Detection of antimicrobial resistance genes in urban air

Ágnes Becsei [a], Norbert Solymosi [b], István Csabai [a], Donát Magyar [c]

a Department of Physics of Complex Systems, Eötvös Loránd University, Pázmány Péter stny. 1/A, 1117, Budapest, Hungary
b Centre for Bioinformatics, University of Veterinary Medicine, István u. 2., 1078, Budapest, Hungary
c National Public Health Center, Albert Flórián út 2-6, 1097, Budapest, Hungary

Corresponding author:
Norbert Solymosi
E-mail: solymosi.norbert@gmail.com

ABSTRACT

To understand antibiotic resistance in pathogenic bacteria, we need to monitor environmental microbes as reservoirs of antimicrobial resistance genes (ARGs). These bacteria are present in the air and can be investigated with the whole metagenome shotgun sequencing approach. This study aimed to investigate the feasibility of a method for metagenomic analysis of microbial composition and ARGs in the outdoor air. Air samples were collected with a Harvard impactor in the PM_{10} range at 50 m from a hospital in Budapest. From the DNA yielded from samples of PM_{10} fraction single-end reads were generated with an Ion Torrent sequencer. During the metagenomic analysis, reads were classified taxonomically. The core bacteriome was defined. Reads were assembled to contigs and the ARG content was analyzed. The dominant genera in the core bacteriome were *Bacillus, Acinetobacter, Leclercia* and *Paenibacillus.* Among the identified ARGs best hits were *vanRA, Bla1, mphL, Escherichia coli* EF-Tu mutants conferring resistance to Pulvomycin*, BcI, FosB,* and *mphM*. Despite the low DNA content of the samples of PM_{10} fraction, the number of detected airborne ARGs was surprisingly high.

Keywords: air metagenomics, Harvard impactor, PM, bacteriome, ARG

1 Introduction

Antimicrobial resistance (AMR) is one of the top global public health threats worldwide. Widespread misuse and overuse of antimicrobial drugs are accelerating the evolution and selection of naturally occurring antimicrobial resistance genes (ARGs) in bacteria. (Vikesland et al., 2019) In hospitals drug-resistant pathogenic bacteria are widespread, however, environmental microbes in the soil, water, air, or animal microbiome act as ample reservoirs of ARGs accumulated over time. The current understanding of AMR is derived from culture-based and phenotypic methods. These methods only aim at a few mostly pathogenic bacteria and do not detect the majority of ARGs in environmental microbes. (Be et al., 2014) Culture-independent

approaches, like 16S rRNA gene sequencing, provide taxonomic identification usually only on the genus level. Quantitative real-time PCR analysis of ARGs only targets a limited number of genes. (Echeverria-Palencia et al., 2017; Fluit et al., 2001) Whole-genome shotgun sequencing provides suitable data for metagenomic analysis. There are many metagenomic studies on the investigation of bacteria and genes in samples like water or soil, but there are only a few studies analyzing air samples (Aalismail et al., 2019; Be et al., 2014; Be et al., 2017, King et al., 2016) Airborne microbes carrying ARGs are attached to solid particles and liquid droplets constituting a mixture called particulate matter (PM). PM circulates in the air for a long time and travels long distances. These particles vary in size, therefore the health risk of $PM_{2.5}$ and PM_{10} is significant because these sizes correspond to human inhalable particles (< 10 μm; Liu et al., 2018) Usually collecting a proper amount of genetic material for air metagenomics studies is challenging. However, there is a wide range of sampling and sequencing methods with various efficacy available for air metagenomics studies, but without standards and best practices interpretation across these studies is difficult. In this study, we aim to investigate the sensitivity of a method to detect airborne ARGs and to examine the airborne microbial community in an outdoor urban environment.

