



Accumulation and expression of multiple antibiotic resistance genes in *Arcobacter cryaerophilus* that thrives in sewage

Jess A. Millar and Rahul Raghavan

Biology Department, Portland State University, Portland, OR, United States

ABSTRACT

We explored the bacterial diversity of untreated sewage influent samples of a wastewater treatment plant in Tucson, AZ and discovered that *Arcobacter cryaerophilus*, an emerging human pathogen of animal origin, was the most dominant bacterium. The other highly prevalent bacteria were members of the phyla Bacteroidetes and Firmicutes, which are major constituents of human gut microbiome, indicating that bacteria of human and animal origin intermingle in sewage. By assembling a near-complete genome of *A. cryaerophilus*, we show that the bacterium has accumulated a large number of antibiotic resistance genes (ARGs) probably enabling it to thrive in the wastewater. We also determined that a majority of ARGs was being expressed in sewage, suggestive of trace levels of antibiotics or other stresses that could act as a selective force that amplifies multidrug resistant bacteria in municipal sewage. Because all bacteria are not eliminated even after several rounds of wastewater treatment, ARGs in sewage could affect public health due to their potential to contaminate environmental water.

Subjects Environmental Sciences, Genomics, Microbiology

Keywords *Arcobacter cryaerophilus*, Sewage, Multidrug resistance

INTRODUCTION

Over the past few decades, based on numerous studies that examined the bacterial composition of wastewater during varying stages of treatment, there is growing evidence that sewage is an important hub for horizontal gene transfer (HGT) of antibiotic resistance genes (e.g., [Baquero, Martínez & Cantón, 2008](#); [Zhang, Shao & Ye, 2012](#); [Rizzo et al., 2013](#); [Pehrsson et al., 2016](#)). Additionally, studies have shown that discharge of treated sewage allows these concentrated communities to spread into environmental water ([Okoh et al., 2007](#); [Varela & Manaia, 2013](#)). The *Arcobacter* genus is commonly detected in sewage treatment plants around the world ([Collado et al., 2008](#); [Zhang, Shao & Ye, 2012](#); [Varela & Manaia, 2013](#)). This sparsely studied Epsilonproteobacteria is frequently associated with veterinary diseases, and is closely related to *Campylobacter*, and is considered an emerging human pathogen that causes enteritis and bacteremia ([Kabeya et al., 2004](#); [Morita et al., 2004](#); [Collado et al., 2008](#)). In addition, *Arcobacter* is known to be resistant to a wide array of commonly used antibiotics, with varying resistance profiles observed in different species ([Houf et al., 2004](#); [Abay et al., 2012](#); [Rahimi, 2014](#)), but the genes that enable antibiotic resistance are mostly unknown ([Abdelbaqi et al., 2007](#); [Miller et al., 2007](#)).

Submitted 30 November 2016

Accepted 3 April 2017

Published 25 April 2017

Corresponding author

Rahul Raghavan,
rahul.raghavan@pdx.edu

Academic editor

A. Murat Eren

Additional Information and
Declarations can be found on
page 10

DOI 10.7717/peerj.3269

© Copyright
2017 Millar and Raghavan

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

In this study, we examined the bacterial diversity and the presence and expression of antibiotic resistance genes (ARGs) and virulence factors in untreated sewage. Our analyses revealed that an *A. cryaerophilus* strain that contained multiple ARGs was a major constituent of the sewage microbiome. In addition, we detected a large number of expressed ARGs and virulence factors in *A. cryaerophilus* and in the rest of the sewage microbiome, which portends potential public health risk if bacteria carrying these genes contaminate public water resources.

MATERIALS & METHODS

Sewage sample collection, DNA and RNA extraction, and deep-sequencing

Three untreated sewage influent samples (50 ml) were collected at the Roger Road Wastewater Reclamation Facility, Tucson, Arizona (March 2012) and immediately transferred on ice to the laboratory and stored at -80°C until further use. The sewage samples were spun down ($12,000 \times G$, 15 min, 4°C) and the pellets were suspended in 1 ml of TRI reagent (Life Technologies). Total RNA from each sample was extracted from the aqueous phase and corresponding DNA was isolated from the interphase using a protocol provided by the manufacturer. RNA samples were treated with TurboDNase (Life Technologies) to remove contaminating DNA, and PCR reactions using 16S rDNA primers were performed to confirm complete DNA removal. Furthermore, RNA samples were depleted of ribosomal RNA using RiboZero Bacteria and RiboZero Human kits (Illumina). Around 100 ηg of RNA from each sample was used to prepare directional mRNA-seq libraries using the Illumina Small RNA Sample Preparation Kit and Directional mRNA-seq Sample Preparation protocol provided by Illumina Inc. DNA samples were further purified using DNeasy kit (Qiagen) and around 5 μg of DNA from each sample was used to prepare paired-end DNA-seq libraries using the Paired-End sample Preparation Kit (Illumina). All RNA libraries were pooled into a single lane, and all DNA samples were pooled into another lane of Illumina HiSeq 2000 and were sequenced at the Yale Center for Genome Analysis (RNA-seq: single-end, 75 cycles; DNA-seq: paired end, 2×75 cycles) using standard adapters. All DNA and RNA reads have been deposited at NCBI (BioProject [PRJNA354077](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA354077)).

Taxonomic classification

DNA reads were cleaned by removing adapters and were filtered by quality ($\geq Q20$) and length (≥ 50 bp) using Trimmomatic v0.32 ([Bolger, Lohse & Usadel, 2014](#)). Each library was assembled into contigs using 17,000,000 reads and IDBA-UD ([Peng et al., 2012](#)), and normalized to the smallest library size (6,000 contigs) using Seqtk (<https://github.com/lh3/seqtk>) ([Perner et al., 2014](#)) (Table S1). All contigs were searched against the NCBI nt database using BLAST and analyzed in MEGAN ([Huson et al., 2007](#)), requiring at least 70% of the query sequence to align with the subject sequence with $\geq 70\%$ identity to be assigned to a given phylum. Contigs classified at the phylum level (48%, 57%, and 54% of contigs from the three samples, respectively) were used to determine their detailed taxonomic positions. Remaining contigs either did not have significant BLAST

hits or mapped to unidentified environmental samples; however, all contigs were used in downstream analyses (detection of antibiotic resistance genes, virulence factors etc.).

