DTU Library

Limited dissemination of the wastewater treatment plant core resistome

Munck, Christian; Albertsen, Mads; Telke, Amar; Ellabaan, Mostafa M Hashim; Nielsen, Per Halkjær; Sommer, Morten Otto Alexander

Link back to DTU Orbit

ARTICLE

       OPEN

Limited dissemination of the wastewater treatment plant core resistome

Christian Munck[1],[*], Mads Albertsen[2],[*], Amar Telke[3],[*], Mostafa Ellabaan[1], Per Halkjær Nielsen[2] & Morten O.A. Sommer[1]

Horizontal gene transfer is a major contributor to the evolution of bacterial genomes and can facilitate the dissemination of antibiotic resistance genes between environmental reservoirs and potential pathogens. Wastewater treatment plants (WWTPs) are believed to play a central role in the dissemination of antibiotic resistance genes. However, the contribution of the dominant members of the WWTP resistome to resistance in human pathogens remains poorly understood. Here we use a combination of metagenomic functional selections and comprehensive metagenomic sequencing to uncover the dominant genes of the WWTP resistome. We find that this core resistome is unique to the WWTP environment, with <10% of the resistance genes found outside the WWTP environment. Our data highlight that, despite an abundance of functional resistance genes within WWTPs, only few genes are found in other environments, suggesting that the overall dissemination of the WWTP resistome is comparable to that of the soil resistome.

---

[1] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Alle 6, DK-2970 Hørsholm, Denmark. [2] Center for Microbial Communities, Department of Chemistry and Bioscience, Aalborg University, 9220 Aalborg, Denmark. [3] Department of Systems Biology, Technical University of Denmark, DK-2800 Lyngby, Denmark. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.O.A.S. (email: msom@bio.dtu.dk).

1

Investigations of environmental sources of antibiotic resistance genes have shown that many clinically relevant resistance genes such as the cephalosporin resistance gene family *ctx-m*, the quinolone resistance gene family *qnr* and the vancomycin resistance gene family *van* have close homologues on the chromosomes of environmental non-pathogenic bacterial species[1–5]. These findings highlight that resistance genes can spread from environmental reservoirs to human pathogens. This has led to a growing interest in uncovering which environmental microbial communities participate most in antibiotic resistance gene exchange with the aim of establishing the most impactful policies that limit the dissemination of antibiotic resistance[3,4,6–8].

To assess the resistance gene load of a given environmental microbial community, researchers have used targeted PCR or metagenomic read mapping to detect the presence of well-characterized resistance genes[7,9]. These studies have shown that representatives of clinically relevant resistance genes can be found in virtually all environments. Although such studies are important and relevant, they rarely consider the abundance of the targeted resistance genes relative to the complete resistome of the studied environment. This is largely due to our incomplete annotation and knowledge of antibiotic resistance genes[10]. Indeed, less biased methods such as metagenomic functional selections have revealed that <1% of the resistance genes found in soil have high identity (>99% nucleotide identity) hits in NCBI databases[8,11]. Accordingly, the majority of the soil resistome has not been found outside the soil environment. On the other hand similar approaches have been used to show that a subset of the human gut microbiome may exchange resistance genes with human pathogens[12]. So far, metagenomic functional selections have largely been applied to study the resistomes of soil and gut microbial communities; yet, other communities are believed to contribute substantially to the dissemination of resistance genes. Hence, to make informed policy decisions addressing the challenges associated with antibiotic resistance, the abundance and dissemination of antibiotic resistance genes in these communities must be characterized.

Wastewater treatment plants (WWTPs) are generally considered to be important hubs for horizontal gene transfer of antibiotic resistance genes[13–15]. In these facilities, wastewater from diverse sources such as hospitals, households and animal production farms is mixed, often containing both pathogenic bacteria and traces of antibiotics[16–18]. Accordingly, the dense microbial communities of WWTPs should provide ideal conditions for horizontal exchange of antibiotic resistance genes between a wide range of bacterial species, including human pathogens.

Several PCR-based studies have shown that clinically relevant resistance genes, including *tem1*, *ampC*, *ndm-1*, *vanA* and *ermB*, can be found in WWTPs[19–22]. Similarly, metagenome sequencing studies characterizing either complete or closed circular metagenomes (plasmids and some phages) have used sequence read mapping to resistance gene databases, to identify several different resistance genes in WWTPs[9,23–26]. Furthermore, traditional cultivation-based studies have also shown that vancomycin-resistant enterococci, methicillin-resistant staphylococci and cephalosporin-resistant Enterobacteriaceae can be isolated from WWTPs[20,27]. However, owing to the specificity of PCR-based screening and database mapping, these approaches can only identify known resistance genes and their close homologues. Consequently, such methods cannot uncover the full spectrum of functional resistance genes within an environment. Likewise, cultivation-based approaches will only recover a small proportion of bacterial species, with estimates indicating that 85%–99% of the bacteria from WWTPs are not readily cultured using current protocols[28]. Here we combine functional selection and metagenomic sequencing to characterize and trace the most abundant antibiotic resistance genes in the WWTP environment. We find that only a small subset of the core WWTP resistome is found outside the WWTP environment.

Results

Metagenomic functional selection. Using metagenomic functional selection, we identified the dominant members of the resistome in a large modern WWTP, Aalborg West (AAW), sampled in the third quarter of 2012. This facility is located in Denmark and receives both hospital and household wastewater, and has been in operation since 1979 (Supplementary Table 1 and Methods). A total of 800 Mbp of DNA was cloned in an *Escherichia coli* expression library (Supplementary Methods) and screened on 15 different antibiotics representing 8 chemical classes: β-lactam, aminoglycoside, macrolide, tetracycline, phenicol, rifamycin, sulphonamide and dihydrofolate reductase inhibitor (Table 1). In total 8,540 resistant clones were identified with an average of 534 colonies per antibiotic ranging from 30 clones for chloramphenicol to more than 2,000 clones for

Table 1 | Selection conditions.

Antibiotic	Concentration (µg ml^{-1})	Code	Class	Number of colonies
Ampicillin	16	AMP	β-Lactam/penicillin	660
Amoxicillin	16	AMX	β-Lactam/penicillin	500
Carbenicillin	64	CAR	β-Lactam/penicillin	1,100
Piperacillin	16	PIP	β-Lactam/penicillin	800
Ceftazidime	1	CTZ	β-Lactam/cephalosporin	350
Amikacin	16	AMK	Aminoglycoside	220
Gentamicin	8	GEN	Aminoglycoside	40
Spectinomycin	32	SPC	Aminoglycoside	330
Azithromycin	16	AZI	Macrolide	120
Erythromycin	100	ERY	Macrolide	200
Tetracycline	4	TET	Tetracycline	100
Chloramphenicol	6	CAM	Phenicol	30
Rifampicin	16	RIF	Rifamycin	210
Trimethoprim	4	TMP	Benzylpyrimidine	2,000
Sulfamethoxazole/trimethoprim	32:4	SXT	Sulfonamide/dihydrofolate reductase inhibitor	1,500
Total				8,160

AMK, amikacin; AMP, ampicillin; AMX, amoxicillin; AZI, azithromycin; CAM, chloramphenicol; CAR, carbenicillin; CTZ, ceftazidime; ERY, erythromycin; GEN, gentamicin; PIP, piperacillin; RIF, rifampicin; SPC, spectinomycin; SXT, sulfamethoxazole/trimethoprim; TET, tetracycline; TMP, trimethoprim.

trimethoprim (Table 1). Of the 8,540 identified clones, the inserts of 749 clones were selected proportionally among the classes of antibiotics and sequenced, resulting in the identification and annotation of 79 unique functionally validated resistance genes (Supplementary Table 2). Rarefaction curve analysis showed that we sampled the majority of the resistance genes in the library (Supplementary Fig. 1). Accordingly, these genes are expected to represent the most abundant genes in the WWTP resistome and we refer to them as the core WWTP resistome.

