**Wide occurrence of putative mobilized colistin resistance genes in the human gut microbiome**

Bruno G. N. Andrade [1,2], Tobias Goris [3], Haithem Afli [1,2], Felipe H. Coutinho [4], Alberto M.R. Davila [5], Rafael R. C. Cuadrat* [6]

1 - Department of Computer Science, Muster Technological University, Cork, Ireland

2 - ADAPTCENTRE, Dublin, Ireland

3 - Department of Molecular Toxicology, Research Group Intestinal Microbiology, German Institute of Human Nutrition Potsdam-Rehbruecke - DIfE, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany

4 - Departamento de producción vegetal y microbiología, Universidad Miguel Hernández, Alicante, Spain

5 - Computational and Systems Biology Laboratory, Oswaldo Cruz Institute, FIOCRUZ. Rio de Janeiro, RJ, Brazil

6 - Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558, Nuthetal, Germany

* Corresponding author: rafaelcuadrat@gmail.com

**Abstract**

**Background:** The high incidence of bacterial genes that confer resistance to last-resort antibiotics, such as colistin caused by MCR genes, poses an unprecedented threat to our civilization's health. To diminish its impact, understanding the spread, evolution, and distribution among human populations should be a priority. To tackle this problem, we investigated the distribution and prevalence of potential mcr genes in the human gut microbiome by using a set of bioinformatics tools to screen for the presence of putative mcr genes (and co-located ARGs) in the Unified Human Gastrointestinal Genome (UHGG) collection, its genomic context, and phylogeny.

**Results:** A total of 2,079 ARGs were classified as different MCR in 2,046 Metagenome assembled genomes (MAGs), present in 1,596 individuals from 41 countries, of which 215 MCRs were identified in plasmidial contigs. The genera that presented the largest number of MCR-like genes were *Suterella* and *Parasuterella*, highly prevalent, and opportunistic pathogens. Other potential pathogens carrying MCR genes belonged to the genus Vibrio and Escherichia. Finally, we identified a total of 22,746 belonging to 21 different classes in the same 2,046 MAGs, suggesting multi-resistance potential, increasing the concern of its impact in the clinical settings.

**Conclusion:** This study uncovers the diversity of MCR-like genes in the human gut microbiome. We showed the cosmopolitan distribution of these genes in patients worldwide and the co-presence of other antibiotic resistance genes, including ESBLs. Also, we described mcr-like genes encoded in the same ORF with PAP2-like in bacteria from the genus Sutterella. Although these novel sequences increase our knowledge about the diversity and evolution of mcr-like genes, their activity has to be experimentally validated in the future.

**keywords:** Colistin, HMP, MCR, Metagenomics, Resistance genes

## Introduction

The prevalence of antibiotic resistance (AR) in clinical pathogens is a significant public health concern, especially in low and middle-income countries (LMICs)[1]. The misuse of antibiotics is the main driving factor for the rise of antibiotic-resistant bacteria. Still, its importance is often underestimated in community infections, as hospitalized infections gain most attention[2]. Many previous studies have addressed the prevalence of AR in clinical environments[3 4 5] and, recently, to also understand AR prevalence in non-hospitalized populations[6 7]. Most of these studies were conducted on cultivable clinical strains using microbiological methods that involve cultivation and antibiogram tests. However, the advances in high-throughput sequencing and bioinformatics enabled the study of metagenomes and access to the so-called human resistome.