Identification of antibiotic resistance genes, virulence factors, transposases, and bacteriophage genes

Contigs were run through MetaProdigal ([Hyatt et al., 2012](#)) to identify encoded ORFs, which were annotated by mapping to antibiotic resistance genes, virulence factors, bacterial transposases, and prophages obtained respectively from CARD, PATRIC, InterPro, and PHAST databases using PHMMER, with an *E*-value of at least $1e-10$ as the cutoff ([Finn, Clements & Eddy, 2011](#); [Zhou et al., 2011](#); [McArthur et al., 2013](#); [Wattam et al., 2014](#); [Mitchell et al., 2015](#)). RNA reads from each sample were filtered by quality ($\geq Q20$) and length (≥ 50 bp) using Trimmomatic v0.32 and normalized using Seqtk to 50,000,000 reads. They were mapped to annotated ORFs using CLC Genomic Workbench v6.5. A strict mapping criterion (at least 95% of each read should map with at least 95% identity to the mapped region) was used in order to minimize non-specific mapping. Genes were filtered and considered expressed based on at least 10 reads mapping to each ORF. Statistical analysis was conducted using SAS Studio v3.4 (SAS Institute, Cary NC).

***Arcobacter* genome assembly**

DNA reads from all three samples were pooled to gain enough coverage depth, and were assembled into contigs using IDBA-UD ([Peng et al., 2012](#)). All contigs were searched against the NCBI nt database using BLASTN and analyzed in MEGAN ([Huson et al., 2007](#)). All *Arcobacter* gene sequences were downloaded from NCBI, and using PHMMER, *Arcobacter* contigs present in our data were identified with at least $1e-10$ *E*-value as the cutoff. These contigs were extracted and run through the differential coverage binning procedure for metagenomic data, as described previously ([Albertsen et al., 2013](#)). In brief, contigs were binned based on coverage, tetranucleotide frequency, GC%, and length, and then examined for presence of essential single copy genes. Phylogenetic analyses were conducted on nucleotide sequences using several housekeeping genes to identify the bins containing *A. cryaerophilus* (marked in blue in [Fig. S1](#)). One genome bin with $\sim 200\times$ coverage that contained all *A. cryaerophilus* housekeeping genes was selected for secondary refinement and finishing (top right cluster in [Fig. S2](#)). This cluster of contigs was isolated and all original trimmed DNA reads were mapped against them using Bowtie2 v2.1.0 ([Langmead & Salzberg, 2012](#)). All mapped reads were then reassembled into contigs using IDBA-UD. These contigs were combined with all original trimmed DNA reads for scaffold extension using SSPACE ([Boetzer et al., 2011](#)) into a final scaffold of ~ 1.8 Mb over 456 contigs. To check for completeness of the assembled *A. cryaerophilus* genome, we used a single-copy gene database ([Albertsen et al., 2013](#)), and as a control we performed the same analysis with the *A. butzleri* (CP000361.1) genome. Visual representation of draft genome was created using Circos ([Krzyszowski et al., 2009](#)). The draft genome has been deposited at NCBI under the accession [LNTC00000000](#).

Detection of HGT

Horizontally acquired genes were identified using HGTector ([Zhu, Kosoy & Dittmar, 2014](#)). *Arcobacter* was set as self-group, and Campylobacterales was set as exclusion group. This method captured HGT events where only *Arcobacter* has acquired a particular gene from outside of Campylobacterales and ignored any events where the genes could also have been transferred elsewhere within the order. This conservative approach was used due to the dearth of annotated genomes within Campylobacterales. BLASTN parameter thresholds were set at 70% identity and an *E*-value of at least $1e-5$. Several putative HGT genes were examined using phylogenetic analysis to validate the HGTector data ([Fig. S3](#)).

Phylogenetic analysis

Nucleotide sequence alignment for all trees was performed using Clustal Omega ([Sievers et al., 2011](#)), and ambiguously aligned regions were removed using Gblocks ([Talavera & Castresana, 2007](#)). The evolution model GTR+I+G (General Time Reversible plus Invariant sites plus Gamma distribution) used for all trees was selected using jModelTest2 ([Darriba et al., 2012](#)). Bayesian trees were constructed using MrBayes as implemented in Geneious ([Huelsenbeck & Ronquist, 2001](#); [Kearse et al., 2012](#)). A chain length of 1,000,000 was used with a burn-in fraction of 25% and sampling every 100 trees. Maximum Likelihood trees were constructed using RAxML ([Stamatakis, Hoover & Rougemont, 2008](#)) as implemented in Geneious with 1,000 bootstrap replicates to confirm Bayesian topologies. *Helicobacter pylori* ([AJ558222.1](#)) was used to root all phylogenetic trees.

RESULTS AND DISCUSSION

A. cryaerophilus thrives in sewage

For the three sewage samples, taxonomic labels were assigned to at least the phylum level for all contigs with significant BLAST hits. There was no significant difference in bacterial distribution between the three samples ([Fig. 1](#)); hence, average values are presented hereafter. Members of the phylum Proteobacteria (67% of total hits) was the most prevalent bacteria, followed by Bacteroidetes (23%) and Firmicutes (9%). A more comprehensive study that examined several sewage samples from across the USA observed a similar pattern of bacterial phyla abundance ([Shanks et al., 2013](#)). However, at the genus level, *Arcobacter* (an Epsilonproteobacteria) was the most dominant bacterium in our study, making up 39% of all annotated contigs, unlike members of Gammaproteobacteria (38% of all pyrotags) in the previous study ([Shanks et al., 2013](#)).

Arcobacter is commonly associated with both humans and farm animals ([Collado et al., 2008](#)), the latter perhaps more relevant to this specific wastewater treatment plant because agriculture accounts for the largest use of water at around 70% of all water demand within the state of Arizona ([ADWR, 2009](#)). In addition, in Tucson, the wastewater treatment plant served both agricultural and municipal areas ([PAG, 2006](#)). Members of the phyla Bacteroidetes and Firmicutes, two of the most abundant bacteria in human gut, were also abundant in the sewage samples ([Cho & Blaser, 2012](#); [Jandhyala et al., 2015](#)). Taken together, these data highlight the important role that sewage systems play as an arena where bacteria of human and animal origin interact, which could promote the exchange of genes

