Journal of Biotechnology 136 (2008) 54–64

Contents lists available at ScienceDirect

Journal of Biotechnology

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

Insight into the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to antimicrobial drugs analysed by the 454-pyrosequencing technology

Rafael Szczepanowski[a], Thomas Bekel[b], Alexander Goesmann[b], Lutz Krause[b], Holger Krömeke[b], Olaf Kaiser[c], Wolfgang Eichler[d], Alfred Pühler[a], Andreas Schlüter[a,*]

[a] Department of Genetics, Bielefeld University, Postfach 100131, D-33501 Bielefeld, Germany
[b] BRF, CeBiTec, Bielefeld University, Postfach 100131, D-33501 Bielefeld, Germany
[c] Roche Diagnostics GmbH, Nonnenwald 2, D-82377 Penzberg, Germany
[d] Landesamt für Natur, Umwelt und Verbraucherschutz NRW, Postfach 101052, D-45610 Recklinghausen, Germany

ARTICLE INFO

Article history:
Received 15 January 2008
Received in revised form 20 March 2008
Accepted 31 March 2008

Keywords:
Replication initiation
Stabilisation
Mobilisation
Conjugative transfer
Antibiotic resistance
Virulence
454-Pyrosequencing
Plasmid metagenome

ABSTRACT

Wastewater treatment plants (WWTPs) are a reservoir for bacteria harbouring antibiotic resistance plasmids. To get a comprehensive overview on the plasmid metagenome of WWTP bacteria showing reduced susceptibility to certain antimicrobial drugs an ultrafast sequencing approach applying the 454-technology was carried out. One run on the GS 20 System yielded 346,427 reads with an average read length of 104 bases resulting in a total of 36,071,493 bases sequence data. The obtained plasmid metagenome was analysed and functionally annotated by means of the Sequence Analysis and Management System (SAMS) software package. Known plasmid genes could be identified within the WWTP plasmid metagenome data set by BLAST searches using the NCBI Plasmid Database. Most abundant hits represent genes involved in plasmid replication, stability, mobility and transposition. Mapping of plasmid metagenome reads to completely sequenced plasmids revealed that many sequences could be assigned to the cryptic pAsa plasmids previously identified in *Aeromonas salmonicida* subsp. *salmonicida* and to the accessory modules of the conjugative IncU resistance plasmid pFBAOT6 of *Aeromonas punctata*. Matches of sequence reads to antibiotic resistance genes indicate that plasmids from WWTP bacteria encode resistances to all major classes of antimicrobial drugs. Plasmid metagenome sequence reads could be assembled into 605 contigs with a minimum length of 500 bases. Contigs predominantly encode plasmid survival functions and transposition enzymes.

1. Introduction

Wastewater treatment plants (WWTPs) represent habitats with ever-changing chemical and biological composition of the sewage. Therefore indigenous microorganisms have to adapt very rapidly to these fluctuating conditions by means of genome rearrangements which require distinctive genetic flexibility. Mobile genetic elements and in particular plasmids which form the horizontally mobile gene pool of bacteria have been recognised to be very important elements for genome restructuring and the dissemination of beneficial properties within the population and the bacterial community (Davison, 1999; Frost et al., 2005; Thomas and Nielsen, 2005). In recent years, many different antibiotic resistance plasmids and plasmids mediating degradative capabilities were isolated from WWTP bacteria and analysed in detail at the genomic level (Schlüter et al., 2003, 2005, 2007a,b,c; Szczepanowski et al., 2004, 2005, 2007; Tauch et al., 2003; Tennstedt et al., 2003, 2005). These studies very clearly showed that there is immense plasmid diversity within the bacterial community of WWTPs. On the other hand they demonstrated that there are fundamental limitations of plasmid genomics on an individual basis. (i) Many plasmids do not replicate in laboratory hosts used for capturing plasmids from WWTP-bacteria. (ii) The accessibility of plasmids which do not carry selectable marker genes is difficult. (iii) Individual genomic analysis is only possible for a selected subset of plasmids. Therefore a metagenomic approach was envisaged to obtain a more comprehensive overview on the genetic information encoded on plasmids residing in WWTP bacteria. Very recently, the ultrafast 454-sequencing methodology has turned out to represent an ideal tool for the analysis of complex microbial communities since DNA-cloning steps are no longer required prior to sequencing (Edwards et al., 2006; Krause et al., 2006).

* Corresponding author. Tel.: +49 521 106 2036; fax: +49 521 106 5626.
E-mail address: Andreas.Schlueter@Genetik.Uni-Bielefeld.DE (A. Schlüter).

0168-1656/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2008.03.020

In this paper, the plasmid metagenome of WWTP bacteria showing reduced susceptibility to certain antimicrobial drugs was sequenced by applying the 454-sequencing technology. Obtained sequence data were analysed by means of the Sequence Analysis and Management System (SAMS) facilitating different BLAST-searches against defined databases to get an overview on the occurrence of plasmids and their accessory genetic elements, especially resistance determinants.

2. Materials and methods

2.1. Isolation of plasmid-DNA from antibiotic resistant bacteria of a wastewater treatment plant

To isolate total plasmid-DNA, an activated sludge sample was taken from a sludge basin of the municipal wastewater treatment plant in Bielefeld-Heepen (Germany) in September 2007. Aliquots of the activated sludge sample (100 μl) were plated (six parallels) on Luria-Broth media supplemented with one of the following antibiotics: 100 μg ampicillin ml^{-1}, 1 μg cefotaxime ml^{-1}, 15 μg cefuroxime ml^{-1}, 1 μg ciprofloxacin ml^{-1}, 200 μg erythromycin ml^{-1}, 20 μg gentamicin ml^{-1}, 50 μg kanamycin ml^{-1}, 1 μg norfloxacin ml^{-1}, 30 μg rifampicin ml^{-1}, 100 μg spectinomycin ml^{-1}, 100 μg streptomycin ml^{-1}, or 5 μg tetracycline ml^{-1}, respectively. Antimicrobial drug concentrations were chosen according to previous publications (Schlüter et al., 2003, 2005, 2007a,b; Szczepanowski et al., 2004, 2005, 2007; Tauch et al., 2003; Tennstedt et al., 2003; Tennstedt et al., 2005). Concentrations for all other antibiotics were determined empirically by means of growth tests of WWTP bacteria on selective media containing different concentrations of the respective antimicrobial compound. Media also contained cycloheximide (at a final concentration of 75 μg ml^{-1}) to reduce growth of fungi. The LB-agar plates were incubated for 36 h at 25 °C. Fungi growing even in the presence of cycloheximide were removed to avoid problems during plasmid isolation. Growing bacteria with reduced susceptibility to the applied antimicrobial drugs were collected and total plasmid-DNA was isolated by using an alkaline lysis method followed by caesium chloride density ultra-centrifugation. Alkaline lysis was done by means of the Macherey–Nagel NucleoBond PC100 Kit on the AX100 columns according to the manufacturer's protocol (Macherey–Nagel, Düren, Germany). Caesium chloride density ultra-centrifugation was performed according to Sambrook and coauthors (Sambrook et al., 1989).

2.2. Sequencing and assembly of a plasmid metagenome

Sequencing of the sampled plasmid metagenome of WWTP bacteria with reduced susceptibility to selected antimicrobial drugs was carried out by 454 Life Sciences (Branford, CT, USA) on the Genome Sequencer (GS) 20 System. For this purpose, 10 μg of the combined total plasmid-DNA preparation was used. Assembly of the obtained nucleotide sequence reads was done by using the Newbler Assembler Software (454 Life Sciences). Obtained plasmid metagenome sequence reads are available via the link: ftp://ftp.cebitec.uni-bielefeld.de/pub/supplements/SzczepanowskiEtAl_Insight_JournalBiotech_2008.zip.

