Computational prediction, rationalization and experimental validation of PCR primers for the detection of antibiotic resistance genes in wastewater treatment plants

Decewicz P.[1,#], Gorecki A.[1,#], Janeczko A.[2], Drewniak L.[3] and Dziewit L.[1,*]

[1]University of Warsaw, Faculty of Biology, Institute of Microbiology, Department of Bacterial Genetics, Miecznikowa 1, 02-096 Warsaw, Poland.

[2] Municipal and Industrial Wastewater Treatment Plant, Nadwislanska 46, 32-600 Oswiecim, Poland

[3] University of Warsaw, Faculty of Biology, Laboratory of Environmental Pollution Analysis, Miecznikowa 1, 02-096 Warsaw, Poland

[#] equal contribution

*corresponding author: e-mail: ldziewit@biol.uw.edu.pl

Abstract

Wastewater treatment plants (WWTPs) are point sources of various emerging pollutants, including antibiotics and antibiotic-resistance genes (ARGs). These reservoirs promote formation of antimicrobial resistance due to natural selection of the multidrug-resistant bacteria and horizontal gene transfer of ARGs. Therefore, there is a great need to effectively monitor the spread, diversity and fate of ARGs in WWTPs. One of the possibilities for such screening is PCR typing. However, efficient and diverse sets of primers specific to various ARGs are needed. In literature, there are plethora of such primers, however their usefulness in environmental studies varies significantly. In this study a dedicated bioinformatic *in silico* PCR (e-PCR) tool was created to validate and calculate various primers specificity and efficacy. Over 300 primer pairs specific to diverse ARGs families were validated against the Comprehensive Antibiotic Resistance Database (CARD) and Antibiotic Resistance Genes Database (ARDB) and various WWTPs metagenomes, which enables calculating primers overall rating and preparation of the database of PCR primers most useful in ARGs environmental screening. Finally, selected primer pairs were used in experimental PCR testing survey to check the presence of the ARGs in waste samples (collected on subsequent stages of the wastewater treatment process) from the Oswiecim WWTP (Poland).

Keywords: Wastewater treatment plant, antimicrobial resistance, antibiotic resistance genes, PCR typing

1. Introduction

Antibiotics are considered one of the most effective pharmaceuticals used in treatment of infectious diseases. Since 1930's antibiotics have saved millions of lives, and the positive influence of these drugs on overall public health cannot be overrated (Ventola, 2015). The improvement of modern technologies and laboratory techniques leads to discovery and production of many synthetically modified derivatives of already known antibiotics, unfortunately in the last 30 years only one really novel antibiotic – teixobactin (Piddock, 2015) has been discovered. The great success of antibiotics treating infectious diseases prompted the drugs to be used in different industrial fields. Plenty of them, either as therapeutic or prophylactic agents, are widely used in agricultural practices or as growth promoters in farm animals (Cromwell, 2002; van Hoek *et al*., 2011). Unfortunately, high marked demands for these pharmaceuticals have had severe environmental consequences. After 60 years since their first introduction, millions of metric tons of antibiotics have been used in variety of applications (Davies and Davies, 2010). Additionally, there are many evidences of overuse and inappropriate prescribing antibiotics in medicine and veterinary. Studies have shown that between 30% and 60% of antibiotics prescribed are incorrect in choice, indication or duration of treatment (Luyt *et al*., 2014). All of these misuses generated new types of emerging contaminants which are antibiotics themselves, antibiotic resistant bacteria and antibiotic resistance genes (ARGs). Nowadays, it is generally known that bacteria exhibit various mechanism to render the antibiotics ineffective (van Hoek *et al*., 2011). The mechanisms of antibiotic resistance are mostly a consequence of a development (via increased mutation rate induced by antibiotic, which leads to evolving of novel enzymes conferring particular antibiotic resistance) or acquiring specific antibiotic resistance genes from other bacteria (Eliopoulos *et al*., 2003). The latter one is a consequence of a presence of various mobile genetic elements within bacterial genomes, which could easily translocate genes via horizontal gene transfer (HGT). The phenomenom of HGT is considered to be the most important factor responsible for the spread of antibiotic resistance in environment. Unfortunately, abundance of ARGs in environment promote formation of multi-resistant pathogens (called superbugs), which is an increasingly serious threat to global public health (von Wintersdorff *et al*., 2016). Wastewater treatment plants (WWTPs) are considered as point source of significantly accumulated diverse emerging pollutants, including