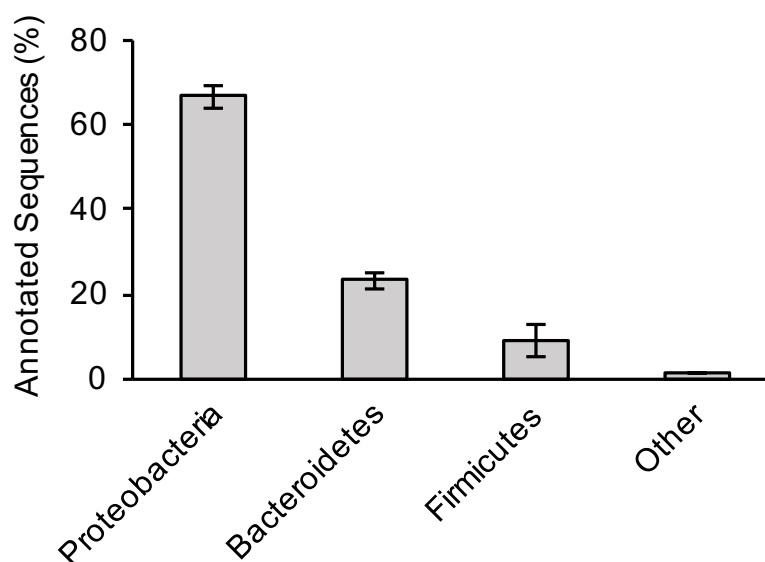


Figure 1 Bacterial composition of sewage samples. Percentage of sewage contigs that were assigned to each bacterial phylum is shown. Data represent means of three samples \pm standard deviations.

between the two groups (Baquero, Martínez & Cantón, 2008; Gaze et al., 2013; Rizzo et al., 2013). After conducting genomic binning of the *Arcobacter* contigs utilizing various factors such as coverage depth, GC%, and tetranucleotide frequency (Albertsen et al., 2013), we were able to identify \sim 80% of these contigs as belonging to *A. cryaerophilus*, an emerging human pathogen that is commonly associated with diseases such as bovine reproductive disorders, diarrhea and hemorrhagic colitis in cattle and sheep (Schroeder-Tucker et al., 1996; Ho, Lipman & Gaastra, 2006).

Presence and expression of multiple ARGs in *A. cryaerophilus*

To better characterize *A. cryaerophilus*, we assembled a near-complete genome from the DNA-seq reads (Fig. 2). Based on the presence of 100 out of 106 single copy genes (Albertsen et al., 2013) with zero redundant copies, we estimate that the *A. cryaerophilus* genome is \sim 95% complete and contains 2,419 ORFs (including partial genes at the ends of contigs) (Table S2). Among these ORFs, 115 (5% of ORFs) encode antibiotic resistance genes (ARGs) belonging to 25 categories as defined by the CARD database (Table S3) (McArthur et al., 2013). Macrolide resistance made up the majority of annotated ARGs (26, 23%) (Table S3), with fluoroquinolones (18, 16%), aminocoumarin (17, 15%) and vancomycin (13, 11%) resistance genes being the next largest groups. Because gene expression is a good representation for functional gene activity, we analyzed the expression of *A. cryaerophilus* genes using RNA-seq and discovered that all 115 putative ARGs genes were being expressed (Fig. 3A; Data Set S1). In comparison, *Helicobacter pylori*, a closely related Epsilonproteobacteria contain 59 ARGs (4% of genes); however, in both bacteria around 50% of ARGs consisted of efflux pumps (Paulsen, Sliwinski & Saier, 1998). It should be noted that although all ARGs were found to be expressed in *A. cryaerophilus*, the median level of expression of single copy genes (3,887 reads mapped) was found to be 10 \times higher

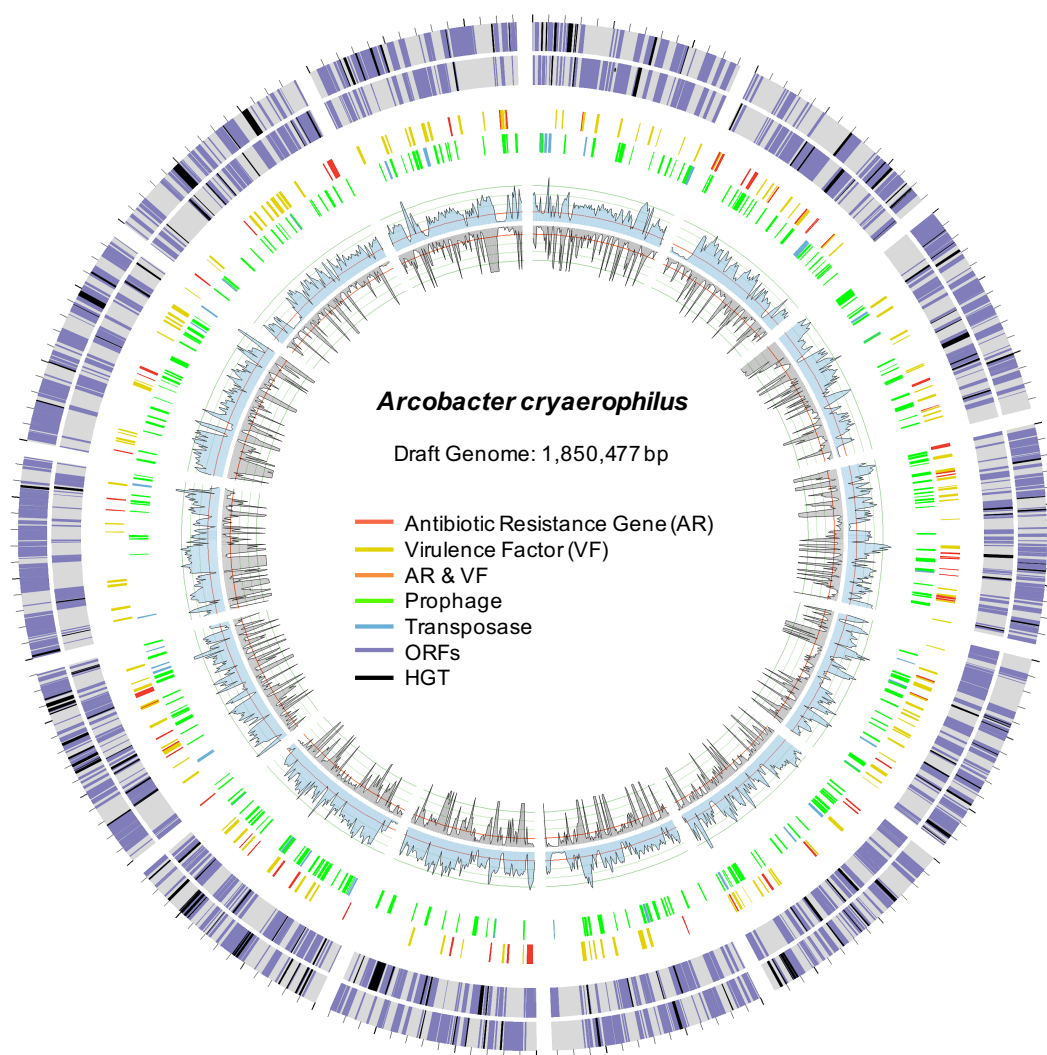


Figure 2 Draft genome of *A. cryaerophilus*. Two outer rings show ORFs (purple) on forward and reverse strands, respectively. Black blocks represent horizontally acquired genes. Each tick mark represents 10,000 bp. Middle two rings show positions of features annotated in the center. Inner blue and grey rings show DNA-seq coverage (mean of three samples) and RNA-seq transcription levels (mean of three samples), respectively. Note that 456 original contigs were randomly assigned to 14 equal fragments for easy visualization.

than the median level of expression of ARGs (363 reads mapped), probably because higher expression of many ARGs requires strong induction.