2.3. Characterisation of plasmid metagenome reads and assembled contigs by means of the SAMS system

The Sequence Analysis and Management System (http://www.cebitec.uni-bielefeld.de/groups/brf/software/sams_info/) was originally designed and implemented for quality control of sequence data obtained during the high throughput phase of genome sequencing projects. SAMS is intrinsically well suited for the analysis of short contigs. Therefore in this study the system was used to characterise plasmid metagenome reads and assembled contigs.

Similar to the analysis of potential coding regions predicted on a bacterial chromosome, in SAMS individual (short) sequences are analysed and functionally annotated. An automated function prediction (Metanor) is computed using a combination of standard bioinformatics tools such as BLAST and InterPro (Mulder et al., 2007). This approach leads to consistent gene annotations, assigning gene names, gene products, EC numbers, functional protein categories (COGs) and GO numbers.

For the functional annotation of unassembled metagenome reads, the Metanor analysis pipeline was applied with four different BLAST tools: Blast2n vs. the MvirDB nucleotide database (*E*-value cut-off of 10^{-10}), Blast2x vs. the MvirDB protein database (*E*-value cut-off of 10^{-10}) and Blast2n vs. the NCBI Plasmid Database (*E*-value cut-off of 10^{-10}).

For the functional annotation of assembled contigs from the plasmid metagenome, the Metanor analysis pipeline was employed with six tools: Blast2n vs. the NCBI nucleotide database (*E*-value cut-off of 10^{-4}); Blast2x vs. the NCBI protein database (*E*-value cut-off of 10^{-4}); Blast2x vs. the KEGG database (Kanehisa et al., 2006) (*E*-value cut-off of 10^{-4}); Blast2x vs. the Swissprot database (Bairoch and Boeckmann, 1993) (*E*-value cut-off of 10^{-4}); Blast2x vs. COG (Tatusov et al., 2003) (*E*-value cut-off of 10^{-4}) and InterPro.

2.4. Mapping of the plasmid metagenome reads by means of the ReadMapper tool

The ReadMapper tool (Krömeke et al., unpublished) was developed to visualise results from similarity searches using the BLAST program. ReadMapper is able to manage an unlimited number of BLAST results and is for this reason used to map plasmid metagenome reads to plasmid sequences. With BLAST for each plasmid metagenome read, a list of plasmid sequences producing significant alignments is received. Alignments with an *E*-value lower or equal to 10^{-10} and more than 90% identity are considered for assigning them to the plasmid sequence for visualisation. For each plasmid sequence and the assigned set of plasmid metagenome reads a plot was drawn. Therefore, the reads were sorted non-overlapping according to start and end position of their alignment into layer, while each read consists of the alignment and possible unmatched overhangs. The plasmid sequence is represented with parts where the set of matching reads show similarity and parts showing no similarity. Furthermore, a coverage-profile is build from the number of alignments at each nucleotide position and known genes are visualised due to their position on the plasmid.

3. Results and discussion

3.1. The plasmid metagenome of antibiotic resistant bacteria residing in a municipal wastewater treatment plant analysed by the 454-sequencing technology

To get an overview on the occurrence of plasmids and their accessory genetic elements residing in WWTP bacteria, a metagenomic approach was envisaged. Since the primary objective of this project was to estimate the diversity and nature of antibiotic resistance plasmids, it was decided to limit the sequencing effort to plasmids harboured by WWTP bacteria that show reduced susceptibility to selected antimicrobial drugs. For this approach activated sludge bacteria of the municipal wastewater treatment plant Bielefeld-Heepen (Germany) were selected on media sup-

Table 1
Categorisation of WWTP plasmid metagenome reads according to clusters of orthologous groups (COGs) of proteins

COG category	Description	Frequency (%)
L	Replication, recombination and repair	37.1
–	No assignment	21.8
V	Defense mechanisms	9.8
R	General function prediction only	8.4
NT	Cell motility; signal transduction mechanisms	4.7
JD	Translation, ribosomal structure and biogenesis; cell cycle control, cell division, chromosome partitioning	2.7
H	Coenzyme transport and metabolism	2.3
P	Inorganic ion transport and metabolism	2.0
J	Translation, ribosomal structure and biogenesis	1.8
S	Function unknown	1.7
K	Transcription	1.4
U	Intracellular trafficking, secretion, and vesicular transport	1.3
PR	Inorganic ion transport and metabolism; general function prediction only	0.9
C	Energy production and conversion	0.8
D	Cell cycle control, cell division, chromosome partitioning	0.8
T	Signal transduction mechanisms	0.7
NU	Cell motility; intracellular trafficking, secretion, and vesicular transport	0.4
O	Posttranslational modification, protein turnover, chaperones	0.3
M	Cell wall/membrane/envelope biogenesis	0.2
KR	Transcription; general function prediction only	0.2
MU	Cell wall/membrane/envelope biogenesis; intracellular trafficking, secretion, and vesicular transport	0.1
Q	Secondary metabolites biosynthesis, transport and catabolism	0.1
I	Lipid transport and metabolism	0.1
G	Carbohydrate transport and metabolism	0.1
E	Amino acid transport and metabolism	0.1

plemented with one of twelve clinically relevant antibiotics each (ampicillin, cefotaxime, cefuroxime, ciprofloxacin, erythromycin, rifampicin, gentamicin, kanamycin, norfloxacin, spectinomycin, streptomycin or tetracycline). Total plasmid-DNA was prepared from bacteria able to grow on these selective media and used as template for an ultrafast 454-sequencing approach on the GS 20 System (454 Life Sciences, Branford, CT, USA). One GS 20 run yielded 346,427 reads with an average sequence length of 104 bases resulting in 36,071,493 bases nucleotide sequence information.

3.2. Annotation of single reads of the WWTP plasmid metagenome

To analyse the genetic information content inhering in the nucleotide sequence data set of the plasmid metagenome, a 'Sequence Analysis and Management System' (SAMS) project was set up. SAMS can manage thousands of shotgun reads, allows for launching different BLAST searches, stores obtained observations and automatically annotates sequence reads which can also be categorized and retrieved for certain criteria. Table 1 provides an overview on categorisation of sequence reads according to Clusters of Orthologous Groups of proteins (COGs). The most prominent proportion of reads could be assigned to the COG category DNA replication, recombination and repair (L) which indicates that genes having functions in plasmid replication and recombination processes (for example homologous recombination, transposition and site-specific recombination) are present in the sequenced sample.

Likewise, plasmid metagenome sequence reads were classified according to the Gene Ontology (GO), which provides a controlled vocabulary for the description of genes and gene products with respect to their associated biological processes, cellular components and molecular functions. The high number of hits to the component term 'extrachromosomal circular DNA' (see Table 2) was expected since the protocol used for plasmid-DNA preparation involved removal of contaminating chromosomal DNA by caesium chloride density-gradient ultra-centrifugation. Assignment of the GO-terms 'DNA replication initiation', 'plasmid maintenance' and 'unidirectional conjugation' indicates that plasmid replication, stabilisation and mobility modules are present in the sequenced plasmid preparation, whereas the GO-terms 'DNA transposition', 'DNA recombination' and 'DNA integration' point to the occurrence of transposable elements and integrons in the sample.

Table 2
Classification of nucleotide sequence reads of a WWTP plasmid metagenome according to gene ontology (GO)

Category	Gene ontology (GO) term	Number of hits[a]
Component	Extrachromosomal circular DNA	11,468
Component	Membrane	1588
Component	Intracellular	963
Component	Integral to membrane	586
Component	Cytoplasm	198
Component	Outer membrane	186
Component	Chromosome	139
Function	DNA binding	19,196
Function	Transposase activity	8754
Function	DNA-directed DNA polymerase activity	4766
Function	Recombinase activity	2310
Function	ATP binding	1301
Function	Sequence-specific DNA binding	1067
Function	Catalase activity	1016
Process	DNA transposition	8931
Process	DNA replication initiation	4889
Process	Plasmid maintenance	4411
Process	DNA recombination	4321
Process	Unidirectional conjugation	2107
Process	DNA replication	1809
Process	DNA integration	1670
Process	Regulation of transcription DNA-dependent	1528
Process	DNA methylation	1218
Process	Electron transport	1195
Process	Response to antibiotic	1168
Process	Response to oxidative stress	1126

[a] Only component GO-terms with more than 100 hits and function and process GO-terms with more than 1000 hits are listed.