The resistome is defined as the collection of the antibiotic resistance genes (ARGs) in a single microorganism, or in a microbial community, and has been investigated in different environments, such as soils[8] or ocean[9], and in host-associated microbiotas such as the animal[10] or human[11] gut. Understanding the human resistome in hospitalized and non-hospitalized populations are essential not only because the commensal microbiota can host and transfer ARGs from and to pathogenic bacteria by horizontal gene transfer (HGT)[12], e.g., during an infection. In addition, HGT can also play a role in ARG mobilization to environmental communities by water and soil contamination[13] or the food we ingest[14 15] . The gut microbiome is of particular interest in the investigation of ARGs in the human microbiota since it is the largest and most diverse[16] and highly exposed to and affected by the intake of antibiotics. A potential influx of ARGs can occur via food intake and/or unhygienic conditions, and the efflux of ARGs to wastewater plants enhances the spread to other environments.As such, the human gut microbiome is thought to be responsible for transferring ARGs[17] to the environment to a large extent. Therefore, the search for ARGs in the human gut microbiome, which is mostly performed using metagenomics and culturomics approaches, is one of the key fields to unravel the transfer of ARGs and the evolution of antibiotic resistance in bacteria. While past metagenomic studies on ARGs relied on shorter contigs, often below 50

kilobases, new assembly methods that allow the recovery of nearly complete bacterial genomes have been developed. Such methods have been applied to many studies regarding the human gut microbiome and allowed the recovery of thousands of metagenome-assembled genomes (MAGs) [18–20]. These datasets were recently combined into one resource, the Human Microbiome Project (HMP), which makes this human gut MAGs collection a valuable resource to screen for ARGs. The advantage of MAGs versus traditional metagenome gene catalogs is manifold; the most apparent is the high accuracy of phylogenetic affiliations and often complete gene clusters, revealing gene synteny. Especially the latter is of high interest when studying ARGs since the genetic environment often shows the genetic mobility of ARGs, e.g., their location on genetic islands or plasmids[21]. Besides, it is also possible to investigate the presence of multi-drug-resistant (MDR) bacteria by detecting more than one ARG in the same bacterial genome or contig[9] when using the MAGs approach.

Recently, colistin (Polymyxin E) has gained attention as the last line of defense drug against MDR bacteria, especially carbapenem-resistant gram-negative pathogens[22]. However, reports of colistin-resistant bacteria are becoming more frequent[23], with its prevalence reaching as high as approximately 20%–40% among Carbapenem-Resistant *Klebsiella pneumoniae* (CRKP) in Italy and Greece[24]. In the past, the only known acquired resistance mechanism for colistin was mediated by chromosomal mutations, mainly in genes regulating the chemical additions of L-Ara4N and pEtN[25]. The first plasmid-mediated polymyxin resistance gene, designated mobilized-colistin resistance-1 (mcr-1), was described for *Enterobacteriaceae* in 2006[26]. Later it was followed by the additional mcrs, mcr-2[27], mcr-3[28], mcr-4[29], mcr-5[30] and, very recently, mcr-6 to mcr-10 [31-35]. An intrinsic mcr-1-like homolog from *Moraxella osloensis* was described, named icr-Mo[36], raising the discussion about possible origins of MCRs in *Moraxella*. The spread of mcrs is of a public health concern as it has been attributed to colistin's over-use, especially in livestock[37] and aquaculture[38,39].

## Methods

### Data retrieval

We retrieved ~171 million non-redundant protein sequences (clustered at 100%) from the Unified Human Gastrointestinal Protein (UHGP) catalog[40] plus the according to the metadata of the 286,997 corresponding metagenomic assembled genomes (MAGs). Also, we obtained the redundant protein identifiers mapped to the non-redundant protein representative for quantification. For the MAGs, including ARGs of interest, we retrieved the fasta sequences and GFF annotation files.

**Antibiotic resistance gene screening**

To search for antibiotic-resistant genes in the human microbiome, we followed the methodology used in our previous study with oceanic samples 8. In short, we used the deepARG tool[41] (model version 2), an in-depth learning approach developed to identify ARGs at both reads or operational reading frames level, to search for ARGs in the non-redundant proteins provided by the UHGP catalog. We then selected all proteins classified as mobilized colistin resistance (MCR) in the deepARG results and explored the prevalence of these putative ARGs in different countries and across diverse taxa.

**Plasmid classification**

To verify if the putative genes are of chromosomal or plasmid origin, we applied the PlasFlow software[42] with the default threshold of 0.7. This software uses neural network models trained on full genomes and plasmid sequences to predict the sequence origin with 96% accuracy.