To determine the prevalence of ARGs in the total sewage, we extended our analysis to all contigs assembled in our study. Out of the 60,723 ORFs encoded in the sewage contigs, 2,606 ORFs matched 42 antibiotic resistance categories (Table S4). Using RNA-seq, we determined that 2,106 (81%) of these ORFs were expressed (Fig. 3B). Of the 2,106 putative antibiotic resistance genes expressed in the sewage samples, macrolide resistance genes made up the largest portion (538, 26%) (Table S4). The next two largest groups were fluoroquinolone resistance (378, 18%) and tetracycline resistance (339, 16%) genes.

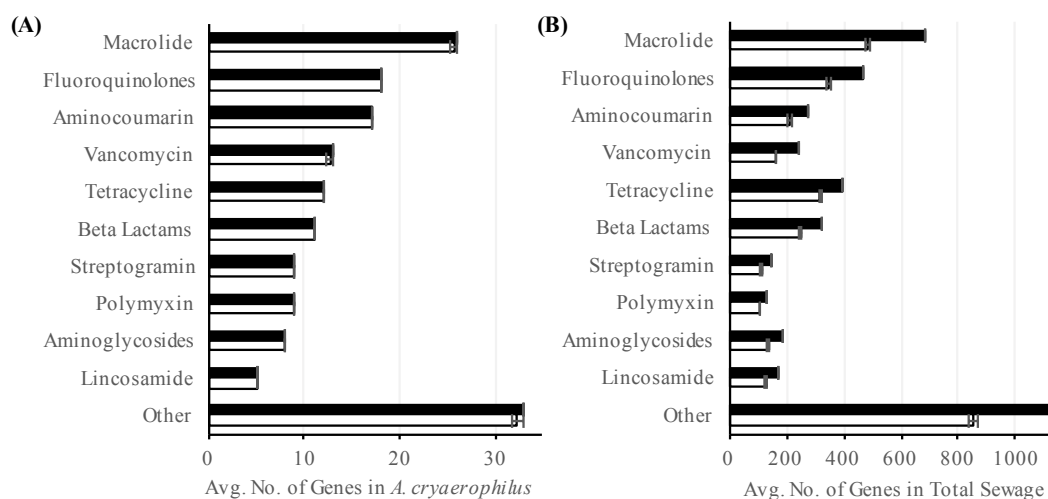


Figure 3 Abundance of antibiotic resistance genes. Number of annotated (black) and expressed (white) antibiotic resistance categories in (A) *A. cryaerophilus* and (B) total sewage sample is shown. Data represent means of three samples \pm standard deviations.

The expression of antibiotic resistance genes could be due to the presence of numerous antibiotics in urban wastewater (Heberer, 2002; Rizzo et al., 2013), which could select for multidrug resistant bacteria, thereby aggravating an already dire situation (Baquero, Martínez & Cantón, 2008; Rizzo et al., 2013; Wellington et al., 2013; Amos et al., 2014). It is also possible that ARGs were being expressed constitutively or in response to stress (Poole, 2012). Additionally, previous studies have shown that ARGs are expressed in a wide variety of environments even in the absence of known anthropogenic antibiotic pressure (Udikovic-Kolic et al., 2014; Versluis et al., 2015; Noyes et al., 2016); hence, further study is required to determine the stimuli for the observed ARG expression. Interestingly, although the total sewage contigs contained 23 \times more ARGs than in *A. cryaerophilus* contigs (2,606 vs. 115), 44% of DNA reads mapped to *A. cryaerophilus* ARGs, indicating that while the sewage contained high diversity of ARGs, most non-*A. cryaerophilus* ARGs were of low abundance.

Signatures of HGT in *A. cryaerophilus* genome

We compared our draft genome of *A. cryaerophilus* to the published genome of *A. butzleri* (CP000361.1), a closely related human and animal pathogen that has been studied much more extensively than *A. cryaerophilus* (Vandenberg et al., 2004; Miller et al., 2007; Collado et al., 2008) (Fig. S4). As observed previously in other members of this genus (Karadas et al., 2013; Merga et al., 2013), the two *Arcobacter* species only shared 1,337 genes (\sim 50%) (Table 1, Data Set S2). A comparison of the two genomes was also conducted using RAST (Overbeek et al., 2014), which showed that merely 846 genes with known functions were shared between *A. butzleri* and *A. cryaerophilus* (Data Set S3). This sizable variation in gene content between the two species indicates that HGT could have played a prominent role in shaping the genomes of *Arcobacter* species. Concomitantly, even after using a very conservative threshold, we detected 209 (9%) and 228 (10%) horizontally acquired genes in *A. cryaerophilus* and *A. butzleri*, respectively (Table 1). While similar in scale, only 73

Table 1 Comparison of *Arcobacter* genomes.

Features	<i>A. cryaerophilus</i>	<i>A. butzleri</i>	Overlap
Total ORFs	2,419 ^a	2,259	1,337
Horizontally Acquired ORFs	209	228	73
Antibiotic Resistance			
Categories	25	29	23
Genes	115	140	54
Virulence Factors			
Categories	24	24	22
Genes	232	185	92
Transposases			
Categories	7	7	5
Genes	61	57	15
Prophages			
Genes	290	320	173
GenBank Accession	LNTC00000000	CP000361.1	

Notes.

^aIncludes partial genes at the ends of contigs.

HGT-origin genes were shared between the two genomes, indicating that parallel HGT events have molded the genomes of the two *Arcobacter* species.