2 Materials and methods

Three air samples were collected with a Harvard impactor (Marple et al., 1987) in the PM_{10} size range, with a flow rate of 10L/min, a total of 44.4 m^3 of air onto 37 mm fiberglass filters (Whatman GF/A, GE Healthcare Life Science) for 3×5 days in 2019. September. The samples were collected in the open air, 50 m from the entrance of a hospital in Budapest, Hungary, which specialized in the treatment of infectious diseases. Total DNA was then extracted from the pellets using the ZR Fecal DNA Kit (Zymo Research). From the yielded DNA in samples, single-end reads were generated by an Ion Torrent Sequencer. Quality-based filtering was performed by Trimmomatic (Bolger et al., 2014) with 20 as a quality threshold for bases and with retaining reads with a minimum length of 50 bp. Replicates were removed by vsearch 2.14.2. (Rognes et al., 2016) Filtered reads were taxonomically classified by Kraken 2 (Wood et al., 2019) (k=35) using the NCBI non-redundant nucleotide database (Pruitt et al., 2005). Bacterial reads were assembled by metaSPADES 3.14.1 (Nurk et al., 2017) with an automatically estimated maximum k-mer size of 127. Protein sequences of open reading frames (ORFs) were predicted by Prodigal setting "meta" mode for metagenome. The ARG content of the ORFs was identified by the Resistance Gene Identifier (RGI) v5.10 (Alcock et al., 2019) using "The Comprehensive Antibiotic Resistance Database" (CARD) v.3.0.6. (Alcock et al., 2019) 'Perfect' hits are protein sequences with 100% match to CARD reference sequences, while the 'strict' category is more flexible allowing some variation from the CARD reference sequence. 'Loose' hits fall out of the detection model cut-offs. All 'Loose' hits with identity ≥95% were nudged to the 'strict' category. All ARGs presented here are classified as "strict" hits. Contigs associated with ARGs with 'strict' or 'perfect' cut-offs were taxonomically classified using Kraken 2 the same way as described above.

3 Results and discussion

3.1 Bacteriome

In this study, sampling procedures for all three samples were alike, except that the samples were collected one after the other indicating a probable diversity of PM content and composition. The average PM_{10} concentration in sample 1 was 24.4 μg/m^3, 25.36 μg/m^3 in sample 2, and 42.9 μg/m^3 in sample 3. A study focusing on bacteria in aerosols showed strong fluctuations that correlated significantly with changes in seasonal temperatures. (Ravva et al., 2012) In another study, relative

humidity and PM_{10} were the key factors that significantly affected the airborne bacterial concentration and community structure. (Li et al., 2018) Therefore, it is reasonable to assume that these were the factors that caused large fluctuations in abundances of airborne particulate matter and thereby cause fluctuations in the concentration of bacteria in our samples.

Figure 1.

Sequencing resulted in 855,654 single-end reads in sample 1, 2,290,392 reads in sample 2, and 527,221 reads in sample 3. By prefiltering steps, 19.57% of sample 1, 21.92% of sample 2, and 42.95% of sample 3 were discarded. Taxonomic classification was successful with 95.52% of the reads in sample 1, 94.11% in sample 2, and 84.79% in sample 3. Taxon classification of reads revealed that most classified reads are aligned to bacterial genomes. Dominant phyla were Firmicutes and Proteobacteria, which are rather common in air samples (Aalismail et al., 2019; Be et al., 2014; Yooseph et al., 2013). The most abundant genera are the *Bacillus*, *Acinetobacter*, *Leclercia* and *Paenibacillus*. (Figure 1.) Members of the genus *Bacillus* are among the most abundant in sample 1 and sample 2 with species of the *Bacillus cereus* group. (Figure 1.) They were also the most abundant inhabitants of urban air in another study. (Be et al., 2014)

In sample 3 the most abundant genus is the *Acinetobacter* genus. (Figure 1.) It is a highly diverse group of mostly non-pathogenic environmental microbes isolated from samples like soil and wastewater. They often carry several ARGs including carbapenemases and extended-spectrum beta-lactamases. (Wong et al., 2017) Other abundant genera in core bacteriome are *Atlantibacter, Citrobacter, Enterobacter, Klebsiella,* and *Pseudoescherichia* (Figure 1.) belonging to the diverse family *Enterobacteriaceae.* (Morales-López et al., 2019)

3.2 Antimicrobial resistance genes

The total number of assembled contigs in sample 1 was 7613, 1137 in sample 2, and 235 in sample 3. In these contigs, 12 ARGs were identified in sample 1, 13 ARGs in sample 2, and 1 ARG in sample 3. The median lengths of assembled contigs are 660 (IQR:510) in sample 1, 698 (IQR:644.5) in sample 2, and 505 (IQR:104.5) in sample 3. The mean coverage of the listed ARGs in sample 1 is 43.27% with a range of 3.91% to 102.27% and mean identity is 97.74% with a range of 87.71% to 100%. In the sample, 2 coverage of ARG hits ranged between 7.46% and 108.7% with a mean value of 49.65%. The range of identity values is between 90.91% and 100% with a mean value of 96.57%. Sample 3 resulted in only one ARG with 100% identity and 3.97% coverage.