**Phylogeny**

Protein sequences classified as MCR present in contigs identified as plasmids were clusterized at 97% of sequence similarity with the software cd-hit v4.7[43] to reduce the number of protein sequences in the tree. The representative sequences of each cluster (and reference sequences obtained at NCBI) were then submitted to NGPhylogeny.fr[44] where the protein sequences were first aligned by Mafft[45]. The informative phylogenetic regions were selected by BMGE[46], and the Maximum likelihood (ML) reconstruction was calculated by

PhyML 3.0[47] with the model selection performed by SMS (AIC method)[48], and 100 bootstrap replicates to infer significance.

**Data visualization**

For data visualization, we used jupyter notebook and python 3, with libraries pandas and matplotlib. For conserved domain visualization on unusual mcr-like sequences, we used NCBI blast and CDD (https://blast.ncbi.nlm.nih.gov/Blast.cgi), running Blastp with default parameters..

**Results and Discussion**

We identified a total of 2079 ARGs classified as MCR (13 MCR-1, 1 MCR-1.2, 9 MCR-2, 634 MCR-3, 456 MCR-4, 966 MCR-5) in 2,046 genomes (166 from isolates and 1880 from MAGs), present in 1,596 individuals (7.2% from the total 21,866 in the study) from 41 different countries. It is important to note the restricted classification of MCR in the deepARG model 2, only assigning up to MCR-5. The highest relative number of mcr-like genes (normalized by the total number of samples) was found in Haiti (>10%), followed by South Korea, Norway, and India (Figure 1). However, it is important to observe that a quantitative comparison is not possible due to the very unbalanced number of genomes from each country, varying from less than 50 to more than 50,000.

**Figure 1:** Relative prevalence of mcr-like genes (bubble area shows number of genes divided by number of genomes from each country) in countries with at least 50 genomes in the study.

The genus with the largest number of Mcr-like genes is *Sutterella* (667 genes), followed by *Parasutterella* (338 genes) and the alphaproteobacterial genus CAG-495 (258 genes) (Figure 2). The genus *Sutterella* is highly prevalent and, mainly represented by *S. wadsworthensis*, abundant in the healthy human gut[49] considered pathogenic in general[50,51].

However, a role in autism in children was suggested[52], and a higher prevalence in prediabetics was observed[53], while their role remains unknown. Due to isolations from several inflammated body parts, *S. wadsworthensis* might be considered an opportunistic pathogen[54]. While some AR occurred in *S. wadsworthensis*, so far, there was no report of Colistin resistance in the *Sutterella* genus. At least one study reported a *S. wadsworthensis* strain as susceptible[55], indicating that colistin susceptibility of *Sutterella* strains from the human gut must be experimentally validated in the future. Similar to *Sutterella*, *Parasutterella* species, while sometimes linked to diseases, appear to be ordinary members of the human gut microbiota[56], pathogenic only in rare cases. The unstudied CAG-495 genus was reported to be abundant in patients diagnosed with Vogt-Koyanagi-Harada disease[57].
The main potentially pathogenic species carrying Mcr-like genes belong to the genera *Vibrio*, *Escherichia* (only a few pathogenic strains, mainly a commensal gut bacterium), and *Campylobacter*. While the relative abundance of mcr-like genes in *Escherichia* is low (< 1%), that for Vibrio (>10%) and *Campylobacter* (>50%) is relatively high and warrants further investigation if the human gut might contribute to spreading colistin-resistant strains of these pathogens.

**Figure 2:** Prevalence of Mcr-like genes in different genera (with 10 or more Mcr-like genes).

Many of the mcr-like gene sequences identified in this study are much larger than the average MCR (~530 amino acids long), with some reaching up to 831 amino acids. We found that there was an extra region encoding PAP2 and PAP2_like domains ( an example can be found in Figure 3). In MCR-1 plasmids, PAP2 is encoded in a separated ORF downstream mcr sequence[58].