antibiotics, antibiotic resistant bacteria and antibiotic resistance genes. WWTPs are also the most important interfaces between the human population and the environment (Akiba *et al*., 2015). Therefore, there is a great need to effectively monitor the occurrence, diversity and spread of ARGs in WWTPs. There are two main approaches for these kinds of analyses, either culture-based or molecular-based. One of most conventional and cheapest molecular-based method is PCR typing. However, efficient and diverse sets of PCR primers specific to various ARGs are needed for such assays (Rizzo *et al*., 2013). In literature, there are plethora of such PCR primers, however their usefulness in environmental studies varies significantly. Certain primers pairs are genus-specific or show relaxed specificity (i.e. they are specific not only to ARGs, but also to other genes). Therefore, such primers are useless in the screening assays, and they do not suppose to be used, as generating error or biased results. Thereupon, in this work, a dedicated bioinformatic *in silico* PCR (ePCR) tool was created to validate and calculate primers specificity, efficacy and taxon specificity. In this study we used a developed ePCR tool to create a database with ranked PCR primers, that could be used for the analyses of the presence and diversity of antibiotic resistance genes in various environments, including WWTPs. Finally, selected primer pairs were used in experimental PCR testing surveys to validate the presence of the ARGs in samples from the Oswiecim WWTP in Poland, collected on subsequent stages of the process. To our knowledge this is the first such complex approach to organize methodology of PCR screening of antibiotic resistance genes in environmental samples.

2. Material and methods

2.1 Samples collection

All samples were collected from the municipal and industrial WWTP located in Oswiecim (Poland). The WWTP has the capacity of about 53 400 m^3 waste per day. The treatment process is divided on 3 main stages: primary treatment, secondary treatment and anaerobic digestion. Samples were collected from all those stages and named in the following order: primary sludge (PS), activated sludge (AS) and anaerobic digestion sludge (ADS).

2.2 DNA extraction

The total DNA was extracted from 100 mg of frozen sludge samples using PowerSoil® DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) according to the manufacturer instructions. The DNA concentrations and purity were determined using Qubit™ 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA).

2.3 Antibiotic resistance genes screening

Occurrence of antibiotic resistance genes in the analysed samples was investigated using PCR assays. The PCRs were performed with 10 pmol of each primer [previously selected using bioinformatic tool (ePCR)], DreamTaq PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) and 150 ng of DNA template. The PCR programs were dependent on recommended annealing temperature for each primer pair. The PCR products were resolved by electrophoresis in an agarose gels.

2.4 Standard molecular biology techniques

The PCR products were cloned into pTZ57R/T vector (Thermo Fisher Scientific, Waltham, MA, USA) and introduced via chemical transformation into *E. coli* DH5α strain (Sambrook and Russell, 2001). Plasmid DNA was purified using Plasmid Miniprep DNA Purification Kit (EURx, Gdansk, Poland) and sequenced using universal primers (M13/pUC) and capillary DNA analyzer ABI-PRISM 377 (Applied Biosystem, Foster City, CA, USA) at the Institute of Biochemistry and Biophysics, Polish Academy of Science (oligo.pl).

2.5 16S rDNA-based metagenomic analysis

For the PCR amplicon preparation the following primer pairs were used: 16S_V3-F: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGC CTACGGGNGGCWGCAG 3' and 16S_V4-R: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GACTACHVGGGTATCTAATCC 3' targeting the variable regions (V3 and V4) of the bacterial 16S rRNA gene. Each reaction was carried using KAPA HiFi polymerase (KAPA Biosystems, Wlimington, MA, USA) in a Mastercycler Nexus GX2 termocycler (Eppendorf, Hamburg, Germany). After 3 minutes of denaturation of DNA in 95°C, 30 cycles including: denaturation (95°C, 30 sec.), primer annealing (60-65°C, 30 sec.) and DNA synthesis (72°C, 30 sec.) were set. Each PCR reaction was repeated in triplicate and then each three probes were mixed and used for the sequencing. Amplicon libraries were sequenced on Illumina MiSeq instrument at oligo.pl with the use of v3 MiSeq chemistry kit in a paired end mode. To perform the diversity analysis raw reads were bioinformatically processed using tools and pipelines wrapped by QIIME v1.9.1 (Caporaso *et al*., 2011). After processing, reads were clustered into OTUs at 97 % identity within centroids with USEARCH v6.1. The taxonomy was assigned using RDP classifier v2.2 (Wang *et al*., 2007) using GreenGenes v13.8.

