochrobactrum* sp. and *Ochrobactrum* sp. using full-length 16S rRNA gene sequencing. Interestingly, a *Pseudochrobactrum* isolate hosted a similar-sized IncN plasmid also encoding complete conjugative transfer systems and multiple ARGs. Further filter mating experiments with a colistin-sensitive recipient strain confirmed the mobility of the IncN plasmid and its associated colistin resistance. However, the responsible determinant could not be readily identified by sequence homology as one of the already characterized mobilized colistin resistance (*mcr*) gene variants. The identification of a potentially novel determinant of colistin resistance transferable to *Escherichia coli* is of significance for surveillance purposes. Together, these results uncover the dynamic nature of self-transmissible IncN plasmids and their role in the dissemination of colistin resistance in the WWTP environment.
Plasmids belonging to the IncN incompatibility group are often associated with the dissemination of clinically relevant ARGs. Their presence in wastewater treatment plants (WWTPs) and effluent waters together with their broad host range, high transfer rate and ability to replicate in Enterobacteriaceae, a group comprising several critical pathogens makes them a critical concern to public health. However, little is known about their transfer dynamics in urban water systems.

In this work, we describe the capture and characterization of conjugative plasmids harbouring clusters of resistance determinants from a municipal WWTP activated sludge sample (Odense, DK). Plasmids were recovered by exogenous isolation using an *Escherichia coli* recipient strain. Nanopore sequencing revealed the presence of a 73Kbp IncN plasmid carrying a complete set of functional conjugative genes and multiple ARGs. Antibiotic susceptibility testing showed that this plasmid can transfer resistance of colistin, a last-resort antibiotic, to a sensitive *Escherichia coli* recipient. Direct isolation from the WWTP community confirmed the presence of colistin-resistant strains in that environment. Most were identified as *Pseudochrobactrum* sp. and *Ochrobactrum* sp. using full-length 16S rRNA gene sequencing. Interestingly, a *Pseudochrobactrum* isolate hosted a similar-sized IncN plasmid also encoding complete conjugative transfer systems and multiple ARGs. Further filter mating experiments with a colistin-sensitive recipient strain confirmed the mobility of the IncN
plasmid and its associated colistin resistance. However, the responsible determinant could not be readily identified by sequence homology as one of the already characterized mobilized colistin resistance ($mcr$) gene variants. The identification of a potentially novel determinant of colistin resistance transferable to *Escherichia coli* is of significance for surveillance purposes. Together, these results uncover the dynamic nature of self-transmissible IncN plasmids and their role in the dissemination of colistin resistance in the WWTP environment.
1.4 DMS 2019

Conference full title: Annual Congress of the Danish Microbiological Society (Copenhagen, Denmark, 2019)

Abstract title: IncN plasmids are vectors disseminating colistin resistance in wastewater microbiota

Zhuofeng Yu¹, Joseph Nesme¹, Jonas Stenløkke Madsen¹, Rafael Pinilla-Redondo¹, Kamille Anna Dam Clasen¹, Asmus Kalckar Olesen¹, Nana McBean Andersen¹, Hanadi Ananbeh²,³, Arnaud Dechesne⁴, Barth Smets⁴ and Søren Sørensen¹

¹ Section of Microbiology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
² Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
³ Central European Institute of Technology, Brno University of Technology, Purkynova 123, CZ-612 00 Brno, Czech Republic
⁴ Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark

Wastewater treatment plants (WWTPs) collect residual pharmaceutical compounds and massive amounts of human-gut associated bacteria, some of which are closely related to human pathogens. As high cell densities and recurrent antibiotic selective pressure favour plasmid-mediated dissemination occur in WWTPs, those are regarded as particularly active hot spots for the spread of plasmid-borne antibiotic resistance genes (ARGs). Plasmids belonging to the IncN incompatibility group are often associated with the dissemination of clinically relevant ARGs. Their presence in WWTPs and effluent waters together with their broad host range, high transfer rate and ability to replicate in Enterobacteriaceae, a group comprising several critical pathogens makes them a critical concern to public health. However, little is known about their transfer dynamics in urban water systems.