**Figure 3:** Conserved domains on one MCR-4 like sequence from the Sutterella genus. This analysis was done on CDD online from NCBI blast tool.

Regarding their genomic context, a total of 215 genes are present in contigs classified as plasmid by PlasFlow, while 1239 were in contigs classified as chromosomes, and 625 genes were in not classified contigs.

We also verified the existence of other ARGs detected by deepARG in genomes containing Mcr-like sequences. We identified a total of 22,746 ARGs (from 21 ARG classes and some unclassified) co-occurring with MCR-like sequences, being the most abundant classes the multidrug resistance (10,008 ARGs), beta-lactam (2,271 ARGs) and glycopeptide (2,261 ARGs) (Figure 4). However, several of those sequences on the multidrug class are efflux proteins, and as discussed in our previous study[9] , those are very hard to distinguish from other transporters that are not involved in antibiotic resistance.

**Figure 4:** Number of ARGs per class co-occurring in the same genome with Mcr-like genes.

Colistin is a last resource antibiotic used against MDR bacteria with extended-spectrum beta-lactamases (ESBL), which makes the investigation of the presence of those ARGs in genomes containing MCR-like sequences so important. Regarding the beta-lactam group, we identified 1138 penA, a penicillin-binding protein 2 (PBP 2) associated with reduced susceptibility to oral cephalosporins[59]. Besides, we also identified 673 pbp1-A and 129 pbp1-B that are also penicillin-binding proteins.

Regarding the ESBL group, we identified 25 blaOXA (class D beta-lactamase capable of destroying 3rd generation cephalosporins)[60], 3 blaCTX-M (a plasmid-encoded ESBL found in Enterobacteriaceae, likely acquired from the environmental bacteria *Kluyvera spp.* by HGT[61]), 2 blaTEM-153, 7 blaTLA-1 and 2 blaCFXA-6. The blaCTX-M enzymes have been found associated with insertion elements (ISEcp1) and transposable elements (for example, Tn402-like transposons). Many conjugative plasmids can transport these mobile elements, and consequently, these enzymes became the most prevalent ESBL [62 63].

The phylogenetic tree (Figure 5) shows the protein sequences classified as MCR-4 and MCR-5 in *Sutterella*, *Parasutterella*, and CAG-521 together in a clade with support value 1,

and clinical MCR-5 sequences in an adjacent clade. The sequences used in the tree are grouped with clinical MCRs instead of the outgroup eptA, providing additional evidence for the annotation of those sequences.

**Figure 5: Phylogenetic tree of MCR-like sequences.** We used only sequences present in contigs classified as plasmids and we clusterized similar sequences with CD-HIT on 97% similarity. The phylogenetic informative regions were selected by BMGE and the Maximum likelihood (ML) phylogenetic tree was calculated by PhyML 3.0 with the model selection performed by SMS (AIC method) and 100 bootstrap replicates to infer significance. We added clinical MCR sequences from NCBI to the analysis.

## Conclusions

This study uncovers the diversity of MCR-like genes in the human gut microbiome. We showed the cosmopolitan distribution of these genes in patients worldwide and the co-presence of other antibiotic resistance genes, including ESBLs. Also, we described mcr-like genes encoded in the same ORF with PAP2-like in bacteria from the genus Sutterella. Although these novel sequences increase our knowledge about the diversity and evolution of mcr-like genes, their activity has to be experimentally validated in the future.

## Data and Code availability

All the code used in this study is available at https://github.com/rcuadrat/human_microbiome_mcr and all the data is available at Zenodo (https://doi.org/10.5281/zenodo.4399676).

**Acknowledgment**

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 801522, by Science Foundation Ireland and co-funded by the European Regional Development Fund through the ADAPT Centre for Digital Content Technology grant number 13/RC/2106. We thank Maria Sorokina and also the colleagues in DIfE for the meaningful discussions during the elaboration of this text.