WWTP community stability. Considering the constant flow of material through the AAW WWTP, up to 50,000 m^3 per day, it could be expected that the microbial community varied greatly over time. However, using a modified version of the Bray–Curtis dissimilarity index ($(1 - \text{Bray–Curtis dissimilarity index}) \times 100$), we calculated the similarity between all pairwise combinations of samples collected from AAW over a 7-year period ($n = 24$) (Fig. 1a)[29]. Sample similarity was indexed from 0 to 100, with 0 indicating completely disjoint sample and 100 indicating identical samples. The comparison revealed that the bacterial communities were remarkably stable (Fig. 1a). Notably, samples collected 7 years apart had an average similarity index of 48 (Fig. 1a) and samples collected < 1 year apart had an average similarity index of 68, highlighting that the WWTP microbial community is highly stable and suggesting that variation in wastewater inflow has a limited impact on the core bacterial communities.

Comparison of the relative abundances of the ten most abundant bacterial orders in the WWTP community across the 7 years of sampling showed that the WWTP community is highly stable and unique to this environment (Fig. 1b). As recently reported, the WWTP environment is dominated by the phyla Actinobacteria (Actinomycetales and Acidimicrobiales), Bacteriodetes (Saprospirales) and Proteobacteria (Rhizobiales,

Burkholderiales and Rhodobacterales), highlighting that WWTPs contain distinct microbial communities[30]. Still, owing to horizontal gene transfer, a stable microbial community at the 16S ribosomal RNA sequence level does not necessarily imply a stable resistome.

WWTP core resistome stability. To investigate the long-term stability of the WWTP core resistome, we sequenced seven metagenomes spanning 2 years from the same WWTP used for the metagenomic functional selections (AAW; Supplementary Table 3). Approximately 230 Gbp of sequence data corresponding to more than 1.5 billion raw reads were sequenced. By mapping the metagenomic sequencing reads to the functional selected resistance genes, we could trace the WWTP core resistome dynamics in the WWTP over the 2-year period (Fig. 2a). Extrapolation of the rarefaction curves predicted that the samples collected over the 2-year sampling period on average contained 70% (s.d. = 13%) of the whole resistance gene set (Fig. 2b). This highlights that despite vast amounts of wastewater passing through the treatment plant the resistome remains stable and we consider this set of genes part of the WWTP core resistome.

To evaluate the impact of sequencing depth, we exhaustively sequenced sample AAW.5.2010, which was collected 2 years before the sample used for the functional selection. In this metagenome, consisting of more than 1.1 billion reads, we identified 67% of the functional selected resistance genes and extrapolations of the rarefaction curve estimated that up to 82% genes could be found in this sample (Fig. 2a,b). This further underscores that the core resistome was largely stable during the 2-year sampling period.

Core resistome hosts. To identify possible hosts carrying the functional selected resistance genes, we assembled contigs from the deeply sequenced metagenome (AAW.5.2010). The relative low coverage of the functionally selected resistance genes challenges their assembly. In addition, the complexity of the WWTP community also challenges the assembly. A total of 43% of all metagenomic reads could be mapped back to the assembled contigs. Seven contigs contained functional selected resistance genes (> 95% identity and > 90% gene coverage). The contigs were blasted against the genbank nt database, to identify host organisms. In general, the query coverage was low, indicating that the contigs assembled from the metagenomes are located on hitherto uncharacterized genomes, which is largely a result of the limited availability of WWTP-specific bacterial genomes[31]. However, a small contig containing CAR_05 had high identity and query coverage to the *tem-1* β-lactamase gene on cloning vectors (Supplementary Table 4). The challenges of assembling the low coverage genes from the metagenome highlight the power of combining functional selection and metagenomic sequencing, to trace functionality through different metagenomes.

The WWTP core resistome is stable across multiple facilities. To explore whether the AAW WWTP core resistome was shared with other WWTP facilities, we sequenced additional eight metagenomes from four different Danish full-scale WWTPs: Egaa, Ejby, Hjoerring and Aalborg East. Two of the plants (Egaa and Ejby) were sampled repeatedly over a 1-year period and all plants, except Aalborg East, received household as well as hospital wastewater (Supplementary Table 3). In total, these additional WWTP metagenomes consisted of > 124 Gbp corresponding to > 800 million reads. Mapping of the metagenomic reads to the AAW core resistome revealed that the majority of the resistance genes were present in the different metagenomic samples (Fig. 2a,c). Extrapolations of the rarefaction curves estimate that
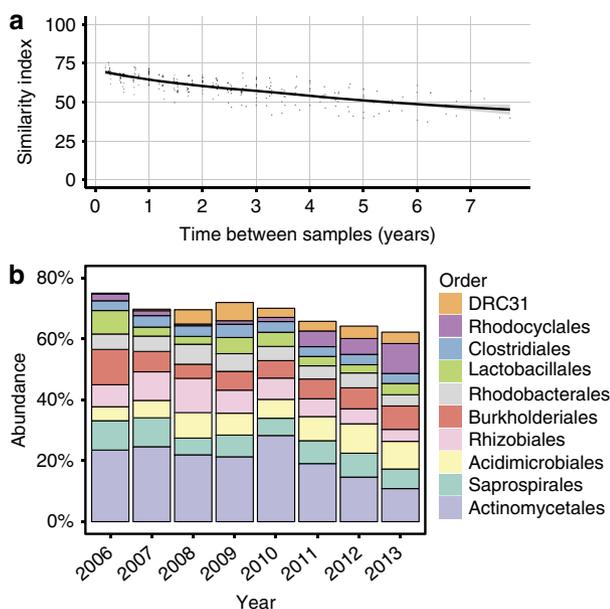