BLASTN search of the data set revealed that approximately 49,000 reads could be assigned to known plasmid genes stored in the NCBI Plasmid Database (http://www.ncbi.nlm.nih.gov/genomes/static/eub_p.html). Most abundant hits to known plasmid genes are listed in Table 3. Replication initiation (*rep*) and mobilisation (*mob*) genes similar to those previously identified on the cryptic *Aeromonas salmonicida* subsp. *salmonicida* plasmids pAsa1, pAsa2, pAsa3, pAsal1, pAsal2 and pAsal3 are among the most prevalent genes present in the WWTP plasmid metagenome sample. The pAsa plasmids contain a ColE-related replication module, a mobilisation module and in some cases a gene pair encoding a toxin–antitoxin system that most probably plays a role for plasmid stability (Boyd et al., 2003). So far, no host-beneficial traits could be identified on any of the pAsa reference plasmids. Other abundant genes are very similar to the partitioning gene *parA* of the mobilisable IncP-6 plasmid Rms149 of *Pseudomonas aeruginosa* (Haines et al., 2005), the replication initiation gene of the *Salmonella enterica* subsp. *enterica* colicigenic plasmid pBERT (NC_001848), the replication gene *rom* of the *Klebsiella pneumoniae* bacteriocin-producing plasmid pKlebB-K17/80 (Riley et al., 2001) and a gene with unknown function of the antibiotic resistance plasmid pRAY of a clinical *Acinetobacter* isolate (Segal and Elisha, 1999). Further prominent sequence reads originate from transposable elements

Table 3
Most abundant hits of WWTP plasmid metagenome sequence reads to plasmid genes deposited in the NCBI plasmid database

Accession no.	Hits[a]	Gene product	Organism	Source	Plasmid	Note[b]	Reference
YP_245435.1	1921	ParA partitioning protein	*Pseudomonas aeruginosa*	Clinical isolate	Rms149	IncP-6, mob, AB-resistance	Haines et al. (2005)
NP_710167.1	1426	Putative RepA protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsal1	Cryptic, mob	NC_004338
NP_861563.1	1280	Replication primase	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa3	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_049460.1	1198	Hypothetical protein pRAY_p01	*Acinetobacter* sp.	Clinical isolate	pRAY	AB-resistance	Segal and Elisha (1999)
NP_861554.1	1108	Replication primase	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa1	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_598138.1	1078	Transposase Tn*5719* (chromate resistance Tn)	Uncultured bacterium	Activated sludge	pB4	IncP-1β, conjugative, bhr, AB-resistance	Tauch et al. (2003)
NP_045376.1	1042	Hypothetical protein pBERT_01	*Salmonella enterica* subsp. *enterica* serovar Berta		pBERT	Colicigenic, pathogenesis-related	NC_001848
NP_710174.1	990	Putative RepA protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsal2	Cryptic, mob	NC_004339
NP_861566.1	812	Mobilisation relaxase	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa3	ColE2-type, cryptic, mob	Boyd et al. (2003)
YP_067858.1	799	Methyl-accepting chemotaxis protein of Tn*1721* (tetracycline resistance Tn)	*Aeromonas punctata*	Hospital effluent isolate	pFBAOT6	IncU, conjugative, In, AB-resistance	Rhodes et al. (2004)
NP_710164.1	769	Putative MobA protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsal1	Cryptic, mob	NC_004338
NP_861573.1	750	Mobilisation relaxase	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa2	ColE1-type, cryptic, mob	Boyd et al. (2003)
NP_861564.1	648	Hypothetical protein pAsa3_05	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa3	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_861557.1	626	Mobilisation relaxase	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa1	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_710178.1	621	Putative MobA protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsal3	Cryptic, mob	NC_004340
NP_068714.1	619	Rom protein	*Klebsiella pneumoniae*		pKlebB-k17/80	ColE1-type, klebicin B plasmid	Riley et al. (2001)
NP_639997.1	613	Transposase of IS*5*	*Proteus vulgaris*	Clinical isolate	Rts1	IncT, conjugative, kanamycin resistance	Murata et al. (2002)
NP_710170.1	576	Putative MobA protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsal2	Cryptic, mob	NC_004339
NP_957548.1	555	TnpA	*Escherichia coli*	Clinical isolate	pC15-1a	R100-like, MDR, cephalosporin resistance	Boyd et al. (2004)
YP_067860.1	554	Transposase for transposon Tn*1721*	*Aeromonas punctata*	Hospital effluent isolate	pFBAOT6	IncU, conjugative, In, AB-resistance	Rhodes et al. (2004)
NP_862996.1	530	Hypothetical protein p165897_084	*Escherichia coli*	Clinical isolate	p1658/97	IncFII, IncFIB, conjugative, β-lactam resistance	Zienkiewicz et al. (2007)
NP_604399.1	468	MbeA	*Salmonella enteritidis* serovar *Enteritidis*		pK	ColE1-type	NC_003456
YP_449010.1	454	Macrolide-specific ABC-type efflux carrier	*Escherichia coli*	Clinical isolate from a pig	pMUR050	IncN, conjugative, AB-resistance	González-Zorn et al. (2005)
YP_209313.1	441	Transposase of a Tn*3*-like Tn	*Salmonella enterica* subsp. *enterica* serovar Choleraesuis	Clinical isolate	pSC138	IncI1, non-conjugative, AB-resistance	Chiu et al. (2005)
YP_067891.1	436	Tn*3*-family transposase	*Aeromonas punctata*	Hospital effluent isolate	pFBAOT6	IncU, conjugative, In, AB-resistance	Rhodes et al. (2004)
YP_009034.1	431	Hypothetical protein pSERF1_p1	*Salmonella enterica* subsp. *enterica* serovar Choleraesuis		pSERF1	Cryptic	NC_005862

Table 3 (*Continued*)

Accession no.	Hits[a]	Gene product	Organism	Source	Plasmid	Note[b]	Reference
NP_861565.1	401	Mobilisation protein	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa3	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_775053.1	395	Probable macrolide efflux protein	*Citrobacter freundii*		pCTX-M3	Transmissible, AB-resistance	NC_004464
NP_710173.1	389	Putative RelE protein	*Aeromonas salmonicida* subsp. *salmonicida*		pAsa12	Cryptic, mob	NC_004339
NP_569374.1	360	Transposase	*Salmonella enterica* subsp. *enterica* serovar *Typhi*	Clinical isolate	pHCM1	IncHI-like, conjugative, AB-resistance, heavy metal resistance	Parkhill et al. (2001)
YP_449009.1	314	Putative macrolide 2′-phosphostransferase	*Escherichia coli*	Clinical isolate from a pig	pMUR050	IncN, conjugative, AB-resistance	González-Zorn et al. (2005)
NP_861556.1	303	Mobilisation protein	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa1	ColE2-type, cryptic, mob	Boyd et al. (2003)
NP_861571.1	277	Hypothetical protein pAsa2_03	*Aeromonas salmonicida* subsp. *salmonicida*	Brown trout, Eure, France	pAsa2	ColE1-type, cryptic, mob	Boyd et al. (2003)
NP_040356.1	261	Colicin E1 protein	*Escherichia coli*		ColE1	Colicin production, mob	Chan et al. (1985)

[a] Only entries with more than 250 hits are listed.
[b] Abbreviations: bhr, broad-host-range; AB, antibiotic; MDR, multidrug resistance; mob, mobilisation; In, integron; Inc, incompatibility.

such as the chromate resistance transposon Tn*5719* (Tauch et al., 2003), the tetracycline resistance transposon Tn*1721* (Allmeier et al., 1992; Rhodes et al., 2004), the transposon Tn*5403* of the extended-spectrum β-lactamase encoding plasmid pC15-1a (Boyd et al., 2004), Tn*3*-related transposons (Rhodes et al., 2004) and an insertion sequence of the IS*5* family. Some of these mobile genetic elements were already identified on antibiotic resistance plasmids that were isolated from activated sludge bacteria of a municipal wastewater treatment plant (Schlüter et al., 2007a). The most abundant resistance gene in the sample is the macrolide efflux gene *mel* that is also present on the WWTP plasmid pRSB105 (Schlüter et al., 2007b) and the *Citrobacter freundii* resistance plasmid pCTX-M3 (AF550415).

