DATABASE
The Journal of Biological Databases and Curation

Database, 2017, 1–7
doi: 10.1093/database/baw165
Original article

Original article

FARME DB: a functional antibiotic resistance element database

James C. Wallace, Jesse A. Port, Marissa N. Smith and Elaine M. Faustman*

Department of Environmental and Occupational Health Sciences, Institute for Risk Analysis and Risk Communication, University of Washington, Seattle, WA, USA

*Corresponding author: Tel: 1 206 685 2269; Fax: 206 685 4696; Email: faustman@u.washington.edu

Present Address: Jesse A. Port, Center for Ocean Solutions, Stanford University, Stanford, CA, USA.

Citation details: Wallace,J.C., Port,J.A., Smith,M.N. *et al.* FARME DB: a functional antibiotic resistance element database. *Database* (2016) Vol. 2016: article ID baw165; doi:10.1093/database/baw165

Received 11 February 2016; Revised 22 November 2016; Accepted 28 November 2016

Abstract

Antibiotic resistance (AR) is a major global public health threat but few resources exist that catalog AR genes outside of a clinical context. Current AR sequence databases are assembled almost exclusively from genomic sequences derived from clinical bacterial isolates and thus do not include many microbial sequences derived from environmental samples that confer resistance in functional metagenomic studies. These environmental metagenomic sequences often show little or no similarity to AR sequences from clinical isolates using standard classification criteria. In addition, existing AR databases provide no information about flanking sequences containing regulatory or mobile genetic elements. To help address this issue, we created an annotated database of DNA and protein sequences derived exclusively from environmental metagenomic sequences showing AR in laboratory experiments. Our Functional Antibiotic Resistant Metagenomic Element (FARME) database is a compilation of publically available DNA sequences and predicted protein sequences conferring AR as well as regulatory elements, mobile genetic elements and predicted proteins flanking antibiotic resistant genes. FARME is the first database to focus on functional metagenomic AR gene elements and provides a resource to better understand AR in the 99% of bacteria which cannot be cultured and the relationship between environmental AR sequences and antibiotic resistant genes derived from cultured isolates.

Database URL: http://staff.washington.edu/jwallace/farme

Introduction

Antibiotic resistance (AR) is a significant and growing public health risk with bacterial mutations and horizontal gene transfer increasingly compromising the efficacy of antibiotic drugs. The Center for Disease Control and Prevention recently issued a report describing AR as a 'serious threat' to public health and estimating that at least 2 million people acquire antibiotic resistant infections each

year (1). One of the four core actions highlighted in this report to prevent AR was tracking resistant bacteria; to do so requires consideration of both clinical and nonclinical (i.e. environmental) exposure pathways (2). In the environmental exposure pathway, antibiotic resistant bacteria are released into the environment through sewage, treated wastewater, agricultural run-off and other methods where they act as reservoirs for antibiotic resistant genes (3, 4). These reservoirs encourage the dissemination of antibiotic resistant genes between nonpathogenic and pathogenic bacteria via horizontal gene transfer (5, 6). In a recent review article, Bush et al. propose that controlling and preventing AR begins with understanding the development of AR in microorganisms found in the environment (7).

AR and the associated public health impacts have traditionally been studied using PCR and culture-based techniques. Because only a small fraction of microorganisms are estimated to be culturable (8, 9), these traditional methods misrepresent the AR potential of environmental microbial communities. Metagenomics has been shown to be a valuable approach for studying the prevalence of AR genes in the environment (10) and in contrast with PCR and culture based methods, allows for the characterization of the genetics of an entire microbial community. Metagenomic AR applications can be either sequence-based or functional, but both rely on DNA sequencing technologies. Sequence-based metagenomics involves the extraction and random sequencing of DNA directly from environmental media (e.g. water, soil, air etc.). Sequences are then compared with reference databases to assign taxonomic identity and functional potential. A diverse group of sequence-based metagenomic datasets are publicly available making global comparisons and analyses possible (11). In terms of AR, this extensive data has allowed for the development of a framework that integrates metagenomics into environmental AR risk assessment (12). However, without a laboratory-based confirmation of AR, it is not possible to determine functional AR through sequence-based metagenomics alone. Functional metagenomics similarly involves extracting DNA from environmental samples, but this metagenomic DNA is then cloned and expressed in a surrogate host (e.g. *Escherichia coli*) and screened for enzymatic activities such as AR (13). Resistant clones can then be selected and sequenced to identify nucleotide sequences conferring resistance, providing experimental validation of functional resistance and identification of novel ARGs (10). Despite the many studies that have used this valuable approach, there are currently no databases compiling functional metagenomic AR genes.