Figure 1 | Phylogenetic analysis of WWTP AAW. (a) The pairwise similarity of the microbial community in samples collected from the WWTP AAW over a 7-year period ($n = 24$). Similarity is calculated as (1 – Bray–Curtis dissimilarity index) × 100 for each pairwise comparison of 16S rRNA amplicons and plotted against the time (years) between the samples. **(b)** Per cent abundance of the ten most abundant bacterial orders in WWTP AAW over a 7-year period. On average, 68% (s.d. = 4%, $n = 24$) of the 16S rRNA gene sequences belong to these orders.
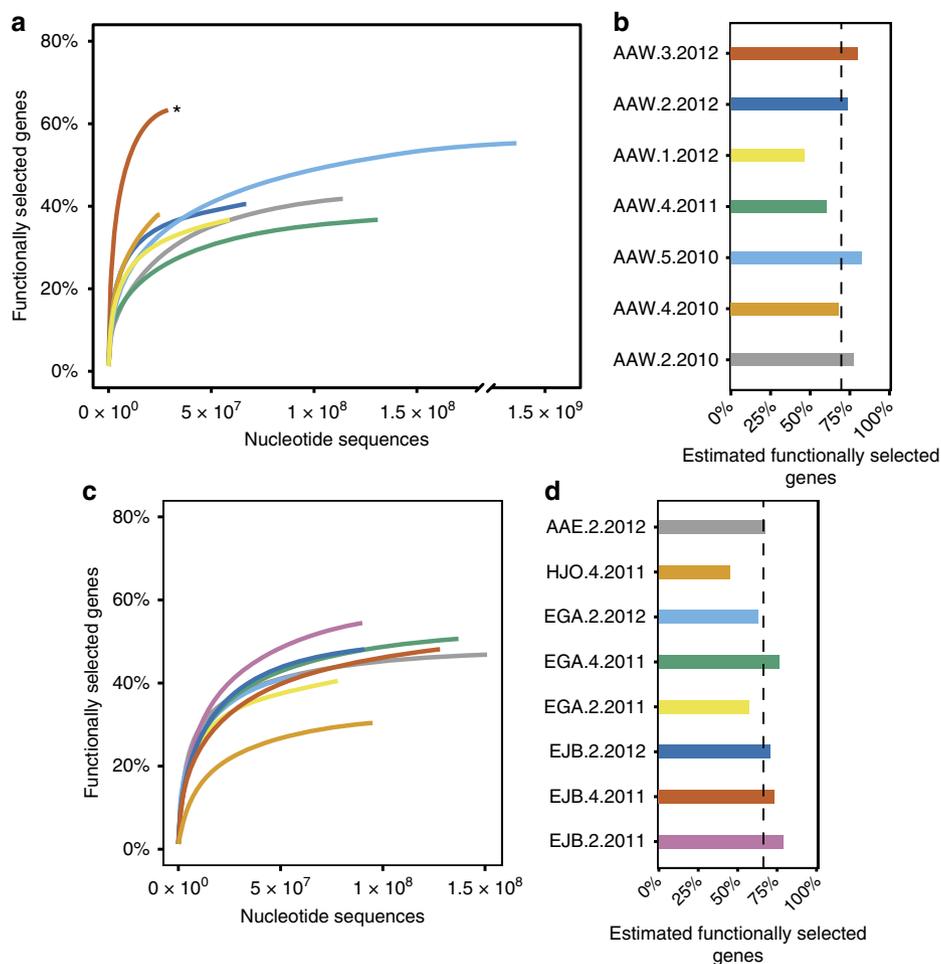
Figure 2 | Tracing the functionally selected resistome. (**a**) Rarefaction curves showing the percentage of the functional selected resistance genes found in the AAW metagenomes sampled over a 2-year period. The curves are based on read mapping ($>95\%$ identity) to the functional selected resistance genes (Methods). For each sample, the corresponding sample name is given by the colour code in **b**. The asterisk denotes the sample used in the functional selection. (**b**) Rarefaction curve-based predictions of the expected percentage of functionally selected resistance genes present in each of the AAW WWTP samples. Samples are named according to the collection time in the format: Location.Quarter.Year (quarter 5 refers to December). The dashed line indicates the average expected abundance. Predictions are made using the model described in Chiu et al.[64] with the R package Vegan[62]. (**c**) Rarefaction curve showing the percentage of the functional selected resistance genes found in the non-AAW WWTP metagenomes generated as in **a**. For each sample, the corresponding sample name is given by the colour code in **d**. (**d**) Rarefaction curve-based predictions of the percentage of functionally selected resistance genes present in each of the non-AAW WWTP samples. Sample names are generated as in **b**. The dashed line indicates the average expected abundance. Despite being distinct from the AAW WWTP, the non-AAW WWTPs are estimated to contain the majority of the resistance genes found in AAW.

on average the non-AAW WWTPs contained 66% (s.d. = 11%) of the functionally selected resistance genes from AAW (Fig. 2c,d).

The stability of the WWTP microbial community and core resistome could support the hypothesis that resistance gene exchange within WWTPs is a rare event as was recently reported for soil microbial communities[11]. However, as activated sludge in WWTPs is continuously mixed and exposed to varying levels of antibiotics, one would expect that horizontal gene transfer of resistance genes would be more frequent in WWTPs compared with soil[16–18]. To investigate whether the WWTP core resistomes correlated with the microbial composition, we used procrustes analysis to correlate resistomes and 16S rRNA sequences from each of the 15 WWTP metagenome samples[11]. Our analysis shows that WWTP resistome and phylogeny correlated ($P < 0.001$, based on 999 permutations), indicating that the WWTP core resistome is linked to the microbial community (Supplementary Fig. 2). This supports the hypothesis that the WWTP resistome is only exchanged to a limited extent within the WWTP environment.

Dissemination of the WWTP core resistome. Recent studies based on metagenomic read mapping showed that the metagenomic sequence data from WWTPs mapped to a distinct set of resistance genes compared with other environments[23]. However, this finding could arise from database biases resulting in an inability to identify shared resistance genes or it could suggest that the WWTP resistome is only disseminated to a limited extent to other environments. To investigate whether the WWTP core resistome is involved in gene exchange networks comprising other environments, as well as human pathogens, we used BLAST to perform a homology search of the WWTP core resistome to genes in the the NCBI nucleotide database. We found that the WWTP resistance genes had an average sequence identity of 62% to the best hit in the database, and that just six genes had an identity $>95\%$, demonstrating that the vast majority of the functional selected resistance genes represent novel sequences not captured by the current antibiotic resistance gene databases (Fig. 3a).
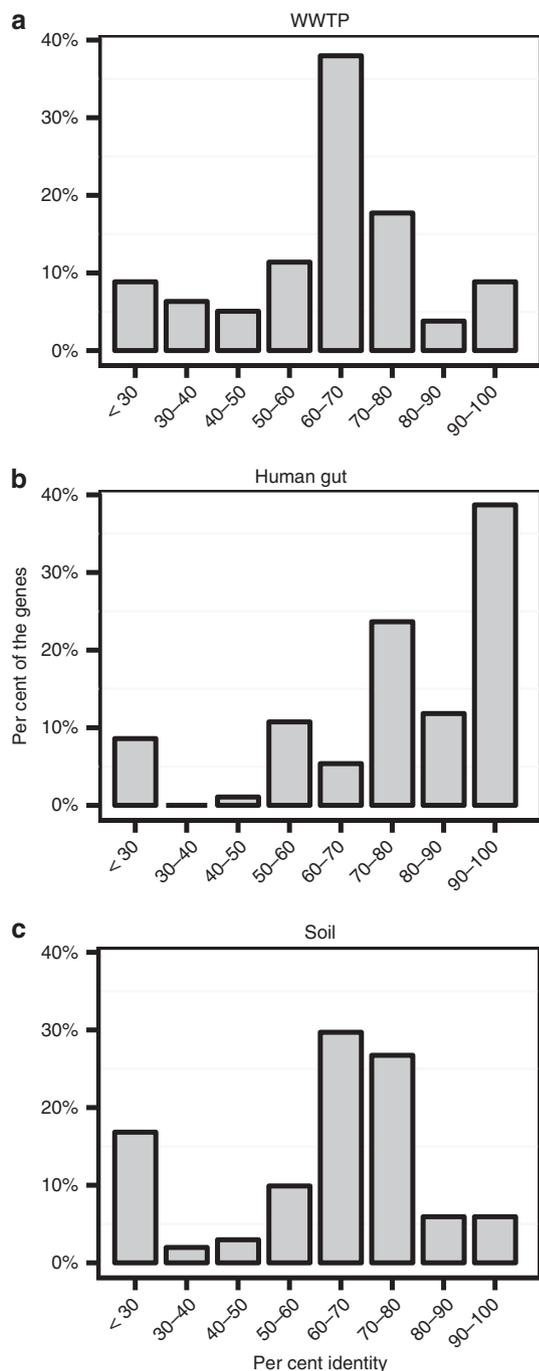
Figure 3 | Genbank nucleotide database hit identity. Functionally selected resistance genes from (**a**) WWTP AAW (*n* = 79), (**b**) human gut (*n* = 94)[12,32] and (**c**) soil (*n* = 101)[33–36] were blasted against the Genbank nt database and the top hit was aligned to the query. The panels show the nucleotide identity distributions of the pairwise alignment for each environment (Methods).