Figure 2.

Among the best hits in samples, *vanRA* (Figure 2.) o*r vans* with a *vanS* protein is part of the regulatory system of the *vanA* resistance gene cluster responsible for peptidoglycan target alteration of the glycopeptide antibiotic, vancomycin. (Courvalin, 2006)

Bla1 and *BcI* are beta-lactamase genes detected with coverage and identity values near 90%. (Figure 2.) *Bla1* codes a penicillinase, first recognized in *Bacillus anthracis*. (Materon et al., 2003) *BcI* codes a zinc Metallo-beta-lactamase associated with *Bacillus cereus* that hydrolyses many penicillins including carbapenems which generally escape from serine beta-lactamases. (Carfi et al., 1995)

With high coverage and identity values, *mphL* and *mphM* (Figure 2.) are expressed as macrolide phosphotransferases which are highly prevalent in members of the *Bacillus cereus* group. (Wang et al., 2015)

Another ARG near 90% coverage and identity is *Escherichia coli* EF-Tu mutants conferring resistance to Pulvomycin. (Figure 2.) Pulvomycin inhibits protein synthesis by acting on elongation factor Tu (EF-Tu). In *E. coli* EF-Tu is very sensible to pulvomycin, but membrane impermeability of Gram-negative bacteria prohibits several antibiotics including pulvomycin to enter the cell. Maybe as a second line of defense in case of increased permeability, non-sensitive EF-Tu mutants are more capable. (Zeef et al., 1994)

The protein coded by *FosB* (Figure 2.) is a Mn^{2+}-dependent enzyme that modifies fosfomycin to a compound with no bactericidal properties. (Thompson et al., 2013)

Other hits with lower coverage and identity values are probably variants of ARG reference sequences in CARD. (Figure 2.)

Figure 3.

In sample 1 33% of the detected ARGs probably originated from the genus *Paenibacillus* and 58.3% from the genus *Bacillus*. In sample 2 revealed ARGs probably originated from *Acinetobacter* (46%), *Bacillus* (38%) genera, and the family *Enterobacteriaceae* (15%). (Figure 3.)

Non-culture-based methods are dominated by PCR techniques (Li et al., 2018; Hu et al., 2018; Xie et al., 2018) and there are only a few studies on airborne ARGs revealed by metagenomic analysis (Fondi et al., 2016; Pal et al., 2016). One of the studies on the investigation of 39 ARG subtypes by PCR in the air of 19 global cities revealed that the most abundant ARGs provide resistance against beta-lactams, quinolones, macrolides, aminoglycosides, and vancomycin. (Li et al., 2018) Similarly, in our result, most ARGs are involved in protection against beta-lactams, tetracyclines, quinolones, macrolides, and diaminopyrimidines. (Figure 3.) Although there were several hospitals with intensive use of antibiotics in the vicinity of the sampling site, the closest one specialized in for treatment of infectious diseases. No contigs containing ARGs associated with pathogenic species or strains were identified in any of the air samples. Environmental bacteria as potential reservoirs of ARGs are worth further investigation as developing a standardized method for the metagenomic analysis of airborne samples would help to estimate the public health risks of airborne ARGs.

Author contributions

Ágnes Becsei: Formal analysis (equal); software (equal); Visualization (equal); Writing - original draft (equal); **Norbert Solymosi**: Conceptualization (equal); Formal analysis (equal); Methodology (equal); Software (equal); Writing-review & editing (equal); **István Csabai**: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Supervision (equal); Writing-review & editing (equal); **Donát Magyar**: Conceptualization (equal); Investigation (equal); Methodology (equal); Supervision (equal); Writing-review & editing (equal).