Current AR databases include the Antibiotic Resistance Genes Database (ARDB) (14) and the Comprehensive Antibiotic Resistance Database (CARD) (15). ARDB is an online tool that aims to provide a centralized resource to facilitate consistent characterization and identification of AR genes and contains ~3000 non-redundant AR genes (14). More recently, CARD was released with ~2200 non-redundant antibiotic resistant gene sequences (15). Both databases are assembled almost exclusively from genomic sequences derived from clinical bacterial isolates and include few, if any, functional metagenomic sequences.

In this article, we present the Functional Antibiotic Resistance Metagenomic Element (FARME) Database. FARME is the first database to focus on functional metagenomic AR gene elements rather than on individual antibiotic resistant genes derived from cultured clinical isolates. We have produced this database by compiling publically available DNA sequences from 20 functional metagenomics projects and their corresponding predicted protein sequences conferring AR. FARME also includes regulatory elements, mobile genetic elements and predicted proteins flanking AR genes. These features have been shown to be conserved between functional metagenomic AR sequences found in soil biomes and pathogenic clinical isolate sequences (16).

We have augmented authors' GenBank (17) annotations with BLAST analysis of DNA and predicted proteins by searching current GenBank non-redundant protein and DNA sequence repositories and annotated conserved protein domains using hidden Markov model (HMM) analysis with the Pfam (18) and Resfams (19) HMM databases. HMM analysis serves as a valuable complement to traditional local similarity searching revealing highly specific AR protein motif conservation suitable for high-resolution analysis and visualization (19).

In addition, a FARME DB website dashboard (http://staff.washington.edu/jwallace/farme) was created to provide interactive evaluation and visualization of FARME AR elements by AR category, biome type and geographic location.

Database Construction and Content

GenBank and MetaGeneMark annotation

Functional metagenomic DNA sequences and annotations from 20 individual functional metagenomics projects (16, 29, 38) were downloaded from GenBank (17) along with their protein sequence predictions and annotations (if available). Results were loaded into MySQL tables created for DNA sequences, protein sequences and HMM predictions (Figure 1).

DNA sequences for three projects without corresponding protein sequence predictions (16, 29, 38) were analyzed with MetaGeneMark software (39) using default settings