To put this finding into perspective, we re-analysed the available functional selected resistance genes from the human gut microbiome[12,32] and from the soil microbiome[33–36] (Fig. 3b,c). We found that functionally selected resistance genes from the human microbiome had an average of 75% identity to genes in the NCBI nucleotide database, whereas functionally selected resistance genes from soil had an average identity of 59%. Notably, 33% of the resistance genes from the human microbiome had > 95% identity to genes found in the NCBI

nucleotide database, whereas none of the soil genes were above this threshold. Owing to database biases, it is expected that genes from the human microbiome have higher identity to the NCBI nucleotide database. However, considering the large extent to which the human microbiota and human clinical isolates have been sequenced, the finding that only a few number of genes from the WWTP core resistome have an identity > 95% to genes in the NCBI nucleotide database indicates that the WWTP core resistome is only disseminated to the human microbiota or human clinical isolates to a limited extent. This probably reflects the fact that the WWTP microbial community is unique to the WWTP environment, and that the majority of the genes in the WWTP core resistome belong to this community.

To further investigate whether the WWTP core resistome was disseminated outside the treatment plant facilities, we mapped reads from online available metagenomes to the functionally selected WWTP resistome. These auxiliary metagenomes contained sequences from the human gut, cow rumen, permafrost and aquifer, in total more than two billions reads[37–40]. Notably, only six (8%) of the WWTP core resistance genes were found in these non-WWTP metagenomes, confirming that the WWTP core resistome is distinct from the resistomes of other environments (Fig. 4a). Rarefaction curve analysis of the non-WWTP metagenomes showed that the low identification frequency of the WWTP resistance genes was not due to undersampling (Supplementary Fig. 3). Four of the six genes found in the non-WWTP metagenomes had a nucleotide identity > 95% to clinically relevant resistance genes, highlighting that a PCR-based approach would probably have successfully identified these genes. However, reliance solely on PCR or metagenomic read mapping to antibiotic resistance gene databases would have neglected the vast majority of the WWTP core resistome, which seems confined to the WWTP environment.

To explore whether the WWTP metagenomes contained important resistance genes, we compiled a list of 27 resistance genes commonly found in Gram-negative and Gram-positive pathogens[41–43] (Supplementary Table 5). This list includes commonly encountered representatives of the *ctx-m* and *erm* gene families, as well as several important tetracycline resistance genes. A total of 578 sequence reads from the deeply sequenced AAW.5.2010 metagenome mapped to 10 of the resistance genes (Supplementary Fig. 4a). Nearly half of the reads (48%) mapped to the *tetA* resistance gene, whereas no reads mapped to the *ctx-m* genes. In contrast, a significantly higher number of 9,401 reads (*P* = 0.008, Mann–Whitney *U*-test) mapped to a total of 53 of the 79 functionally selected resistance genes (Supplementary Fig. 4b). These mapping results confirm that clinically important resistance genes are indeed present in the WWTP environment, yet again they also highlight how reliance on targeted methodologies can be misleading when assessing the contribution of the WWTP core resistome to the dissemination of resistance genes outside of the WWTP environment.

As our study is focused on dissemination of the WWTP core resistome, we cannot rule out the possibility that low abundance species engage in important resistance gene exchange in WWTPs. Indeed, mobile plasmids carrying known resistance genes have been isolated from WWTPs[26,44]. However, our results show that only a small fraction of the WWTP core resistome participate in gene exchange, suggesting that WWTP microbiomes appear as probably as soil microbiomes to be sources for resistance gene dissemination. This conclusion should be cautioned by the fact that we only sampled WWTPs in one country. It is possible that these findings may be different in other parts of the world that deploy simpler WWTPs at higher temperatures and higher load of antibiotics.
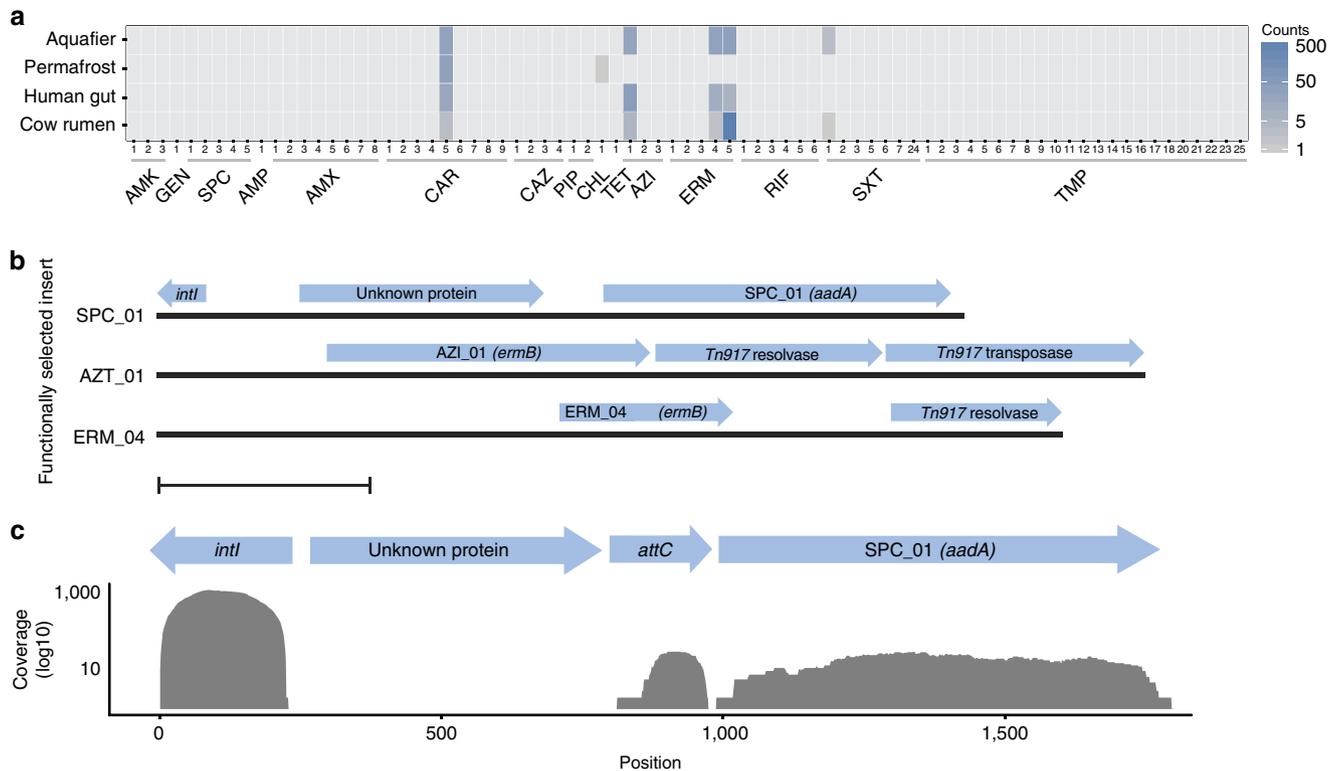
Figure 4 | WWTP resistome dissemination. (**a**) Mapping of metagenomic reads from aquifer, permafrost, human gut and cow rumen, to the functionally selected resistance genes. Mapping reads have > 95% identity and > 95% coverage. The functionally selected resistance genes are grouped by the antibiotic they were selected on as denoted by the three-letter code (AMK (amikacin), AMP (ampicillin), AMX (amoxicillin), AZI (azithromycin), CAM (chloramphenicol), CAR (carbenicillin), CTZ (ceftazidime), ERY (erythromycin), GEN (gentamicin), PIP (piperacillin), RIF (rifampicin), SPC (spectinomycin), SXT (sulfamethoxazole/trimethoprim), TET (tetracycline), TMP (trimethoprim)). (**b**) The flanking regions of three functional selected resistance genes contained elements, indicating that the genes were either located in an integron (SPC_01) or on a transposon (AZT_01 and ERM_04). Blue arrows denote the different genes and their orientation. Scale bar, 500 bp. Genes were annotated using BLAST. (**c**) Read coverage (> 95% identity and > 95% coverage) of the integrase gene (*intI*) and the attachment site sequence (*attC*) found in the integron carrying SPC_01 in sample AAW.5.2010. The integrase was 50 times more abundant than the SPC_01 resistance gene *aadA*. The blue arrows denoted genes and functional regions within the integron as annotated by BLAST.

