Environ Sci Pollut Res
DOI 10.1007/s11356-017-0021-3

RESEARCH ARTICLE

Monitoring of genetically modified *Escherichia coli* in laboratory wastewater

Fabienne Wichmann[1] · **Ines Wyrsch**[1] · **Jörg Frank**[2] · **Matthias Müller**[1] · **Nicole Bertschi**[1] · **Peter Brodmann**[1] · **Claudia Bagutti**[1]

Received: 16 February 2017 / Accepted: 23 August 2017

Abstract Containment of genetically modified (GM) microorganisms such as *Escherichia coli* is a legal requirement to protect the environment from an unintended release and to avoid horizontal gene transfer (HGT) of recombinant DNA to native bacteria. In this study, we sampled the laboratory wastewater (LWW) at a large Swiss university from three sources over 2 years and cultured ampicillin-resistant, presumptive GM *E. coli*. From a total of 285 samples, 127 contained presumptive GM *E. coli* (45%) at a mean concentration of 2.8×10^2 CFU/ml. Plasmid DNA of 11 unique clones was partially or entirely sequenced. All consisted of cloning vectors harboring research-specific inserts. To estimate the chance of HGT between GM *E. coli* and native bacteria in LWW, we identified taxa representative for the bacterial community in LWW using 16S rRNA amplicon sequencing and measured conjugation frequencies of *E. coli* with five LWW isolates. At optimal conjugation conditions, frequencies were between 3.4×10^{-3} and 2.4×10^{-5}. Given the absence of transferable broad-host range plasmids and suboptimal conjugation conditions in the LWW system, we conclude that the chance of HGT is relatively low. Still, this study shows that the implementation of robust containment measures is key to avoid the escape of GM microorganisms.

Responsible editor: Gerald Thouand

✉ Claudia Bagutti
claudia.bagutti@bs.ch

[1] State Laboratory Basel-City, Biosafety Laboratory, Basel, Switzerland

[2] Department for Safety, Security and Environment, University of Zurich, Zurich, Switzerland

Keywords Genetically modified · *E. coli* · Wastewater · Laboratory · Release · Containment · Horizontal gene transfer

Introduction

Genetically modified (GM) microorganisms derived from e.g., *Escherichia coli* are frequently used in laboratories for research or for the production of pharmaceuticals (reviewed by Way et al. (2014)). GM microorganisms usually harbor recombinant plasmid DNA, to express a desired product or to study a particular DNA segment. Although recombinant DNA offers immense benefits for research and medicine, it also poses risks for the environment. These concerns were already recognized in 1974 by a group of forward-looking researchers headed by Paul Berg. They proposed that scientists should "voluntarily defer" to introduce antibiotic resistance genes into bacteria that naturally do not carry them and equally defer to construct plasmids harboring genes conferring resistance to clinically relevant antibiotics (Berg et al. 1974). After a release into the environment, recombinant DNA may be transferred to the native microbiome through horizontal gene transfer (HGT) either by conjugation or by transformation. Whether these processes occur, depends on the persistence of the GM microorganisms and their recombinant DNA in the environment (reviewed in Davison (1999); Lorenz and Wackernagel (1994)). According to previous studies, GM *E. coli* survive significantly longer in filtered (Amy and Hiatt 1989; Chao and Feng 1990) or sterilized water (Awong et al. 1990; Bogosian et al. 1996; Brettar and Hofle 1992), i.e., in the absence of potentially competing microorganisms or grazing eukaryotes. Still, survival times of up to 20 days were observed in a membrane chamber diffused with untreated lake water (Awong et al. 1990). In contrast to loosely drifting *E. coli*, cells associated with particles or integrated

Published online: 01 September 2017

into a biofilm matrix exhibited a resilience of up to 20 weeks (Abberton et al. 2016; Banning et al. 2003; Brettar and Hofle 1992), likely resulting from the protection against nucleases, toxins, phagocytic organisms or shear forces and an increased availability of nutrients. Furthermore, biofilm matrices are significant for the persistence of extracellular DNA and plasmids, which bacteria may take up through transformation (Lorenz and Wackernagel 1994; Mao et al. 2014). Mao et al. showed that extracellular DNA encoding antibiotic resistance genes is assimilated by native bacteria (Mao et al. 2014). Nevertheless, conjugation is the most relevant exchange mechanism between introduced and native bacteria and has been investigated intensively in the context of dissemination of antibiotic resistance genes (Davison 1999; von Wintersdorff et al. 2016).

The laboratory wastewater (LWW) system represents an important escape route for GM microorganisms harboring recombinant DNA. It is noteworthy that already in the 1980s mobilization of plasmids such as pBR325 from GM *E.coli* to LWW bacteria was demonstrated even under conditions not supporting bacterial growth (Gealt et al. 1985; McPherson and Gealt 1986; Sagik et al. 1981). Considering the potential consequences, it is imperative to limit the escape of GM microorganisms and to protect the environment. These concerns are met by national containment regulations.

In Switzerland, most activities handling GM *E. coli* belong to class 1. The Swiss "Containment Ordinance" states that in the case of activities in class 1 any escape by these organisms "[…]should be limited to the extent that human beings, animals, and the environment as well as biological diversity and its sustainable use cannot be endangered[…]" (The Federal Authorities of the Swiss Confederation 2012). Furthermore, a zero tolerance applies to the release of waste knowingly contaminated with living GM microorganisms. To accomplish this objective, the scientific communities of both academia and industry have to implement robust containment measures. Even though containment efforts are taken within common laboratory practices, it is likely that some GM microorganisms escape from laboratories into the environment. In a blind questionnaire, users have reported that carelessness, minor accidents, or incorrect procedures are the most frequent causes (Claudia Ruprecht, WWEA Canton Zurich, personal communication). Several possible scenarios for an escape of GM microorganisms via the LWW system were stated, including the lack of inactivation of residual culture fluid before washing flasks, spillage, broken culture flasks, or insufficient inactivation using chemicals beyond the expiration date, at an insufficiently effective concentration or by not inactivating long enough. However, the extent to which GM microorganisms escape the laboratory setting via LWW has not been elucidated to date.

Therefore, this study aimed at (i) testing whether GM microorganisms are released from the containment and can be detected in LWW, (ii) isolating plasmids of presumptive GM *E. coli* and identifying their vectors and inserts, (iii) examining the bacterial community of LWW by using 16S rRNA gene amplicon sequencing, and (iv) using a selection of representative native bacterial isolates from LWW to study transfer of a broad-host range plasmid through HGT.

