City-wide metagenomics uncover antibiotic resistance reservoirs in urban beach and sewage waters

Pablo Fresia[1,2], Verónica Antelo[1,2,#], Cecilia Salazar[1,2,#], Matías Giménez[1,3], Bruno D'Alessandro[4], Ebrahim Afshinnekoo[5,6], Christopher Mason[5,6,7], Gastón H Gonnet[2,8,9], Gregorio Iraola[1,2,10,*]

1 Unidad de Bioinformática, Institut Pasteur Montevideo, Montevideo, Uruguay.
2 Proyecto "Centro de Metagenómica", Institut Pasteur Montevideo, Montevideo, Uruguay.
3 Unidad de Microbiología Molecular, Instituto de Investigaciones Biológicas Clemente Estable, Montevideo, Uruguay.
4 Laboratorio de Calidad Ambiental, Intendencia Municipal de Montevideo, Montevideo, Uruguay.
5 Department of Physiology and Biophysics, Weill Cornell Medicine, New York, USA.
6 The HRH Prince Alwaleed Bin Talal Bin Abdelaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, USA.
7 The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, USA.
8 ETH Zurich, Computer Science, Switzerland.
9 SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland.
10 Centro de Biología Integrativa, Universidad Mayor, Santiago de Chile, Chile.

* Corresponding author: Gregorio Iraola, Unidad de Bioinformática, Institut Pasteur Montevideo. Mataojo 2020 (PO 11400), Montevideo, Uruguay. E-mail: giraola@pasteur.edu.uy.

# Author's equal contribution

Abstract

Background: Microbial communities present in environmental waters constitute a reservoir for antibiotic-resistant pathogens that impact human health. For this reason a diverse variety of water environments are being analyzed using metagenomics to uncover public health threats. However, the composition of these communities along the coastal environment of a whole city where sewage and beach waters are mixed, is poorly understood.

Results: We shotgun-sequenced 20 coastal areas from the city of Montevideo (capital of Uruguay) including beach and sewage water samples to characterize bacterial communities and their virulence and antibiotic resistance repertories. We found that sewage and beach environments presented significantly different bacterial communities. Sewage waters harbored a higher prevalence and a more diverse repertory of virulence and antibiotic resistant genes mainly from well-known enterobacteria, including carbapenemases and extended-spectrum betalactamases reported in hospital infections in Montevideo. Additionally, we were able to genotype the presence of both globally-disseminated pathogenic clones as well as emerging antibiotic-resistant bacteria in sewage waters.

Conclusions: Our study represents the first in using metagenomics to jointly analyze beaches and the sewage system from an entire city, allowing us to characterize antibiotic-resistant pathogens circulating in urban waters. The data generated in this initial study represent a baseline metagenomic exploration to guide future longitudinal (time-wise) studies, whose systematic implementation will provide useful epidemiological information to improve public health surveillance.

Keywords: Sewage, Beach, Metagenomics, Taxonomy, Antimicrobial resistance, Bacterial pathogens.

Introduction

Human activity shapes the microbial communities residing in urban environments. In particular, urban sewage systems are designed to evacuate human wastes from the houses to areas of low human exposure and gradually reinstate them into natural watercourses such as creeks, beaches or the sea. This cycle is of tremendous importance for public health as waste waters can be reservoir and vehicle for the transmission of pathogenic bacteria and antibiotic resistance mechanisms. Indeed, the rapid emergence and spread of pathogenic bacteria with extensive antibiotic resistance has been recognized by the World Health Organization as a top health issue [1], since water can easily move microorganisms between humans and other animal species. Accordingly, the analysis of environmental waters is being adopted as an effective method to monitor the dynamics of antibiotic resistant pathogens [2], as this kind of environments can play a role as important as clinical settings for the selection of antibiotic resistance [3].

Recent advances in high-throughput sequencing (HTS) and computational biology now allow the exploration of microbial communities based on culture-independent approaches using metagenomics. This enables us to quantify and functionally characterize environmental microbiomes with unprecedented precision and comprehensiveness [4]. Indeed, the very recent implementation of this methodology to explore the microbial diversity in the built environment is providing a completely new layer of information to be integrated in the management of cities, potentially assisting decisions that range from urban design to public health [5,6]. In particular, urban sewage or beach water systems have been previously characterized using metagenomics not only focused on uncovering ecological patterns [7] but also in characterizing pathogenic and antibiotic resistant bacteria [8,9]. However, the joint analysis of bacterial communities present at the same time in the sewage and beach waters from the same metropolitan area remains to be explored in depth.

Sewage waters have been shown to accurately reflect the population's gut microbiota composition [10], raising the possibility of using metagenomics to directly gain information about infection dynamics [11]. Additionally, beaches are important for recreational use but also are frequently recognized as risky environments for the contagion and transmission of bacterial infections [12], particularly if they are constantly or sporadically impacted by sewage spillovers. Accordingly, we performed a cross-sectional shotgun metagenomic analysis along the urban coast of Montevideo, the capital of Uruguay, aiming to characterize bacterial communities present in the sewage and beach water. Our study represents the first of that kind in a South American city and is the kick-off towards the incorporation of metagenomics in the surveillance of microbiological risks at city scale.

Results

Composition of sewage and beach communities. First, we explored the structure of microbial communities present in our beach and sewage samples using a multiset k-mer counting approach. This strategy provides an unbiased view that is not affected by taxonomic or functional assignment, conversely, it just evaluates the differential abundance of unique DNA segments [13]. Figure 1A shows a clustering analysis based on this methodology that shows a complete discrimination between sewage and beach samples, suggesting substantial differences in the composition of communities in these environments. Second, we confirmed the observed discrimination from a taxonomic point of view by calculating relative abundances of bacterial species present in each sample using an approach based on the identification and quantification of marker genes [14]. Figure 1B shows a dendrogram based on beta diversities (between samples) calculated using the Bray-Curtis dissimilarity distance from the taxonomic profiles, showing a complete discrimination between beach and sewage. Beta diversity was 0.42 within sewage samples (SD = 0.23) and 0.41 within beach samples (SD = 0.22) but increased to 0.63 (SD = 0.12) when comparing sewage against beach samples. Alpha diversity (within samples) was calculated using the Shannon index and averaged 3.65 (SD = 0.64) for sewage and 3.7 (SD = 0.4) for beach samples (Additional file 1: Fig. S1). These results indicate that taxonomic composition of bacterial communities from these environments are substantially different and can discriminate between beach or sewage origin (geographic location of each sample along the coast of Montevideo is displayed in Fig. 1C).