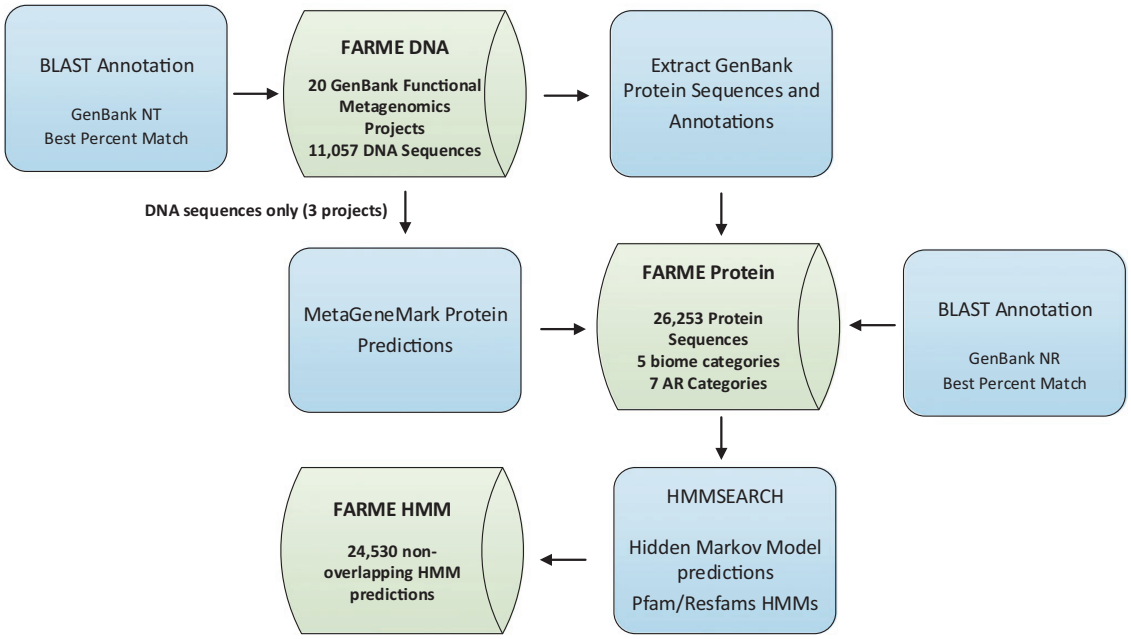


Figure 1. FARME database workflow showing analysis software components in blue and main database tables in green.

and predicted proteins were included in the FARME protein sequence table. Only sequences from uncultured metagenomic sources were considered excluding sequences derived from 'mixed' culture isolates.

GenBank and HMM sequence analysis annotation

Sequences were searched against the August 2015 GenBank non-redundant DNA and protein sequence databases using BLAST software (40) in order to annotate the best GenBank match with a minimum *e*-value threshold of 10^{-5} and excluding self-matches. The AR category for each DNA-sequenced clone was assigned using data from the methods section for each project.

Version 29 of the Pfam protein families database (18) and version 1.1 of the Resfams (19) AR HMM database were searched against the FARME predicted protein sequences FASTA file using the HMMSEARCH module of the HMMER version 3.1B2 software package (41) and the 'trusted cutoff' threshold. 'Trusted cutoff' is the score of the lowest scoring known true positive included in a full HMM profile alignment (42). Overlapping HMMs within the same protein region were resolved by choosing the HMM model with the highest HMMSEARCH bitscore for annotation. Non-overlapping HMMs within a predicted protein were annotated in the FARME HMM table along with the predicted gene sequence and position within the predicted gene. HMMs were designated as resistance elements for FARME website visualization based on their annotations in Pfam and Resfams databases. AR elements were partitioned into AR genes, transcriptional regulators and mobile genetic elements.

Web Framework

The FARME database website is built on a 'LAMP' (Linux, Apache, MySQL and PHP) open-source architecture. PHP scripts query MySQL tables returning JSON format data to Google Charts with Google Maps using AJAX (Asynchronous Javascript and XML) and HTML for geographical representation and drilldown into FARME projects as shown in Figure 2. Google map markers represent geographical sample collection coordinates, if available. Otherwise, markers represent project laboratory site coordinates.

The HMM visualization browser tool (Figure 3) uses custom Pfam Javascript libraries (http://pfam.xfam.org). DNA and protein browser table entries link to their respective NCBI GenBank records. The complete set of records for each project are loaded into the Google Charts interface featuring sortable individual table fields and searching using native browser search features. We have also provided a BLAST interface (40, 43) as part of the web infrastructure to allow user-provided DNA or protein sequences to be searched against the FARME DB programs. Compatible web browsers tested include Apple Safari, Microsoft Edge, Microsoft Internet Explorer, Google Chrome and Mozilla Firefox.