Materials and methods

Laboratory wastewater sampling

From July 2012 until June 2014, LWW was sampled from a large Swiss university campus twice per day (at 10:00 and 12:00) approximately every 14 days. Different days of the week were sampled in order to avoid a weekday-specific bias. Sampling took place five-times on Monday and Wednesday, respectively, 12-times on Tuesday, 24-times on Thursday, and once on Friday (total number of sampling days = 47). Samples were collected from three LWW sources: downpipes of two different buildings harboring mainly laboratories (building downpipe B1, building downpipe B2) and the mixing tank M (volume = 16 m^3) where the LWW from the entire campus is collected. The LWW at all sites is devoid of fecal wastewater. B1, B2, and M were chosen because they represent diverse sources with respect to location on the campus (B1 and B2 being situated in two different buildings), flow rate, and dilution factor. To select ampicillin-resistant *E. coli*, LWW samples (200 μl) were spread immediately after collection onto chromogenic Tryptone Bile Glucuronic (TBX; Oxoid, Pratteln, Switzerland) agar supplemented with ampicillin (100 μg/ml) and the plates were incubated overnight at 37 °C. In a pre-study, the specificity of the TBX agar to select *E. coli* was confirmed by MALDI-TOF MS (Matrix-assisted laser desorption/ionization time of flight mass spectrometry) (Axima™ Confidence, Shimadzu-Biotech, Reinach, Switzerland) using the SARAMIS™ Database ((Welker and Moore 2011), Spectral Archive And Microbial Identification System, AnagnosTec, Potsdam-Golm, Germany), and the PAPMID Database ((Toh et al. 2015), Mabritec SA, Riehen, Switzerland) according to the methods and parameters described by Ziegler et al. (2015). Alpha-cyano-4 hydroxycinnamic acid was used as a matrix and 50 protein mass fingerprints were averaged and processed of two duplicate spots. The blue colonies corresponding to ampicillin-resistant, presumptive GM *E. coli* were enumerated by hand (lower detection limit: 5 CFU/ml; upper detection limit: 1500 CFU/ml). Between February 2014 and July 2014, three blue colonies from sampling source B1, 18 blue colonies from sampling source B2, and 24

blue colonies from sampling source M were selected for further analyses.

Characterization of GM *E. coli* by multiplex PCR

In order to identify the most frequently used laboratory *E. coli* genotypes (i.e., K-12, B, C, and W), the isolated *E. coli* were analyzed by multiplex PCR according to Bauer et al. (2007) and Kuhnert et al. (1995) with minor modifications. Single colonies were picked and resuspended in a 20 μl reaction volume containing the following: 1× Hotstar PCR buffer (Qiagen, Basel, Switzerland), 50 mM $MgCl_2$, 10 mM (each) dNTP, 20 μM (each) of primer pairs 22-4, Tn7, P27, K12, and control ECPAL (BioTez Berlin-Buch GmbH, Berlin, Germany; 22-4, Tn7, P27 (Bauer et al. 2007); K12, ECPAL (Kuhnert et al. 1995)) and 1.6 U Hotstar Taq DNA Polymerase (Qiagen). The cycling parameters were as follows: Initial denaturation for 15 min at 94 °C; 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min; followed by a final elongation step at 72 °C for 7 min. PCR fragments were analyzed on a 1% TBE agarose gel stained with ethidium bromide. Fragment sizes were confirmed using a 100 bp ladder (New England BioLabs, Allschwil, Switzerland).

Characterization of plasmid DNA

E.coli isolates were incubated overnight at 37 °C and 200 rpm in Luria Bertani (LB) broth (Oxoid) containing ampicillin (100 μg/ml) and plasmids were extracted using the QIAprep Spin miniprep kit (Qiagen) following the manufacturer's instructions. Plasmid DNA (200 ng) was visually inspected by electrophoresis on a 1% TBE agarose gel. A multiplex PCR approach, using standard primers T7 (T7-F TAATACGACTCACTATAGGG, T7-R TGCTAGTTATTGCTCAGCGG), M13 (M13-F GTAAAACGACGGCCAGTG, M13-R GGAAACAGCTATGACCAT), T3 (T3-R CAATTAACCCTCACTAAAGG), and SP6 (SP6-R TACGATTTAGGTGACACTATAG) was performed in a 20 μl reaction volume containing the following: 1× Hotstar Taq PCR buffer (Qiagen); 50 mM $MgCl_2$ (50 mM); 10 mM (each) dNTPs; 20 μM (each) of primers T7, M13, SP6, and T3 (Microsynth, Balgach, Switzerland); 1.6 U Hotstar Taq DNA Polymerase (Qiagen); and 2 ng of template plasmid DNA. PCR amplifications were performed with the following parameters: Initial denaturation for 15 min at 94 °C; 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 C for 1 min; followed by a final elongation step at 72 °C for 7 min. PCR fragments were analyzed on a 1% TBE agarose gel. Plasmids, from which a specific PCR fragment had been amplified, were sequenced by Sanger sequencing (Microsynth) with standard primers T7-F and M13-R.

Illumina sequencing of entire plasmids

Visually unique plasmids, for which no specific PCR product had been obtained, were entirely sequenced by Illumina sequencing. The *E.coli* isolates were grown overnight at 37 °C and 200 rpm in LB broth containing ampicillin (100 μg/ml) and plasmids were extracted with the alkaline lysis method using QIAprep buffers (P1, P2, and N3) from the QIAprep Spin miniprep kit (Qiagen), followed by chloroform extraction and isopropanol precipitation. Contaminating bacterial genomic DNA was removed by Plasmid-Safe™ ATP-Dependent DNase (Epicenter, Hessisch Oldendorf, Germany) according to manufacturer's instructions. The DNA concentrations were adjusted to 50 ng/μl and verified by gel electrophoresis. All samples passing quality assessments were subjected to the sequencing process, in which the plasmids were tagged, pooled, and sequenced using the Illumina miseq technology (Microsynth). In brief, library preparation was performed using the Nextera XT DNA Library Preparation Kit (Illumina) and sequenced using the MiSeq Reagent Kit v2 (Illumina) with the micro configuration (2 × 250 bp) following the manufacturer's recommendations. Reads were mapped to the reference *E. coli* K12 DH1 strain (Genbank Accession: CP001637.1) using bowtie2 (Langmead and Salzberg 2012) in order to remove remaining host DNA reads. Filtered raw data was then subsampled to 10% of the original amount and de novo assembled using the MIRA (v4.0.1) software (Chevreux et al. 2004). Sequence analyses were performed using the Geneious software (v6.1.6).