Occurrence of antibiotic resistance genes (ARGs). The presence of different microbial communities in sewage and beach samples led us to hypothesize they could also encode distinct ARG repertories. Hence, metagenomic assemblies (Additional file 2: Table S1) were screened against the Comprehensive Antibiotic Resistance Database (CARD) [15] because is currently the most up-to-date and manually curated resource for ARGs detection. We found that 108 out of 2177 (~5%) ARGs had hits in our samples and they belong to 10 different antibiotic classes (Additional file 3: Table S2), being the clinically relevant TEM-4 and TEM-33 beta-lactamases the top occurring genes but aminoglycoside-modifying enzymes (like acetyl- or phosphotransferases) the most abundant class of ARGs. In particular, a significant difference ($p = 0.002$, Mann-Whitney U test) in ARGs abundance and a significantly higher diversity ($p = 0.0024$, Mann-Whitney U test) of ARGs according to the Simpson's index were found in favor of sewage compared to beach samples (Additional file 4: Fig. S2). Furthermore, when inspecting antibiotic classes we observed that sewage samples encoded ARGs belonging to 90% of antibiotic classes found in these environments while beach samples only encoded 40% of antibiotic classes, evidencing a more complex composition of antibiotic resistance mechanisms in the urban sewage waters (Fig. 2A). Indeed, only elfamycin resistance was present in beach but absent in sewage samples. On the other side, the occurrence of resistance mechanisms for aminoglycosides, betalactams, tetracyclines, sulfonamides, macrolides and streptogramins was significantly greater ($p < 0.01$, Mann-Whitney U test) in sewage compared to beach samples (Fig. 2B).

Distribution of ARGs in mobile elements. As a general trend, we found that ARGs present in our samples were more prevalent in plasmids than in bacterial chromosomes. Specifically, ARGs for sulfonamides, betalactams, aminoglycosides, phenicols, macrolides and streptogramins were more prevalent in plasmids than in bacterial chromosomes (Fig. 3A). Additionally, plasmids carrying ARGs found in our samples resulted to be extensively distributed among many clinically relevant enterobacteria like *E. coli*, *Salmonella*, *Klebsiella*, *Enterobacter*, *Citrobacter* and *Acinetobacter*, among others (Fig. 3B). ARGs for tetracyclines, lincosamides, fluoroquinolones and elfamycins were more frequently encoded in chromosomes.

We also looked for integrons and found a higher prevalence of them in sewage (74%) than in beach (24%) samples. Furthermore, clinically relevant integron classes 1, 2 and 3 were almost exclusively found in sewage samples (~90%) (Fig. 3C). Also, we were able to identify cassette ARGs with conserved *attC* sites associated to 24 out of 39 integrons (61%). These cassette genes mostly coded for multi-drug efflux pumps, but also we found carbapenemases (OXA family), GES extended-spectrum betalactamases (ESBLs), *aadA1* aminoglycoside nucleotidyltransferases, *cat* chloramphenicol acetyltransferases and *aac6-Ib* amikacin resistance. These cassette genes were exclusively found in sewage integrons.

Occurrence of virulence factors (VFs). To complement the characterization of ARGs we also screened metagenomic assemblies against the Virulence Factor Database (VFdb) [16]. Ninety nine out of 451 (~22%) VFs were detected in our samples. Specifically, VFs were found in 7 out of 8 (87.5%) sewage samples and 4 out of 12 (33%) beach samples. Sewage samples also presented a higher count and diversity of VFs compared to beach samples (Additional file 5: Fig. S3A). The functional classification of these VFs showed that those involved in bacterial motility, cell adherence, iron uptake and secretion were predominant among sewage samples. Interestingly, we found the presence of the ShET2 enterotoxin (*senB*) in a single sewage sample. We also examined the taxonomic distribution of these VFs and found that they are predominantly distributed in *Pseudomonas*, *Salmonella*, *Escherichia*, *Yersinia* and *Shigella*, among others (Additional file 5: Fig. S3B). These results highlight the importance of urban waters as reservoir and vehicle for VFs responsible of determining well-known pathogenic mechanism in clinically relevant bacteria.

Taxa carrying mobile ARGs and VFs are sewage biomarkers. We also aimed to identify bacterial taxa associated to sewage or beach that can explain the observed differences in the composition of the overall bacterial community and their ARGs and VFs repertories. Then, we applied a linear discriminant analysis (LDA) and effect size estimation [17] to determine statistically significant taxa associated to beach or sewage. We summarized the results at the genus level and found 6 genera associated to the beach environment while 47 genera were characteristic of the sewage environment (Fig. 4A). Interestingly, 10 out of 47 (~21%) sewage-associated genera comprise bacterial species that are well-known human pathogens, including *Aeromonas*, *Acinetobacter*, *Arcobacter*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Pseudomonas*,

Sphingobacterium, *Stenotrophomonas* and *Streptococcus* (Fig. 4B). Most of these taxa match with those found carrying mobile ARGs and VFs.

Identification of pathogenic genotypes. To gain further insight on the microbiological risks in these environments, we attempted to resolve the presence of previously reported pathogens at the strain level using metagenome-derived multilocus sequence typing (MLST) [18,19]. Despite the fact that this method was unable to identify complete previously reported sequence types (STs), the detection of partial allele combinations allowed us to infer the presence of several genotypes of clinical importance. In sewage samples we found three alleles of *Citrobacter freundii* whose combination determines the ST-209, which has been detected from isolates recovered from diarrheal patients [20]. Also, we detected three alleles from *Escherichia coli* defining the clonal complex ST-131, which is a globally disseminated multidrug-resistant clone associated with human extraintestinal (urinary and bloodstream) infections [21]. Furthermore, the detection of two alleles of *Arcobacter cryaerophilus* defined the presence of ST-392, a recently characterized genotype causative of persistent diarrhea [22]. We also identified several *Salmonella* alleles that diverged from known genotypes, preventing the inference of putative STs. Many unknown alleles of *Pseudomonas fluorescens* were also identified both in sewage and beach samples. This species is mainly associated to food spoilage but has been sporadically reported as an opportunistic human pathogen causing systemic infections associated to the consumption of animal byproducts [23]. Overall, these results indicate the presence of pathogenic genotypes in urban waters of Montevideo.