HGT is known to promote ARG dissemination between bacteria (Hawkey & Jones, 2009; Gaze et al., 2013; Pehrsson et al., 2016); hence, we compared ARGs present in *A. cryaerophilus* to those present in *A. butzleri* in order to identify those that are of possible HGT origin. We identified 140 putative genes belonging to 29 antibiotic resistance categories in *A. butzleri*, and out of the 25 antibiotic resistance categories present in *A. cryaerophilus*, 23 were present in *A. butzleri*, with two categories (Glycylcycline and Roxithromycin resistance) found only in *A. cryaerophilus*, and six categories (Bicyclomycin, Elfamycin, Isoniazid, Kanamycin, Streptomycin, and Teicoplanin resistance) exclusive to *A. butzleri* (Table 1). However, within each category large differences in gene content was observed between the two bacteria, with only 54 genes shared between *A. cryaerophilus* and *A. butzleri*. These data show that even though the antibiotic resistance capabilities of both bacteria overlap, their respective gene repertoires were largely assembled through independent HGT events. Transposons and bacteriophages are important agents of HGT in bacteria, and we found several transposases and bacteriophage ORFs in *A. cryaerophilus* (61 transposase ORFs, 290 phage ORFs) and *A. butzleri* (57, 320) (Fig. 2 and Table 1, Table S2). Additionally, we discovered that several of the ARGs in *A. cryaerophilus* were located in close proximity to transposases or bacteriophage genes (Fig. 4), suggestive of a role for these mobile genetic elements in the accumulation of ARGs in this emerging pathogen.

Presence and expression of virulence factors in sewage

In addition to ARGs, another class of genes in *A. cryaerophilus* that could potentially impact human health is virulence factors. Most of the previous work at the molecular level has focused on nine putative virulence genes first described in *Arcobacter butzleri*

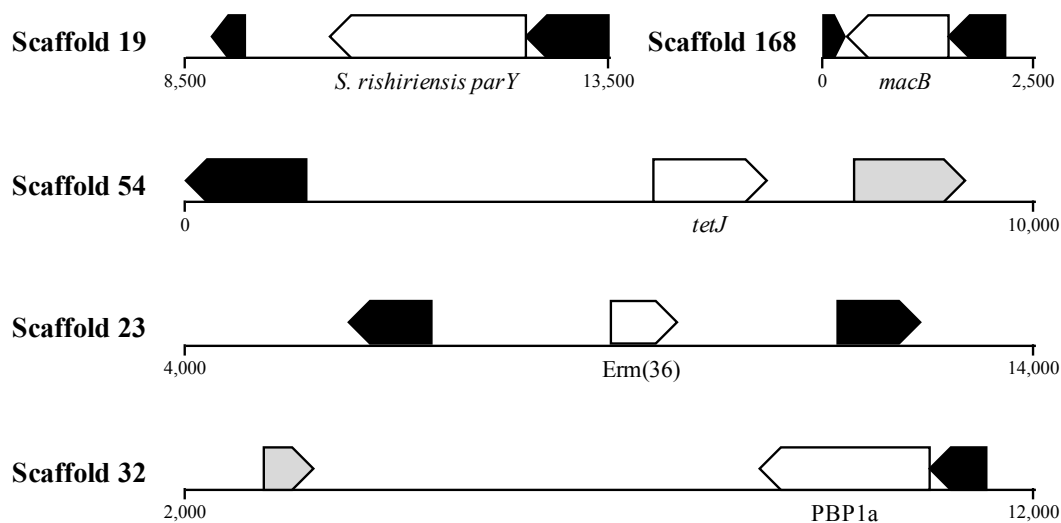


Figure 4 Location of ARGs indicates horizontal acquisition. Several antibiotic resistance genes (white) in *A. cryaerophilus* that are flanked by prophage genes (black) and transposases (grey) are shown. Nucleotide positions within each contig are also provided.

strain RM4018. The presence of these nine virulence genes in *Arcobacter* genomes is highly variable and are all rarely found together in the same genome (Miller *et al.*, 2007; Doudah *et al.*, 2012). In general, the ability to adhere to and invade cells varies widely between *Arcobacter* species, with some of the most invasive strains isolated from feces or sewage samples (Ho *et al.*, 2007; Karadas *et al.*, 2013; Levican *et al.*, 2013). Using the PATRIC database we identified 232 putative virulence genes (24 virulence categories) in *A. cryaerophilus* (Table S5), out of which 231 were expressed (Data Set S4). In PATRIC, virulence factors are assigned the category “virulence” if their mode of action is not specified in an associated study. Among the expressed virulence factor genes, 101 were annotated with a category other than “virulence.” Of these, “intracellular survival and replication” was the largest group (30, 30%) (Table S5). The next largest groups present were “cellular metabolism” (22, 22%), “adhesion” (18, 18%), and “invasion” (11, 11%) (Fig. 5A). In the total sewage contigs, we identified 4,440 putative virulence factor genes (38 virulence categories (Table S6), out of which, 3,776 were expressed (Data Set S4)). Excluding the “virulence” category, 1,812 genes belonging to 37 other virulence categories were identified in the sewage microbiome. Of these, 1,589 genes from 35 categories were expressed, with “intracellular survival and replication” (548, 35%), “invasion” (318, 20%) and “adhesion” (229, 14%) being the top three categories (Fig. 5B; Table S6).

Our data suggest that untreated sewage contains several genes that potentially promote bacterial antibiotic resistance and virulence, and that *A. cryaerophilus*, a potential human pathogen that contains multiple drug resistance and virulence factors is a major component of this sewage system. Because we analyzed only a limited number of samples, further study is required to determine whether the dominance of *A. cryaerophilus* was a short-term phenomenon or whether this bacterium is a long-term resident of this sewage system (McLellan *et al.*, 2010; Shanks *et al.*, 2013). Although its cause is not understood, as observed