2.6 ePCR tool

To test primer pairs ePCR tool was developed in Python v3.4 which wraps nucleotide BLAST program to search primers against selected databases or fasta files. Selected BLASTN parameters can be modified by the user, however default ones were already optimized for short query sequences i.e. word size: 7, e-value: 10000, reward for positive match: 1, penalty for mismatch: -2, penalty for gap opening: 5, penalty for gap extension: 3, maximum number of aligned sequences to keep: 10000, minimum query coverage per subject: 75. The program takes as an input a tab-delimited mapping file that consists of primer pairs with their names and corresponding sequences, acquired product size and its possible deviation. This file is parsed and each primer from the pair is used as an individual query during BLASTN search. Next, the results are parsed and only proper *in silico* amplification products are retained and provided in the resulting summary file.

2.7 Calculation of primer pairs parameters

All of the primer pairs collected during data mining were checked against NCBI nucleotide NT database. At first, the results were checked manually to verify specificity of each primer pair (S) – whether the obtained product amplified

desired gene. Taxon specificity (TS) was based on the ratio of the number of products amplified from different genera and all possible genera in which a homologs of reference gene from CARD database were identified. The last parameter, efficacy (E), was calculated as a ratio of all good *in silico* products and all homologs identified in NT database.

2. 8 Analysis of wastewater treatment plant metagenomes

The validation of all the primer pairs were conducted with the use of selected WWTP metagenomes from the SRA database: SRR866802 from Vermont (USA), SRR1144841 from Shanghai and SRR1106773 from Nanjing (China) and SRR1611146 from Shifflange (Luxemburg). These were all assembled using MEGAHIT v1.1.1-2 with meta-large preset settings (Li *et al*., 2015).The obtained contigs were afterwards searched with ePCR tool to check the presence of selected ARGs.

3. Results and discussion

3.1 Rationalization and validation of PCR primer pairs specific to antibiotic resistance genes and construction of a database of ARGs-specific primers