In this work, we describe the capture and characterization of conjugative plasmids harbouring clusters of resistance determinants from a municipal WWTP activated sludge sample (Odense, DK). Plasmids were recovered by exogenous isolation using an *Escherichia coli* recipient strain. Nanopore sequencing revealed the presence of a 73Kbp IncN plasmid carrying a complete set of functional conjugative genes and multiple ARGs. Antibiotic susceptibility testing showed that this plasmid can
transfer resistance of colistin, a last-resort antibiotic, to a sensitive *Escherichia coli* recipient. Direct isolation from the WWTP community confirmed the presence of colistin-resistant strains in that environment. Most were identified as *Pseudochrobactrum* sp. and *Ochrobactrum* sp. using full-length 16S rRNA gene sequencing. Interestingly, a *Pseudochrobactrum* isolate hosted a similar-sized IncN plasmid also encoding complete conjugative transfer systems and multiple ARGs. Further filter mating experiments with a colistin-sensitive recipient strain confirmed the mobility of the IncN plasmid and its associated colistin resistance. However, the responsible determinant could not be readily identified by sequence homology as one of the already characterized mobilized colistin resistance (*mcr*) gene variants. The identification of a potentially novel determinant of colistin resistance transferable to *Escherichia coli* is of significance for surveillance purpose. Together, these results uncover the dynamic nature of self-transmissible IncN plasmids and their role in the dissemination of colistin resistance in the WWTP environment.
2. Conference posters

2.1 DMS 2018

Danish wastewater harbors multiple mobilized colistin resistance (mcr) genes: a preliminary study on the environmental mcr reservoir

INTRODUCTION

Why mcr
- Colistin: constitutes a last-resort antibiotic against multidrug-resistant Gram-negative pathogens (introduced resistance) despite its neuro- and nephrotoxicity
- mcr: consequently, the fast-paced mobilized colistin resistance (mcr) becoming a serious public health concern
- Current mcr Research is primarily focusing on clinical and veterinary isolates while overlooking environmental reservoirs
- The recent emergence of mcr-1 incocidentally calls for epidemiological studies of the dynamics of plasmid borne mcr-genes

mcr Globalization
- Databank: Enrichment of mcr-genes in the environment is of high concern
- Current: mcr-genes are spreading within clinical and veterinary isolates while overlooking environmental reservoirs

Why WWTP
- The urban wastewater continuum collects both residual pharmaceutical compounds and massive amounts of human-gut associated bacteria, some of which are closely related to human pathogens. This creates an ideal situation for horizontal gene transfer of antibiotic resistance, such as mcr, as high cell densities and recurrent antibiotic-selected pressure favor plasmid-mediated dissemination of antibiotic resistance determinants
- Activated sludge are of particular interest because of their very high biomass and longer retention time in the treatment process. As a result, they can be a key point of antibiotic resistance in wastewater treatment plants (WWTP).

AIM

1) Revealing environmental mcr reservoir by multiplex PCR targeting 5 different mcr genes variants (mcr-1; mcr-2; mcr-3; mcr-4; mcr-5).
2) Uncovering the array of plasmids vectors harbouring mcr genes and their host association in an urban WWTP by plasmid DNA extraction from mcr PCR positive samples and subsequent sequencing and assembly.
3) Infer horizontal transfer dissemination potential by analyzing regions flanking antibiotic resistance genes for the presence of mobility genes.

METHODS

1) Isolation
Collistin resistant isolates were obtained from Odense municipality WWTP activated sludge by selective plating on 100 violet red bile glucose agar plates supplemented with either 2 or 8 µg/mL of colistin. A total of 88 isolates were purified by successive isolation from single colony restreaking and finally plated on LB agar with corresponding colistin concentrations.

2) Screening
All isolates were screened for mcr genes by multiplex colony PCR. The protocol used is based on a recently developed multiplex PCR assay in DTU that allows measuring five mcr-genes variants (i.e. mcr-1; mcr-2; mcr-3; mcr-4; mcr-5) simultaneously.