3.3. Mapping of WWTP plasmid metagenome reads to plasmid genes and genomes deposited in databases

To visualise the distribution of plasmid metagenome reads on selected plasmid genes, the mapping tool ReadMapper (Krömeke et al., unpublished) was used. Inspection of mapping data revealed that the genes mentioned in the previous section are densely covered by sequence reads. For example, reads mapping to the *Aeromonas salmonicida* subsp. *salmonicida* pAsa3 *rep* gene (Boyd et al., 2003) are shown in Fig. 1. These results confirm high abundance of the pAsa3 replication module.

Sequence reads were also mapped to complete plasmid genomes extracted from the NCBI Plasmid Database. Fig. 1 shows the mapping results for the cryptic plasmid pAsa3 (Boyd et al., 2003). Uneven distribution of reads over the pAsa3 genome reflects the modular composition of the plasmid. The module represented by *orf4* on pAsa3 seems to be absent from some pAsa3-related plasmids residing in WWTP bacteria. Likewise, Fig. 2 shows the coverage of the antibiotic resistance plasmid pRAY originating from the clinical isolate *Acinetobacter* strain SUN (Segal and Elisha, 1999). The pRAY module containing mobilisation genes and three *orf*s of unknown function is covered by more plasmid metagenome reads than the rest of the plasmid indicating higher abundance of the corresponding module in the sequenced sample. Mapping of reads to the conjugative IncU resistance plasmid pFBAOT6 from *Aeromonas punctata* (Rhodes et al., 2004) revealed that the coverage of the conjugative transfer region and a module of unknown function is very low, whereas many reads could be assigned to the accessory modules of the plasmid (see Fig. 3). These segments contain different transposons, insertion sequences and remnants of these elements. Thus, the modular composition of some pFBAOT6-related plasmids present in the WWTP plasmid metagenome sample seems to be different compared to the reference plasmid. It is very likely that mobile genetic elements identified in the pFBAOT6 genome are also integral parts of other plasmids residing in WWTP bacteria.

3.4. Assignment of plasmid metagenome reads to known antibiotic resistance determinants

To identify sequence reads representing antibiotic resistance genes, different approaches were followed. First, the nucleotide sequences of all 346,427 plasmid metagenome reads stored in the SAMS-project were blasted to genes deposited in the MvirDB database (Zhou et al., 2007). This database integrates nucleotide sequence and protein sequence information for protein toxins, virulence factors and antibiotic resistance genes. Entries of the ARGO (Antibiotic Resistance Genes Online) database of vancomycin and β-lactam antibiotic resistance genes (Scaria et al., 2005) provide the set of resistance determinants in MvirDB. Accordingly, MvirDB is not a comprehensive database including all known antibiotic resistance genes and proteins. A total of 918 reads matched to

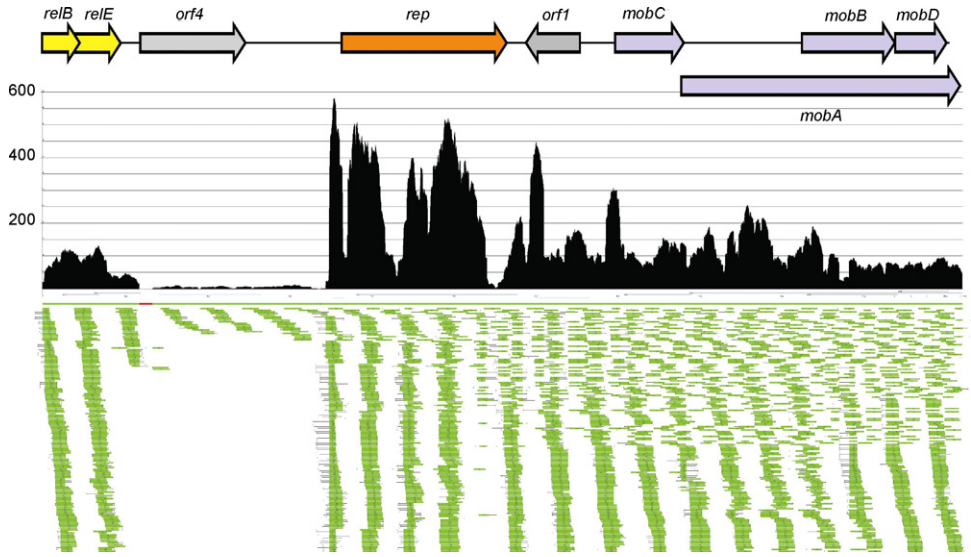

Fig. 1. Mapping of WWTP plasmid metagenome nucleotide sequence reads to the cryptic plasmid pAsa3 of *Aeromonas salmonicida* subsp. *salmonicida*. Arrows symbolise the genes of plasmid pAsa3 that is 5616 bp in size. The plasmid is composed of a toxin-antitoxin gene pair (*relB–relE*), a replication initiation module (*rep*) and a mobilisation module (*mob*). The functions of *orf1* and *orf4* are unknown. Metagenome reads were aligned to pAsa3 using BLAST with an *E*-value cut-off of 10^{-10}. Coverage of the pAsa3 genome by reads is plotted below the genetic map. Aligned reads are shown as green bars. The *y*-axis denotes the coverage per nucleotide of the reference sequence. The bottom of the figure was cut.

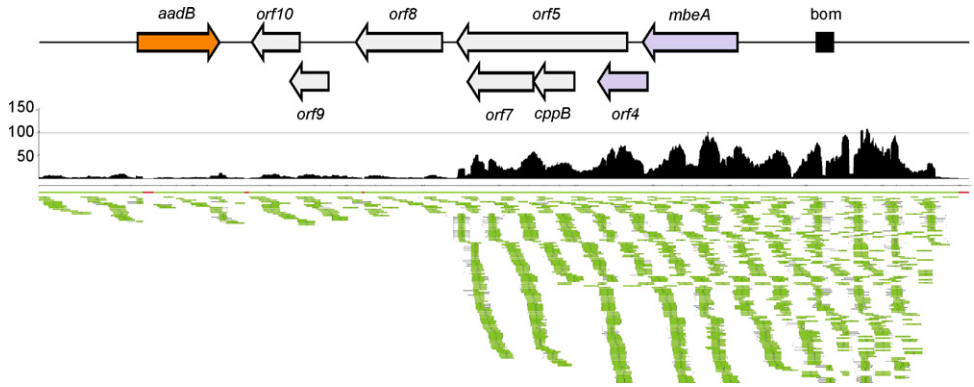

Fig. 2. Mapping of WWTP plasmid metagenome nucleotide sequence reads to the antibiotic resistance plasmid pRAY of the clinical isolate *Acinetobacter* strain SUN. Arrows symbolise the genes of plasmid pRAY that is 6076 bp in size. The plasmid is composed of a mobilisation module (*mbeA*, *orf4*), *orfs* of unknown function and the aminoglycoside resistance gene *aadB*. The black rectangle indicates the basis of mobility (bom). Metagenome reads were aligned to pRAY using BLAST with an *E*-value cut-off of 10^{-10}. Coverage of the pRAY genome by reads is plotted below the genetic map. The *y*-axis denotes the coverage per nucleotide of the reference sequence. Aligned reads are shown as green bars. The bottom of the figure was cut.