The first aim of the study was to create a database of PCR primers useful in detection of various ARGs. Such a database was created based on the thorough scientific literature review. Currently, this database contains 302 primer pairs which can be used for amplification of genes encoding proteins responsible for resistance to various antibiotics, including (i) aminoglycosides (38 primer pairs), (ii) β-lactamases (78), (iii) chloramphenicols (22), (iv) macrolides (30), (v) fluoroquinolones (10), (vi) sulfonamides (8), (vii) tetracycline (58). Moreover, there are 17 primer pairs encoding multidrug resistance pumps and 41 other types of antibiotic resistance proteins. The most important feature of the database is its internal ranking of the PCR primer pairs, based on 3 parameters calculated for all tested primers, i.e. efficacy (E), specificity (S) and taxon specificity (TS), which allowed selection of the most appropriate primers (by the database user) for particular purposes. As an example we present the comprehensive analysis of selected 18 PCR primer pairs (Table. 1). The selected primers (comprising various S, E and TS scores) were designed for *ermA*, *ermB*, *ermC*, *ermF*, *sul1*, *sul2*, *sul3*, *tetA*, *tetC*, *tetD*, *tetT*, *tetL*, *tetM*, *tetS*, *tetX*, *qnrA*, *qnrB* and *qnrS* genes which encode proteins providing resistance to macrolides (*erm*), sulfonamides (*sul*), tetracyclines (*tet*) and fluoroquinolones (*qnr*) respectively. The *erm*A was the only gene for which all primer parameters equal 0% indicating the lowest possible statistics, whereas the remaining primers show 100% sequence specificity (S) to appropriate ARGs (Tab. 1). Moreover, the E rate ranges from 6.8% (*tetS*) to 100% (*tetC*, *sul1*), while for the other primers it equals: 37.5% (*tetA*), 87.45% (*ermB*), 39.64% (*tetD*), 44.58% (*sul3*), 44.81% (*tetL*), 55.4% (*tetM*), 59.76% (*tetX*), 64.0% (*tetA*), 66.67% (*ermC*), 67.86% (*ermF*), 68.0% (*qnrS*), 69.62% (*qnrB*), 82.84% (*qnrA*), 83.61% (*sul2*) and 87.45% (*ermB*). The TS parameter was higher than 50% in 15 out of 18 tested primer pairs. Among these were *tetA*, *tetC*, *tetD*, *tetT*, *tetM*, *tetX*, *qnrA*, *qnrB*, *qnrS*, *sul*1, *sul*2, *sul*3, *ermB*, *ermC* and *ermF* genes. The remaining primers (specific to *tetL*, *tet*S, *ermA* genes) showed TS below 50%. The conducted analysis clearly demonstrated that not all published PCR primers which should be specific to particular ARGs are in fact useful in environmental analyses, and especially as a screening tools. On the other hand, a created database of ARGs specific primers enables fast and easy selection of appropriate primers, with best ratings. Due to calculated parameters we obtained crucial information about primers usefulness in detection of antibiotic resistance genes, constituting abovementioned novel group of emerging pollutants. We believe that in environmental screening procedures it is extremely important to choose primers with the highest ratings, which allow to minimize false negative results (due to low efficacy and/or taxon specificity of chosen primers), but also false positive results (due to low specificity rate of primers).

3.2 Occurrence of antibiotics resistance genes in wastewater samples from Oswiecim (Poland) and various WWTPs metagenomes

To verify calculated statistics and test chosen PCR primers, an *in silico* analysis of 4 various WWTP metagenomes and an experimental analysis of samples collected from 3 subsequent stages of the waste purification process in the WWTP in Oswiecim (Poland) were conducted. The results obtained from analysis of metagenomes allowed preliminary insight into usefulness of tested primers, as well as the correlation of their statistics and detection of relevant ARGs. Results showed that detected ARGs were "*in silico* amplified" in ePCR only by primers with S, E and TS higher than 50% (Tab. 1). The analysis revealed the prevalence of *tetA*, *tetC*, *tetX*, *qnrS*, *sul1*, *sul2*, *ermB* and *ermF* genes. Interesting results were obtained for metagenome from ecological WWTP in Vermont (United Stated) where none of ARGs was detected. In the next step PCR reactions for 18 ARGs were performed with total DNAs isolated from collected in Oswiecim WWTP samples. The results showed that *tetA*, *tetX*, *qnrS*, *sul1* and *sul2* genes are prevalent in all tested probes (which stays in a good agreement with the results for the WWTPs metagenomes analysis). Moreover, *tetT* and *tetM* genes were detected exclusively in anaerobic digestion sludge while *qnrB* gene was identified only in activated sludge. Other resistance genes, i.e. *tetD*, *tetL*, *tetS*, *qnrA*, *sul*3, *ermA*, *erm*C were not detected in any stage of the Oswiecim WWTP. Those results show that at least one type of antibiotic resistance gene conferring resistance to each of analyzed antibiotic groups was detected in analyzed WWTP. All together obtained results stay in good agreement with the results of previous analyses showing that WWTPs are in fact reservoirs of antibiotic resistance genes and specific anthropogenic "bioreactors" in which emerging biological pollutants (i.e. antibiotic resistance genes) can evolve and easily spread among various bacterial hosts (Pruden *et al.*, 2006). Moreover, performed analyses clearly demonstrated that for the studies of the prevalence of antibiotic resistance genes in various environments and therefore assessing their epidemiological risk, only PCR primers with highest possible E, S and TS rates should be used, as only in those cases one can be sure that the performed analysis reflects the real state.

.