3) Plasmids Sequencing
Plasmid extraction was performed for four isolates growing on plates supplemented with 8 µg/mL of colistin and positive for at least one mcr gene by multiplex PCR. Extracted plasmids DNA were sequenced on a MiONi portable single-molecule nanopore sequencer (Oxford Nanopore Technologies, Oxford, UK) using Rapid Barcoding sequencing kit. Raw reads signals were first sorted by barcode using DeepPltinase® and basecalled using Albacore 2.2.1 before assembly using Unicycler (miniasm2 +ucasoc polishing). Circular contigs were visually selected using de novo assembly graphs visualisation software Bandage 0.8.1. Circular-like sequences were then annotated with PlasmidFinder 2.0.0, CARD, ResFinder 3.1, VirulenceFinder 2.0.0, and orgnFinder 1.0.0 to identify antibiotic resistance, virulence factors and mobility regions on each plasmid.

RESULTS

We confirmed the presence of colistin resistant Enterobacteriaceae on plates and the occurrence and/or co-occurrence of the responsible mcr genes variants in WWTP mcr isolates by multiplex PCR.

Only one putative plasmid sequence from our WWTP mcr PCR-positive isolates (4 tested in total) hits the PlasmidFinder 2.0 database and is identified as an IncN plasmid (99.81% coverage, 98.83% identity).

BLAST result indicates the plasmid has strong sequence homology to plasmids from the family Enterobacteriaceae as expected, and mostly hit the genera Klebsiella and Echerichia coli.

The 73 kb plasmid shows multiple resistance to aminoglycoside, beta-lactam, fluoroquinolone, phenicol, sulphonamide, trimethoprim.

Interestingly, despite the WWTP mcr PCR positive isolate indeed shows resistance to colistin, its putative functional plasmid seems not involving any known mcr genes in database (Fig. 6). Matting experiments with colistin sensitive E.coli will be performed to verify pColB2F034 can transfer colistin resistance phenotype by conjugation. Further sequence polishing is required to correct frameshifts caused by indels errors common in Nanopore sequencing.

REFERENCES

IncN plasmids are vectors disseminating colistin resistance in wastewater microbiota

Zhuoeng Yu1, Joseph Nesme1, Kamille Anna Dam Clausen1, Hanadi Ananbeh2, Rafael Pinilla-Redondo1, Jonas Stenenløkke Madsen3, Nana McBean Andersen1, Asmus Kalkar Olsen1, Arnaud Dechesne1, Barth Smet2 and Søren Sørensen1

1Section of Microbiology, University of Copenhagen, Universitetsparken 17, DK-2100 Copenhagen, Denmark
2Department of Chemistry and Biochemistry, Faculty of Life Sciences, Mødalsvej 30, 8200 Aarhus N, Denmark
3Central European Institute of Technology, Brno University of Technology, Pilsenová 123, CZ-616 00 Brno, Czech Republic

INTRODUCTION

Emergence of Global Colistin Resistance

- IncN plasmids are widespread in colistin-resistant clones isolated from environmental samples, and their prevalence has increased in the last decade.
- The spread of colistin-resistant plasmids has become a significant concern.

Why Studying WWTP?

- WWTPs function as hotspots for the development and dissemination of antibiotic resistance, including colistin resistance.
- Studying WWTPs provides insights into the mechanisms driving the dissemination of colistin-resistant plasmids.

AIMS

- Reveal environmental transmissible colistin resistance reservoir by exogenous plasmid isolation, followed by a direct colistin screening of the wastewater bacterial communities and multiplex PCR targeting 5 different mcr genes variants.
- Infer horizontal transfer dissemination potential by conjugation experiment for a wastewater-derived colistin resistant isolate and a colistin-sensitive recipient strain, and also by analyzing regions flanking antibiotic resistance genes for the presence of mobility genes.

METHODS

Exogenous Plasmid Isolation: capture the transmissible ESB resistance plasmids from wastewater

Filtration of wastewater

LIVE/DEAD staining

 LIVE cells + E. coli

 Filter mating

Verify with FACS

COL5 TC-1

COL7 TC-2

Calculation of transfer frequency

Plasmid DNA extraction

Data analysis

RESULTS

Characteristics of Isolated Plasmids

<table>
<thead>
<tr>
<th>Species</th>
<th>Potential mcr gene variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>mcr-1, mcr-2, mcr-3, mcr-4, mcr-5</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>mcr-1, mcr-2, mcr-3, mcr-4, mcr-5</td>
</tr>
</tbody>
</table>

Perspectives

- Next generation sequencing will provide insights into the mechanisms driving the dissemination of colistin-resistant plasmids.
- The observed prevalence of IncN plasmids in wastewater could have implications for the dissemination of colistin resistance in aquatic ecosystems.