Discussion

Pharmaceutical products, including antibiotics, can be only partially metabolized by humans so these compounds or their derived metabolites are excreted [24] lastly reaching the environment. Sewage pipes have been largely considered as passive water transport systems, but recent studies uncovered hotspots of microbial diversity and activity in these urban environments. Indeed, the sewage system is the first step of city wastewater cycle and thus the most likely place where excreted antibiotic residues can induce the emergence of antibiotic resistant bacteria [25]. This is particularly relevant since recent studies have demonstrated that microbial communities present in the sewage recapitulate those found in the human gut microbiome [10]. Consequently, the exposure of human-derived bacteria to environmental pressures facilitates the emergence and spread of antibiotic resistant pathogens that can transmit and impact population's health [26].

The characterization of urban waters in our capital city revealed a remarkably distinct taxonomic composition of bacterial communities found in sewage and beach environments, suggesting that the sewage system is efficient in evacuating most hazardous microorganisms far away from areas of human exposure. Indeed, the vast majority of virulence and antibiotic resistance mechanisms associated to clinically relevant

pathogens were found in the sewage but not in beach samples. However, a more dense and longitudinal sampling is necessary to further characterize the dynamics of hazardous microorganisms circulating in these environments. Also, the comparison with metagenomes from hospital effluents along the city would provide a detailed picture of how nosocomial pathogens are being dispersed through the environment.

Indeed, many clinically relevant ARGs that we found in the city environment such as carbapenemases and ESBLs, have been frequently reported in nosocomial infections in Uruguay during the last decade [27–30]. This indicates that important antibiotic-resistant pathogens are being somehow transmitted among clinical settings and the urban environment, representing a public health threat. However, other ARGs such as metallo-β-lactamases that have also been reported in Uruguay [31–34] were not found in the sewage or beaches. Beyond technical biases are possible, this can be attributed to a differential capacity between distinct antibiotic-resistant clones to survive and spread in the environment; given that selective fitness of antibiotic-resistant pathogens (adapted to high antibiotic pressures in hospital settings) may be lower in less-exposed environmental waters [35].

In this sense, despite the urban environment may not be directly exposed to similar concentrations of antibiotics than those used to treat infections, sewage systems have been recognized as ARG reservoirs. So, considering that we found most environmental ARGs typically widespread in enterobacterial plasmids and that clinically relevant integrons were fundamentally recovered from sewage samples, genetic platforms for horizontal gene transfer can be playing a relevant role as reservoir of ARGs. Additionally, plasmids and integrons are prone to recombination [36] and genetic plasticity of certain bacteria has been proved to increase under subinhibitory pressures (as those probably found in the environment) with certain antibiotics [37], so the city sewage should be also considered as a birthplace for new antibiotic resistance mosaics mediated by recombination and horizontal gene transfer.

Regarding this, we were able to identify internationally disseminated pathogens like the *E. coli* ST-131 clonal complex, which encompasses multidrug-resistant genotypes with great capacity of recruiting new resistance genes. Also, we uncovered the presence of clinically underestimated bacteria like *Arcobacter cryaerophilus*, which is today considered an emerging waterborne pathogen and whose resistance to third generation cephalosporins has been already reported [38]. So, the compositional complexity of urban waters where different genotypes and gene repertories can coexist within a fluctuating bacterial community, opens the possibility of using environmental samples to monitor population's health. Accordingly, beyond providing a detailed characterization of individual virulence and antibiotic resistance mechanisms, our study supports the possibility of using high-resolution metagenomics to study the epidemiological dynamics of antibiotic-resistant pathogens using urban waters as a proxy at the population-level.

Conclusion

Our study represents a cross-sectional analysis of a metropolitan area encompassing more than 2.2 million inhabitants and, to the best of our knowledge, constitutes the first work using metagenomics to jointly characterize bacterial communities found in the sewage and beach waters of an entire city.

Our approach demonstrated its usefulness to identify antibiotic resistance determinants which were known to be present in nosocomial infections, as well as to uncover the presence of globally-widespread or underestimated pathogens with strain-level resolution. We consider that future longitudinal studies (time-wise) will be useful to monitor the fluctuations of bacterial communities, allowing the development of associative models with relevant metadata like outbreak information, rainfall or antibiotic prescription and stewardship.

The data generated in this initial study represent a baseline metagenomic characterization of environmental waters of Montevideo, which will be useful to guide future efforts to implement systematic studies aiming to evaluate antibiotic-resistant pathogen dynamics through time and space across different cities. This information can be later incorporated to improve public health surveillance for antibiotic-resistant pathogens.

Methods

Sample collection. We collected 20 water samples from 12 beaches and 8 sewage pipes or creeks where wastewater from the sewage system is poured directly. These sampling points are distributed along the whole coastal line of Montevideo, an uninterrupted system of sandy beaches that spans more than 20 kms (Fig. 1C, Additional file 2: Table S1). The samples were collected all the same day (around 3 hours elapsed from the first to the last). All samples were collected using sterile 200 mL plastic bottles and conserved in ice until they were processed within the same day.

DNA purification and metagenomic sequencing. Each sample was centrifuged at 10,000 g for 15 min at 4 ºC. Supernatants we discarded and pellets were processed using the FastDNA™ Spin Kit (MP Biomedicals) following the manufacturer's protocol. Paired end (2 × 125 bp) sequencing reads were generated on the Illumina HiSeq2500 machine, yielding in average 39.2 M (SD ± 5.4 M) reads per sample. On average 30.85 M (SD ± 4.07 M) reads passed an initial filter and were used in all further analyses. Sequencing data was deposited at the Sequence Read Archive (SRA) repository under BioProject number XXXXX.