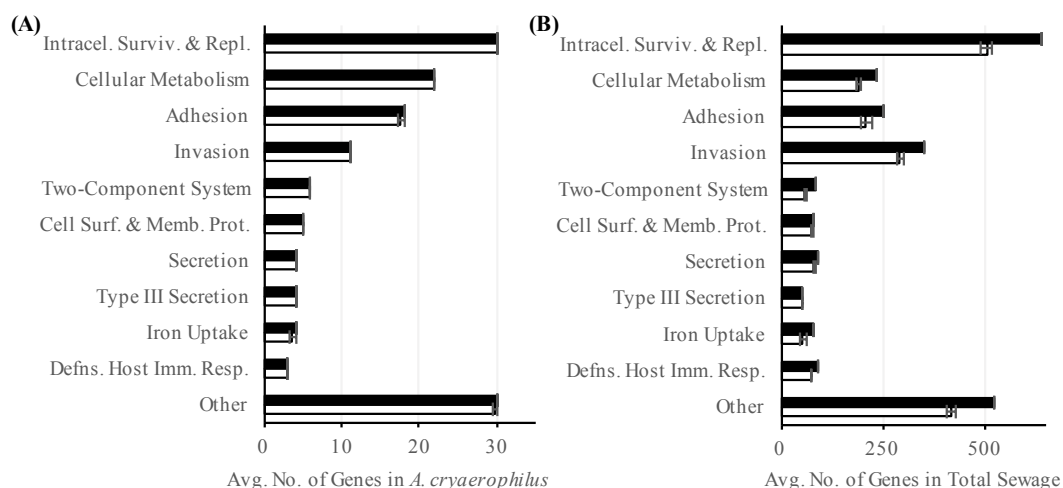


Figure 5 Abundance of virulence factors. Number of annotated (black) and expressed (white) virulence factor categories in (A) *A. cryaerophilus* and (B) total sewage sample is shown. Data represent means of three samples \pm standard deviations.

in our study, *Arcobacter* has been shown to be highly prevalent in other sewage systems (Fisher *et al.*, 2014). A possible explanation is the formation of biofilm on pipe surfaces and in deposited sediments along the sewer system (Chen, Leung & Hung, 2003), another possibility is that the presence of multiple antibiotics, heavy metals or xenobiotics in wastewater, even at very low concentrations is selecting for *A. cryaerophilus*, which contains multiple ARGs (Heberer, 2002; Hawkey & Jones, 2009; Gullberg *et al.*, 2014; Jutkina *et al.*, 2016). Similar to our observation, selection for antibiotic resistant bacteria has been described from other wastewater treatment plants (Goñi Urriza *et al.*, 2000; Czekalski *et al.*, 2012; Mao *et al.*, 2015); consequently, constant monitoring of both pre- and post-treatment sewage is warranted because of the risk of reintroducing bacteria replete with ARGs and virulence factors into natural environments (Fahrenfeld *et al.*, 2013; Czekalski, Gasco & Burgmann, 2014; Mao *et al.*, 2015; Pehrsson *et al.*, 2016).

ACKNOWLEDGEMENTS

We thank Howard Ochman for collecting the sewage samples and Peter King for assistance with data analysis.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by Portland State University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Portland State University.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Jess A. Millar conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Rahul Raghavan conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

GenBank: [LNTC00000000](#) and [PRJNA354077](#).

Data Availability

The following information was supplied regarding data availability:

The raw data has been supplied as a [Supplementary File](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.3269#supplemental-information>.

REFERENCES

- Abay S, Kayman T, Hizlisoy H, Aydin F. 2012.** *In vitro* antibacterial susceptibility of *Arcobacter butzleri* isolated from different sources. *Journal of Veterinary Medical Science* 74:613–616 DOI [10.1292/jvms.11-0487](#).
- Abdelbaqi K, Menard A, Prouzet-Mauleon V, Bringaud F, Lehours P, Megraud F. 2007.** Nucleotide sequence of the *gyrA* gene of *Arcobacter* species and characterization of human ciprofloxacin-resistant clinical isolates. *FEMS Immunology & Medical Microbiology* 49:337–345 DOI [10.1111/j.1574-695X.2006.00208.x](#).
- Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. 2013.** Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnology* 31:533–538 DOI [10.1038/nbt.2579](#).
- Amos GCA, Zhang L, Hawkey PM, Gaze WH, Wellington EM. 2014.** Functional metagenomic analysis reveals rivers are a reservoir for diverse antibiotic resistance genes. *Veterinary Microbiology* 171:441–447 DOI [10.1016/j.vetmic.2014.02.017](#).
- Arizona Department of Water Resources (ADWR). 2009.** Arizona's water supplies and water demands. Available at <http://www.azwater.gov/AzDWR/PublicInformationOfficer/documents/supplydemand.pdf>.
- Baquero F, Martínez J-L, Cantón R. 2008.** Antibiotics and antibiotic resistance in water environments. *Current Opinion in Biotechnology* 19:260–265 DOI [10.1016/j.copbio.2008.05.006](#).

- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011.** Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579 DOI [10.1093/bioinformatics/btq683](https://doi.org/10.1093/bioinformatics/btq683).
- Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120 DOI [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Chen GH, Leung DHW, Hung JC. 2003.** Biofilm in the sediment phase of a sanitary gravity sewer. *Water Research* 37:2784–2788 DOI [10.1016/S0043-1354\(03\)00083-6](https://doi.org/10.1016/S0043-1354(03)00083-6).
- Cho I, Blaser MJ. 2012.** The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* 13:260–270 DOI [10.1038/nrg3182](https://doi.org/10.1038/nrg3182).
- Collado L, Inza I, Guarro J, Figueras MJ. 2008.** Presence of *Arcobacter* spp. in environmental waters correlates with high levels of fecal pollution. *Environmental Microbiology* 10:1635–1640 DOI [10.1111/j.1462-2920.2007.01555.x](https://doi.org/10.1111/j.1462-2920.2007.01555.x).
- Czekalski N, Berthold T, Caucci S, Egli A, Bürgmann H. 2012.** Increased levels of multiresistant bacteria and resistance genes after wastewater treatment and their dissemination into lake Geneva, Switzerland. *Frontiers in Microbiology* 3:Article 106 DOI [10.3389/fmicb.2012.00106](https://doi.org/10.3389/fmicb.2012.00106).
- Czekalski N, Gasco E, Burgmann H. 2014.** Wastewater as a point source of antibiotic-resistance genes in the sediment of a freshwater lake. *International Society for Microbial Ecology Journal* 8:1381–1390 DOI [10.1038/ismej.2014.8](https://doi.org/10.1038/ismej.2014.8).
- Darriba D, Taboada GL, Doallo R, Posada D. 2012.** jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9:772 DOI [10.1038/nmeth.2109](https://doi.org/10.1038/nmeth.2109).
- Doudah L, De Zutter L, Baré J, De Vos P, Vandamme P, Vandenberg O, Van den Abeele AM, Houf K. 2012.** Occurrence of putative virulence genes in *Arcobacter* species isolated from humans and animals. *Journal of Clinical Microbiology* 50:735–741 DOI [10.1128/JCM.05872-11](https://doi.org/10.1128/JCM.05872-11).
- Fahrenfeld N, Ma Y, Brien MO, Pruden A. 2013.** Reclaimed water as a reservoir of antibiotic resistance genes: distribution system and irrigation implications. *Frontiers in Microbiology* 4:Article 130 DOI [10.3389/fmicb.2013.00130](https://doi.org/10.3389/fmicb.2013.00130).
- Finn RD, Clements J, Eddy SR. 2011.** HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* 39:W29–W37 DOI [10.1093/nar/gkr367](https://doi.org/10.1093/nar/gkr367).
- Fisher JC, Levican A, Figueras MJ, McLellan SL. 2014.** Population dynamics and ecology of *Arcobacter* in sewage. *Frontiers in Microbiology* 5:Article 525 DOI [10.3389/fmicb.2014.00525](https://doi.org/10.3389/fmicb.2014.00525).
- Gaze WH, Krone SM, Joakim Larsson DG, Li XZ, Robinson JA, Simonet P, Smalla K, Timinouni M, Topp E, Wellington EM, Wright GD, Zhu YG. 2013.** Influence of humans on evolution and mobilization of environmental antibiotic resistance. *Emerging Infectious Diseases* 19:e120871 DOI [10.3201/eid1907.120871](https://doi.org/10.3201/eid1907.120871).
- Goñi Urriza M, Pineau L, Capdepuuy M, Roques C, Caumette P, Quentin C. 2000.** Antimicrobial resistance of mesophilic *Aeromonas* spp. isolated from two European rivers. *Journal of Antimicrobial Chemotherapy* 46:297–301 DOI [10.1093/jac/46.2.297](https://doi.org/10.1093/jac/46.2.297).