ACKNOWLEDGEMENTS
Part 4: Appendix
Appendix 1: PhD project briefings

I. The DARWIN Project (JPI-AMR, 7044-00004B)

Project title
Dynamics of Antimicrobial Resistance in the Urban Water Cycle in Europe, DARWIN

Project abstract (adapted and modified from the original proposal)
Therapeutic and domestic antibiotic uses are known to impact the evolution of antibiotic resistance (AR), while the environmental dimension of AR is also of great concern. We propose that urban wastewater treatment systems (UWS) are receptacles for excreted antibiotics, antibiotic-resistant organisms, and antibiotic resistance genes (ARGs), which are central conduits of AR to and from pathogens and environmental microbes. This is because UWS are normally of high microbial densities, and the co-mingling of different wastes (co-occurrence of antibiotics, biocides, metals, and microbes) promotes/accelerate the horizontal gene transfer (HGT) of ARGs and multiple drug resistances. In the DARWIN project, we undertake a never-before-performed pan-European examination of the fate of key antibiotic-resistant organisms and genetic determinants in the UWS resulting from the discharged hospital and community wastes, including transmission mechanisms in different stages of sewer catchments, and receiving waters. We focus on the spread of ARGs encoding clinically relevant extended-spectrum beta-lactam (ESBL) and carbapenem resistance in three countries with distinct AR profiles and sewage management practices. We posit that ARGs readily transmit in the UWS from pathogens and commensal hosts in human wastes (after antibiotic use) to environmental strains better adapted to migrate through the sewer environment. And this is driven by local ecologies, conjugal plasmid transfer and phage mediated transduction. We determine specific bacterial hosts that carry ARGs across the UWS and identify where key HGT events occur with the goal of assessing the relative risk of ARGs returning to humans due to environmental exposure. To guide risk assessments, a predictive dynamic mathematical model for UWS will be developed to assist in health and sewage management decisions.

My role in the DARWIN project
(1) System sampling & characterization: pre-treatment and DNA extraction for the samples (executor); 16S-rRNA sequencing (operator).
(2) Inventory and quantification of ARGs and mobile genetic elements: qPCR of ARGs and MGEs (minor role); exogenous plasmid isolation to capture conjugative plasmids from the UWS
communities that confer ESBL and/or carbapenem resistance (executor); mobilome sequencing and analysis (executor); metagenome sequencing (operator).

(3). Transfer potential: permissiveness assays for the DARWIN UWS-derived ESBL and/or carbapenem plasmids using the DARWIN sewer and biological treatment process wastewater samples (executor).

Project-associated collaborative publication

Citation: Environ. Sci. Technol. 2021, 55, 9, 5939–5949 https://doi.org/10.1021/acs.est.0c08548

Title: Extended-spectrum β-lactamase and carbapenemase genes are substantially and sequentially reduced during conveyance and treatment of urban sewage

Liguan Li¹,§, Joseph Nesme²,§, Marcos Quintela-Baluja³,§, Sabela Balboa⁴, Syed Hashsham⁵, Maggie R. Williams⁶, Zhuofeng Yu², Søren J. Sørensen², David W. Graham³, Jesus L. Romalde⁴, Arnaud Dechesne¹, Barth F. Smets¹,*

¹Department of Environmental Engineering, Technical University of Denmark, Lyngby, DK
²Department of Biology, University of Copenhagen, Copenhagen, Denmark
³School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK
⁴Department of Microbiology and Parasitology, CIBUS-Faculty of Biology & Institute CRETUS, Universidade de Santiago de Compostela, Santiago de Compostela 15782, Spain
⁵Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI USA
⁶School of Engineering and Technology, Central Michigan University, Mt. Pleasant, MI USA