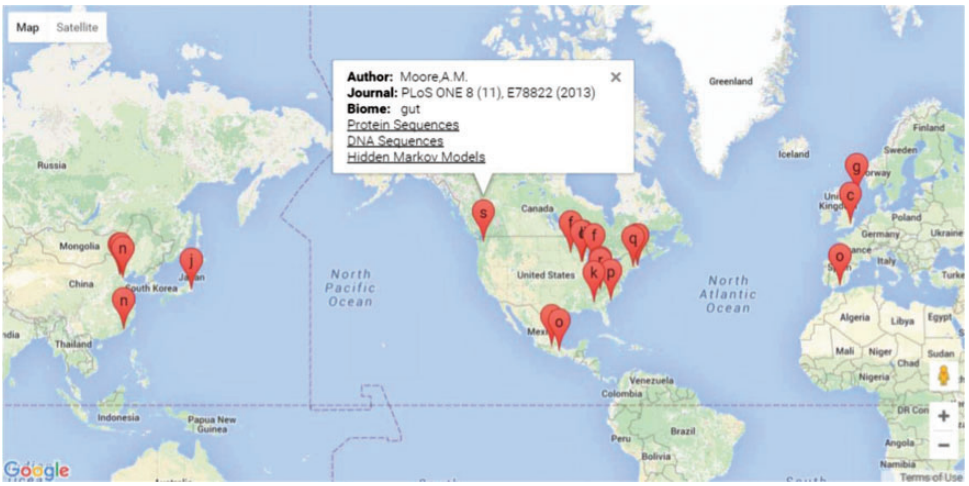
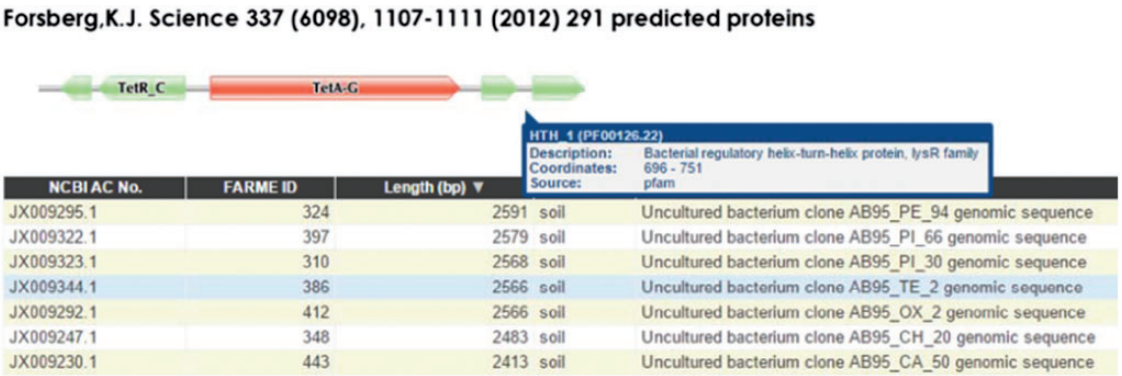


Figure 2. Screen shot of the FARME website project drill down feature showing geographical location of projects, if available, or project experiment sites with links to protein and DNA sequences and HMMs.



NCBI AC No.	FARME ID	Length (bp) ▼		
JX009295.1	324	2591	soil	Uncultured bacterium clone AB95_PE_94 genomic sequence
JX009322.1	397	2579	soil	Uncultured bacterium clone AB95_PI_66 genomic sequence
JX009323.1	310	2568	soil	Uncultured bacterium clone AB95_PI_30 genomic sequence
JX009344.1	386	2566	soil	Uncultured bacterium clone AB95_TE_2 genomic sequence
JX009292.1	412	2566	soil	Uncultured bacterium clone AB95_OX_2 genomic sequence
JX009247.1	348	2483	soil	Uncultured bacterium clone AB95_CH_20 genomic sequence
JX009230.1	443	2413	soil	Uncultured bacterium clone AB95_CA_50 genomic sequence

Figure 3. FARME website screen shot for a tetracycline resistant clone showing a tetracycline resistance HMM prediction flanked by TetR and LysR family transcriptional regulator HMMs. Users can interactively drill down into each sequence assembly for any of 20 FARME projects to help visualize genomic neighborhood HMM features including mouse over tooltips describing HMM feature details as shown above for LysR family transcriptional regulator.