- Gullberg E, Albrecht LM, Karlsson C, Gullberg E, Albrecht LM, Karlsson C, Sandegren L, Andersson DI. 2014. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *mBio* 5:e01918–14 DOI 10.1128/mBio.01918-14.Editor.
- Hawkey PM, Jones AM. 2009. The changing epidemiology of resistance. *Journal of Antimicrobial Chemotherapy* 64:i3–i10 DOI 10.1093/jac/dkp256.
- Heberer T. 2002. Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data. *Toxicology Letters* 131:5–17 DOI 10.1016/S0378-4274(02)00041-3.
- Ho HTK, Lipman LJA, Gaastra W. 2006. *Arcobacter*, what is known and unknown about a potential foodborne zoonotic agent! *Veterinary Microbiology* 115:1–13 DOI 10.1016/j.vetmic.2006.03.004.
- Ho H.T, Lipman L.J, Hendriks H.G, Tooten P.C, Ultee T, Gaastra W. 2007. Interaction of *Arcobacter* spp. with human and porcine intestinal epithelial cells. *FEMS Immunology & Medical Microbiology* 50:51–58 DOI 10.1111/j.1574-695X.2007.00230.x.
- Houf K, Devriese LA, Haesebrouck F, Vandenberg O, Butzler J-P, Van Hoof J, Vandamme P. 2004. Antimicrobial susceptibility patterns of *Arcobacter butzleri* and *Arcobacter cryaerophilus* strains isolated from humans and broilers. *Microbial Drug Resistance* 10:243–247 DOI 10.1089/mdr.2004.10.243.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755 DOI 10.1093/bioinformatics/17.8.754.
- Huson DH, Auch AF, Qi J, Schuster SC. 2007. MEGAN analysis of metagenomic data. *Genome Research* 17:377–386 DOI 10.1101/gr.5969107.
- Hyatt D, LoCascio PF, Hauser LJ, Uberbacher EC. 2012. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28:2223–2230 DOI 10.1093/bioinformatics/bts429.
- Jandhyala SA, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. 2015. Role of the normal gut microbiota. *World Journal of Gastroenterology* 21:8787–8803 DOI 10.3748/wjg.v21.i29.8787.
- Jutkina J, Rutgersson C, Flach CF, Larsson DG. 2016. An assay for determining minimal concentrations of antibiotics that drive horizontal transfer of resistance. *Science of the Total Environment* 548–549:131–138 DOI 10.1016/j.scitotenv.2016.01.044.
- Kabeya H, Maruyama S, Morita Y, Ohsuga T, Ozawa S, Kobayashi Y, Abe M, Katsube Y, Mikami T. 2004. Prevalence of *Arcobacter* species in retail meats and antimicrobial susceptibility of the isolates in Japan. *International Journal of Food Microbiology* 90:303–308 DOI 10.1016/S0168-1605(03)00322-2.
- Karadas G, Sharbati S, Hänel I, Messelhäuser U, Glocker E, Alter T, Gözl G. 2013. Presence of virulence genes, adhesion and invasion of *Arcobacter butzleri*. *Journal of Applied Microbiology* 115:583–590 DOI 10.1111/jam.12245.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform

- for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649 DOI [10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199).
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* **19**:1639–1645 DOI [10.1101/gr.092759.109](https://doi.org/10.1101/gr.092759.109).
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**:357–359 DOI [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923).
- Levican A, Alkeskas A, Günter C, Forsythe SJ, Figueras MJ. 2013. Adherence to and invasion of human intestinal cells by *Arcobacter* species and their virulence genotypes. *Applied and Environmental Microbiology* **79**:4951–4957 DOI [10.1128/AEM.01073-13](https://doi.org/10.1128/AEM.01073-13).
- Mao D, Yu S, Rysz M, Luo Y, Yang F, Li F, Hou J, Mu Q, Alvarez PJJ. 2015. Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants. *Water Research* **85**:458–466 DOI [10.1016/j.watres.2015.09.010](https://doi.org/10.1016/j.watres.2015.09.010).
- McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJV, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013. The comprehensive antibiotic resistance database. *Antimicrobial Agents and Chemotherapy* **57**:3348–3357 DOI [10.1128/AAC.00419-13](https://doi.org/10.1128/AAC.00419-13).
- McLellan SL, Huse SM, Mueller-Spitz SR, Andreihcheva EN, Sogin ML. 2010. Diversity and population structure of sewage derived microorganisms in wastewater treatment plant influent. *Environmental Microbiology* **12**:378–392 DOI [10.1111/j.1462-2920.2009.02075.x.Diversity](https://doi.org/10.1111/j.1462-2920.2009.02075.x.Diversity).
- Merga JY, Williams NJ, Miller WG, Leatherbarrow AJH, Bennett M, Hall N, Ashelford KE, Winstanley C. 2013. Exploring the diversity of *Arcobacter butzleri* from cattle in the UK using MLST and whole genome sequencing. *PLOS ONE* **8**(2):e55240 DOI [10.1371/journal.pone.0055240](https://doi.org/10.1371/journal.pone.0055240).
- Miller WG, Parker CT, Rubenfield M, Mendz GL, Wösten MMSM, Ussery DW, Stolz JF, Binnewies TT, Hallin PF, Wang G, Malek JA, Rogosin A, Stanker LH, Mandrell RE. 2007. The complete genome sequence and analysis of the Epsilonproteobacterium *Arcobacter butzleri*. *PLOS ONE* **2**(12):e1358 DOI [10.1371/journal.pone.0001358](https://doi.org/10.1371/journal.pone.0001358).
- Mitchell A, Chang H-Y, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong S-Y, Bateman A, Punta M, Attwood TK, Sigrist CJA, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD. 2015. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Research* **43**:D213–D221 DOI [10.1093/nar/gku1243](https://doi.org/10.1093/nar/gku1243).
- Morita Y, Maruyama S, Kabeya H, Boonmar S, Nimsuphan B, Nagai A, Kozawa K, Nakajima T, Mikami T, Kimura H. 2004. Isolation and phylogenetic analysis of *Arcobacter* spp. in ground chicken meat and environmental water in Japan and