Metagenomic data analysis. Initial data quality inspection was performed with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and then reads were filtered and trimmed using Trimmomatic [39] with the following parameters: LEADING:20, TRAILING:20, SLIDINGWINDOW:5:20, AVGQUAL:20 MINLEN:90. Resulting reads were used in downstream analyses. Characterization of

bacterial pathogens at the strain-level was performed with metaMLST [19], that tries to identify multilocus sequence typing alleles directly from metagenomic sequences.

First, an unbiased description of the variability among communities in sewage and beach was obtained by running Simka [13] with default parameters. Second, MetaPhlan2 [14] was used to identify species and to determine their relative abundances across samples. Beta diversities were calculated over taxonomic profiles using the Bray-Curtis distance as implemented in the Vegan R package [40] and alpha diversities were calculated using the Shannon index in the base R package [41].

Metagenomes were *de novo* assembled for each sample with MEGAHIT [42]. Then, contigs over 1 kb. were retained and merged at 99% of identity using CD-HIT-EST [43]. Resulting contigs were secondary assembled using Minimus2 [44] requiring a minimum overlap of 100 bp. with at least 95% of identity at contig boundaries. Genes were predicted on the resulting contigs using Prodigal [45]. Antibiotic resistance and virulence genes were identified with Abricate (https://github.com/tseemann/abricate) by comparing contigs against CARD [15] and VFdb [16], respectively. Only hits with query coverage > 90% and sequence identity > 70% were kept.

Integrons were identified using IntegronFinder [46]. We used MAFFT [47] (with the L-INS-i option) to align the amino acid sequences of IntI genes recovered from metagenomes together with reference sequences of class 1 (IntI1, AAQ16665.1), class 2 (IntI2, AAT72891.1), class 3 (IntI3, AAO32355.1), class 4 (IntI4, AAC 38424) and class 5 (IntI5, AAD 55407.2) integrases. The resulting alignment was used to build a phylogenetic tree with RAxML [48].

References

1. Robinson TP, Bu DP, Carrique-Mas J, Fèvre EM, Gilbert M, Grace D, et al. Antibiotic resistance is the quintessential One Health issue. Trans R Soc Trop Med Hyg. 2016;110:377-80.

2. Amos GCA, Zhang L, Hawkey PM, Gaze WH, Wellington EM. Functional metagenomic analysis reveals rivers are a reservoir for diverse antibiotic resistance genes. Vet Microbiol. 2014;171:441-7.

3. Murray A, Zhang L, Yin X, Zhang T, Buckling A, Snape J, et al. Novel insights into selection for antibiotic resistance in complex microbial communities. bioRxiv. 2018;323634.

4. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017;551:457-63.

5. The MetaSUB International Consortium. The Metagenomics and Metadesign of the Subways and Urban Biomes (MetaSUB) International Consortium inaugural meeting report. Microbiome. 2016;4:24.

6. Afshinnekoo E, Meydan C, Chowdhury S, Jaroudi D, Boyer C, Bernstein N, et al. Geospatial resolution of human and bacterial diversity with city-scale metagenomics. Cell Syst. 2015;1:72-87.

7. Hervé V, Leroy B, Pires ADS, Lopez PJ. Aquatic urban ecology at the scale of a capital: community structure and interactions in street gutters. ISME J. 2018;12:253-66.

8. Amos GCA, Ploumakis S, Zhang L, Hawkey PM, Gaze WH, Wellington EMH. The widespread dissemination of integrons throughout bacterial communities in a riverine system. ISME J. 2018;12:681-91.

9. Mohiuddin MM, Salama Y, Schellhorn HE, Golding GB. Shotgun metagenomic sequencing reveals freshwater beach sands as reservoir of bacterial pathogens. Water Res. 2017;115:360-9.

10. Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, et al. Sewage reflects the microbiomes of human populations. mBio. 2015;6:e02574-14.

11. Iraola G, Kumar N. Surveying what's flushed away. Nat Rev Microbiol. 2018;1.

12. Solo-Gabriele HM, Harwood VJ, Kay D, Fujioka RS, Sadowsky MJ, Whitman RL, et al. Beach sand and the potential for infectious disease transmission: observations and recommendations. J Mar Biol Assoc UK. 2016;96:101-20.

13. Benoit G, Peterlongo P, Mariadassou M, Drezen E, Schbath S, Lavenier D, et al. Multiple comparative metagenomics using multiset k-mer counting. PeerJ Comput Sci. 2016;2:e94.

14. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods. 2015;12:902.

15. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45:D566-73.

16. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 2012;40:D641-5.

17. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12:R60.

18. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA. 1998;95:3140-5.

19. Zolfo M, Tett A, Jousson O, Donati C, Segata N. MetaMLST: multi-locus strain-level bacterial typing from metagenomic samples. Nucleic Acids Res. 2017;45:e7.

20. Liu L, Lan R, Liu L, Wang Y, Zhang Y, Wang Y, et al. Antimicrobial resistance and cytotoxicity of *Citrobacter* spp. in Maanshan Anhui province, China. Front Microbiol. 2017;8:1357.

21. Petty NK, Zakour NLB, Stanton-Cook M, Skippington E, Totsika M, Forde BM, et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. Proc Natl Acad Sci USA. 2014;111:5694-9.

22. Figueras MJ, Levican A, Pujol I, Ballester F, Quilez MR, Gómez-Bertomeu F. A severe case of persistent diarrhoea associated with *Arcobacter cryaerophilus* but attributed to *Campylobacter* sp. and a review of the clinical incidence of *Arcobacter* spp. New Microbes New Infect. 2014;2:31-7.

23. Gershman MD, Kennedy DJ, Noble-Wang J, Kim C, Gullion J, Kacica M, et al. Multistate outbreak of *Pseudomonas fluorescens* bloodstream infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. Clin Infect Dis. 2008;47:1372-9.