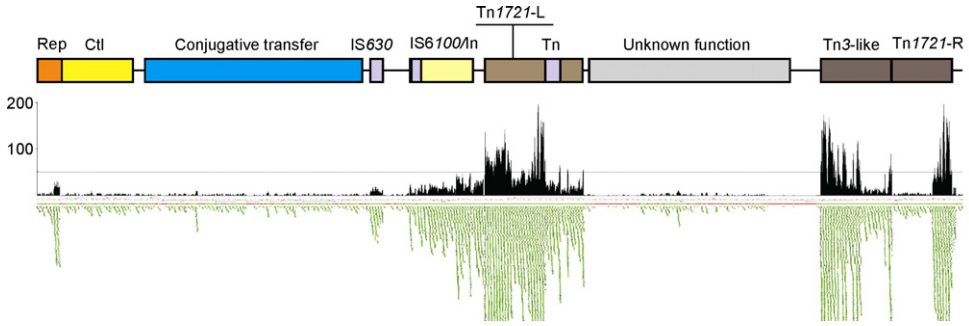

Fig. 3. Mapping of WWTP plasmid metagenome nucleotide sequence reads to the conjugative IncU resistance plasmid pFBAOT6 of *Aeromonas punctata*. Bars symbolise the modular composition of the plasmid, which is 84,749 bp in size. It is composed of a replication module (Rep), a control module (Ctl), a conjugative transfer module and a region of unknown function. Insertion sequences (IS*630*, IS*6100*), a class 1 integron (In) and transposons (Tn*1721* and Tn*3*-like) constitute the accessory modules of the plasmid. Metagenome reads were aligned to pFBAOT6 using BLAST with an *E*-value cut-off of 10^{-10}. Coverage of the pFBAOT6 genome by plasmid metagenome reads is plotted below the map of the plasmid. Aligned reads are displayed as green bars. The *y*-axis denotes the coverage per nucleotide of the reference sequence. The bottom of the figure was cut.

Table 4
Resistance gene specific sequences identified within the WWTP plasmid metagenome sequence data set by BLASTn analyses to the microbial virulence database (Mvir-DB) and by comparison with a manually established antibiotic resistance gene database (ARG-DB)

Antibiotic class	Resistance genes[a]	Hits to Mvir-DB[b]	Hits to ARG-DB[c]
Aminoglycosides	*aacA*[c], *aacA29b*[b,c], *aacA4*[b,c], *aacC1*[c], *aacC2*[c], *aacC3*[c], *aadA1*[c], *aadA2*[c], *aadA5*[c], *aadA6/aadA10*[c], *aadA8*[c], *aadA11*[c], *aadA12*[c], *aadA13*[c], *aadA23*[c], *aadB*[b,c], *aphA1*[c], *aphA*[c], *aphA-6*[c], *orf5/aadB*[b], *strA*[b,c], *strB*[c]	378	1075
β-Lactams	*ampC*[b], bla_{CMY-1}[b], bla_{CMY-8}[b], bla_{CMY-9}[c], bla_{CMY-10}[b,c], bla_{CMY-11}[b], bla_{FOX-1}[b], bla_{FOX-2}[b], bla_{FOX-3}[b], bla_{FOX-5}[b], bla_{FOX-6}[b], bla_{GES-2}[b], bla_{GES-3}[b,c], bla_{GES-4}[b], bla_{IBC-1}[b], bla_{IBC-2}[b], bla_{MOX-1}[b], bla_{TEM-1}[c], $bla_{TEM-105}$[b], $bla_{TEM-107}$[b], bla_{TLA-2}[c], bla_{OXA-1}[b,c], bla_{OXA-2}[b,c], bla_{OXA-5}[c], bla_{OXA-10}[c], bla_{OXA-56}[c], bla_{OXA-58}[c], bla_{PER-1}[c], bla_{SHV-34}[c]	255	277
Chloramphenicol	*cat*[b], *cat2*[c], *catA-1*[c], *catB2*[b,c], *catB4*[c], *cmlA1*[b,c], *cmlA5*[c]	31	59
Macrolides	*erm*[b], *erm(B)*[c], *erm(GT)*[c], *mel*[c], *mph2*[c], *mph*(A)[b,c]	17	1437
Quinolones	*qnrS2*[c]	–	40
Rifampicine	*arr2*[c]	–	3
Sulfonamides/trimethoprim	*dfr1*[b], *dfrV*[c], *dfrVII*[c], *dfrVIII*[c], *dhfr*[b], *dhfr1*[c], *sul1*[b,c], *sul2*[b,c]	113	238
Tetracycline	*tetA*(A)[b,c], *tetA*(B)[b], *tetA*(C)[b,c], *tetA*(D)[c], *tetA*(E)[b,c], *tetA*(X)[b,c], *tet*(39)[c]	75	534
Multidrug	*acrB*[b,c], *mexB*[b,c], *orf11*[c]	5	54
Quaternary ammonium compounds (QACs)	*qacEΔ1*[bc], *qacF*[c], *qacH*[c]	44	70

[a] Accession numbers for resistance gene hits are given in supplementary Table S1.
[b] Resistance genes identified by BLASTn analyses against the Mvir-DB database.
[c] Resistance genes identified by BLASTn analyses against the ARG-DB database.

nucleotide sequences of 47 different resistance genes present in MvirDB. Table 4 summarises the identified resistance genes with respect to the antimicrobial drug class to which corresponding reference genes confer resistance. Different resistance genes known to play a role in resistance to β-lactam antibiotics (20), tetracyclines (9), aminoglycosides (5), chloramphenicol (3), macrolides (2), sulfonamide and trimethoprim (5) and quaternary ammonium compounds used as disinfectants (1) could be identified. In addition, genes specifying multidrug-resistance efflux systems are listed in Table 4.

In a second approach, an *A*ntibiotic *R*esistance *G*ene *D*ata*b*ase (ARG-DB) was set up manually by extracting nucleotide and protein sequence data of antibiotic resistance genes from GenBank. Only those resistance genes were uploaded into ARG-DB for which experimental evidence exists that the respective gene confers resistance to (a) specific antimicrobial compound(s). ARG-DB contains 192 different entries including 26 aminoglycoside, 48 β-lactam, 20 chloramphenicol, 1 methicillin, 20 macrolide, 4 quinolone, 1 rifampicin, 36 tetracycline, 19 trimethoprim and sulfonamide and 8 quaternary ammonium compound resistance genes, as well as 9 multidrug-resistance efflux genes (ARG-DB will be described in more detail elsewhere). In total, 3787 reads could be assigned to 65 different resistance genes (see Table 4). As expected, many resistance genes detected in the WWTP plasmid metagenome, namely *aacA4*, *aacA29b*, *aacC1*, *aadA1*, *aadA2*, *aadA5*, *aadA8*, *aadB*, *aphA1*, *aphA* (on the IncP-1α plasmid pTB11), *strA*, *strB*, bla_{TEM-1}, bla_{TLA-2}, bla_{OXA-1}, bla_{OXA-2}, bla_{OXA-10}, *catA*-1, *catB2*, *cmlA1*, *cmlA5*, *mel*, *mph2*, *qnrS2*, *dfrA1*, *dfrV*, *dfrVII*, *sul1*, *sul2*, *tetA*(C), *qacEΔ1*, *qacF* and *qacH* were previously identified on completely sequenced plasmids originating from WWTP bacteria (Bönemann et al., 2006; Heuer et al., 2004; Schlüter et al., 2003, 2005, 2007c; Szczepanowski et al., 2004; Tennstedt et al., 2005) and on plasmid-borne class 1 integrons (Tennstedt et al., 2003) which were also isolated from WWTP bacteria. Other resistance genes found by this approach, like for example bla_{CMY-10} (Lee et al., 2004), bla_{CMY-11} (Kim et al., 2004), bla_{GES-3} and bla_{GES-4} (Vourli et al., 2004), bla_{OXA-58} (Poirel et al., 2005), which confer resistance to β-lactam antibiotics and the tetracycline resistance gene *tet*(B) (Lancashire et al., 2005), were recently isolated from clinical pathogenic bacteria which caused infections or from the swine pathogen *Haemophilus parasuis*. These results indicate that WWTP bacteria play a role for the dissemination of clinically relevant antibiotic resistance determinants. Evidence for the statement that WWTP bacteria serve as reservoir for very different and clinically relevant resistance genes to antimicrobial drugs was provided by analysing the coding capacity of the plasmid metagenome obtained from these bacteria.