An explanation for the limited dissemination of the WWTP core resistome could be that the resistance genes are not selected for or mobilized. However, studies have shown that even very low concentrations of antibiotics, probably present in WWTPs, can select for resistance determinants[45,46]. Therefore, we analysed the functionally selected inserts for the presence of canonical indicators of mobilization such as transposases and phage elements. We found that three (4%) of the functionally selected inserts carried canonical indicators of mobility including integrases associated with class I integrons, transposes and resolvases (Fig. 4b). Two of the resistance genes carried on these inserts (SPC_01 and AZI_01) had a high identity hit (> 95% identity) to the NCBI nucleotide database. Furthermore, the resistance genes located on the AZI_01 and ERM_04 inserts could be found in the auxiliary metagenomes (Fig. 4a). This supports the hypothesis that mobilization is a major limiting factor for the dissemination of antibiotic resistance genes; for example, once a resistance gene is mobilized, it easily becomes disseminated across ecosystems. It should be noted that, owing to the relative short insert size (~ 2 kb), some mobility elements may be overlooked. However, previous functional selections from human faecal samples using the same insert size found mobilization elements in 14% of the inserts, which is significantly higher than what is observed in this study ($P = 0.027$, χ^2-test)[12].

Of particular interest was the spectinomycin resistance gene *aadA* in the SPC_01 insert. This gene was located in a class 1

integron and mapping of reads from the deeply sequenced metagenome to the insert containing the *aadA* gene showed that the integrase gene was roughly 50 times more abundant than the *aadA* gene, highlighting that the integron was not always associated with the *aadA* gene (Fig. 4c). This gene is part of the WWTP core resistome and has previously been found to be associated with wastewater environments including a French and a Chinese WWTP, effluent wastewater from a French hospital, pig manure sampled in Denmark and on a plasmid isolated from a Danish soil/manure sample[47–51] (Supplementary Fig. 5). Interestingly, the gene had also been found in an *E. coli* isolate from a bloodstream infection (99.9% identity). This confirms that once mobilized, members of the core WWTP resistome can be found outside this environment and even in clinical isolates (Supplementary Fig. 5). Yet, only a small fraction of the core WWTP resistome has crossed this mobilization barrier.

Discussion

Integrating metagenomic functional selection with deep metagenomic sequencing represents a powerful approach to tracing functionally validated antibiotic resistance genes across different environments. Using functional selections, we identified a core WWTP resistome consisting of mostly novel resistance genes that confer resistance to all antibiotics tested, highlighting that

WWTPs is a reservoir of previously uncharacterized resistance genes.

16S rRNA sequencing of longitudinal samples collected over a 7-year period from the same WWTP revealed that the microbial community within the WWTP is remarkably stable. In addition, longitudinal metagenomic sequencing from the same WWTP demonstrated that the majority of the genes within the WWTP core resistome were also stably maintained within the WWTP. This indicates that the resistome is linked to the community, which was confirmed with procrustes analysis. Integration of metagenomic data sets from other WWTPs showed that a majority of the functionally verified resistance genes are also maintained in those microbial communities. These data point towards the existence of a shared WWTP core resistome harboured by the dominant members of the WWTP microbial community.

Notably, <10% of the WWTP core resistome sampled in this study is found in other metagenomic data sets or in sequences deposited at NCBI. Although database bias towards human pathogens and commensals could explain these findings, this also support that members of the WWTP core resistome rarely take part in gene exchange networks with human pathogens. Instead, the WWTP microbial community can be considered a unique and stable microbial community harbouring a rich resistome that is difficult to access by human pathogens or other bacteria. This conclusion is consistent with a previous metagenomic study based on read mapping to existing antibiotic resistance gene databases, showing that the WWTP resistance gene profiles are distinct from other environments[23].

Our findings do not preclude the existence of clinically important resistance gene exchange in the WWTP. In fact, we do detect a subset of clinically important resistance genes in the deeply sequenced metagenome. Consequently, species found in low abundance in WWTPs, including human pathogens, may exchange resistance genes within this environment contributing to clinical problems of antibiotic resistance. However, our findings are consistent with the hypothesis that the abundant WWTP core resistome is only disseminated to a very limited extent to other microbial communities. The few resistance genes identified in this study to be disseminated to other microbial communities are associated with canonical mobility elements, suggesting that mobilization, rather than functionality, is the main barrier preventing spread of resistance genes.

Although our study demonstrates that the WWTP core resistome harbours an unappreciated diversity of resistance genes that are unique to this environment, the genetic diversity captured using functional selection is still limited, compared with the total genetic diversity within the WWTP environment. Consequently, the functional selection only allows us to analyse the more abundant resistance genes. Larger studies are needed to characterize the full WWTP resistome and its relation to clinical problems of antibiotic resistance. However, application of metagenomic methods such as those presented in this study should improve the risk classification of various environments, as it allows relating the number of clinically relevant resistance genes to the overall number of functional resistance genes. This will provide a useful tool for establishing control policies for environments that contribute substantially in the dissemination of antibiotic resistance genes.

Methods

DNA extraction. A total of 15 activated sludge samples were collected over a period of 3 years from five different Danish WWTPs with biological nitrogen and phosphorus removal. DNA was extracted using the FastDNA spin kit for soil (MP Biomedicals) according to the manufacturer's instructions, except an initial incubation for 3 min in 90 °C phenol. Briefly, 1 ml sample aliquots were centrifuged at 13,000 r.p.m. (17,000g) for 5 min and the supernatant discarded. The pellets were redissolved in 250 µl phosphate buffer (included in the FastDNA kit) and

transferred to the FastDNA bead beating tubes. To each tube, 750 µl preheated phenol (Biological grade, pH 8, 0.1 M EDTA, Sigma-Aldrich) was added and the samples incubated for 3 min at 90 °C with occasional shaking. The samples were then bead beaten using the manufacturer's instructions and afterwards centrifuged at 13,000 r.p.m. for 10 min, to separate the phenol phase from the top liquid phase containing the DNA. The top phase was extracted and the DNA further purified using the FastDNA kit according to the manufacturer's instructions. DNA concentrations were measured using Qubit (Life Technologies) and DNA integrity evaluated using gel electrophoresis.

Library for functional selection. DNA from WWTP AAW sampled in the third quarter of 2012 was used for functional selection. The DNA was sheared by Covaris E220 instrument (USA). Two hundred microlitres of DNA solution was sheared in a Covaris mini-tube according to Covaris shearing protocol for 2 kbp DNA fragments. Sheared DNA was end repaired using the End-It end repair kit (Epicentre) with the following protocol:

1. Combine:

   a. 68 µl of sheared DNA (90 ng µl⁻¹)
   b. 10 µl 10 × end repair buffer
   c. 10 µl 2.5 mM dNTP mix
   d. 10 µl 10 mM ATP
   e. 2 µl of End It enzyme mix.