Acknowledgments

This work was financed by EU Horizon 2020 program Grant agreement “VEO” No. 874735. The authors are thankful to Dr. Tamás Pándics, Dr. Tamás Szigeti and Dr. Tamás Tirczka for their suggestions, and Boglárka Sára Balogh, Vivien Pál, Gábor Résі and Zsófia Tischner for their assistance in sample collection (National Public Health Center, Budapest, Hungary).

Ethics statement

None required.

Conflict of interest

None declared.

Data availability statement

The datasets generated and analyzed during the current study are available in the NCBI repository at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA747808

References

1. Aalismail, N. A., Ngugi, D. K., Díaz-Rúa, R., Alam, I., Cusack, M. and Duarte, C. M. (2019) Functional metagenomic analysis of dust-associated microbiomes above the Red Sea. *Sci Rep* **9 (1)**, 13741. https://doi.org/10.1038/s41598-019-50194-0
2. Alcock, B. P., Raphenya A. R., Lau T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A. *et al.* (2019) CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research* **48 (D1)**, D517–D525. https://doi.org/10.1093/nar/gkz935
3. Be, N. A., Avila-Herrera, A., Allen, J. E., Singh, N., Sielaff, C., Jaing, C. *et al.* (2017) Whole metagenome profiles of particulates collected from the international space station. *Microbiome* **5 (1),** 81. https://doi.org/10.1186/s40168-017-0292-4
4. Be, N. A., Thissen, J. B., Fofanov, V., Y., Allen, J. E., Rojas, M., Golovko, G. *et al.* (2014) Metagenomic analysis of the airborne environment in urban spaces. *Microbial ecology* **69 (2),** 346-55. https://doi.org/10.1007/s00248-014-0517-z
5. Carfi, A., Pares, S., Duée, E., Galleni, M., Duez, C., Frère, J. M. and Dideberg, O. (1995) The 3-D structure of a zinc metallo-beta-lactamase from bacillus cereus reveals a new type of protein fold. *The EMBO journal* **14 (20)**, 4914—4921.
6. Bolger, A. M., Lohse, M. & Usadel, B. (2014) Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30 (15)**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170
7. Courvalin, P. (2006) Vancomycin resistance in gram-positive cocci. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **42 Suppl 1**, S25–34. https://doi.org/10.1086/491711
8. Echeverria-Palencia, C. M., Thulsiraj, V., Tran, N., Ericksen, C. A., Melendez, I., Sanchez, M. G. *et al.* (2017) Disparate antibiotic resistance gene quantities revealed across 4 major cities in California: A survey in drinking water, air, and soil at 24 public parks. *ACS Omega* **2 (5),** 2255–2263. https://doi.org/10.1021/acsomega.7b00118
9. Fluit, A. D. C., Visser, M. and Schmitz, F.-J. (2001) Molecular detection of antimicrobial resistance. *Clinical microbiology reviews* **14 (4),** 836–71. https://doi.org/10.1128/microbiolspec.ARBA-0011-2017
10. Fondi, M., Karkman, A., Tamminen, M. V., Bosi, E., Virta, M., Fani, R. *et al.* (2016) 'Every gene is everywhere but the environment selects': Global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biology and Evolution* **8 (5)**, 1388–1400. https://doi.org/10.1093/gbe/evw077
11. Hu, J., Zhao, F., Zhang, X., Li, K., Li, C., Ye, L. *et al.* (2018) Metagenomic profiling of ARGs in airborne particulate matters during a severe smog event. *Science of The Total Environment* **615**, 1332–1340. https://doi.org/10.1016/j.scitotenv.2017.09.222