3.5. Overview on assembled contigs of the WWTP plasmid metagenome

Nucleotide sequence reads were assembled by means of the Newbler Assembler Software (Roche, Penzberg, Germany) which resulted in 605 contigs with a size larger than 500 bases. These contigs nearly include 223,000 reads; the longest contig is 4.5 kb in size and the average contig length is approximately 1 kb. Annotation results for contigs larger than 2.5 kb are shown in Table 5. Contigs mainly contain plasmid replication initiation, stabilisation and mobility genes. Comparison of mapping results (see Section 3.3) with contig annotations revealed that plasmid genes covered densely by sequence reads were also identified on contigs. This is for example very evident for genes present on the pAsa reference plasmids from *Aeromonas salmonicida* subsp. *salmonicida* (see Table 5). It is obvious that plasmid modules encoding survival functions are conserved since natural selection ensures integrity, completeness and functional preservation of these modules. Nature and type of plasmid backbone modules present in the sequenced sample are described in more detail in an accompanying paper (Schlüter et al., this issue). As outlined above, many reads of the sequenced plasmid metagenome represent genes originating from transposable elements, so that it is not astonishing that several contigs contain transposons, insertion sequences, integrons and remnants of these elements. For example, the 4133-bp contig 13,830 encodes a complete transposase (TnpA, 989 amino acids) that is 93% identical to TnpA of the pB4 chromate resistance Transposon Tn*5719* and a resolvase (TnpR, 187 amino acids, 91% identity to pB4 TnpR) (Tauch et al., 2003). Transposable elements are important for evolution and adaptive rearrangements of accessory plasmid modules. The largest contig is 4519 bp in length and encodes a putative transposase of the Tn*3*-family that is related (88% identity) to a corresponding transposase of *Marinobacter aquaeolei* (YP_958534) and a DNA-invertase that is 94% identical to a site-specific recombinase of the *Shewanella oneidensis* megaplasmid pMR-1 (NP_720358). Assembly of plasmid genomes from metagenomic nucleotide sequence reads seems to be difficult which is most probably due to the modular composition of plasmids. Plasmids evolved from component parts representing plasmid survival functions (Thomas, 2000) which means that

Table 5
Gene content of large contigs (>2.5 kb) assembled from the obtained WWTP plasmid metagenome sequence reads

Contig no.	Size (bp)	Gene content, similarity to reference gene product (%)	Gene product	Pfam/COG assignment[a]	Organism	Encoded on plasmid	Accession no.[b]
00149	4143	plu0499 (40%)	–	pfam07395	*Photorhabdus luminescens* subsp. *laumondii* TTO1	–	NP_927852
		CKO_pCKO3p0614 (49%)	Hypothetical protein	pfam05713	*Citrobacter koseri* ATCC BAA-895	pCKO390	YP_001456624
		mobC (69%)	Bacterial mobilisation protein	–	*Escherichia coli*	pColD-157	CAA71437
		mobA (60%)	Relaxase	pfam03432	*Citrobacter freundii*	ColA	S04790
		rop (78%)	Rom-like protein; RNA I modulator	pfam01815	*Escherichia coli*	pEC157	AAL27615
		plu0449 (48%)	–	pfam07395	*Photorhabdus luminescens* subsp. *laumondii* TTO1	–	NP_927852
01082	3749	*rep* (75%)	Replication primase	pfam03090, pfam08780	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa3	NP_861563
		orf36 (71%)	Hypothetical protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsal2	NP_710168
		mobC (63%)	Mobilisation protein	pfam05713	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa3	NP_861565
		mobA (52%)	Mobilisation relaxase	pfam03432	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa3	NP_861566
12352	2534	*orf* (56%)	Phage integrase family protein	–	*Pseudomonas putida* GB-1	–	ZP_01715677
		orf36-like (77%)	Hypothetical protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsal2	NP_710168
13582	3216	GSU2773[c] (45%)	Hypothetical protein	COG0616	*Geobacter sulfurreducens* PCA	–	NP_953817
		mobA (37%)	Mobilisation relaxase	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsal2	NP_710170
		mobD (52%)	Mobilisation protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa1	NP_861559
		mobB (34%)	Mobilisation protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsal2	NP_710171
13678	2663	*orf1*-like (75%)	Hypothetical protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa3	NP_861564
		mobC (36%)	Auxiliary mobilisation protein C	–	–	pGNB2	ABE98193
		orf1 (42%)	Relaxase	–	–	pGNB2	ABE98194
13830	4133	*tnpA* (93%)	Transposase	pfam01526, COG4644	–	pB4	NP_598138
		tnpR (91%)	TnpR protein	pfam00239, COG1961	–	pB4	NP_598137
14061	2586	*repA* (59%)	Replication initiation protein A	pfam03090	*Klebsiella* sp. KCL-2	pMGD2	NP_620615
		mobA (64%)	Mobilisation relaxase	pfam03432	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa1	NP_861557
		mobD (60%)	Mobilisation protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsa3	NP_861568
		mobB (73%)	Putative MobB protein	–	*Aeromonas salmonicida* subsp. *salmonicida* A449	pAsal2	NP_710171
14154	2872	PputGB1DRAFT_5429 (57%)	Phage integrase	pfam00589	*Pseudomonas putida* GB-1	–	ZP_01715677
14444	2527	*vktA* (100%)	Catalase	pfam00199, COG0753	*Vibrio rumoiensis* S-1	–	BAB12412
14455	2780	*repA* (44%)	Putative RepA protein	pfam01446	*Neisseria flavescens*	pMIDG2830	AAO42608

Table 5 (*Continued*)