Assessment of native flora in the analyzed sources

16S rRNA amplicon sequencing from the V3/V4 regions of bacterial genes was used to assess the bacterial community structure in the LWW samples of B1, B2, and M. Two samples of each source were collected on two different sampling days (March 16th and March 22nd 2016). Samples were temporarily stored at – 20 °C. For DNA extraction, samples were filtered through a 47 mm-PES filter (polyethersulfon, 0.22 μm, Millipore, Merck Schaffhausen, Switzerland) and subsequently subjected to DNA extraction using the PowerWater DNA isolation kit (MoBio Laboratories, Oxoid) according to the manufacturer's recommendations. Bacterial 16S RNA genes were amplified using a Nextera two-step PCR containing the primers 341F_ill (CCTACGGGNGGCWGCAG) and 802R_ill (GACTACHVGGGTATCTAATCC) and sequenced using the MiSeq Reagent Kit v2 with the micro configuration (2 × 250 bp; Illumina).

Reads were demultiplexed and trimmed of Illumina adaptors using the Illumina MiSeq Reporter (v2.5.1.3) and the cutadapt software (v1.8.1) (Martin 2011). Read stitching was performed using the FLASH software (v1.2.11) (Magoč and

Salzberg 2011). Stitched reads with an average quality score of 25 or higher were used for downstream analyses. The default workflow in QIIME (v1.9.1) (Caporaso et al. 2010) was used for operational taxonomic unit (OTU) picking, chimera detection, and removal and taxonomic assignments. In brief, chimeric sequences were detected and removed using Uchime (v4.2) (Edgar et al. 2011) and Usearch (v8.1.1861) (Edgar 2010). Clustering and taxonomic assignment of OTUs was performed based on 97% sequence identity using Uclust (Edgar 2010) and the SILVA rRNA database (v111) (Quast et al. 2013). Singletons were removed and the community was rarefied to 31,706 sequences per sample for alpha and beta diversity analyses. Alpha diversity was described for each sample using Chao1 richness and Shannon diversity index, Faith's phylogenetic diversity and the number of OTUs. For a heat map visualization of relative abundances of OTUs, only the OTUs that each comprised more than 3% of the relative abundance across all samples (top 29 OTUs) were included. This was done using the heatmap.2 function in *R* using the "gplots" package.

HGT of recombinant plasmids to native wastewater bacteria

Serial dilutions of fresh LWW samples were grown on LB plates and incubated for 24 h at 28 °C. Approximately 30 visually unique colonies were isolated, patched onto fresh LB plates, and further characterized by MALDI-TOF MS as described above. Five isolates yielded good spectra and were identified to the species level. These included *Pseudomonas putida*, *Citrobacter freundii*, *Brevundimonas vesicularis*, *Aeromonas hydrophilia*, and *Pseudomonas fluorescens*. To test HGT by conjugation, we used the *E. coli* strain S17-1 harboring the broad-host range plasmid pProbe-gfp (tagless) (Miller et al. 2000) purchased from DSMZ (Braunschweig, Germany; DSM 9079) as a donor cell. The *E. coli* JM109 strain (Promega, Dübendorf, Switzerland) served as a positive control recipient.

For the conjugation experiments, isolates were scraped from LB plates, resuspended in sterile PBS and adjusted to an OD_{600} corresponding to four (i.e., approximately 3×10^9 cells/ml). Donor and recipient cells were mixed in a ratio of 1:100 in 1 ml volume and incubated for 24 h at 28 °C. The conjugation mixtures were serially diluted and selected to grow on LB agar plates containing kanamycin (50 μg/ml, for selection of the donor), ampicillin (100 μg/ml), or nalidixic acid (25 μg/ml; for selection of the recipient) or a combination of kanamycin with either ampicillin or nalidixic acid (for the selection of the transconjugants) (Table 3). After incubation for 24 h at 28 °C, recovered colonies were counted and the efficiency of conjugation was calculated as the number of transconjugants per donor cell. The presence of the plasmid pProbe-gfp in transconjugants was confirmed by amplification of aminoglycoside-phosphotransferase gene (*nptII*) by PCR. To this end, single colonies were picked and resuspended in a 20 μl reaction volume containing the following: 1× Hotstar PCR buffer (Qiagen), 50 mM $MgCl_2$, 10 mM (each) dNTP, 20 μM (each) of primer pairs ECPAL and ntpII (nptII-F CCTTGAGCCTGGCGAACA, nptII-R GCATCGAGCGAGCACGTA) and 1.6 U Hotstar Taq DNA Polymerase (Qiagen). The cycling parameters were as follows: initial denaturation for 15 min at 94 °C; 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min; followed by a final elongation step at 72 °C for 7 min. In addition, the species identity of the transconjugants was verified by MALDI-TOF MS as described above.

Results

Analysis of LWW for the presence of ampicillin-resistant *E. coli*

From July 2012 until June 2014, the LWW of the two downpipes B1 and B2 and the mixing tank M was monitored to capture the escape of GM *E. coli* from laboratories at a university campus. Ampicillin-resistant *E. coli* were examined as an indication of the presence of GM *E. coli*, since the LWW was devoid of fecal waste.

In total, 285 samples were analyzed (95 from each source; Table 1). Ampicillin-resistant presumptive GM *E. coli* were detected in 127 (45%) samples. The average concentration of presumptive *E. coli* over all positive samples was 2.8×10^2 CFU/ml. Throughout the sampling period, we constantly detected presumptive GM *E. coli* in M, with 85 out of 95 (89%) being positive (Fig. 1, Table 1). In contrast, the occurrence of presumptive GM *E. coli* in the LWW of B1 and B2 was more sporadic. In each source, 21 out of 95 samples were positive (22%) and the occurrence of presumptive GM *E. coli* seemed to follow a temporal trend. In B1, the majority of the positive samples (19 out of 21) were detected in three time blocks: July–September 2012; April–September 2013; February–June 2014 (Fig. 1). In B2, all positive samples were found within the time blocks September 2012–May 2013 and November 2013–June 2014.