24. Jelic A, Rodriguez-Mozaz S, Barceló D, Gutierrez O. Impact of in-sewer transformation on 43 pharmaceuticals in a pressurized sewer under anaerobic conditions. Water Res. 2015;68:98-108.

25. Auguet O, Pijuan M, Borrego CM, Rodriguez-Mozaz S, Triadó-Margarit X, Giustina SVD, et al. Sewers as potential reservoirs of antibiotic resistance. Sci Total Environ. 2017;605:1047-54.

26. Baquero F, Martínez J-L, Cantón R. Antibiotics and antibiotic resistance in water environments. Curr Opin Biotechnol. 2008;19:260-5.

27. Vignoli R, Cordeiro NF, García V, Mota MI, Betancor L, Power P, et al. New TEM-derived extended-spectrum β-lactamase and its genomic context in plasmids from *Salmonella enterica* serovar derby isolates from Uruguay. Antimicrob Agents Chemother. 2006;50:781-4.

28. Vignoli R, Varela G, Mota MI, Cordeiro NF, Power P, Ingold E, et al. Enteropathogenic *Escherichia coli* strains carrying genes encoding the PER-2 and TEM-116 extended-spectrum β-lactamases isolated from children with diarrhea in Uruguay. J Clin Microbiol. 2005;43:2940-3.

29. García-Fulgueiras V, Bado I, Mota MI, Robino L, Cordeiro NF, Varela A, et al. Extended-spectrum β-lactamases and plasmid-mediated quinolone resistance in enterobacterial clinical isolates in the paediatric hospital of Uruguay. J Antimicrob Chemother. 2011;66:1725-9.

30. Medina-Presentado JC, Seija V, Vignoli R, Pontet J, Robino L, Cordeiro NF, et al. Polyclonal endemicity of *Acinetobacter baumannii* in ventilated patients in an intensive care unit in Uruguay. Int J Infect Dis. 2013;17:e422-7.

31. Ingold AJ, Castro M, Nabón A, Borthagaray G, Márquez C. VIM-2 metallo-ß-lactamase gen detection in a class 1 integron associated to bla(CTX-M-2) in a *Pseudomonas aeruginosa* clinical isolate in Uruguay: first communication. Rev Argent Microbiol. 2011;43:198-202.

32. Bado I, Papa-Ezdra R, Cordeiro N, Outeda M, Caiata L, García-Fulgueiras V, et al. Detection of qnrVC6, within a new genetic context, in an NDM-1-producing *Citrobacter freundii* clinical isolate from Uruguay. J Glob Antimicrob Resist. 2018;14:95-8.

33. Seija V, Medina Presentado JC, Bado I, Papa Ezdra R, Batista N, Gutierrez C, et al. Sepsis caused by New Delhi metallo-β-lactamase (blaNDM-1) and qnrD-producing *Morganella morganii*, treated successfully with fosfomycin and meropenem: case report and literature review. Int J Infect Dis. 2015;30:20-6.

34. Vignoli R, Cordeiro N, Seija V, Schelotto F, Radice M, Ayala J, et al. Genetic environment of CTX-M-2 in *Klebsiella pneumoniae* isolates from hospitalized patients in Uruguay. Rev Argent Microbiol. 2006;38:84-8.

35. Hall AR, Angst DC, Schiessl KT, Ackermann M. Costs of antibiotic resistance – separating trait effects and selective effects. Evol Appl. 2015;8:261-72.

36. Mazel D. Integrons: agents of bacterial evolution. Nat Rev Microbiol. 2006;4:608-20.

37. López E, Blázquez J. Effect of subinhibitory concentrations of antibiotics on intrachromosomal homologous recombination in *Escherichia coli*. Antimicrob Agents Chemother. 2009;53:3411-5.

38. Ferreira S, Oleastro M, Domingues F. *Arcobacter* spp. in food chain—from culture to omics. In: Foodborne Pathogens and Antibiotic Resistance. Wiley-Blackwell. 2017;73-117.

39. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-20.

40. Oksanen J. Vegan: R functions for vegetation ecologists. 2005. Available from: https://cran.r-project.org/web/packages/vegan/index.html.

41. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statitical Computing. 2008. Available from: http://www.R-project.org.

42. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674-6.

43. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150-2.

44. Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M. Next generation sequence assembly with AMOS. Curr Protoc Bioinformatics. 2011;11.

45. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.

46. Cury J, Jové T, Touchon M, Néron B, Rocha EP. Identification and analysis of integrons and cassette arrays in bacterial genomes. Nucleic Acids Res. 2016;44:4539-50.

47. Katoh K, Standley DM. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772-80.

48. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312-3.

309

310 **Figure legends**

311

312 **Figure 1. Community composition of beach and sewage waters of Montevideo.** A) Heatmap showing a
313 clustering analysis based on k-mer distances evidencing a complete separation between sewage (red) and
314 beach (blue) samples. B) Clustering analysis separating sewage (red) from beach (blue) samples obtained by
315 comparing beta diversities (dissimilarity between samples) calculated from relative abundance profiles of
316 bacterial species. C) Sampling points along the coast of Montevideo (grey shade). Sewage water samples are
317 in red and beach water samples are in blue.

318

319 **Figure 2. Occurrence of antibiotic resistance mechanisms.** A) Circos representation showing the presence
320 of ARGs across sewage (red) or beach (blue) samples. Links are drawn when a certain ARG (right blocks) is
321 found in a certain sample (left blocks). Genes are colored according to antibiotic classes. Barplots above each
322 left side block indicate the alpha diversity of ARGs within each sample. B) Boxplots showing ARG counts for
323 different antibiotic classes in beach (blue) and sewage (red) samples.

324

325 **Figure 3. Distribution of ARGs in mobile genetic elements.** A) Barplot showing the frequency of ARGs in
326 bacterial chromosomes (yellow) or plasmids (violet) summarized by antibiotic class. B) Taxonomic
327 distribution of bacterial plasmids carrying ARGs found in urban metagenomes. C) Phylogeny of reference
328 integrase genes (grey) and those recovered from beach (blue) and sewage (red) samples.