Results

The FARME DNA table consists of 11,057 DNA sequences totaling 19,856,189 bases. The FARME protein table contains 26,253 corresponding predicted protein sequences with a total of 5,441,301 amino acids. The FARME HMM table is populated with 24,530 total non-overlapping HMM models and 2,250 different models found within 21,172 protein sequences. This includes 8,478 (35%) predicted AR HMMs, 1,369 (5.6%) transcriptional regulator HMMs and 360 (1.5%) mobile genetic element HMMs.

Figure 4 shows the AR categories present in FARME DB. Overall, functional AR gene elements including all predicted proteins were derived from two main biome types: soil and gut (fecal matter). Other biome types represented included wastewater treatment plants, oral and aquatic biomes.

Many FARME predicted proteins match known sequences in the GenBank protein sequence repository at a high percent identity: over 28% of FARME predicted proteins match sequences in GenBank at 100% identity and over two-thirds of FARME predicted genes (69%) match sequences in GenBank at >80% identity. Figure 5 shows the sequence similarity of 8,280 FARME predicted protein sequences containing AR resistance elements compared to GenBank and ARDB. FARME DB sequences have much higher similarity to GenBank sequences than to ARDB, thus illustrating the value of maintaining an up-to-date repository of functional metagenomic AR sequences not included in AR databases derived from cultured isolates. In

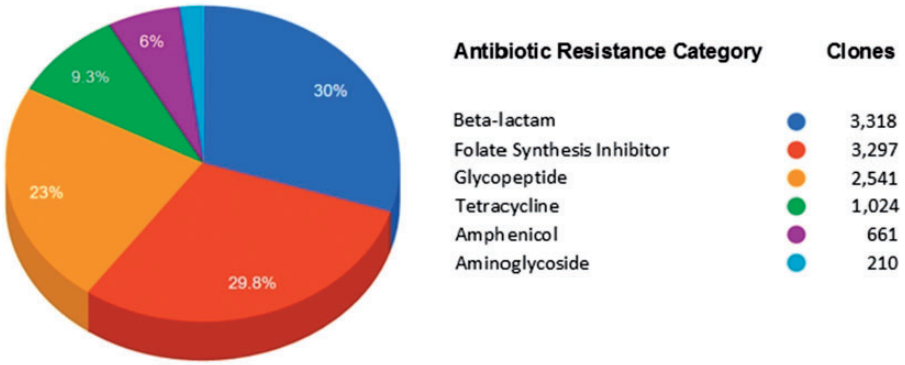


Figure 4. Main antibiotic categories used to select clones containing FARME DNA sequences for 20 individual projects.

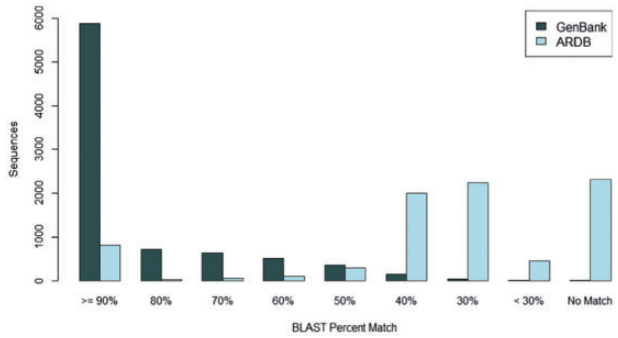


Figure 5. The number of FARME sequences within a percent identity bin from BLAST searching 8,280 FARME protein sequences containing AR resistance elements when compared with GenBank non-redundant protein and ARDB databases. FARME sequences show a dramatically higher percent identity with the GenBank non-redundant protein database than with the ARDB database illustrating the value of maintaining an up-to-date functional metagenomics AR database as a complement to clinically derived AR databases.

addition, there are 1,334 FARME protein sequences with <80% similarity to both GenBank and ARDB suggesting potentially novel AR elements derived from environmental samples.