Confirmation of GM *E. coli*

In order to verify that the ampicillin-resistant *E. coli* isolates were indeed GM *E. coli*, we performed genotyping and sequencing of plasmid DNA. To this end, we isolated a total of 45 ampicillin-resistant *E. coli* from B1, B2, and M between February and June 2014 and screened them for their uniqueness regarding genotype, plasmids, and inserts. Out of three isolates from B1, two were unique; out of 18 isolates from B2, seven were unique and out of 24 isolates from M, 10 were

Table 1 Summary of the number (*n*) of samples harvested, samples containing ampicillin-resistant *Escherichia coli*, isolates used for genotyping and sequencing, respectively, and number of sequenced vectors and inserts

Processing step	Source of isolates			Total (*n*)
	B1	B2	M	
Samples collected (*n*)	95	95	95	285
Samples containing ampicillin-resistant *E.coli* (*n*)	21	21	85	127
Isolates chosen for genotyping (*n*)	3	18	24	45
Unique isolates (*n*)	2	7	10	19
Unique isolates with *E. coli* genotype B/K12 (laboratory strain) (*n*)	2	7	3	12
Unique isolates used for vector sequencing (*n*)	2	5	3	10
Vectors identified (*n*)	2	5	3	10
Unique isolates used for insert sequencing (*n*)	2	6	3	11
Inserts identified (*n*)	2	6	1	9

unique (Table 1). All isolates of B1 and B2 (100%) were identified as typical laboratory *E. coli* genotypes, and were either derived from the K-12 or the B strains (Tables 1 and 2). In contrast, only three out of 10 (30%) isolates from M were *E. coli* laboratory strains. All 10 sequenced vectors were highly similar (99–100%) to cloning and expression vectors

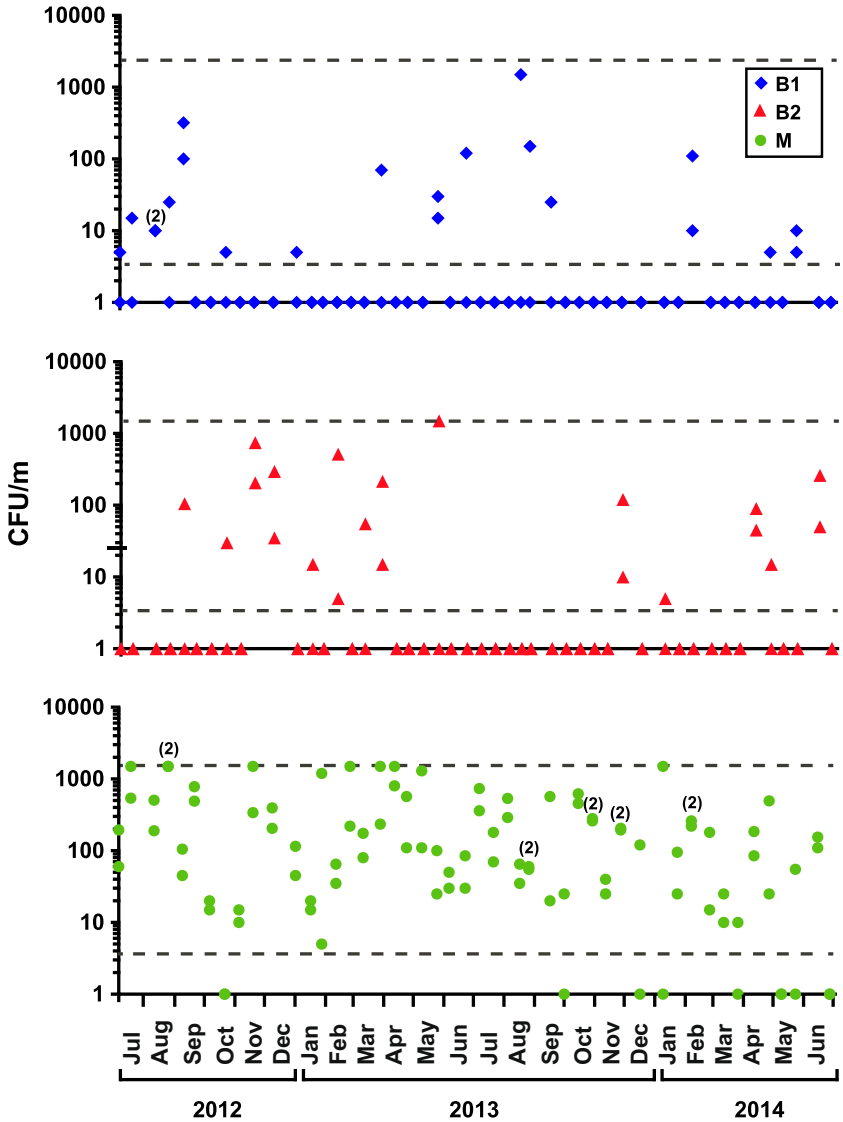

Fig. 1 Number of colony forming units (CFU) per milliliter detected in the building downpipe B1 (top graph, diamonds), building downpipe B2 (middle graph, triangles), and the mixing tank M (bottom graph, circles) between July 2012 and June 2014. On each sampling day, two samples were taken at 10:00 and 12:00. Upper and lower detection limits are indicated with dashed gray lines. Superimposed incidences of non-zero data points are labeled by (2)

Table2 Summary of unique genetically modified *Escherichia coli* found in the two building downpipes B1 and B2, and the mixing tank M. The *E.coli* genotypes were identified by multiplex PCR (Bauer etal. 2007) and the plasmids and their inserts were characterized by either Sanger or entire plasmid sequencing