Table 1. Resistance genes detected in various metagenomes (ePCR) and samples from WWTP in Oswiecim (Poland)

gene name	gene product	rate [%]			*in silico* analysys				*in vivo* analysys		
		S	E	TS	M1	M2	M3	M4	PS	AS	ADS
tetA	MFS tetracycline efflux	100	64	78.9	+	-	+	-	+	+	+
tetC	MFS tetracycline efflux	100	100	88	+	-	+	-	+	+	+
tetD	MFS tetracycline efflux	100	39.6	54.1	-	-	-	-	-	-	-
tetT	ribosomal protection protein	100	37.5	50	-	-	-	-	-	-	+
tetL	ribosomal protection protein	100	44.8	43.4	-	-	-	-	-	-	-
tetM	ribosomal protection protein	100	55.4	72.4	-	-	+	-	-	-	+
tetS	ribosomal protection protein	100	6.8	18.7	-	-	-	-	-	-	-
tetX	inactivation of tetracycline	100	59.7	75	-	+	+	-	+	+	+
qnrA	DNA- gyrase and topoisomerase IV protection	100	82.8	93.7	-	-	-	-	-	-	-
qnrB	DNA- gyrase and topoisomerase IV protection	100	69.6	68.4	-	-	-	-	-	+	-
qnrS	quinolone resistance determinant	100	68	84.6	+	-	+	-	+	+	+
sul1	dihydropteroate synthetase	100	100	100	+	+	-	-	+	+	+
sul2	dihydropteroate synthetase	100	83.6	87.5	+	-	+	-	+	+	+
sul3	dihydropteroate synthetase	100	44.5	66.6	-	-	-	-	-	-	-
ermA	rRNA adenine N6-methyltransferase	0	0	0	-	-	-	-	-	-	-
ermB	rRNA adenine N6-methyltransferase	100	87.4	92.5	+	+	+	-	+	+	+
ermC	rRNA adenine N6-methyltransferase	100	66.6	54.5	-	-	-	-	-	-	-
ermF	rRNA adenine N6-methyltransferase	100	67.8	75	+	-	+	-	+	+	+

Abbreviations: S - specificity; E - efficacy; TS - taxon specificity; M1 - Nanjing (China); M2 - Shifflange (Luxemburg); M3 - Shanghai (China); M4 - Vermont (United States); PS - primary sludge (Poland); AS - activated sludge (Poland); ADS - anaerobic digestion sludge (Poland)

.

3.3 Microbial diversity at subsequent stages of the Oswiecim WWTP

Nucleotide sequences of the obtained PCR products were used as a query against GenBank database. Based on the sequence identity (best BLSAT hits), putative hosts of the PCR detected ARGs could be proposed. This was compared with the results of the V3-V4 region of 16S rRNA gene analysis (Fig. 1). Interestingly, majority of the best BLAST hits for the detected ARGs suggested that they originated from *Enterobacteriaceae*, which were identified exclusively in activated sludge (AS; 0.03%). This finding indicates that the identified in PCR screening ARGs rather have hosts belonging to other (than *Enterobacteriaceae*) taxonomic groups, and they could be transferred via horizontal gene transfer. It only emphasizes the role and need of antibiotic resistance analyses in WWTPs. It is also worth mentioning, that analysis of the correlation between bacterial diversity in studied samples and PCR primers ratings clearly demonstrated that for the most detailed studies the PCR primers with the highest TS rates should be chosen, as only in this case we can expect that any taxonomical group of bacteria will be omitted.