Functional antibiotic resistant metagenomics clone sequences with little or no identity to GenBank sequences included in FARME have been recently shown to contain genes which confer AR via new resistance models. Allen *et al.* (44) isolated beta-lactamase resistant clones from pristine Alaskan soil containing no known AR genes. They discovered a DNA binding response regulator gene with a 57% amino acid identity in GenBank enhancing resistance to the Beta-Lactam carbenicillin in *E. coli*. Forsberg *et al.* (45) used functional metagenomics clones isolated from multiple soil types to discover a family of tetracycline resistance genes dubbed 'tetracycline destructases'. This new family of tetracycline genes shows no amino acid identity to known tetracycline resistance genes. However, tetracycline inactivating clones in this project share a common ResFams HMM motif, 'TE Inactivator', which is also contained within tetracycline resistant clones found in another FARME project which characterizes the pediatric gut resistome (29).

Discussion

FARME is the first database to focus on functional metagenomic AR genes and gene elements sourced from environmental samples rather than on individual antibiotic resistant genes derived from cultured clinical isolates. FARME contains over seven times the number of non-redundant protein sequences compared with other AR databases such as ARDB and CARD and includes essential information about the 'genomic neighborhood' proximal to functionally tested genes, e.g. well-known regulatory elements identified by HMM analysis (e.g. TetR, LysR) (16) (Figure 3). As such, FARME provides a basis for analyzing the sequence similarity between functional AR genes derived from the environment and future sequences from clinical and metagenomic studies. Although the majority of FARME sequences share similarity to known AR genes from clinical studies, a number of FARME sequences share

little to no similarity to known AR genes, suggesting novel AR genes from the environment.

Four projects out of 20 in FARME utilize high-throughput second-generation DNA sequence technology (16, 24, 25, 29) and one of the newest projects (36) utilizes third-generation sequencing which combines high-throughput with long-read sequencing reads (>7 kb). In the future, next-generation sequencing (NGS) will provide greater throughput, lower cost and higher resolution for functional metagenomic sequencing experiments. Adoption of NGS will necessitate new search strategies providing maximum speed, sensitivity and specificity for gene annotation. As such, researchers utilizing NGS will require easy to use analysis frameworks and up-to-date databases like FARME to help achieve goals such as providing timely global AR surveillance. In addition, leveraging HMMs to identify specific AR genes and mechanisms of action within predicted protein sequences offers the potential of high-performance screening of thousands of metagenomic sequences in a fraction of the time taken by traditional similarity searching methods. Recent improvements in the Hidden-Markov model software HMMER, used to generate the FARME HMM table in our database, can provide search speeds four orders of magnitude faster than BLAST (46).

In summary, FARME is the first AR database to focus exclusively on environmentally derived metagenomic genes, and as such provides an opportunity for researchers to access and analyze AR genes found outside of the clinical setting. The database, annotation schema and browser interface provide a valuable and needed resource to better understand and characterize AR elements in the majority of uncultured bacteria and their genetic similarity to AR elements derived from cultured isolates.

Funding

This work was supported by the National Oceanic and Atmospheric Administration (NOAA)-funded Pacific Northwest Consortium for Pre- and Post-doctoral Traineeships in Oceans and Human Health [grant number S08-67883 MOD03] and was also supported by the National Science Foundation (NSF) (grant numbers 0910624 and 1128883). This publication was also made possible by United States Environmental Protection Agency (US EPA) grant 8357380. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NOAA, NSF or the US EPA. Further, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication. This work was also supported by The University of Washington, Center for Exposures, Diseases, Genomics and Environment, of the National Institutes of Health under award number: P30ES007033.