- Thailand. *Microbiology and Immunology* **48**:527–533
DOI 10.1111/j.1348-0421.2004.tb03548.x.
- Noyes NR, Yang X, Linke LM, Magnuson RJ, Cook SR, Zaheer R, Yang H, Woerner DR, Geornaras I, McArt JA, Gow SP, Ruiz J, Jones KL, Boucher CA, McAllister TA, Belk KE, Morley PS. 2016. Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems. *Scientific Reports* **6**:24645
DOI 10.1038/srep24645.
- Okoh AI, Odjadjare EE, Igbiosa EO, Osode AN. 2007. Wastewater treatment plants as a source of microbial pathogens in receiving watersheds. *Journal of Biotechnology* **6**:2932–2944 DOI 10.5897/AJB2007.000-2462.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research* **42**:D206–D214 DOI 10.1093/nar/gkt1226.
- Paulsen IT, Sliwinski MK, Saier MH. 1998. Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *Journal of Molecular Biology* **277**:573–592 DOI 10.1006/jmbi.1998.1609.
- Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G. 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533**:212–216 DOI 10.1038/nature17672.
- Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: A *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**:1420–1428 DOI 10.1093/bioinformatics/bts174.
- Perner M, Gonnella G, Kurtz S, LaRoche J. 2014. Handling temperature bursts reaching 464C: different microbial strategies in the sisters peak hydrothermal chimney. *Applied and Environmental Microbiology* **80**:4585–4598 DOI 10.1128/AEM.01460-14.
- Pima Association of Governments (PAG). 2006. Existing wastewater treatment facilities and other point source NPDES discharges. Available at https://www.pagnet.org/documents/Water/PC208/Ch5_Apr06.pdf.
- Poole K. 2012. Bacterial stress responses as determinants of antimicrobial resistance. *Journal of Antimicrobial Chemotherapy* **67**:2069–2089 DOI 10.1093/jac/dks196.
- Rahimi E. 2014. Prevalence and antimicrobial resistance of *Arcobacter* species isolated from poultry meat in Iran. *British Poultry Science* **55**:174–180
DOI 10.1080/00071668.2013.878783.
- Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, Michael I, Fatta-Kassinos D. 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Science of the Total Environment* **447**:345–360 DOI 10.1016/j.scitotenv.2013.01.032.
- Schroeder-Tucker L, Wesley IV, Kiehlbauch JA, Larson DJ, Thomas LA, Erickson GA. 1996. Phenotypic and ribosomal RNA characterization of *Arcobacter* species isolated from porcine aborted fetuses. *Journal of Veterinary Diagnostic Investigation* **8**:186–195 DOI 10.1177/104063879600800208.

- Shanks OC, Newton RJ, Kelty CA, Huse SM, Sogin ML, McLellan SL. 2013. Comparison of the microbial community structures of untreated wastewaters from different geographic locales. *Applied and Environmental Microbiology* 79:2906–2913 DOI 10.1128/AEM.03448-12.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7:Article 539 DOI 10.1038/msb.2011.75.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57:758–771 DOI 10.1080/10635150802429642.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56:564–577 DOI 10.1080/10635150701472164.
- Udikovic-Kolic N, Wichmann F, Broderick NA, Handelsman J. 2014. Bloom of resident antibiotic-resistant bacteria in soil following manure fertilization. *Proceedings of the National Academy of Sciences of the United States of America* 111:15202–15207 DOI 10.1073/pnas.1409836111.
- Vandenberg O, Dediste A, Houf K, Ibekwem S, Souayah H, Cadranel S, Douat N, Zisis G, Butzler JP, Vandamme P. 2004. Arcobacter species in humans. *Emerging Infectious Diseases* 10:1863–1867 DOI 10.3201/eid1010.040241.
- Varela AR, Manaia CM. 2013. Human health implications of clinically relevant bacteria in wastewater habitats. *Environmental Science and Pollution Research* 20:3550–3569 DOI 10.1007/s11356-013-1594-0.
- Versluis D, D’Andrea MM, Ramiro Garcia J, Leimena MM, Hugenholtz F, Zhang J, Öztürk B, Nylund L, Sipkema D, Van Schaik W, De Vos WM, Kleerebezem M, Smidt H, Van Passel MWJ. 2015. Mining microbial metatranscriptomes for expression of antibiotic resistance genes under natural conditions. *Scientific Reports* 5:11981 DOI 10.1038/srep11981.
- Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJC, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research* 42:D581–D591 DOI 10.1093/nar/gkt1099.
- Wellington EMH, Boxall ABA, Cross P, Feil EJ, Gaze WH, Hawkey PM, Johnson-Rollings AS, Jones DL, Lee NM, Otten W, Thomas CM, Williams AP. 2013. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. *The Lancet Infectious Diseases* 13:155–165 DOI 10.1016/S1473-3099(12)70317-1.
- Zhang T, Shao M-F, Ye L. 2012. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *The ISME Journal* 6:1137–1147 DOI 10.1038/ismej.2011.188.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Research* **39**:W347–W352 DOI [10.1093/nar/gkr485](https://doi.org/10.1093/nar/gkr485).

Zhu Q, Kosoy M, Dittmar K. 2014. HGTector: An automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* **15**:717 DOI [10.1186/1471-2164-15-717](https://doi.org/10.1186/1471-2164-15-717).