§ These authors contributed equally to this work.
*Corresponding author

Abstract: Urban wastewater systems (UWSs) are a main receptacle of excreted antibiotic resistance genes (ARGs) and their host microorganisms. However, we lack integrated and quantitative observations of the occurrence of ARGs in the UWS to characterize the sources and identify processes that contribute to their fate. We sampled the UWSs from three medium-size cities in Denmark, Spain, and the United Kingdom and quantified 70 clinically important extended-spectrum β-lactamase and carbapenemase genes along with the mobile genetic elements and microbial communities. Results from all three countries showed that sewage—especially from hospitals—carried substantial loads of
ARGs ($10^6$–$10^7$ copies per person equivalent), but these loads progressively declined along sewers and through sewage treatment plants, resulting in minimal emissions ($10^1$–$10^4$ copies per person equivalent). Removal was primarily during sewage conveyance (65 ± 36%) rather than within sewage treatment (34 ± 23%). The extended-spectrum β-lactamase and carbapenemase genes were clustered in groups based on their persistence in the UWS compartments. The less-persistent groups were associated to putative host taxa (especially Enterobacteriaceae and Moraxellaceae), while the more persistent groups appeared horizontally transferred and correlated significantly with total cell numbers and mobile genetic elements. This documentation of a substantial ARG reduction during sewage conveyance provides opportunities for antibiotic resistance management and a caution for sewage-based antibiotic resistance surveillance.
II. The SandBAR Project (DFF-Research Project 2 Grants from the Danish Council for Independent Research | Technology and Production, 7017-00210A)

**Project title**
Strategies and Barriers to avoid the spread of Antibiotic Resistance genes during wastewater treatment, SandBAR

**Project abstract (adapted and modified from the original proposal)**
Currently, AR is one of the global public health threats and human wastes bring abundant ARGs into the sewer systems. The SandBAR project is the first comprehensive initiative to investigate how wastewater treatment plants (WWTPs) affect the transfer and occurrence of AR to and in human pathogens. SandBAR examines how plasmids and genes encoding AR spread, persist or are eliminated from WWTPs and how this affects their occurrence in pathogens. We will identify traits of the WWTPs microbial communities that promote HGT and the spread of ARGs. SandBAR will reveal microbial groups and mobile genetic elements that facilitate the spread of AR traits from and to human pathogens. Finally, SandBAR will identify WWTP conditions and configurations that mitigate transfer. Overall, SandBAR is a critical step in evaluating the potential burden or mitigation opportunities of wastewater treatment towards AR dissemination.

**My role in the SandBAR project**
Transfer potential: permissiveness assays for the DARWIN UWS-derived ESBL and/or carbapenem plasmids using the SandBAR WWTP samples (executor).
Appendix 2: Sampling information of the UWSs

Table A1 Information on the catchments

<table>
<thead>
<tr>
<th>City</th>
<th>Approx. population in sampled community sewer</th>
<th>Hospital beds</th>
<th>WWTP population equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durham (UK)</td>
<td>-</td>
<td>~ 450</td>
<td>35,000</td>
</tr>
<tr>
<td>Odense (DK)</td>
<td>14,200</td>
<td>~ 1,000</td>
<td>Capacity: 385,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(<a href="http://ouh.dk/wm397745">http://ouh.dk/wm397745</a>)</td>
<td>Actual load (2018): 235,000</td>
</tr>
<tr>
<td>Santiago de Compostela (SP)</td>
<td>9,600</td>
<td>1,395</td>
<td>Capacity: 103,000</td>
</tr>
</tbody>
</table>

Table A2 Approx. GPS coordinate of sampling points

<table>
<thead>
<tr>
<th>City</th>
<th>Hospital sewer</th>
<th>Residential sewer</th>
<th>Mixed sewer and the WWTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durham (UK)</td>
<td>54°47'22.2&quot;N, 1°35'31.5&quot;W</td>
<td>54°47'35.2&quot;N, 1°34'38.1&quot;W</td>
<td>54°46'58.6&quot;N, 1°34'10.1&quot;W</td>
</tr>
<tr>
<td>Odense (DK)</td>
<td>55°23'04.8&quot;N, 10°22'21.6&quot;E</td>
<td>55°24'24.9&quot;N, 10°23'15.3&quot;E</td>
<td>55°23'57.8&quot;N, 10°25'01.8&quot;E</td>
</tr>
<tr>
<td>Santiago de Compostela (SP)</td>
<td>42°87'01.0&quot;N, 8°56'93.2&quot;W</td>
<td>42°88'11.2&quot;N, 8°52'26.2&quot;W</td>
<td>42°87'03.2&quot;N, 8°59'79.8&quot;W</td>
</tr>
</tbody>
</table>