Conflict of interest. None declared.

References

1. Center for Disease Control and Prevention. (2013) Antibiotic resistance threats in the United State. http://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf
2. Allen,H.K., Donato,J., Wang,H.H. *et al*. (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol*., 8, 251–259.
3. Baquero,F., Martínez,J.L., and Cantón,R. (2008) Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol*., 19, 260–265.
4. Martínez,J.L. (2008) Antibiotics and antibiotic resistance genes in natural environments. *Science*, 321, 365–367.
5. Frost,L.S., Leplae,R., Summers,A.O., and Toussaint,A. (2005) Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol*., 3, 722–732.
6. Wright,G.D. (2010) Antibiotic resistance in the environment: a link to the clinic?. *Curr. Opin. Microbiol*., 13, 589–594.
7. Bush,K., Courvalin,P., Dantas,G. *et al*. (2011) Tackling antibiotic resistance. *Nat. Rev. Microbiol*., 9, 894–896.
8. Amann,R.I., Ludwig,W., and Schleifer,K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Revi*., 59, 143–169.
9. Kolar,M., Urbanek,K., and Latal,T. (2001) Antibiotic selective pressure and development of bacterial resistance. *Int. J. Antimicrob. Agents*, 17, 357–363.
10. Schmieder,R., and Edwards,R. (2012) Insights into antibiotic resistance through metagenomic approaches. *Future Microbiol*., 7, 73–89.
11. Meyer,F., Paarmann,D., D'souza,M. *et al*. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9, 386.
12. Port,J.A., Cullen,A.C., Wallace,J.C. *et al*. (2013) Metagenomic frameworks for monitoring antibiotic resistance in aquatic environments. *Environ. Health Perspect*, 122(3), 222–228.
13. Mullany,P. (2014) Functional metagenomics for the investigation of antibiotic resistance. *Virulence*, 5, 14–13.
14. Liu,B., and Pop,M. (2009) ARDB–antibiotic resistance genes database. *Nucleic Acids Res*., 37, D443–D447.
15. McArthur,A.G., Waglechner,N., Nizam,F. *et al*. (2013) The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother*., 57, 3348–3357.
16. Forsberg,K.J., Reyes,A., Wang,B. *et al*. (2012) The shared antibiotic resistome of soil bacteria and human pathogens. *Science*, 337, 1107–1111.
17. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J. *et al*. (2008) GenBank. *Nucleic Acids Res*., 36, D25–D30.
18. Bateman,A., Coin,L., Durbin,R. *et al*. (2004) The Pfam protein families database. *Nucleic Acids Res*., 32, D138–D141.
19. Gibson,M.K., Forsberg,K.J., and Dantas,G. (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *isme J*., 9, 207–216.
20. Allen,H.K., Moe,L.A., Rodbumrer,J. *et al*. (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *isme J*., 3, 243–251.
21. Cheng,G., Hu,Y., Yin,Y. *et al*. (2012) Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol. Lett*., 336, 11–16.