**References**

1. Nadimpalli, M. *et al.* Combating Global Antibiotic Resistance: Emerging One Health Concerns in Lower- and Middle-Income Countries. *Clin. Infect. Dis.* **66**, 963–969 (2018).
2. van Hecke, O., Wang, K., Lee, J. J., Roberts, N. W. & Butler, C. C. Implications of Antibiotic Resistance for Patients' Recovery From Common Infections in the Community: A Systematic Review and Meta-analysis. *Clin. Infect. Dis.* **65**, 371–382 (2017).
3. Gerver, S. M. *et al.* Surveillance of Bloodstream Infections in Intensive Care Units in England, May 2016-April 2017: Epidemiology and Ecology. *J. Hosp. Infect.* S0195670120302413 (2020) doi:10.1016/j.jhin.2020.05.010.
4. Stohr, J. J. J. M., Verweij, J. J., Buiting, A. G. M., Rossen, J. W. A. & Kluytmans, J. A. J. W. Within-patient plasmid dynamics in Klebsiella pneumoniae during an outbreak of a carbapenemase-producing Klebsiella pneumoniae. *PLOS ONE* **15**, e0233313 (2020).
5. Jean, S.-S., Ko, W.-C. & Hsueh, P.-R. Susceptibility of clinical isolates of methicillin-resistant Staphylococcus aureus and phenotypic non-extended-spectrum β-lactamase-producing Klebsiella pneumoniae to ceftaroline in Taiwan: Results from the Antimicrobial Testing Leadership and Surveillance (ATLAS) in 2012–2018 and Surveillance of Multicentre Antimicrobial Resistance in Taiwan (SMART) in 2018–2019. *Int. J. Antimicrob. Agents* 106016 (2020) doi:10.1016/j.ijantimicag.2020.106016.
6. Martínez, E. P. *et al.* Trends, seasonality and the association between outpatient antibiotic use and antimicrobial resistance among urinary bacteria in the Netherlands. *J.*

*Antimicrob. Chemother.* dkaa165 (2020) doi:10.1093/jac/dkaa165.

7. Bitew, A. & Tsige, E. High Prevalence of Multidrug-Resistant and Extended-Spectrum *β* -Lactamase-Producing Enterobacteriaceae: A Cross-Sectional Study at Arsho Advanced Medical Laboratory, Addis Ababa, Ethiopia. *J. Trop. Med.* **2020**, 1–7 (2020).
8. Cytryn, E. The soil resistome: The anthropogenic, the native, and the unknown. *Soil Biol. Biochem.* **63**, 18–23 (2013).
9. Cuadrat, R. R. C., Sorokina, M., Andrade, B. G., Goris, T. & Dávila, A. M. R. Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. *GigaScience* **9**, giaa046 (2020).
10. Lim, S.-K., Kim, D., Moon, D.-C., Cho, Y. & Rho, M. Antibiotic resistomes discovered in the gut microbiomes of Korean swine and cattle. *GigaScience* **9**, giaa043 (2020).
11. Carr, V. R. *et al.* Abundance and diversity of resistomes differ between healthy human oral cavities and gut. *Nat. Commun.* **11**, 693 (2020).
12. Liu, L. *et al.* The human microbiome: A hot spot of microbial horizontal gene transfer. *Genomics* **100**, 265–270 (2012).
13. Boto, L., Pineda, M. & Pineda, R. Potential impacts of horizontal gene transfer on human health and physiology and how anthropogenic activity can affect it. *FEBS J.* **286**, 3959–3967 (2019).
14. Capita, R. & Alonso-Calleja, C. Antibiotic-Resistant Bacteria: A Challenge for the Food Industry. *Crit. Rev. Food Sci. Nutr.* **53**, 11–48 (2013).
15. Baron, S. A., Diene, S. M. & Rolain, J.-M. Human microbiomes and antibiotic resistance. *Hum. Microbiome J.* **10**, 43–52 (2018).
16. Dave, M., Higgins, P. D., Middha, S. & Rioux, K. P. The human gut microbiome: current knowledge, challenges, and future directions. *Transl. Res.* **160**, 246–257 (2012).
17. Bengtsson-Palme, J. *et al.* The Human Gut Microbiome as a Transporter of Antibiotic Resistance Genes between Continents. *Antimicrob. Agents Chemother.* **59**, 6551–6560 (2015).
18. Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S. & Kyrpides, N. C. New insights from uncultivated genomes of the global human gut microbiome. *Nature* **568**, 505–510 (2019).
19. Pasolli, E. *et al.* Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* **176**, 649-662.e20 (2019).
20. Almeida, A. *et al.* A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504 (2019).
21. Zhao, R. *et al.* Deciphering the mobility and bacterial hosts of antibiotic resistance genes under antibiotic selection pressure by metagenomic assembly and binning