Isolate name	*E.coli* geno-type[a]	Best blastn hit[b] (vector)	*E* value	Score[c]	ID (%)	Best blastn hit[b] (insert)	E-value	Score[c]	ID (%)	Best blastx hit[b] (insert)	*E* value	Score[c]	ID (%)	Isolation date
B1.01[d]	B	pGEX-6P-3 cloning vector	0	4591	99	Sirtuin 6 [*Mus musculus*]	0	1295	100	NAD-dependent protein deacetylase sirtuin-6 [*Mus musculus*]	4.00 E-168	480	100	24.04.2014
B1.02[d]	K-12	T7RNA polymerase vector pGemT7cat	0	5204	99	Synthetic construct R-CaMP1.07	0	2715	100	R-CaMP1.07 protein [synthetic construct]	0	885	100	20.05.2014
B2.01	K-12	Not sequenced	–	–	–	No significant hit	–	–	–	No significant hit	–	–	–	09.04.2014
B2.03[e]	K-12	Cloning vector U6p-gRNA1	0	1053	99	Scaffold0000777[f] [*Caenorhabditis elegans*]	1.00E-122	449	100	No significant hit	–	–	–	09.04.2014
B2.04[e]	K-12	Cloning vector U6p-gRNA1	0	1351	99	Scaffold0000269 [*Caenorhabditis elegans*]	6.00E-58	233	99	No significant hit	–	–	–	09.04.2014
B2.05[e]	K-12	Cloning vector p5E-smyhc1	2.00 E-76	298	99	Scaffold0000051 [*Caenorhabditis elegans*]	0	1572	100	No significant hit	–	–	–	09.04.2014
B2.07[e]	K-12	Not sequenced	ND	ND	ND	Scaffold0000056 [*Caenorhabditis elegans*]	0	1179	99	No significant hit	–	–	–	09.04.2014
B2.09[d]	B	Cloning vector pGATA	0	4481	99	B cell CLL/lymphoma 9 (Bcl9) [*Mus musculus*]	0	981	99	B cell CLL/lymphoma 9 protein isoform X2 [*Mus musculus*]	2.00 E-132	423	92	24.04.2014
B2.13[e]	K-12	Cloning vector pACW-hyg	3.00 E-109	405	100	Cosmid K09B11 [*Caenorhabditis elegans*]	0	896	98	Hypothetical protein T06C10.5 [*Caenorhabditis elegans*]	6.00 E-65	217	89	11.06.2014
M.11[e]	K-12	Expression vector pBGH	0	5280	99	No significant hit	–	–	–	Metal ion transporter (Nramp) family [*Eremococcus coleocola* ACS-139-V-Col8]	0	783	100	09.04.2014
M.13[d]	B	pGEX-6P-3 cloning vector	0	4591	99	Sirtuin 6 [*Mus musculus*]	0	1295	100	NAD-dependent protein deacetylase sirtuin-6 isoform 1 [*Mus musculus*]	4.00 E-168	480	100	09.04.2014
M.17[d]	K-12	Expression vector pBGH	0	5240	99	No significant hit	–	–	–	Membrane protein [*Sphingobacterium spiritivorum*]	4.00 E-165	466	99	20.05.2014

[a] Identified by multiplex PCR as described in (Bauer et al. 2007) and (Kuhnert et al. 1995)

[b] The blast search was performed on 07.07.2016

[c] Maximum score indicated (where different from the total score)

[d] Identified by sequencing the entire plasmid

[e] Identified by Sanger sequencing of the insert

[f] From genome assembly C_elegans_Bristol_N2_v1_5_4

based on blastn sequence identity (Table 2). The identified plasmids predominantly harbored inserts with sequences from the *Caenorhabditis elegans* genome (Table 2) and were all found in LWW from B2 (n = 5). Three plasmids harbored inserts with highly similar sequences to *Mus musculus* (sirtuin 6 and B cell CLL/lymphoma 9 (Bcl9) transcripts). Two harbored inserts with high similarity to bacterial genes (i.e., a metal ion (Mn2+/Fe2+) transporter (Nramp) of *Eremococcus coleocola* and a membrane protein of *Sphingobacterium spiritivorum*). One plasmid consisted of a T7 RNA polymerase vector harboring a synthetic construct (R-CaMP1.07). Surprisingly, we found identical isolates in B1 (B1.01) and M (M.13). As they were isolated on different days, this indicates that detection of GM *E.coli*, originally released in downpipes, is sensitive enough to be detected M despite a strong dilution.

Assessment of bacterial communities in LWW

The LWW monitoring revealed that B1, B2, and M contained between 2.7×10^4 and 4.5×10^4 CFU/ml (total viable counts). In a next step, we identified the most abundant bacterial taxa and addressed the question to what extend recombinant plasmids can be transferred to native LWW bacteria via HGT. We deployed 16S rRNA amplicon sequencing of isolated DNA of six LWW samples (two of each source B1, B2, and M). This resulted in a total of 1,173,044 raw reads and 577,174 stitched and merged reads of 419 bp average length. Rarefaction to 31,706 sequences yielded 2245 operational taxonomic units (OTUs), each containing sequences that share at least 97% identity. The Bray–Curtis distance matrix revealed that the bacterial communities were specific to their environment, since samples from the same source clustered together (Fig. 2a). Representation of the OTU abundance patterns (top 29) revealed that M was dominated by two bacterial taxa belonging to the *Citrobacter* and *Acinetobacter* genera, comprising 63 and 58%, respectively, of the total observed taxa in the two samples (Fig. 2b). In contrast to the communities of M, communities of B1 and B2 contained more diverse taxa with lower relative abundances.

HGT of recombinant plasmids to native wastewater bacteria

We addressed the question to what extend recombinant plasmids can be transferred to representative species of native LWW bacteria via HGT. For this purpose, we isolated five LWW isolates corresponding to frequently identified species (Fig. 2), namely *P. putida*, *C. freundii*, *B. vesicularis*, *A. hydrophilia*, and *P. fluorescens*. None of the GM *E. coli* isolates qualified as donors in a conjugation experiment due to the lack of containing a broad-host plasmid and of a suitable antibiotic resistance panel for the selection of the resulting transconjugants, respectively. We therefore used the *E. coli* S-17 donor strain and examined whether its broad-host range mobilizable plasmid pProbe-gfp (Miller et al. 2000) can be transferred to the abovementioned five native LWW isolates. Conjugation experiments were conducted according to the ideal conditions previously assessed using the positive control recipient strain *E. coli* JM109 (data not shown). We show that four out of five representative LWW isolates served as recipients of mobilized plasmid DNA. Conjugation frequencies were between 3.4×10^{-3} and 2.4×10^{-5}, which is comparable to frequencies observed for the positive control conjugation (Table 3). The highest transfer frequency was obtained with *A. hydrophila*. No plasmid transfer was observed for *B. vesicularis*, the only recipient which is not a Gammaproteobacterium.