12. King, P., Pham, L. K., Waltz, S., Sphar, D., Yamato, R. T., Conrad, D. *et al.* (2016) Longitudinal metagenomic analysis of hospital air identifies clinically relevant microbes. *PLOS ONE* **11 (8)**, 1–14. https://doi.org/10.1371/journal.pone.0160124
13. Li, J. Cao, J., Zhu, Y., Chen, Q. Shen, F., Wu, Y. *et al.* (2018) Global survey of antibiotic resistance genes in air. *Environmental science & technology* **52 (19)**, 10975–10984. https://doi.org/10.1021/acs.est.8b02204
14. Liu, H., Zhang, X., Zhang, H., Yao, X., Zhou, M., Wang, J. *et al.* (2018) Effect of air pollution on the total bacteria and pathogenic bacteria in different sizes of particulate matter. *Environmental pollution* **233**, 483—493. https://doi.org/10.1016/j.envpol.2017.10.070
15. Marple, V. A., Rubow, K. L., Turner, W. and Spengler, J. D. (1987) Low flow rate sharp cut impactors for indoor air sampling: Design and calibration. *JAPCA* **37 (11)**, 1303–1307. https://doi.org/10.1080/08940630.1987.10466325
16. Materon, I. C., Queenan, A. M., Koehler, T. M., Bush, K. and Palzkill, T. (2003) Biochemical characterization of β-lactamases Bla1 and Bla2 from bacillus anthracis. *Antimicrobial Agents and Chemotherapy* **47 (6)**, 2040–2042. https://doi.org/10.1128/AAC.47.6.2040-2042.2003
17. Morales-López, S., Yepes, J., Prada-Herrera, J. and Torres-Jiménez, A. (2019) Enterobacteria in the 21st century: A review focused on taxonomic changes. *The Journal of Infection in Developing Countries* **13 (4)**, 265–273. https://doi.org/10.3855/jidc.11216
18. Nurk, S., Meleshko, D., Korobeynikov, A. and Pevzner, P. (2017) metaSPAdes: A new versatile de novo metagenomics assembler. *Genome Research* **27 (5)**, 824–834. https://doi.org/10.1101/gr.213959.116
19. Pal, C., Bengtsson-Palme, J., Kristiansson, E. and Larsson, J. (2016) The structure and diversity of human, animal and environmental resistomes. *Microbiome* **4 (1)**, 54. https://doi.org/10.1186/s40168-016-0199-5
20. Pruitt, K. D., Tatusova, T. and Maglott, D. R. (2005) NCBI reference sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *NUCLEIC ACIDS RES* **33 (Database issue)**, 501–504. https://doi.org/10.1093/nar/gkl842
21. Ravva, S. V., Hernlem, B. J., Sarreal, C. Z. and Mandrell, R. E. (2012) Bacterial communities in urban aerosols collected with wetted-wall cyclonic samplers and seasonal fluctuations of live and culturable airborne bacteria. *J. Environ. Monit.* **14 (2)**, 473–481. https://doi.org/10.1039/C1EM10753D
22. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016) VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **4,** e2584 https://doi.org/10.7717/peerj.2584
23. Thompson, M. K., Keithly, M. E., Harp, J., Cook, P. D., Jagessar, K. L., Sulikowsky, G. A. *et al.* (2013) Structural and chemical aspects of resistance to the antibiotic fosfomycin conferred by FosB from bacillus cereus. *Biochemistry* **52 (41)**, 7350–7362. https://doi.org/10.1021/bi4009648
24. Vikesland, P., Garner, E., Gupta, S., Kang, S., Maile-Moskowitz, Ayella and Zhu, Ni. (2019) Differential drivers of antimicrobial resistance across the world. *Accounts of Chemical Research* **52 (4),** 916–924. https://doi.org/10.1021/acs.accounts.8b00643
25. Wang, C., Sui, Z., Leclercq, S. O., Zhang, G., Zhao, M., Chen W. *et al.* (2015) Functional characterization and phylogenetic analysis of acquired and intrinsic macrolide phosphotransferases in the bacillus cereus group. *Environmental microbiology* **17 (5)**, 1560—1573. https://doi.org/10.1111/1462-2920.12578
26. Wong, D. Nielsen T. B., Bonomo, R. A., Pantapalangkoor, P., Luna, B. and Spellberg, B. (2017) Clinical and Pathophysiological Overview of Acinetobacter Infections: a Century of Challenges. Clinical M*icrobiology Reviews* **30 (1)**, 409–447. https://doi.org/10.1128/CMR.00058-16
27. Wood, D., Lu, J. and Langmead, B. (2019) Improved metagenomic analysis with kraken 2. *Genome Biology* **20 (1),** 257 https://doi.org/10.1186/s13059-019-1891-0