Table A3 Dates of sampling (dd/mm)

<table>
<thead>
<tr>
<th>City</th>
<th>2017*</th>
<th>Winter 2018*</th>
<th>Summer 2018*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durham (UK)</td>
<td>23/06 and 06/07</td>
<td>13/03; 15/03; and 20/03</td>
<td>18/09; 25/09; and 28/09</td>
</tr>
<tr>
<td>Odense (DK)</td>
<td>31/05 and 01/06</td>
<td>30/01; 20/02; and 21/02</td>
<td>12/09; 13/09; and 18/09</td>
</tr>
<tr>
<td>Santiago de Compostela (SP)</td>
<td>24/10</td>
<td>19/02; 21/02; and 23/02</td>
<td>17/10; 19/10; and 21/10</td>
</tr>
</tbody>
</table>

* The sampling campaign in 2017 was the preliminary sampling construction. Hospital sewer and residential sewer in Denmark of the sampling campaign 2017 were used for the exogenous plasmid isolation in Manuscript 1.

* Danish samples of the sampling campaign Winter 2018 were used for the permissiveness test in Manuscript 1.

* The sampling campaigns Winter 2018 and Summer 2018 were used for Manuscript 2 and Manuscript 3.
Appendix 3: List of collaborations

1. **External Partner:** Prof. Barth F. Smets (Department of Environmental Engineering, Technical University of Denmark)
   **Description:** Key collaborator in the DARWIN and SandBAR projects
   **Contact:** Barth F. Smets, bfsm@env.dtu.dk

2. **External Partner:** Dr. Arnaud Dechesne (Department of Environmental Engineering, Technical University of Denmark)
   **Description:** Key collaborator in the DARWIN and SandBAR projects
   **Contact:** Arnaud Dechesne, arde@env.dtu.dk

3. **External Partner:** Dr. Liguan Li (Department of Environmental Engineering, Technical University of Denmark; now at Department of Civil Engineering, The University of Hong Kong)
   **Description:** Collaborator in the DARWIN project
   **Contact:** Liguan Li, liliguan@hku.hk

4. **External Partners:** Prof. David Graham (School of Engineering, Newcastle University) and Dr. Marcos Quintela Baluja (School of Engineering, Newcastle University)
   **Description:** Collaborators in the DARWIN project
   **Contact:** David Graham, david.graham@ncl.ac.uk; Marcos Quintela Baluja, quintela.baluja@ncl.ac.uk

5. **External Partners:** Prof. Jesús L Romalde (Department of Microbiology and Parasitology, University of Santiago de Compostela) and Dr. Sabela Balboa (Department of Microbiology and Parasitology, University of Santiago de Compostela)
   **Description:** Collaborators in the DARWIN project
   **Contact:** Jesús L Romalde, jesus.romalde@usc.es; Sabela Balboa, sabela.balboa@usc.es

6. **External Partners:** Dr. Syed Hashsham (Department of Civil and Environmental Engineering, Michigan State University) and Asst. Prof. Maggie R. Williams (School of Engineering and Technology, Central Michigan University)
   **Description:** Collaborators in the DARWIN project
   **Contact:** Syed Hashsham, hashsham@egr.msu.edu; Maggie R. Williams, willi32m@cmich.edu
7. **External Partners:** Prof. Lars Hestbjerg Hansen (Department of Plant and Environmental Sciences, University of Copenhagen) and Asst. Prof. Witold Piotr Kot (Department of Plant and Environmental Sciences, University of Copenhagen)

**Description:** Key collaborators in the DARWIN project

**Contact:** Lars Hestbjerg Hansen, lhha@plen.ku.dk; Witold Piotr Kot, wk@plen.ku.dk

8. **External Partners:** Prof. Jan Pravsgaard Christensen (Department of Immunology and Microbiology, University of Copenhagen), Dr. Karen Hogg (Department of Biology, University of York) and cytometrists at the Royal Microscopical Society (RMS)