Discussion

The containment of GM microorganisms in laboratories presents a challenge for researchers and industry worldwide. It is crucial to ensure efficient containment measures and be aware of the predominant escape routes of GM microorganisms into the environment. By monitoring GM *E. coli* in LWW and sequencing their recombinant plasmids, we took an unbiased approach to examine whether GM microorganisms are released into the LWW systems from a university campus. Indeed, the LWW represented a repeated escape route of GM *E. coli* and our approach proved highly suitable to isolate and characterize presumptive GM *E. coli*. We found that 45% of all samples contained ampicillin-resistant, presumptive GM *E. coli* at an average concentration of 2.8×10^2 CFU/ml. Furthermore, by estimating the risk of plasmid transfer to five members of the native bacterial community, we assessed the environmental risks associated with an unintended release.

All ampicillin-resistant *E. coli* from the two building downpipes B1 and B2 proved to be GM *E. coli* and harbored research-specific plasmids. Since GM *E. coli* contamination varied greatly between time periods, we conclude that our findings represent actual escapes from the laboratory containment into LWW. A release from a biofilm attached to a pipe wall is expected to occur in a continuous manner and was therefore considered unlikely. Even if GM *E. coli* were sporadically shed from a biofilm matrix, they still represented earlier laboratory escapes. Indeed, it was shown by an experimental biofilm system with groundwater that escapes can date back as far as 20 days (Banning et al. 2003). Partial or complete sequencing of 11 plasmids of the building downpipe isolates pointed towards laboratory-specific activities. Interestingly, most of the unique GM *E. coli* isolates (7 out of 9) harbored inserts from *C. elegans* and mouse sirtuin, indicating that they were released from only very few

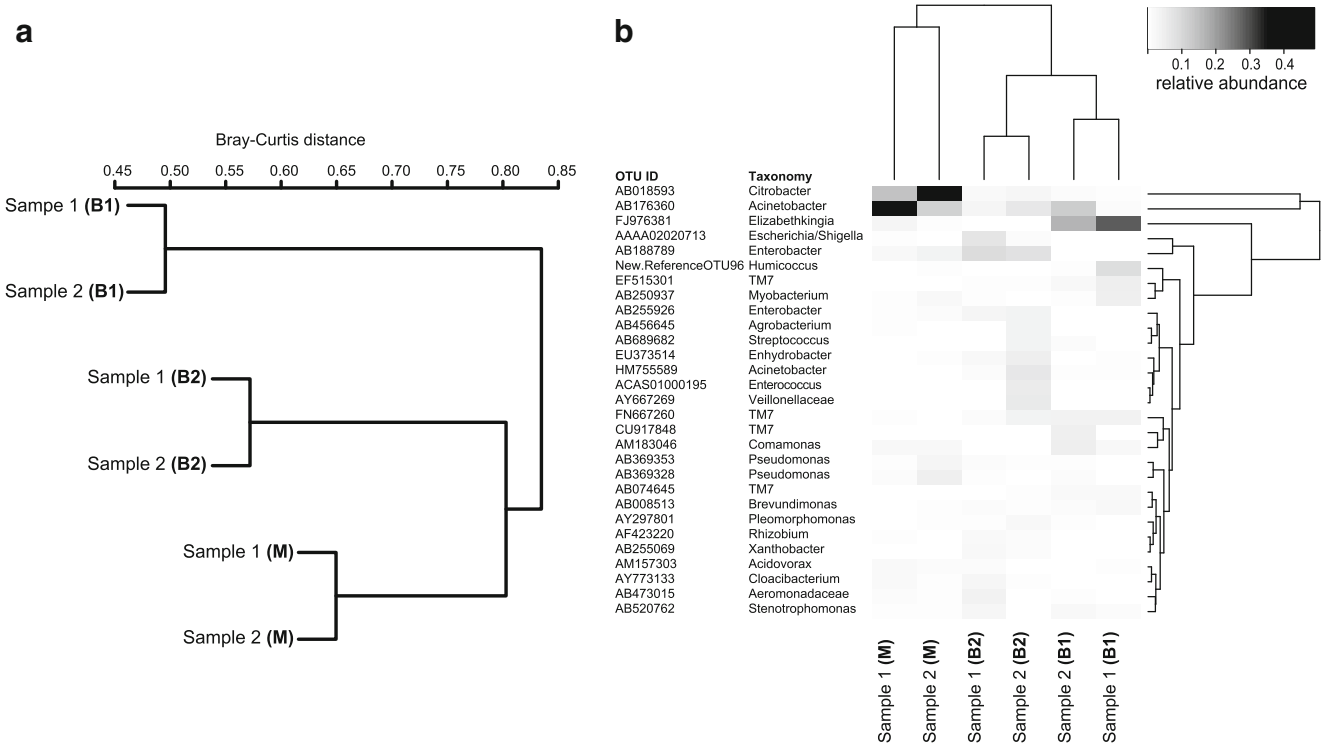

Fig. 2 **a** Hierarchical cluster of bacterial communities in laboratory wastewater (LWW) based on Bray–Curtis similarities of 16S rRNA amplicon sequencing. Bray–Curtis similarity coefficients were calculated from relative operational taxonomic unit (OTU) abundances in bacterial communities of LWW of the building downpipes B1 and B2, and the mixing tank M. Two samples were obtained from each source. **b** Most abundant OTUs comprising over all samples more than 3% of the total bacteria community in B1, B2 and M

individual laboratories on the campus (activities at the campus indeed contain *C. elegans* and mouse sirtuin research groups).

In contrast to the occurrence of ampicillin-resistant *E. coli* in B1 and B2, they were continuously detected in the mixing tank (M) throughout the sampling period. This observation could suggest an actual constant release of these bacteria via additional building downpipes (which were not sampled) or an enrichment of previously escaped organisms in M. After genotyping 24 isolates from M, only 30% of the unique isolates were identified as GM *E. coli* harboring a research-specific plasmid. However, this number is not absolute. Instead, it represents the ratio of the GM *E. coli* among the analyzed unique ampicillin-resistant *E. coli*. We only followed up on unique isolates ($n = 10$), and the ones that were indeed GM *E. coli* ($n = 3$) very often occurred as clonal replicates. Yet, the number is comparable to the ratio of the downpipes where 100% of the isolated ampicillin-resistant *E. coli* were found to be GM. Alternatively, the presence and abundance of ampicillin-resistant, but non-GM *E. coli* in M might be explained by joining LWW pipes from the mouse cage washing stations leading to M. Indeed, Shimoda et al. (2002) showed that approximately 7% of the *E. coli* strains from laboratory animal feces are resistant to ampicillin.