Contig no.	Size (bp)	Gene content, similarity to reference gene product (%)	Gene product	Pfam/COG assignment[a]	Organism	Encoded on plasmid	Accession no.[b]
14469	4421	V12G01_21273 (30%)	Hypothetical protein	–	*Vibrio alginolyticus* 12G01	–	ZP_01260508
		V12G01_21268 (42%)	Hypothetical protein	–	*Vibrio alginolyticus* 12G01	–	ZP_01260507
		parD-like (82%)	Hypothetical protein	pfam01402, COG3905	–	pRSB101	YP_133838
		parE-like (89%)	Hypothetical protein	pfam05016, COG3668	–	pRSB101	YP_133839
		finO (partial) (26%)	–	pfam05286	*Soladis glossinidius*	pSG2	YP_257093
		*AHA*_2030 (31%)	Hypothetical protein	–	*Aeromonas hydrophila* subsp. *hydrophila* ATCC7966	–	YP_856560
15082	3250	*tnpA* (100%)	Transposase	pfam01526, COG4644	*Corynebacterium striatum* M82B	pTP10	NP_862256
		tnpR (100%)	Resolvase	pfam00239, COG1961	*Pseudomonas aeruginosa*	RPL11	ABK33455
		strA[c] (100%)	Aminoglycoside-3″-phosphotransferase	pfam01636, COG3231	–	pRSB107	CAH64747
		strB (100%)	Aminoglycoside-6-phosphotransferase	pfam04655, COG3570	–	pB10	NP_858030
15413	3840	*mobA*[c] (98%)	Mobilisation protein	–	–	pRSB105	ABI20469
		orf8 (100%)	Hypothetical protein	pfam04365, COG2929	–	pRSB105	ABI20470
		orf9 (98%)	Hypothetical protein	–	–	pRSB105	ABI20471
		orf10 (100%)	Hypothetical protein	–	–	pRSB105	ABI20472
		orf14 (98%)	Hypothetical protein	COG3041	–	pRSB105	ABI20476
		pin (100%)	Invertase/recombinase-like protein	pfam00239, COG1961	–	pRSB101	YP_133858
15792	3972	*pin* (100%)	Resolvase/site-specific recombinase	pfam00239, COG1961	–	pRSB105	ABI20507
		mph (100%)	Macrolide 2′-phosphotransferase	pfam01636, COG3173	–	pRSB105	ABI20452
		mel (100%)	Efflux protein	pfam00005, COG0488	–	pRSB105	ABI20451
16358	4519	*tnpR* (94%)	Invertase	pfam00239, pfam02796, COG1961, COG2452	*Shewanella oneidensis* MR-1	megaplasmid pMR-1	NP_720358
		orf6[c] (83%)	Hypothetical protein	pfam01934, COG2361, COG01669	*Delftia acidovorans*	pUO1	BAC82011
		Maqu_1257 (88%)	Transposase	pfam01526, COG4644	*Marinobacter aquaeolei* VT8	–	YP_958534
17659	2543	*misL* (40%)	Putative autotransporter	pfam03797, COG3468	*Salmonella typhimurium* LT2	–	NP_462656
		Pnap_3707′ (42%)	Glycosyl transferase, family 2	pfam00535, COG0463	*Polaromonas naphthalenivorans* CJ2	–	YP_983924

[a] Conserved domains were identified by protein blast.

[b] Protein accession numbers correspond to the NCBI protein database.

[c] Genes containing frameshifts in comparison to a reference gene.

different survival modules can be integral parts of different plasmids. In addition, comparative analysis of completely sequenced plasmids has shown that accessory plasmid segments are highly mosaic. Certain parts of these regions were acquired from different sources and incorporated into the plasmid replicon by illegitimate, site-specific or homologous recombination. Consequently, accessory plasmid regions are highly diverse and potentially can be translocated into other plasmid molecules with the consequence that distinct transposable elements and adjacent DNA segments can occur on different plasmids. These circumstances complicate or in some cases even preclude correct assembly of plasmid genomes from metagenomic data. Artificial assembly of contigs cannot be excluded. Interpretation of contig sequences is also complicated by the existence of frameshifts within several genes (see Table 5). It remains unclear whether these frameshifts represent sequencing errors or different gene variants within the plasmid metagenome sample.

3.6. Concluding remarks

The ultrafast 454-sequencing technology proved to be a powerful tool to generate plasmid metagenome nucleotide sequence data of WWTP bacteria. Obtained data allow for insights into the genetic traits encoded on plasmids residing in WWTP bacteria that show reduced susceptibility to antimicrobial drugs. Especially, the identification of sequence signatures specific for antibiotic resistance genes indicate that WWTP bacteria are an important reservoir for clinically relevant resistance determinants and might contribute to further rapid dissemination of antibiotic resistances.

Saturation of the sequencing effort that has been carried out was not achieved since many detected genes are only covered by very few reads so that it would be appropriate to perform additional sequencing-runs on the sampled plasmid preparation. The composition of the analysed plasmid pool was biased due to limitation of the sequencing approach to WWTP bacteria that showed reduced susceptibility to certain antimicrobial drugs. Hence, the next obvious project should aim towards the establishment of the complete plasmid metagenome of non-selected and non-cultured WWTP bacteria. Likewise, it would be informative to compare plasmid metagenomes obtained from WWTPs that receive effluents from hospitals and those that are not connected to any medical facilities.

Acknowledgement

Sequencing of the WWTP plasmid metagenome was financed by the Landesamt für Natur, Umwelt und Verbraucherschutz des Landes Nordrhein-Westfalen (Düsseldorf, Germany).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jbiotec.2008.03.020.

References

Allmeier, H., Cresnar, B., Greck, M., Schmitt, R., 1992. Complete nucleotide sequence of Tn*1721*: gene organization and a novel gene product with features of a chemotaxis protein. Gene 111, 11–20.

Bairoch, A., Boeckmann, B., 1993. The SWISS-PROT protein sequence data bank, recent developments. Nucleic Acids Res. 21, 3093–3096.

Bönemann, G., Stiens, M., Pühler, A., Schlüter, A., 2006. Mobilizable IncQ-related plasmid carrying a new quinolone resistance gene, *qnrS2*, isolated from the bacterial community of a wastewater treatment plant. Antimicrob. Agents Chemother. 50, 3075–3080.

Boyd, J., Williams, J., Curtis, B., Kozera, C., Singh, R., Reith, M., 2003. Three small, cryptic plasmids from *Aeromonas salmonicida* subsp. *salmonicida* A449. Plasmid 50, 131–144.

Boyd, D.A., Tyler, S., Christianson, S., McGeer, A., Muller, M.P., Willey, B.M., Bryce, E., Gardam, M., Nordmann, P., Mulvey, M.R., 2004. Complete nucleotide sequence of a 92-kilobase plasmid harboring the CTX-M-15 extended-spectrum beta-lactamase involved in an outbreak in long-term-care facilities in Toronto, Canada. Antimicrob. Agents Chemother. 48, 3758–3764.

Chan, P.T., Ohmori, H., Tomizawa, J., Lebowitz, J., 1985. Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. 260, 8925–8935.

Chiu, C.H., Tang, P., Chu, C., Hu, S., Bao, Q., Yu, J., Chou, Y.Y., Wang, H.S., Lee, Y.S., 2005. The genome sequence of *Salmonella enterica* serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. Nucleic Acids Res. 33, 1690–1698.

Davison, J., 1999. Genetic exchange between bacteria in the environment. Plasmid 42, 73–91.

Edwards, R.A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D.M., Saar, M.O., Alexander, S., Alexander Jr., E.C., Rohwer, F., 2006. Using pyrosequencing to shed light on deep mine microbial ecology. BMC Genomics 7, 57.

Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. 3, 722–732.

González-Zorn, B., Teshager, T., Casas, M., Porrero, M.C., Moreno, M.A., Courvalin, P., Domínguez, L., 2005. *armA* and aminoglycoside resistance in *Escherichia coli*. Emerg. Infect. Dis. 11, 954–956.

Haines, A.S., Jones, K., Cheung, M., Thomas, C.M., 2005. The IncP-6 plasmid Rms149 consists of a small mobilizable backbone with multiple large insertions. J. Bacteriol. 187, 4728–4738.

Heuer, H., Szczepanowski, R., Schneiker, S., Pühler, A., Top, E.M., Schlüter, A., 2004. The complete sequences of plasmids pB2 and pB3 provide evidence for a recent ancestor of the IncP-1β group without any accessory genes. Microbiology 150, 3591–3599.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., Hirakawa, M., 2006. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 34, D354–D357.

Kim, J.Y., Song, J.S., Bak, S.H., Cho, Y.E., Kim, D.W., Jeong, S.H., Park, Y.M., Lee, K.J., Lee, S.H., 2004. Dissemination of *Escherichia coli* producing AmpC-type beta-lactamase (CMY-11) in Korea. Int. J. Antimicrob. Agents 24, 320–326.

Krause, L., Diaz, N.N., Bartels, D., Edwards, R.A., Pühler, A., Rohwer, F., Meyer, F., Stoye, J., 2006. Finding novel genes in bacterial communities isolated from the environment. Bioinformatics 22, e281–e289.

Lancashire, J.F., Terry, T.D., Blackall, P.J., Jennings, M.P., 2005. Plasmid-encoded Tet B tetracycline resistance in *Haemophilus parasuis*. Antimicrob. Agents Chemother. 49, 1927–1931.