28. Xie, J., Jin, L., Luo, X., Zhao, Z. and Li, X. (2018) Seasonal disparities in airborne bacteria and associated antibiotic resistance genes in $PM_{2.5}$ between urban and rural sites. *Environmental Science & Technology Letters* **5 (2)**, 74–79. https://doi.org/10.1021/acs.estlett.7b00561
29. Yooseph, S., Andrews-Pfannkoch, C., Tenney, A., McQuaid J., Williamson, S., Thiagarajan, M. *et al.* (2013) A metagenomic framework for the study of airborne microbial communities. *PLOS ONE* **8 (12)**, 1–13. https://doi.org/10.1371/journal.pone.0081862
30. Zeef, L. A., Bosch, L., Anborgh, P. H., Cetin, R., Parmeggiani, A. and Hilgenfeld, R. (1994) Pulvomycin-resistant mutants of E. coli elongation factor Tu. *The EMBO journal* **13 (21)**, 5113—5120.

Figure legends

Figure 1. Relative abundances of the most common bacterial genera. Core bacteriome was defined as the relative abundance of agglomerated counts at the genus level above 0.0005 and with prevalence above 0.6. Core bacteriome is dominated by the genera *Bacillus*, *Acinetobacter*, *Leclercia* and *Paenibacillus*.

Figure 2. Identity and coverage of detected ARGs. All ARGs presented here are classified as "strict" hits. Different symbols stand for nudged or non-nudged hits from the 'loose' to 'strict' category. The color of the symbols corresponds to the percentage of identity of the top ARG hit. The size of the points corresponds to the ratio of length between contig and the CARD reference sequence. 'EF-Tu mutants' refers to *Escherichia coli* EF-Tu mutants conferring resistance to pulvomycin and 'R. fascians cmr' to *Rhodococcus fascians cmr*.

Figure 3. Identified ARGs, their most probable taxon of origin, and the group of antibiotics they protect against. One gene probably originated from a member of the *Enterobacteriaceae* family, four genes from *Paenibacillus alvei*, three from unknown members of the *Bacillaceae* family, six from *Bacillus cereus*, from *B. paraanthracis* and *B. thuringiensis,* one each and four from *Acinetobacter gyllenbergii.* Most of the detected genes protect beta-lactams (penams, carbapenems, cephalosporins).