In samples containing ampicillin-resistant *E. coli*, an average concentration of 2.8×10^2 CFU/ml was found. Notably, under suboptimal conditions many bacteria including *E. coli*

Table 3 Conjugation efficiencies between *Escherichia coli* S17-1 containing the broad-host range plasmid pProbe-gfp (Miller et al. 2000) and bacterial isolates from laboratory wastewater or the control strain *E. coli* JM109. The donor:recipient ratio was 1:100, which was optimal for conjugation with the control strain *E. coli* JM109

Recipient strain	Antibiotics for selection (μg/ml)	Conjugation efficiency
Escherichia coli JM109	Kanamycin (50), Nalidixic acid (25)	1.9×10^{-3}
Citrobacter freundii	Kanamycin (50), ampicillin (100)	2.4×10^{-5}
Pseudomonas fluorescence	Kanamycin (50), ampicillin (100)	1.0×10^{-5}
Pseudomonas putida	Kanamycin (50), ampicillin (100)	7.0×10^{-4}
Aeromonas hydrophila	Kanamycin (50), ampicillin (100)	4.3×10^{-3}
Brevundimonas vesicularis	Kanamycin (50), nalidixic acid (25)	No conjugation/plasmid transfer

have been reported to enter a dormant, i.e., viable-but-nonculturable state (Oliver 2005). Thus, the number of released viable GM *E. coli* might be underrepresented by the culturing-dependent approach used in this study. The Swiss "Containment Ordinance" (The Federal Authorities of the Swiss Confederation 2012) requires a so called "harmless disposal" and "minimization of escape" of GM organisms at the biosafety level 1. Although a legal limit is not provided, one can assume that the dispositions of the "Containment Ordinance" are not being respected in case of a deliberate disposal of contaminated material without prior inactivation. We could, however, not elucidate to what extent deliberate disposal is the cause for the presence of the detected GM *E. coli* in the LWW under investigation. To address this, further investigations regarding the potential causes of the releases of GM *E. coli* into LWW systems are necessary. The focus of this work was not to point out individual shortcomings, but rather to show the feasibility of detecting GM *E. coli* in LWW.

The most significant environmental risk associated with an unintended release is HGT of recombinant DNA to native bacteria present in LWW. Indeed, numerous studies reported the acquisition of plasmids in environmental isolates by conjugation and transformation (reviewed in Davison (1999); Lorenz and Wackernagel (1994); von Wintersdorff et al. (2016)). Under experimental conditions standing for a worst case scenario, we show that four representative LWW isolates from the monitoring sites can serve as recipients of plasmids mobilized from a laboratory *E. coli* K12 derivative.

Given the suboptimal conditions for conjugations in the environment of the downpipes, in particular the high flow rate, we assume that the actual occurrence of HGT in building downpipes is probably rare. M could be far more relevant in the context of HGT since areas of locally increased bacterial concentrations may arise. Furthermore, M contained high relative abundances (47%) of a Citrobacter OTU. We showed that HGT of a broad-host range plasmid from *E. coli* to *Citrobacter freundii* is possible, and might therefore play a role in HGT. However, taking into account that M contained up to 4.5×10^4 total viable CFU per ml and using a ratio of 1:100 as a working hypothesis of culturable to nonculturable bacteria (Gonzalez et al. 1992), we conclude that approximately 10^6 competent bacteria per milliliter are present in M. This is significantly lower than the concentration generally used for conjugation experiments (10^8 to 10^9/ml; this work and (Gealt et al. 1985; Leverstein-van Hall et al. 2002; McPherson and Gealt 1986; Sandt and Herson 1991; Zurfluh et al. 2015)). Furthermore, the formation of transconjugants with a *Citrobacter* OTU or other bacterial taxa different from *E. coli* requires that the GM *E. coli* harbor a transferable broad-host range plasmid. This is a major point, since none of the identified plasmids in B1, B2, or M were actually self-transmissible broad-host range plasmids carrying the genes necessary for transfer initiation at the origin of transfer (*oriT*) (Guiney and Yakobson 1983).

In conclusion, our data show that GM *E. coli* were released into the LWW system and the majority could be traced to their origin of release. Nevertheless, the detected release of GM *E. coli* represents individual cases and the causes of a deliberate or unintended release need to be investigated further. Even if GM *E. coli* were released into LWW, the chance of survival and HGT seems low. An additional and suitable containment measure could include the use of genomically recoded *E. coli*, which depend on nutrients not available in nature, rendering them nonviable in the environment when released (Kato 2015; Mandell et al. 2015; Rovner et al. 2015). In our view, this approach is feasible and desirable, but should not replace the implementation of effective technical and organizational containment strategies, including a strong focus on training.

Acknowledgements This work was financially supported by the Federal Office of Public Health (FOPH) in Switzerland. We thank Monica Alt for technical assistance, Dirk Hamburger for critical reading of the manuscript, and the Office of Waste, Water, Energy and Air WWEA, Canton Zurich for providing the results of the Questionnaire. We are grateful to the member of the working group "Safe disposal of Risk Group 1 organisms" headed by the FOPH including Basil Gerber (Federal Office for the Environment FOEN), Manuela Ocaña and Samuel Roulin (Federal Office of Public Health FOPH), Claudia Ruprecht and Christina Stadler (Office of Waste, Water, Energy and Air WWEA, Canton Zurich) for constructive discussions as well as valuable comments on the manuscript.