Lee, J.H., Jung, H.I., Jung, J.H., Park, J.S., Ahn, J.B., Jeong, S.H., Jeong, B.C., Lee, S.H., 2004. Dissemination of transferable AmpC-type beta-lactamase (CMY-10) in a Korean hospital. Microb. Drug Resist. 10, 224–230.

Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Buillard, V., Cerutti, L., Copley, R., Courcelle, E., Das, U., Daugherty, L., Dibley, M., Finn, R., Fleischmann, W., Gough, J., Haft, D., Hulo, N., Hunter, S., Kahn, D., Kanapin, A., Kejariwal, A., Labarga, A., Langendijk-Genevaux, P.S., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Orengo, C., Petryszak, R., Selengut, J.D., Sigrist, C.J., Thomas, P.D., Valentin, F., Wilson, D., Wu, C.H., Yeats, C., 2007. New developments in the InterPro database. Nucleic Acids Res. 35, D224–D228.

Murata, T., Ohnishi, M., Ara, T., Kaneko, J., Han, C.G., Li, Y.F., Takashima, K., Nojima, H., Nakayama, K., Kaji, A., Kamio, Y., Miki, T., Mori, H., Ohtsubo, E., Terawaki, Y., Hayashi, T., 2002. Complete nucleotide sequence of plasmid Rts1: implications for evolution of large plasmid genomes. J. Bacteriol. 184, 3194–3202.

Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrell, B.G., 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature 413, 848–852.

Poirel, L., Marqué, S., Héritier, C., Segonds, C., Chabanon, G., Nordmann, P., 2005. OXA-58, a novel class D β-lactamase involved in resistance to carbapenems in *Acinetobacter baumannii*. Antimicrob. Agents Chemother. 49, 202–208.

Rhodes, G., Parkhill, J., Bird, C., Ambrose, K., Jones, M.C., Huys, G., Swings, J., Pickup, R.W., 2004. Complete nucleotide sequence of the conjugative tetracycline resistance plasmid pFBAOT6, a member of a group of IncU plasmids with global ubiquity. Appl. Environ. Microbiol. 70, 7497–7510.

Riley, M.A., Pinou, T., Wertz, J.E., Tan, Y., Valletta, C.M., 2001. Molecular characterization of the klebicin B plasmid of *Klebsiella pneumoniae*. Plasmid 45, 209–221.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Labolatory Manual. Cold Spring Harbor Laboratory Press, New York.

Scaria, J., Chandramouli, U., Verma, S.K., 2005. Antibiotic resistance genes online (ARGO): a database on vancomycin and beta-lactam resistance genes. Bioinformation 1, 5–7.

Schlüter, A., Heuer, H., Szczepanowski, R., Forney, L.J., Thomas, C.M., Pühler, A., Top, E.M., 2003. The 64 508 bp IncP-1β antibiotic multiresistance plasmid pB10 isolated from a waste-water treatment plant provides evidence for recombination between members of different branches of the IncP-1β group. Microbiology 149, 3139–3153.

Schlüter, A., Heuer, H., Szczepanowski, R., Poler, S.M., Schneiker, S., Pühler, A., Top, E.M., 2005. Plasmid pB8 is closely related to the prototype IncP-1β plasmid R751

but transfers poorly to *Escherichia coli* and carries a new transposon encoding a small multidrug resistance efflux protein. Plasmid 54, 135–148.

Schlüter, A., Szczepanowski, R., Pühler, A., Top, E.M., 2007a. Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. FEMS Microbiol. Rev. 31, 449–477.

Schlüter, A., Szczepanowski, R., Kurz, N., Schneiker, S., Krahn, I., Pühler, A., 2007b. Erythromycin resistance-conferring plasmid pRSB105, isolated from a sewage treatment plant, harbors a new macrolide resistance determinant, an integron-containing *Tn402*-like element, and a large region of unknown function. Appl. Environ. Microbiol. 73, 1952–1960.

Schlüter, A., Krahn, I., Kollin, F., Bönemann, G., Stiens, M., Szczepanowski, R., Schneiker, S., Pühler, A., 2007c. IncP-1β plasmid pGNB1 isolated from a bacterial community from a wastewater treatment plant mediates decolorization of triphenylmethane dyes. Appl. Environ. Microbiol. 73, 6345–6350.

Schlüter, A., Krause, L., Szczepanowski, R., Goesmann, A., Pühler, A., 2008. Genetic diversity and composition of a plasmid metagenome from a wastewater treatment plant. Journal of Biotechnology 136, 65–72.

Segal, H., Elisha, B.G., 1999. Characterization of the acinetobacter plasmid, pRAY, and the identification of regulatory sequences upstream of an *aadB* gene cassette on this plasmid. Plasmid 42, 60–66.

Szczepanowski, R., Krahn, I., Linke, B., Goesmann, A., Pühler, A., Schlüter, A., 2004. Antibiotic multiresistance plasmid pRSB101 isolated from a wastewater treatment plant is related to plasmids residing in phytopathogenic bacteria and carries eight different resistance determinants including a multidrug transport system. Microbiology 150, 3613–3630.

Szczepanowski, R., Braun, S., Riedel, V., Schneiker, S., Krahn, I., Pühler, A., Schlüter, A., 2005. The 120, 592 bp IncF plasmid pRSB107 isolated from a sewage-treatment plant encodes nine different antibiotic-resistance determinants, two iron-acquisition systems and other putative virulence-associated functions. Microbiology 151, 1095–1111.

Szczepanowski, R., Krahn, I., Bohn, N., Pühler, A., Schlüter, A., 2007. Novel macrolide resistance module carried by the IncP-1β resistance plasmid pRSB111, isolated from a wastewater treatment plant. Antimicrob. Agents Chemother. 51, 673–678.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4, 41.

Tauch, A., Schlüter, A., Bischoff, N., Goesmann, A., Meyer, F., Pühler, A., 2003. The 79,370-bp conjugative plasmid pB4 consists of an IncP-1β backbone loaded with a chromate resistance transposon, the *strA-strB* streptomycin resistance gene pair, the oxacillinase gene bla_{NPS-1}, and a tripartite antibiotic efflux system of the resistance-nodulation-division family. Mol. Genet. Genomics 268, 570–584.

Tennstedt, T., Szczepanowski, R., Braun, S., Pühler, A., Schlüter, A., 2003. Occurrence of integron-associated resistance gene cassettes located on antibiotic resistance plasmids isolated from a wastewater treatment plant. FEMS Microbiol. Ecol. 45, 239–252.

Tennstedt, T., Szczepanowski, R., Krahn, I., Pühler, A., Schlüter, A., 2005. Sequence of the 68,869 bp IncP-1α plasmid pTB11 from a waste-water treatment plant reveals a highly conserved backbone, a Tn*402*-like integron and other transposable elements. Plasmid 53, 218–238.

Thomas, C.M., 2000. Paradigms of plasmid organization. Mol. Microbiol. 37, 485–491.

Thomas, C.M., Nielsen, K.M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 3, 711–721.

Vourli, S., Giakkoupi, P., Miriagou, V., Tzelepi, E., Vatopoulos, A.C., Tzouvelekis, L.S., 2004. Novel GES/IBC extended-spectrum β-lactamase variants with carbapenemase activity in clinical enterobacteria. FEMS Microbiol. Lett. 234, 209–213.

Zhou, C.E., Smith, J., Lam, M., Zemla, A., Dyer, M.D., Slezak, T., 2007. MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res. 35, D391–D394.

Zienkiewicz, M., Kern-Zdanowicz, I., Golebiewski, M., Zylinska, J., Mieczkowski, P., Gniadkowski, M., Bardowski, J., Ceglowski, P., 2007. Mosaic structure of p1658/97, a 125-kilobase plasmid harboring an active amplicon with the extended-spectrum beta-lactamase gene *bla*SHV-5. Antimicrob. Agents Chemother. 51, 1164–1171.