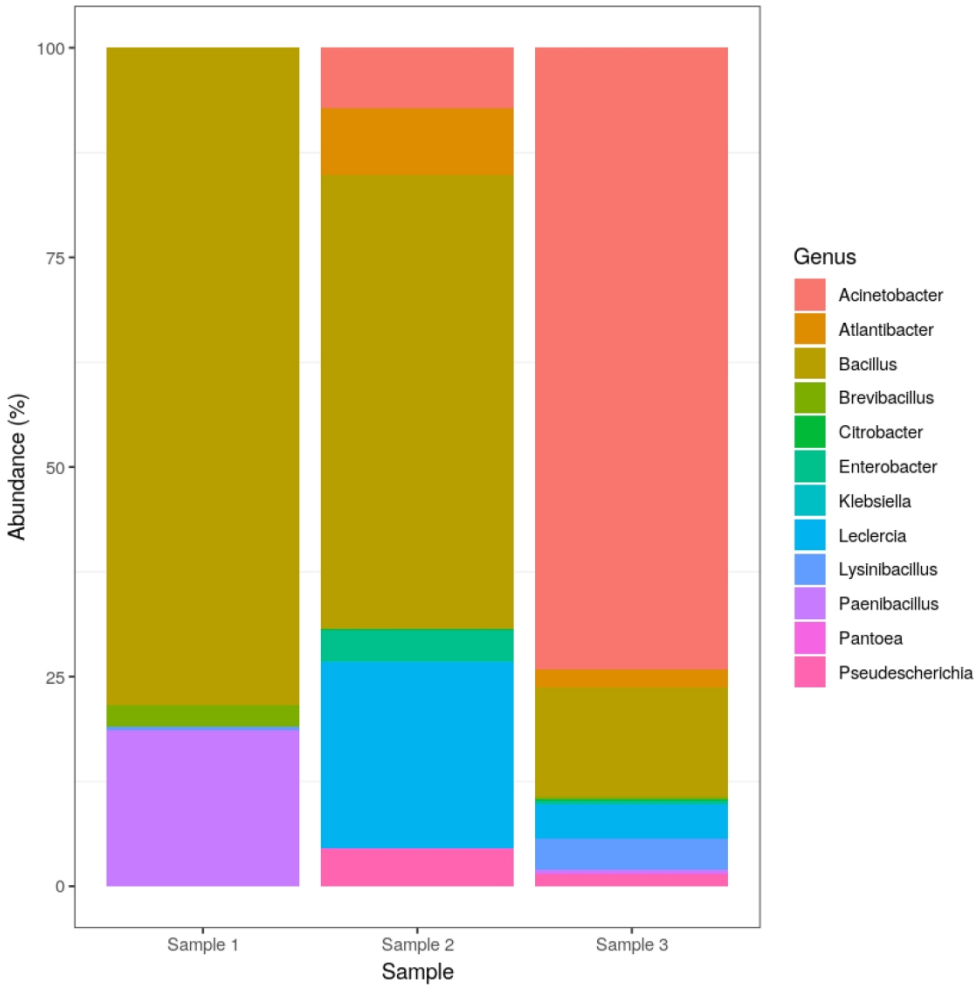

Figure 1.

296x296mm (600 x 600 DPI)

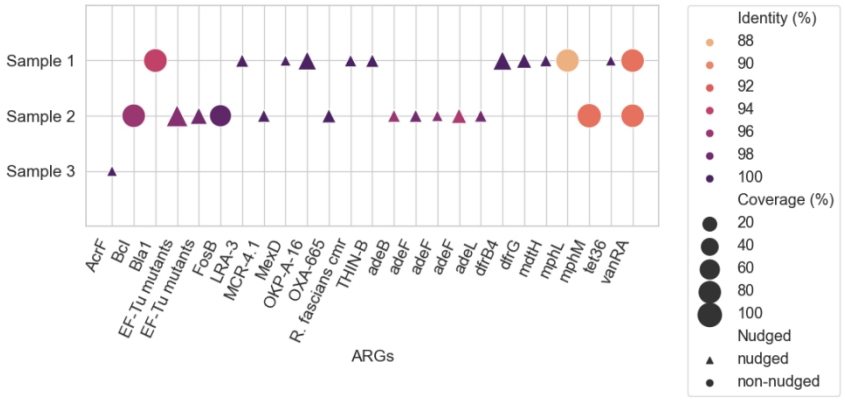

Figure 2.

451x217mm (600 x 600 DPI)

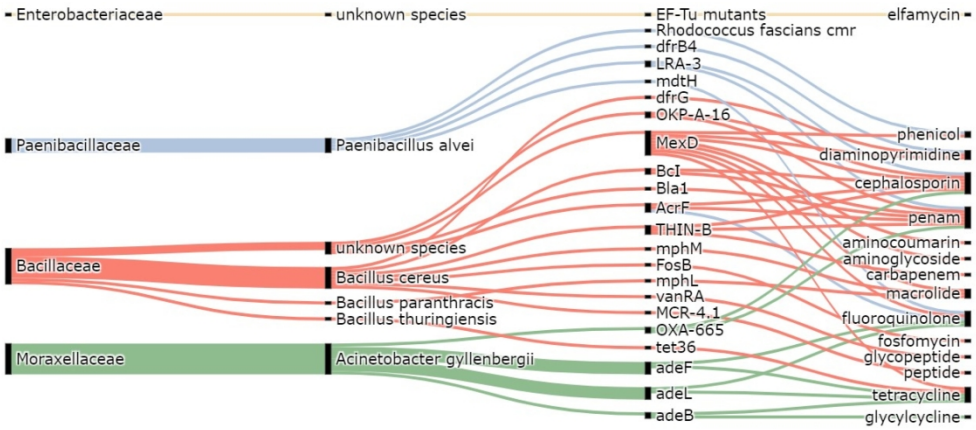

Figure 3.

450x195mm (600 x 600 DPI)