References

Abberton CL, Bereschenko L, van der Wielen PWJJ, Smith CJ (2016) Survival, biofilm formation, and growth potential of environmental and enteric *Escherichia coli* strains in drinking water microcosms. Appl Environ Microbiol 82:5320–5331. https://doi.org/10.1128/aem.01569-16

Amy PS, Hiatt HD (1989) Survival and detection of bacteria in an aquatic environment. Appl Environ Microbiol 55:788–793

Awong J, Bitton G, Chaudhry GR (1990) Microcosm for assessing survival of genetically engineered microorganisms in aquatic environments. Appl Environ Microbiol 56:977–983

Banning N, Toze S, Mee BJ (2003) Persistence of biofilm-associated *Escherichia coli* and *Pseudomonas aeruginosa* in groundwater and treated effluent in a laboratory model system. Microbiology 149:47–55. https://doi.org/10.1099/mic.0.25938-0

Bauer AP, Dieckmann SM, Ludwig W, Schleifer KH (2007) Rapid identification of *Escherichia coli* safety and laboratory strain lineages based on multiplex-PCR. FEMS Microbiol Lett 269:36–40. https://doi.org/10.1111/j.1574-6968.2006.00594.x

Berg P et al (1974) Potential biohazards of recombinant DNA molecules. Science 185:303

Bogosian G, Sammons LE, Morris PJ, O'Neil JP, Heitkamp MA, Weber DB (1996) Death of the *Escherichia coli* K-12 strain W3110 in soil and water. Appl Environ Microbiol 62:4114–4120

Brettar I, Hofle MG (1992) Influence of ecosystematic factors on survival of Escherichia coli after large-scale release into lake water mesocosms. Appl Environ Microbiol 58:2201–2210

Caporaso JG et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303

Chao WL, Feng RL (1990) Survival of genetically engineered *Escherichia coli* in natural soil and river water. J Appl Bacteriol 68:319–325

Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WE, Wetter T, Suhai S (2004) Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res 14:1147–1159. http://www.genome.org/cgi/doi/10.1101/gr.1917404. Accessed 28. August 2017

Davison J (1999) Genetic exchange between bacteria in the environment. Plasmid 42:73–91. https://doi.org/10.1006/plas.1999.1421

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Gealt MA, Chai MD, Alpert KB, Boyer JC (1985) Transfer of plasmids pBR322 and pBR325 in wastewater from laboratory strains of *Escherichia coli* to bacteria indigenous to the waste disposal system. Appl Environ Microbiol 49:836–841

Gonzalez JM, Iriberri J, Egea L, Barcina I (1992) Characterization of culturability, protistan grazing, and death of enteric bacteria in aquatic ecosystems. Appl Environ Microbiol 58:998–1004

Guiney DG, Yakobson E (1983) Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. Proc Natl Acad Sci U S A 80:3595–3598

Kato Y (2015) An engineered bacterium auxotrophic for an unnatural amino acid: a novel biological containment system. PeerJ 3:e1247. https://doi.org/10.7717/peerj.1247

Kuhnert P, Nicolet J, Frey J (1995) Rapid and accurate identification of *Escherichia coli* K-12 strains. Appl Environ Microbiol 61:4135–4139

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

Leverstein-van Hall MA, Box AT, Blok HE, Paauw A, Fluit AC, Verhoef J (2002) Evidence of extensive interspecies transfer of integron-mediated antimicrobial resistance genes among multidrug-resistant Enterobacteriaceae in a clinical setting. J Infect Dis 186:49–56. https://doi.org/10.1086/341078

Lorenz MG, Wackernagel W (1994) Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev 58:563–602

Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Mandell DJ et al (2015) Biocontainment of genetically modified organisms by synthetic protein design. Nature 518:55–60. https://doi.org/10.1038/nature14121

Mao D et al (2014) Persistence of extracellular DNA in river sediment facilitates antibiotic resistance gene propagation. Environ Sci Technol 48:71–78. https://doi.org/10.1021/es404280v

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12

McPherson P, Gealt MA (1986) Isolation of indigenous wastewater bacterial strains capable of mobilizing plasmid pBR325. Appl Environ Microbiol 51:904–909

Miller WG, Leveau JH, Lindow SE (2000) Improved gfp and inaZ broad-host-range promoter-probe vectors. Mol Plant-Microbe Interact 13:1243–1250. https://doi.org/10.1094/MPMI.2000.13.11.1243

Oliver JD (2005) The viable but nonculturable state in bacteria. J Microbiol 43 spec no:93-100

Quast C et al (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219

Rovner AJ et al (2015) Recoded organisms engineered to depend on synthetic amino acids. Nature 518:89–93. https://doi.org/10.1038/nature14095

Sagik BP, Sorber CA, Moore BE (1981) The survival of EK1 and EK2 systems in sewage treatment plant models. Molecular Biology, Pathogenicity, and Ecology of Bacterial Plasmids. Springer, In, pp 449–460

Sandt CH, Herson DS (1991) Mobilization of the genetically engineered plasmid pHSV106 from *Escherichia coli* HB101(pHSV106) to *Enterobacter cloacae* in drinking water. Appl Environ Microbiol 57:194–200

Shimoda K, Kimura J, Oguri S, Nakamura N, Urano T, MAEJIMA K (2002) Penetration of drug resistance in *Escherichia coli* isolated from laboratory animals. J Vet Med Sci 64:1133–1135. https://doi.org/10.1292/jvms.64.1133

The Federal Authorities of the Swiss Confederation (2012) Ordinance on handling organisms in contained systems (Containment ordinance, CO) of 9 May 2012 (Status as of 1 June 2014). SR. 814.912. Available via https://www.admin.ch/opc/en/classified-compilation/20100803/index.html. Accessed 28. August 2017

Toh BEW et al (2015) Differentiation of *Acinetobacter* genomic species 13BJ/14TU from *Acinetobacter haemolyticus* by use of matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS). J Clin Microbiol 53:3384–3386. https://doi.org/10.1128/JCM.03468-14

von Wintersdorff CJ et al (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Front Microbiol 7:173. https://doi.org/10.3389/fmicb.2016.00173

Way JC, Collins JJ, Keasling JD, Silver PA (2014) Integrating biological redesign: where synthetic biology came from and where it needs to go. Cell 157:151–161. https://doi.org/10.1016/j.cell.2014.02.039

Welker M, Moore ERB (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 34:2–11. https://doi.org/10.1016/j.syapm.2010.11.013

Ziegler D et al. (2015) Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS. Appl Microbiol Biotechnol 99:5547-5562. https://doi.org/10.1007/s00253-015-6515-3

Zurfluh K, Glier M, Hächler H, Stephan R (2015) Replicon typing of plasmids carrying blaCTX-M-15 among *Enterobacteriaceae* isolated at the environment, livestock and human interface. Sci Total Environ 521–522:75–78. https://doi.org/10.1016/j.scitotenv.2015.03.079