Figure 4. Identification of sewage biomarker taxa. A) Barplots showing LDA (Linear Discrimination Analysis) scores for bacteria genera that distinguish beach (blue) from sewage (red) samples. B) Boxplots comparing relative abundances between beach (blue) and sewage (red) samples for bacterial genera enclosing pathogenic species.

Declarations

Ethics approval and consent to participate. Not applicable.

Consent for publication. Not applicable.

Availability of data and material. All data generated during this study is available at the Sequence Read Archive (SRA) under BioProject number XXXXX.

Competing interests. The authors declare that they have no competing interests.

Funding. This work was funded by the Agencia Nacional de Investigación e Innovación (ANII) grant number OPR_X_2016_1_1006944.

Author's contributions. GI and GG conceived the idea. BD, EA and CM generated the data. PF, VA, CS, MG and GI analyzed and interpreted the data. PF, VA, CS, GG and GI wrote the manuscript. All authors read and approved the final manuscript.

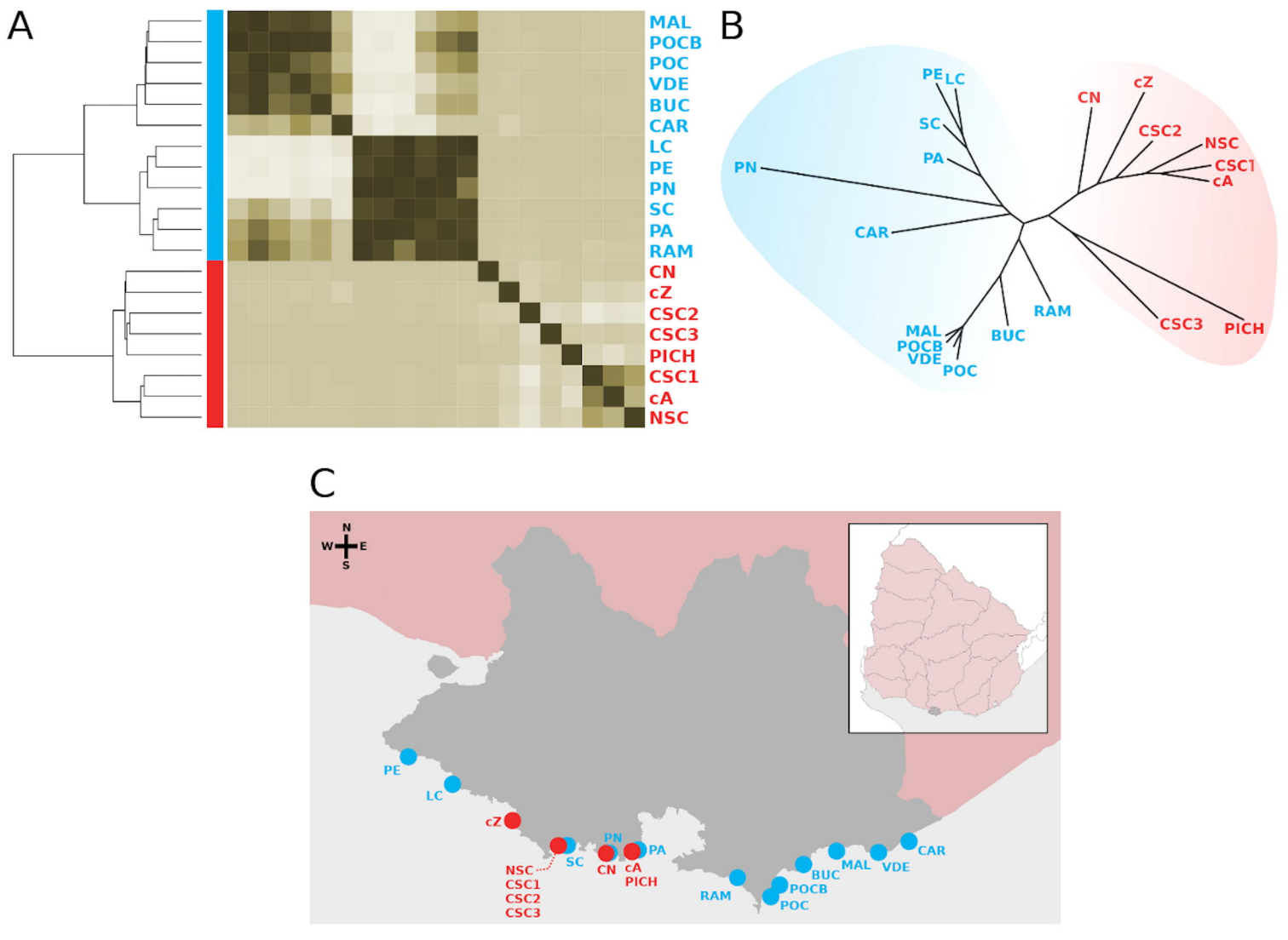
Acknowledgements. We thank technicians from Intendencia de Montevideo for valuable assistance during sampling and data collection. We thank Hugo Naya for providing lab space and computational resources at the Bioinformatics Unit in the Institut Pasteur Montevideo during the development of this work.