**Description:** Knowledge sharing in the learning and application of flow cytometry

**Contact:** Jan Pravsgaard Christensen, jpc@sund.ku.dk; Karen Hogg, karen.hogg@york.ac.uk; RMS, https://www.rms.org.uk/community/science-sections/flow-cytometry.html

9. **Internal Partner:** Dr. Rafael Pinilla Redondo (Department of Biology, University of Copenhagen)

**Description:** Collaborator in the DARWIN and SandBAR projects

**Contact:** Rafael Pinilla Redondo, rafael.pinilla@bio.ku.dk

10. **Internal Partner:** Wanli He (Department of Biology, University of Copenhagen)

**Description:** Key collaborator in the DARWIN project

**Contact:** Wanli He, wanli.he@bio.ku.dk

11. **Internal Partners:** Kamille Anna Dam Clasen (Department of Biology, University of Copenhagen), Asmus Kalckar Olesen (Department of Biology, University of Copenhagen), Dr. Qin Qin Wang (Department of Biology, University of Copenhagen), Dr. Nan Yang (Department of Biology, University of Copenhagen), and Zhuang Gong (Department of Biology, University of Copenhagen)

**Description:** Collaborators in the DARWIN and SandBAR projects, and key helper hands in daily lab work

**Contact:** Kamille Anna Dam Clasen, kamilleclasen@me.com; Asmus Kalckar Olesen, asmus.olesen@bio.ku.dk; Qin Qin Wang, wang.qinquin@bio.ku.dk; Nan Yang, nan.yang@bio.ku.dk; Zhuang Gong, nrm685@alumni.ku.dk
12. **Other Partners**: Dr. Hanadi Ennab (Department of Chemistry and Biochemistry, Mendel University in Brno and also Central European Institute of Technology, Brno University of Technology; now at Institute of Animal Physiology and Genetics, Academy of the Sciences of the Czech Republic), Xiao Peng (Institute of Microbiology, Chinese Academy of Sciences) and Bruce Wu (School of Environmental Science and Engineering, Southern University of Science and Technology and also Department of Biology, University of Copenhagen)

**Description**: Collaborators in the DARWIN project and helper hands in daily lab work

**Contact**: Hanadi Ennab, hanadi_ennab@yahoo.co.uk; Xiao Peng, alan201666@yeah.net; Bruce Wu, ziqi.wu@bio.ku.dk
Acknowledgement

People and Affiliations

👨‍🏫 Prof. Søren Johannes Sørensen

👩‍🏫 Asst. Prof. Joseph Nesme
Asst. Prof. Jonas Stenløkke Madsen

👩‍🏫 Prof. Barth Smets, Dr. Arnaud Dechesne
Prof. Lars Hestbjergh Hansen, Asst. Prof. Witold Kot

👨‍👩‍👧‍👦 Anette Hårdum Løth, Ayoe Lüchau, Jannie Christensen

Projects

Darwin
Sandbar

Wanli He, Dr. Rafael Pinilla-Redondo, Dr. Qingin Wang,
Asmus Kalckar Olesen, Dr. Franziska Klincke,
Kamille Anna Dam Clasen, Zhuang Gong, Dr. Nan Yang

Prof. David Graham, Dr. Marcos Quintela Baluña Kalckar,
Dr. Liguan Li, Prof. Jesús L. Romalde, Dr. Sobela Balboa,
Dr. Syed Hashsham, Asst. Prof. Maggie R. Williams,
Prof. Jan Pravsgaard Christensen, Dr. Karen Hogg,
Xiao Peng, Bruce Wu, Dr. Hanadi Ennab

DTU Environment

Artworks

Adobe Creative Cloud
Adobe Stock
ASTUTE GRAPHICS
BioRender
Basiselemente
500px
Sketch
WWDC20 & 22

Design Resources

Apple Developer
Dropbox Design
Google Design
Microsoft Design
nature
SONY Design Philosophy

People

👨‍🏫: Principal supervisor
👩‍🏫: Co-supervisors
👩‍👩‍👧‍👦: Key collaborators in PhD projects
👨‍🔬: Technicians
👩‍🔬: Wet lab and dry lab helpers
👩‍🔬: ‘Encyclopaedia’
Projects

**DARWIN**, the DARWIN project (JPI-AMR, 7044-00004B): Dynamics of antimicrobial resistance in the urban water cycle in Europe (see Appendix 1.1).