2. Incubate at room temperature for 45 min.
3. Stop reaction by heating at 70 °C for 20 min.

End-repaired DNA was size selected by electrophoresis through a 1% low melting point agarose gel in 0.5 × Tris-Borate-EDTA (TBE) buffer. A gel slice corresponding to 1,500–3,000 bp was excised from the gel and DNA was extracted using a QIAquick Gel Extraction Kit (Qiagen). DNA was ligated into pZE21 at the HincII site using the Fast Link ligation kit (Epicenter) with the following ligation protocol:

1. Combine:

   a. 1.0 µl 10 × fast ligation buffer
   b. 0.5 µl 10 mM ATP
   c. 1.0 µl HincII-cut pZE21 MCS 1 vector (65 ng µl⁻¹)
   d. 2.50 µl sheared, end-repaired, gel-purified DNA insert (154 ng µl⁻¹)
   e. 1.0 µl Fast-Link DNA ligase (2 units per µl).

2. Incubate at room temperature for 16 h.
3. Heat inactivate at 70 °C for 20 min.

Then, 3 µl of the fresh ligation mixture was used for electro-transformation (Btex electroporator) into 50 µl of electrocompetent *E. coli* TOP10 cells (Invitrogen). After transformation using standard protocols for a 2-mm electroporation cuvette (Biorad), cells were recovered in 1 ml super optimal broth with catabolite repression (SOC) medium for 1 h at 37 °C. Libraries were titred by plating out 1, 0.1 and 0.01 µl of recovered cells onto lysogeny broth (LB) agar plates containing 50 µg ml⁻¹ kanamycin. For each library, insert size distribution was estimated by gel electrophoresis of colony PCR products obtained by amplifying the insert using primers flanking the multiple cloning site of the pZE21 MCS1 vector (which contains a selectable marker for kanamycin resistance). The average insert size for all libraries was found to be 2.07 kbp. The total size of metagenomic library was 8.28 × 10⁸ bp as determined by multiplying average PCR-based insert size with the number of colony forming units. The rest of the recovered cells were inoculated into 10 ml of LB containing 50 µg ml⁻¹ kanamycin and grown overnight. The overnight culture was frozen down with 15% glycerol and kept at −70 °C for subsequent screening.

Functional selections of antibiotic-resistant clones from metagenomic library. 100 µl of the insert library freezer stock corresponding to 5 × 10⁶ colony forming units, were plated out on LB agar plates containing binary combinations of kanamycin (50 µg ml⁻¹) and 1 of 15 different antibiotics (Table 1), and incubated at 37 °C for 16 h. Assay drug concentrations were chosen to prevent growth of overnight control libraries with the empty vector. Based on the original library titres as well as titres of the freezer stocks, the average amplification of a given library can be estimated. On average, each unique clone in the library was plated out with 10 × coverage. To minimize the redundancy of clones in subsequent analysis, on average, the number of colonies picked for sequencing corresponded to approximately the number of resistant clones on an agar plate divided by the estimated average library amplification.

Each of the clones picked was inoculated into 96 deep-well plates containing liquid LB medium supplemented with kanamycin (50 µg ml⁻¹) and the antibiotic

to which resistance had been selected (Table 1), and grown overnight to verify resistance phenotype.

Sequencing and analysis of metagenomic inserts. Functional selected metagenomic inserts were sequenced using Sanger sequencing. A total of 749 clones were sequenced bidirectionally using the following primers:

Forward primer_pZE21_81_104_57C: 5′-GAA TTC ATT AAA GAG GAG AAA GGT-3′

Reverse primer_pZE21_151_174rc_58C: 5′-TTT CGT TTT ATT TGA TGC CTC TAG-3′.

All reads (≥500 bp) obtained after sequencing were assembled into contigs using CLC Main Workbench 6. The sequence from each contig corresponds to single clone containing unique insert. The full insert sequence for each unique insert was obtained by merging the forward primer sequence reads and reverse primer sequence reads using EMBOSS:Merger server (bioinfo.nhri.org.tw/cgi-bin/emboss/merger, accessed: 20 March 2013). Open reading frames were identified and annotated using ORFfinder (http://www.ncbi.nlm.nih.gov/projects/gorf/, accessed: 20 April 2013). Annotated resistance genes were compared with the GenBank non-redundant nucleotide database (accessed: 1 May 2013) using tblastx, which computes local sequence alignment between the nucleotide query translated in all six frames and the non-redundant nucleotide database translated in all six frames. For each query, the genbank ID and the alignment coordinate for the top scoring tblastx hit was obtained. For each of these top-scoring hits, any annotated feature that overlapped the alignment coordinates in the specific Genbank file was obtained from GenBank. Global sequence alignments and corresponding percentage identities between the query and the obtained sequences were computed using EMBOSS:Stretcher at the nucleotide level in the annotated frame. When multiple annotated features were obtained for a query sequence, only the sequence most similar to the query at the nucleotide level was retained (Supplementary Table 1).

Metagenome sequencing. Samples were prepared for sequencing using the Nextera DNA Sample Preparation Kits (Illumina Inc.) with 50 ng of DNA. The library DNA concentration was measured using the QuantIT kit (Molecular Probes) and paired-end sequenced (2 × 151 bp) on an Illumina HiSeq2000 using the TruSeq PE Cluster Kit v3-cBot-HS and TruSeq SBS kit v.3-HS sequencing kit (Illumina Inc.). The sample AAW-2012-3 was paired-end sequenced (2 × 301 bp) on the Illumina MiSeq platform using the MiSeq Reagents kit v2 (Illumina Inc.) (Supplementary Table 4).

Metagenome read trimming and mapping. Metagenome reads in fastq format were imported to CLC Genomics Workbench v. 7.03 (CLC Bio) and trimmed using a minimum phred score of 20, a minimum length of 50 bp, allowing no ambiguous nucleotides and trimming off Illumina Nextera sequencing adapters, if found. The trimmed metagenome reads were mapped to all antibiotic contigs using CLC Genomics Workbench v. 7.03, using a minimum of 95% similarity over the full read length. The number of reads hitting within the putative antibiotic gene were counted and further data analysis and visualization was conducted using R[52].

The trimmed metagenome reads were also mapped to the Greengenes 16S rRNA database version 13_5 using the map reads to reference function in CLC genomics Workbench v. 7.03 requiring 70% similarity over the full read length and random assignments of reads, which mapped to two sequences equally well.

Metagenome de novo assembly. The trimmed metagenome reads from the AAW-5-2010 sample were assembled using CLC's de novo assembly algorithm, using a kmer of 63 and a minimum scaffold length of 2 kbp.

Metagenomes from other environments. Metagenome reads from four other environments, human gut[37] (ERS006497), permafrost[38], cow rumen[39] (SRP004875) and an aquifer[40] (SRA050978), were downloaded from the NCBI SRA and used for comparison. Trimming and read mapping to the antibiotic contigs were performed as previously described.

16S rRNA gene sequencing. Activated sludge was sampled from aeration tanks at AAW WWTP. Sampling were carried out up to four times a year from 2006 to 2013, resulting in a total of 24 samples.