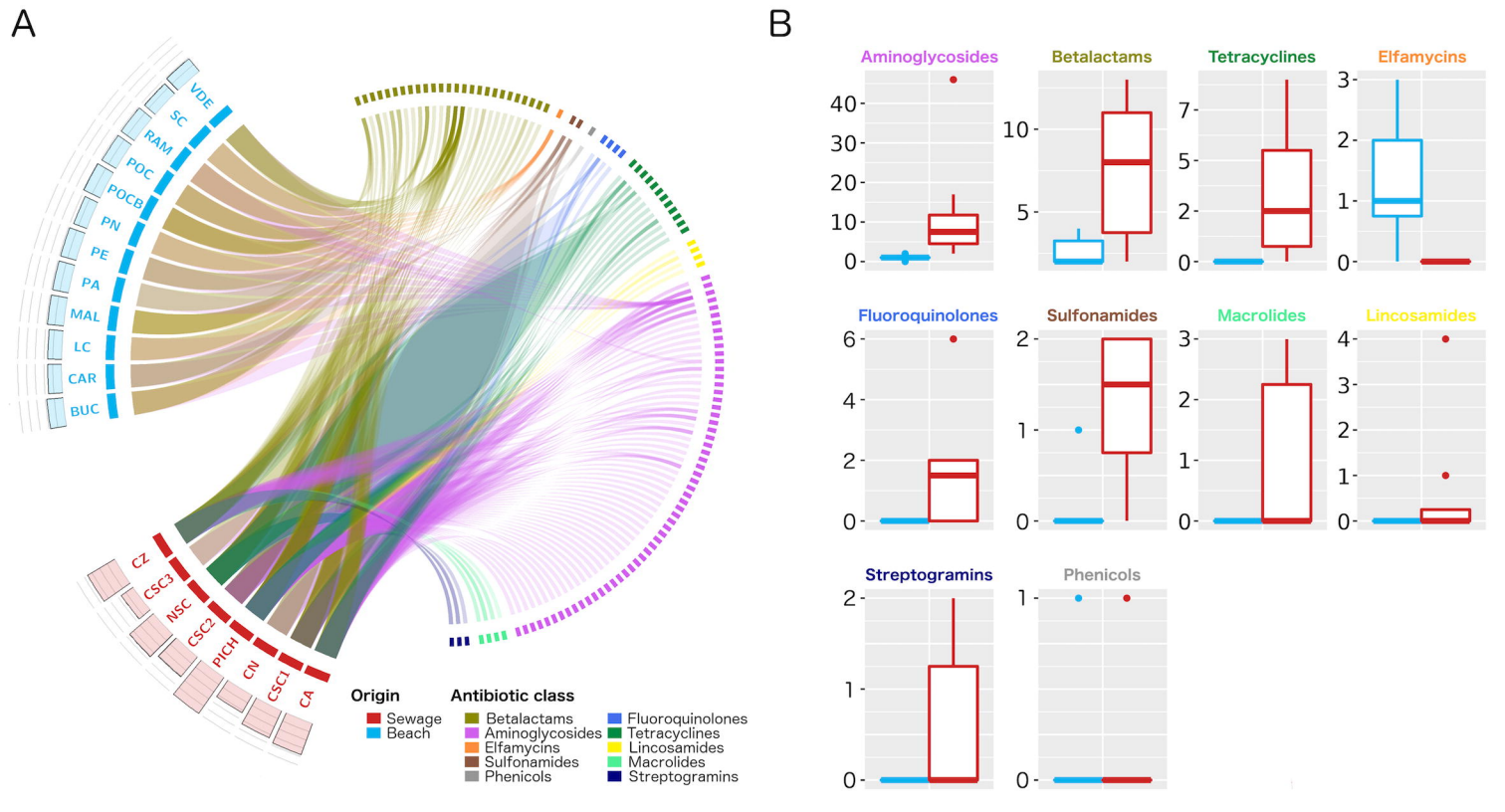
365

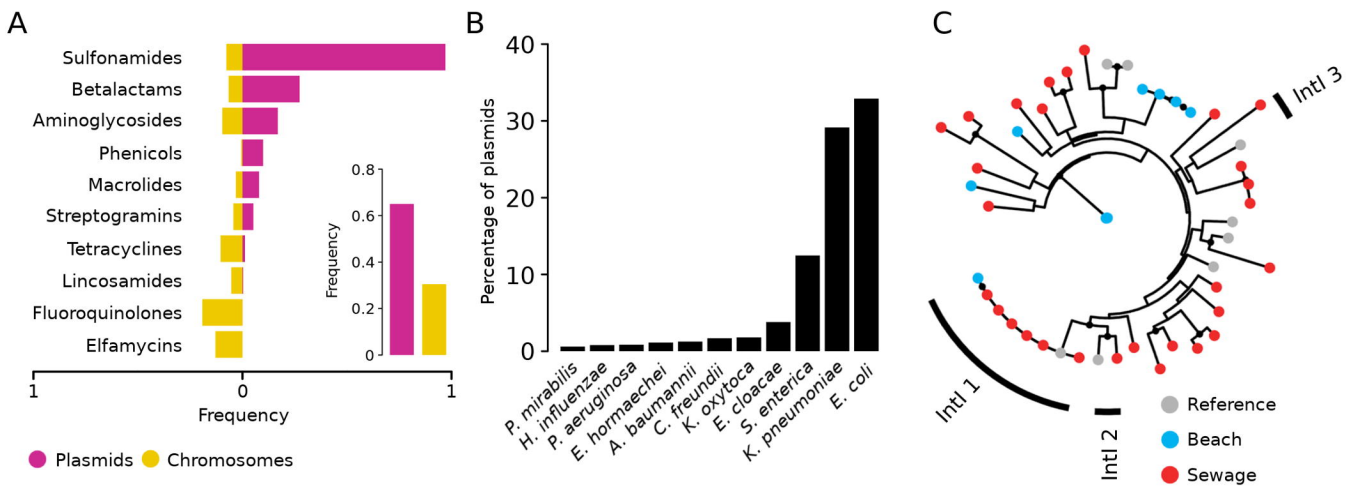
366

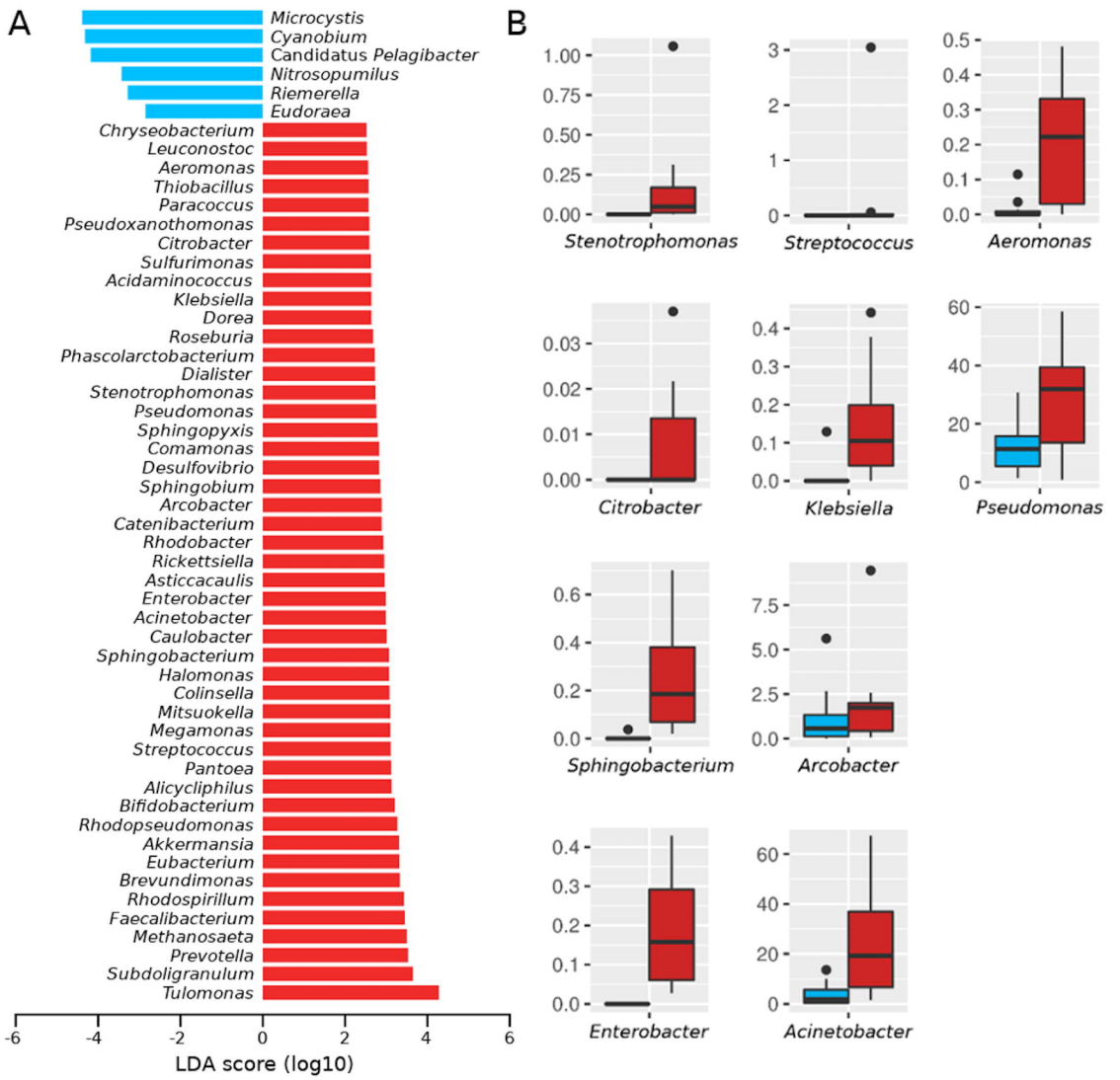
367

368


A
MAL
POCB
POC
VDE
BUC
CAR
LC
PE
PN
SC
PA
RAM
CN
cZ
CSC2
CSC3
PICH
CSC1
cA
NSC
B
PE
LC
SC
PA
PN
CAR
CN
cZ
CSC2
NSC
CSC1
cA
RAM
MAL
POCB
VDE
POC
BUC
CSC3
PICH
C
N
W
E
S
PE
LC
cZ
NSC
CSC1
CSC2
CSC3
SC
PN
CN
cA
PICH
PA
RAM
BUC
POCB
POC
MAL
VDE
CAR


A
VDE
SC
RAM
POC
POCB
PN
PE
PA
MAL
LC
CAR
BUC
CZ
CSC3
NSC
CSC2
PICH
CN
CSC1
CA
Origin
Sewage
Beach
Antibiotic class
Betalactams
Aminoglycosides
Elfamycins
Sulfonamides
Phenicols
Fluoroquinolones
Tetracyclines
Lincosamides
Macrolides
Streptogramins
B
Aminoglycosides
Betalactams
Tetracyclines
Elfamycins
Fluoroquinolones
Sulfonamides
Macrolides
Lincosamides
Streptogramins
Phenicols


A
Sulfonamides
Betalactams
Aminoglycosides
Phenicols
Macrolides
Streptogramins
Tetracyclines
Lincosamides
Fluoroquinolones
Elfamycins
0.8
0.6
0.4
0.2
0
Frequency
1
0
1
Frequency
Plasmids
Chromosomes
B
40
30
20
10
0
Percentage of plasmids
P. mirabilis