**SANDBAR**, the SandBAR project (DFF-Research Project 2 Grants from the Danish Council for Independent Research | Technology and Production, 7017-00210A): Strategies and barriers to avoid the spread of antibiotic resistance genes during wastewater treatment (see Appendix 1.2).

Artworks


Artworks were then used and/or modified in Adobe Illustrator v26.3 with the plugins functioned by Astute Graphics v3.0 (https://astutegraphics.com/), Adobe Photoshop v23.4 and Adobe XD v51.0 served by Adobe Creative Cloud (https://www.adobe.com/creativecloud.html). Meanwhile, Sketch v89 (https://www.sketch.com/) was used for some vectors design.

Design Resources

Postscript

These four years have witnessed me going around many places, meeting different kinds of people, settling down and working hard on my PhD project at MME (KU) and METLab (DTU). There are mountains of stories I can tell. However, I wouldn’t write some of them down since they were all precious to me and picking up one or two scenarios appears not fair to other memories. The journey to the peak of PhD wasn’t full of smiles, happiness, and joy. But that’s how it is. You may get stuck and fall into a loop at one point while you think the road to the goal is always straightforward in the beginning. At many times, I recalled a saying I saw on the desk of my mother’s colleagues when I was in primary school – “life is short, while suffering is long”. This saying would not be so true for everyone, and it also appears a bit pessimistic. Nonetheless, I remember this saying up to now as it indicates that we shall treasure our time in some sense, and we shall face adversities bravely as difficulties/problems/troubles occur constantly. In addition, pains make one stronger, tears make one braver, and failures make one wiser. The misfortune of the past calls for a better future.

Life is short, while suffering is long.

Eileen Chang

As is known, time flies and waits for no one, and I can feel the time during the different phases of my Bachelor's to Master's then to PhD’s speeds up rapidly. Nowadays I hardly do something yet half of the day has already passed. At this stage, I always recall the day I chose to come to KU for my PhD study while giving up other options at the Université Paris-Sud in France and the North Carolina State University in the USA. I think I would still make the same choice even if time goes back. It’s not because I’m afraid to face new challenges or am used to the life in Copenhagen. I think it’s more attributed to the fact that I like the scientific environment and culture here and if time plays back, I will tell myself to enjoy rather than adapt.

During my PhD time, I acquired much knowledge, gained many skills, and learned fruitful information on top of my research question. I really appreciate all the ones from different countries, institutions and universities who cared, helped, and guided me throughout the way. In addition, I
would like to express a special thanks to all the designers who create the artworks and design resources (listed in the Acknowledgement chapter) I have used and/or got inspiration from to produce this PhD thesis and my PhD defence keynote. Indeed, it was a ‘pain’ to generate all the figures for the manuscripts during my writing, but “no pain, no gain”, in fact, I enjoyed the period of having my own ideas to create a meaningful and remarkable plot, put the diagrams together neatly and pretty, and draw clear graphic abstracts and interesting scientific discoveries during the process. I think I gradually love this learning (the dataset and research background backwards and forwards) and producing (figures and manuscripts) process. As said in the article by Krause K., *Nature*, 2019 – “Should science be ugly? I disagree. Science sorely needs best practices in visual communication as well as in information design, a mature field with quantitative methods.”, a good scientist is also a good presenter disseminating his knowledge and introducing his research using brilliant and eye-catching illustrations. Without doubts, “life imitates art far more than art imitates life”, and “beauty is everywhere, it is not that she is lacking to our eye, but our eyes which fail to perceive her”. I do consider conducting scientific research as a form of performing art and most of the time what we lack is the ability to discover the ‘beauties’ (e.g., interesting phenomena, underappreciated experimental results and overlooked data analysis) in our surroundings. Pursuing aesthetics in science is now a ‘creed’ to me.

At this point, I have completed my PhD study, but of course, after climbing a great hill, one only finds that there are many more hills to climb. And it is probably true that more work/issues/problems will come to me after my PhD graduation. Nevertheless, every step towards a dream today is a step away from regret tomorrow. Remember these and why you started for any tasks in your heart, Zhuofeng. You can do it!