DNA extraction was conducted using the FastDNA spin kit for soil (MP Biomedical) according to the manufacturer's instructions, except the bead beating was increased to 4 × 40 s at 6 m s⁻¹ using a FastPrep FP120 (MP Biomedicals). The procedure for bacterial 16S rRNA amplicon sequencing targeting the V1-3 variable region was modified from ref. 53. Briefly, 10 ng of extracted DNA was used as template and the PCR reaction (25 µl) contained dNTPs (400 nM of each), MgSO₄ (1.5 mM), Platinum Taq DNA polymerase HF (2 mU), 1× Platinum High Fidelity buffer (Thermo Fisher Scientific) and a pair of barcoded library adaptors (400 nM), V1-3 primers 27F: 5′-AGAGTTTGATCCTGGCTCAG-3′ and 534R: 5′-ATTACCGCGGCTGCTGG-3′. Thermo cycler settings: initial denaturation at 95 °C for 2 min, 30 cycles of 95 °C for 20 s, 56 °C for 30 s, 72 °C for 60 s and final elongation at 72 °C for 5 min. All PCR reactions were run in duplicate and pooled afterwards. The amplicon libraries were purified using the Agencourt AMPure XP

bead protocol (Beckmann Coulter, Brea, CA, USA) with the following exceptions: the sample/bead solution ratio was 5/4 and the purified DNA was eluted in 33 µl nuclease-free water. Library concentration was measured with Quant-iT HS DNA Assay (Thermo Fisher Scientific) and quality validated with a Tapestation 2200 using D1K ScreenTapes (Agilent). Based on library concentrations and calculated amplicon sizes, the samples were pooled in equimolar concentrations and diluted to 4 nM. The library pool was sequenced on a MiSeq (Illumina, San Diego, CA, USA) using a MiSeq Reagent kit v3 (2 × 300 PE) following the procedure in ref. 54, with exception of 10% phiX control library (Illumina) spike-in and final library loading concentration of 20 pM. A detailed DNA extraction and amplicon sequencing protocol can be obtained from www.midasfieldguide.org/en/protocols.

16S rRNA data processing. All sequenced sample libraries were subsampled to 50,000 raw reads and low-quality reads removed using Trimmomatic v. 0.32 (ref. 55) with the settings SLIDINGWINDOW:1:3 and MINLEN:275. All high-quality PE reads were screened for phiX contamination using bowtie2 v. 2.1.0 (ref. 56), with standard settings and all matching reads removed. The potential phiX contamination is due to the use of an un-indexed phiX as a quality control, which can result in index carryover from nearby clusters with indexes.

Forward and reverse reads were merged using FLASH v. 1.2.7 (ref. 57), with the setting -r 300 -f 475 -s 50 -m 25 -M 200 and afterwards merged reads smaller than 425 bp or larger than 525 bp were discarded.

The merged reads were de-replicated and formatted for use in the uparse workflow[58]. The merged reads were clustered using the usearch v. 7.0.1090–cluster_otus with default settings. Operation taxonomic units (OTU) abundance was estimated using the usearch v. 7.0.1090–usearch_global with–id 0.97. Taxonomy was assigned using the RDP classifier[59] as implemented in the parallel_assign_taxonomy_rdp.py script in QIIME[53] using the Greengenes database version 13_5 (ref. 60).

16S rRNA data analysis and visualization. All data analysis and visualizations were conducted using R[52] through the Rstudio IDE (http://www.rstudio.com/). Briefly, OTU counts and associated taxonomic assignments were imported and merged to a phyloseq object[61]. All samples were rarefied to 10,000 reads and subsequently OTUs with a maximum count of <5 in the 24 samples were removed.

Community stability within WWTP AAW as a function of time was calculated by comparing Bray–Curtis β-diversity between all samples.

Rarefaction curve analysis. The rarefaction curves are based on the number of reads that map (>95% identity and >95% coverage) to the functional selected resistance genes. The curves are computed using the R package Vegan (step size = 1 × 10⁶) and ggplot2 (refs 62,63). Abundance estimates are calculated using the Vegan implementation of the species accumulation model described by Chiu et al.[64].

BLAST of functionally selected resistance genes. Functionally selected resistance genes from WWTP AAW, human gut[12,32] and soil[33–36] were blasted against NCBI genbank nt database, excluding self-hits. The top hit was aligned to the functionally selected resistance gene using EMBOSS Matcher and the identity was calculated[65].

Mapping to clinically relevant genes. The trimmed metagenome reads were mapped to reference database (Supplementary Table 1) using CLC Genomics Workbench v. 7.03, using a minimum of 95% similarity over the full read length. The number of reads hitting within the putative antibiotic gene were counted, and further data analysis and visualization were conducted using R.

Procrustes analysis. Square-root-transformed metagenome abundance counts of the antibiotic resistance genes were used for a principal component analysis (PCA) analysis in R through the vegan package[62]. For comparison, matching samples (for a few samples, the closest time point had to be used as matching samples were not available) from the larger 16S rRNA data set were extracted and square-root-transformed OTU abundance counts used for a PCA analysis. The two PCAs were compared using the procrustes function implemented in vegan and significance tested using the protest function with 999 permutations.

References

1. Poirel, L., Kampfer, P. & Nordmann, P. Chromosome-encoded ambler class a betalactamase of *Kluyvera georgiana*, a probable progenitor of a subgroup of CTX-M extended spectrum betalactamases. *Antimicrob. Agents Chemother.* **46,** 4038–4040 (2002).

2. Wright, G. D. Antibiotic resistance in the environment: a link to the clinic? *Curr. Opin. Microbiol.* **13,** 589–594 (2010).

3. Wellington, E. M. H. *et al.* The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect. Dis.* **13,** 155–165 (2013).

4. Forsberg, K. J. *et al.* The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337,** 1107–1111 (2012).

5. Jacoby, G. A. & Hooper, D. C. Phylogenetic analysis of chromosomally determined qnr and related proteins. *Antimicrob. Agents Chemother.* **57,** 1930–1934 (2013).

6. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74,** 417–433 (2010).

7. Allen, H. K. *et al.* Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* **8,** 251–259 (2010).

8. Gibson, M. K., Forsberg, K. J. & Dantas, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* **9,** 207–216 (2015).

9. Nesme, J. *et al.* Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr. Biol.* **24,** 1096–1100 (2014).

10. Dantas, G. & Sommer, M. O. Context matters — the complex interplay between resistome genotypes and resistance phenotypes. *Curr. Opin. Microbiol.* **15,** 577–582 (2012).

11. Forsberg, K. J. *et al.* Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509,** 612–616 (2014).

12. Sommer, M. O. A., Dantas, G. & Church, G. M. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325,** 1128–1131 (2009).

13. Baquero, F., Martínez, J.-L. & Cantón, R. Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol.* **19,** 260–265 (2008).

14. Walsh, T. R., Weeks, J., Livermore, D. M. & Toleman, M. A. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect. Dis.* **11,** 355–362 (2011).

15. Rizzo, L. *et al.* Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Sci. Total Environ.* **447,** 345–360 (2013).

16. Renew, J. E. & Huang, C.-H. Simultaneous determination of fluoroquinolone, sulfonamide, and trimethoprim antibiotics in wastewater using tandem solid phase extraction and liquid chromatography–electrospray mass spectrometry. *J. Chromatogr. A* **1042,** 113–121 (2004).

17. Brown, K. D., Kulis, J., Thomson, B., Chapman, T. H. & Mawhinney, D. B. Occurrence of antibiotics in hospital, residential, and dairy effluent, municipal wastewater, and the Rio Grande in New Mexico. *Sci. Total Environ.* **366,** 772–783 (2006).

18. Watkinson, A. J., Murby, E. J., Kolpin, D. W. & Costanzo, S. D. The occurrence of antibiotics in an urban watershed: From wastewater to drinking water. *Sci. Total Environ.* **407,** 2711–2723 (2009).