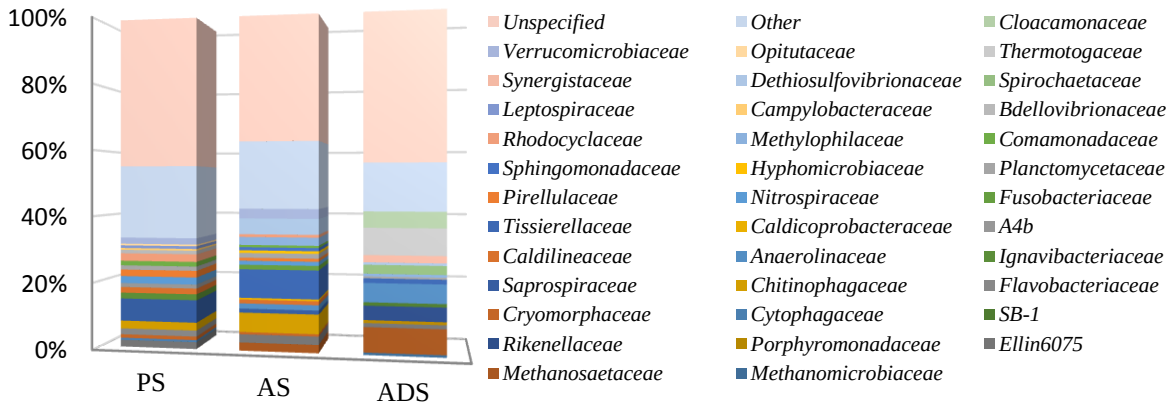

Figure 1. Family-level microbial diversity in subsequent stages of the process in the Oswiecim WWTP. Abbreviations: PS - primary sludge; AS - activated sludge; ADS - anaerobic digestion sludge.

Acknowledgments

This work was supported by the Gekon 2 (Generator of Ecological CONcepts) programme of the National Research and Development Centre and National Fund for Environmental Protection and Water Management [Project number GEKON2/O2/266405/7/2015]

References

Akiba, M., H. Senba, H. Otagiri, V. P. Prabhasankar, S. Taniyasu, N. Yamashita, K. Lee, T. Yamamoto, T. Tsutsui, D. Ian Joshua, K. Balakrishna, I. Bairy, T. Iwata, M. Kusumoto, K. Kannan and K. S. Guruge (2015), Impact of wastewater from different sources on the prevalence of antimicrobial-resistant *Escherichia coli* in sewage treatment plants in South India. *Ecotoxicol. Environ. Saf.*, **115**: 203-208.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer and R. Knight (2011), Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA*, **108 Suppl 1**: 4516-4522.

Cromwell, G. L. (2002), Why and how antibiotics are used in swine production. *Anim. Biotechnol.*, **13**(1): 7-27.

Davies, J. and D. Davies (2010), Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.*, **74**(3): 417-33.

Eliopoulos, G. M. and J. Blazquez (2003), Hypermutation as a factor contributing to the acquisition of antimicrobial resistance. *Clin. Infect. Dis.*, **37**(9): 1201-1209.

Li, D., C. M. Liu, R. Luo, K. Sadakane and T. W. Lam (2015), MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.*, **31**(10): 1674-6.

Luyt, C. E., N. Brechot, J. L. Trouillet and J. Chastre (2014), Antibiotic stewardship in the intensive care unit. *Crit. Care*, **18**(5): 480.

Piddock, L. J. (2015), Teixobactin, the first of a new class of antibiotics discovered by iChip technology? *J. Antimicrob. Chemother.*, **70**(10): 2679-80.

Pruden, A., R. Pei, H. Storteboom and K. H. Carlson (2006), Antibiotic resistance genes as emerging contaminants: studies in northern Colorado. *Environ. Sci. Technol.*, **40**(23): 7445-7450.

Rizzo, L., C. Manaia, C. Merlin, T. Schwartz, C. Dagot, M. C. Ploy, I. Michael and D. Fatta-Kassinos (2013), Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Sci. Total. Environ.*, **447**: 345-360.

Sambrook, J. and D. W. Russell (2001). Molecular cloning: A laboratory manual. New York, Cold Spring Harbor Laboratory Press.

van Hoek, A. H., D. Mevius, B. Guerra, P. Mullany, A. P. Roberts and H. J. Aarts (2011), Acquired antibiotic resistance genes: an overview. *Front. Microbiol.*, **2**: 203.

Ventola, C. L. (2015), The antibiotic resistance crisis: part 1: causes and threats. *P. T.*, **40**(4): 277-83.

von Wintersdorff, C. J. H., J. Penders, J. M. van Niekerk, N. D. Mills, S. Majumder, L. B. van Alphen, P. H. M. Savelkoul and P. F. G. Wolffs (2016), Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front. Microbiol.*, **7**: 173.

Wang, Q., G. M. Garrity, J. M. Tiedje and J. R. Cole (2007), Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*, **73**(16): 5261-5267.