H. influenzae
P. aeruginosa
E. hormaechei
A. baumannii
C. freundii
K. oxytoca
E. cloacae
S. enterica
K. pneumoniae
E. coli
C
IntI 3
IntI 1
IntI 2
Reference
Beach
Sewage


A
Microcystis
Cyanobium
Candidatus Pelagibacter
Nitrosopumilus
Riemerella
Eudoraea
Chryseobacterium
Leuconostoc
Aeromonas
Thiobacillus
Paracoccus
Pseudoxanothomonas
Citrobacter
Sulfurimonas
Acidaminococcus
Klebsiella
Dorea
Roseburia
Phascolarctobacterium
Dialister
Stenotrophomonas
Pseudomonas
Sphingopyxis
Comamonas
Desulfovibrio
Sphingobium
Arcobacter
Catenibacterium
Rhodobacter
Rickettsiella
Asticcacaulis
Enterobacter
Acinetobacter
Caulobacter
Sphingobacterium
Halomonas
Colinsella
Mitsuokella
Megamonas
Streptococcus
Pantoea
Alicycliphilus
Bifidobacterium
Rhodopseudomonas
Akkermansia
Eubacterium
Brevundimonas
Rhodospirillum
Faecalibacterium
Methanosaeta
Prevotella
Subdoligranulum
Tulomonas
-6
-4
-2
0
2
4
6
LDA score (log10)
B
Stenotrophomonas
Streptococcus
Aeromonas
Citrobacter
Klebsiella
Pseudomonas
Sphingobacterium
Arcobacter
Enterobacter
Acinetobacter