22. Diaz-Torres,M.L., McNab,R., Spratt,D.A. *et al.* (2003) Novel tetracycline resistance determinant from the oral metagenome. *Antimicrob. Agents Chemother.*, 47, 1430–1432.
23. Donato,J.J., Moe,L.A., Converse,B.J. *et al.* (2010) Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. *Appl. Environ. Microbiol.*, 76, 4396–4401.
24. Forsberg,K.J., Patel,S., Gibson,M.K. *et al.* (2014) Bacterial phylogeny structures soil resistomes across habitats. *Nature*, 509, 612–616.
25. Gibson,M.K., Wang,B., Ahmadi,S. *et al.* (2016) Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.*, 1, 16024.
26. Kazimierczak,K.A., Scott,K.P., Kelly,D., and Aminov,R.I. (2009) Tetracycline resistome of the organic pig gut. *Appl. Environ. Microbiol.*, 75, 1717–1722.
27. Lang,K.S., Anderson,J.M., Schwarz,S. *et al.* (2010) Novel florfenicol and chloramphenicol resistance gene discovered in Alaskan soil by using functional metagenomics. *Appl. Environ. Microbiol.*, 76, 5321–5326.
28. Lopez-Perez,M., Mirete,S., Jardon-Valadez,E., and Gonzalez-Pastor,J.E. (2013) Identification and modeling of a novel chloramphenicol resistance protein detected by functional metagenomics in a wetland of Lerma, Mexico. *Int. Microbiol.*, 16, 103–111.
29. Moore,A.M., Patel,S., Forsberg,K.J. *et al.* (2013) Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. *PloS One*, 8, e78822.
30. Mori,T., Mizuta,S., Suenaga,H., and Miyazaki,K. (2008) Metagenomic screening for bleomycin resistance genes. *Appl. Environ. Microbiol.*, 74, 6803–6805.
31. Riesenfeld,C.S., Goodman,R.M., and Handelsman,J. (2004) Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol*, 6, 981–989.
32. Sommer,M.O., Dantas,G., and Church,G.M. (2009) Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science*, 325, 1128–1131.
33. Su,J.Q., Wei,B., Xu,C.Y. *et al.* (2014) Functional metagenomic characterization of antibiotic resistance genes in agricultural soils from China. *Environ. Int.*, 65, 9–15.
34. Torres-Cortes,G., Millan,V., Ramirez-Saad,H.C. *et al.* (2011) Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. *Environ. Microbiol.*, 13, 1101–1114.
35. Uyaguari,M.I., Fichot,E.B., Scott,G.I., and Norman,R.S. (2011) Characterization and quantitation of a novel beta-lactamase gene found in a wastewater treatment facility and the surrounding coastal ecosystem. *Appl. Environ. Microbiol.*, 77, 8226–8233.
36. Wichmann,F., Udikovic-Kolic,N., Andrew,S., and Handelsman,J. (2014) Diverse antibiotic resistance genes in dairy cow manure. *mBio*, 5, e01017.
37. Zhou,W., Wang,Y., and Lin,J. (2012) Functional cloning and characterization of antibiotic resistance genes from the chicken gut microbiome. *Appl. Environ. Microbiol.*, 78, 3028–3032.
38. Parsley,L.C., Consuegra,E.J., Kakirde,K.S. *et al.* (2010) Identification of diverse antimicrobial resistance determinants carried on bacterial, plasmid, or viral metagenomes from an activated sludge microbial assemblage. *Appl. Environ. Microbiol.*, 76, 3753–3757.
39. Zhu,W., Lomsadze,A., and Borodovsky,M. (2010) Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.*, 38, e132.
40. Altschul,S.F., Madden,T.L., Schaffer,A.A. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 25, 3389–3402.
41. Eddy,S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput. Biol.*, 7, e1002195.
42. Eddy,S. (2010) HMMER User's Guide Version 3.0 http://hmmer.org.
43. Deng,W., Nickle,D.C., Learn,G.H. *et al.* (2007) ViroBLAST: a stand-alone BLAST web server for flexible queries of multiple databases and user's datasets. *Bioinformatics*, 23, 2334–2336.
44. Allen,H.K., An,R., Handelsman,J., and Moe,L.A. (2015) A response regulator from a soil metagenome enhances resistance to the beta-lactam antibiotic carbenicillin in Escherichia coli. *PLoS One*, 10, e0120094.
45. Forsberg,K.J., Patel,S., Wencewicz,T.A., and Dantas,G. (2015) The tetracycline destructases: a novel family of tetracycline-inactivating enzymes. *Chem. Biol.*, 22, 888–897.
46. Sunagawa,S., Mende,D.R., Zeller,G. *et al.* (2013) Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods*, 10, 1196–1199.