19. Szczepanowski, R. *et al.* Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiology* **155,** 2306–2319 (2009).

20. Schwartz, T., Kohnen, W., Jansen, B. & Obst, U. Detection of antibiotic-resistant bacteria and their resistance genes in wastewater, surface water, and drinking water biofilms. *FEMS Microbiol. Ecol.* **43,** 325–335 (2003).

21. Volkmann, H., Schwartz, T., Bischoff, P., Kirchen, S. & Obst, U. Detection of clinically relevant antibiotic-resistance genes in municipal wastewater using real-time PCR (TaqMan). *J. Microbiol. Methods* **56,** 277–286 (2004).

22. Luo, Y. *et al.* Proliferation of multidrug-resistant New Delhi metallo-β-lactamase genes in municipal wastewater treatment plants in Northern China. *Environ. Sci. Technol. Lett.* **1,** 26–30 (2014).

23. Yang, Y., Li, B., Ju, F. & Zhang, T. Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. *Environ. Sci. Technol.* **47,** 10197–10205 (2013).

24. Wang, Z. *et al.* Metagenomic profiling of antibiotic resistance genes and mobile genetic elements in a tannery wastewater treatment plant. *PLoS ONE* **8,** e76079 (2013).

25. Zhang, T., Zhang, X.-X. & Ye, L. Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. *PLoS ONE* **6,** e26041 (2011).

26. Rahube, T. O. & Yost, C. K. Antibiotic resistance plasmids in wastewater treatment plants and their possible dissemination into the environment. *Afr. J. Biotechnol.* **9,** 9183–9190 (2010).

27. Shannon, K. E., Lee, D. Y., Trevors, J. T. & Beaudette, L. A. Application of real-time quantitative PCR for the detection of selected bacterial pathogens during municipal wastewater treatment. *Sci. Total Environ.* **382,** 121–129 (2007).

28. Amann, R. I., Ludwig, W. & Schleifer, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59,** 143–169 (1995).

29. Bray, J. R. & Curtis, J. T. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27,** 325–349 (1957).

30. Albertsen, M., Karst, S. M., Ziegler, A. S., Kirkegaard, R. H. & Nielsen, P. H. Back to basics - the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PLoS ONE* **10,** e0132783 (2015).

31. Albertsen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* **31,** 533–538 (2013).

32. Cheng, G. *et al.* Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol. Lett.* **336,** 11–16 (2012).

33. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A. & Handelsman, J. Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *ISME J.* **3,** 243–251 (2008).

34. Donato, J. J. *et al.* Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. *Appl. Environ. Microbiol.* **76,** 4396–4401 (2010).

35. Lang, K. S. *et al.* Novel florfenicol and chloramphenicol resistance gene discovered in Alaskan soil by using functional metagenomics. *Appl. Environ. Microbiol.* **76,** 5321–5326 (2010).

36. Riesenfeld, C. S., Goodman, R. M. & Handelsman, J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* **6,** 981–989 (2004).

37. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464,** 59–65 (2010).

38. Mackelprang, R. *et al.* Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* **480,** 368–371 (2011).

39. Hess, M. *et al.* Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* **331,** 463–467 (2011).

40. Wrighton, K. C. *et al.* Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337,** 1661–1665 (2012).

41. Livermore, D. M. *et al.* CTX-M: changing the face of ESBLs in Europe. *J. Antimicrob. Chemother.* **59,** 165–174 (2006).

42. Schmitz, F. J. *et al.* Prevalence of macrolide-resistance genes in Staphylococcus aureus and Enterococcus faecium isolates from 24 European university hospitals. *J. Antimicrob. Chemother.* **45,** 891–894 (2000).

43. Schmitz, F. J. *et al.* The prevalence of aminoglycoside resistance and corresponding resistance genes in clinical isolates of staphylococci from 19 European hospitals. *J. Antimicrob. Chemother.* **43,** 253–259 (1999).

44. Sentchilo, V. *et al.* Community-wide plasmid gene mobilization and selection. *ISME J.* **7,** 1173–1186 (2013).

45. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat. Rev. Microbiol.* **12,** 465–478 (2014).

46. Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L. & Andersson, D. I. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *mBio* **5,** e01918–14 (2014).

47. Girlich, D., Poirel, L., Szczepanowski, R., Schlüter, A. & Nordmann, P. Carbapenem-hydrolyzing GES-5-encoding gene on different plasmid types recovered from a bacterial community in a sewage treatment plant. *Appl. Environ. Microbiol.* **78,** 1292–1295 (2012).

48. Stalder, T. *et al.* Quantitative and qualitative impact of hospital effluent on dissemination of the integron pool. *ISME J.* **8,** 768–777 (2014).

49. Agersø, Y. & Sandvang, D. Class 1 integrons and tetracycline resistance genes in alcaligenes, arthrobacter, and Pseudomonas spp. isolated from pigsties and manured soil. *Appl. Environ. Microbiol.* **71,** 7941–7947 (2005).

50. Li, D. *et al.* Antibiotic-resistance profile in environmental bacteria isolated from penicillin production wastewater treatment plant and the receiving river. *Environ. Microbiol.* **11,** 1506–1517 (2009).

51. Sengeløv, G., Kristensen, K. J., Sørensen, A. H., Kroer, N. & Sørensen, S. J. Effect of genomic location on horizontal transfer of a recombinant gene cassette between Pseudomonas strains in the rhizosphere and spermosphere of barley seedlings. *Curr. Microbiol.* **42,** 160–167 (2001).

52. Core Team, R. *R: A Language and Environment for Statistical Computing.* Available at < http://www.R-project.org/ > .

53. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7,** 335–336 (2010).

54. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6,** 1621–1624 (2012).

55. Lohse, M. *et al.* RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* **40,** W622–W627 (2012).

56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

57. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27,** 2957–2963 (2011).

58. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10,** 996–998 (2013).

59. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

60. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6,** 610–618 (2012).

61. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8,** e61217 (2013).

62. Oksanen, J., Kindt, R. & Legendre, P. *vegan: Community Ecology Package* 1–190 (2007).

63. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer Publishing Company, Incorporated, 2009).

64. Chiu, C.-H., Wang, Y.-T., Walther, B. A. & Chao, A. An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics* **70,** 671–682 (2014).

65. McWilliam, H. *et al.* Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41,** W597–W600 (2013).

Acknowledgements

Author contributions

P.H.N. and M.O.A.S. designed the study. A.T. performed functional selection. M.A. collected samples and performed metagenomic sequencing. C.M., M.A., A.T. and M.E. performed data analysis. C.M., M.A., P.H.N. and M.O.A.S. wrote the manuscript.

Additional information

Accession codes: Nucleotide sequences of the functionally selected resistance genes have been deposited in the Genbank database with accession codes KT387137 to KT387215.

The raw metagenome reads have been deposited in the European Nucleotide Archive under study PRJEB8087. The assembled metagenome scaffolds from AAW.5.2010 have been deposited in the MG-RAST database with accession code MG-RAST4611649.3. The raw V13 16S rRNA amplicon sequences are part of the MiDAS data set (www.midasfieldguide.org) and have been deposited in the European Nucleotide Archive under study PRJEB8105. The accession codes for the specific samples used are: ERS634988, ERS634989, ERS634990, ERS634991, ERS635004, ERS635036, ERS635037, ERS635038, ERS635362, ERS635433, ERS635452, ERS635453, ERS635454, ERS635455, ERS635477, ERS635478, ERS635521, ERS635529, ERS635530, ERS635539, ERS635540, ERS635541, ERS635549 and ERS635550.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Munck, C. *et al.* Limited dissemination of the wastewater treatment plant core resistome. *Nat. Commun.* 6:8452 doi: 10.1038/ncomms9452 (2015).