approaches. *Water Res.* **186**, 116318 (2020).

22. Zafer, M. M. *et al.* Emergence of colistin resistance in multidrug-resistant Klebsiella pneumoniae and Escherichia coli strains isolated from cancer patients. *Ann. Clin. Microbiol. Antimicrob.* **18**, 40 (2019).
23. Rossi, F. *et al.* Emergence of colistin resistance in the largest university hospital complex of São Paulo, Brazil, over five years. *Braz. J. Infect. Dis.* **21**, 98–101 (2017).
24. Giamarellou, H. Epidemiology of infections caused by polymyxin-resistant pathogens. *Int. J. Antimicrob. Agents* **48**, 614–621 (2016).
25. Berglund, B. Acquired Resistance to Colistin via Chromosomal And Plasmid-Mediated Mechanisms in Klebsiella pneumoniae. *Infect. Microbes Dis.* **1**, 10–19 (2019).
26. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).
27. Xavier, B. B. *et al.* Identification of a novel plasmid-mediated colistin-resistance gene, *mcr-2* , in *Escherichia coli* , Belgium, June 2016. *Eurosurveillance* **21**, 30280 (2016).
28. Yin, W. *et al.* Novel Plasmid-Mediated Colistin Resistance Gene mcr-3 in Escherichia coli. *mBio* **8**, mBio.00543-17, e00543-17 (2017).
29. Carattoli, A. *et al.* Novel plasmid-mediated colistin resistance mcr-4 gene in Salmonella and Escherichia coli, Italy 2013, Spain and Belgium, 2015 to 2016. *Euro Surveill. Bull. Eur. Sur Mal. Transm. Eur. Commun. Dis. Bull.* **22**, (2017).
30. Borowiak, M. *et al.* Identification of a novel transposon-associated phosphoethanolamine transferase gene, mcr-5, conferring colistin resistance in d-tartrate fermenting Salmonella enterica subsp. enterica serovar Paratyphi B. *J. Antimicrob. Chemother.* **72**, 3317–3324 (2017).
31. AbuOun, M. *et al.* mcr-1 and mcr-2 (mcr-6.1) variant genes identified in Moraxella species isolated from pigs in Great Britain from 2014 to 2015. *J. Antimicrob. Chemother.* **72**, 2745–2749 (2017).
32. Yang, Y.-Q., Li, Y.-X., Lei, C.-W., Zhang, A.-Y. & Wang, H.-N. Novel plasmid-mediated colistin resistance gene mcr-7.1 in Klebsiella pneumoniae. *J. Antimicrob. Chemother.* **73**, 1791–1795 (2018).
33. Wang, X. *et al.* Emergence of a novel mobile colistin resistance gene, *mcr-8* , in NDM-producing *Klebsiella pneumoniae*. *Emerg. Microbes Infect.* **7**, 1–9 (2018).
34. Carroll, L. M. *et al.* Identification of Novel Mobilized Colistin Resistance Gene *mcr-9* in a Multidrug-Resistant, Colistin-Susceptible *Salmonella enterica* Serotype Typhimurium Isolate. *mBio* **10**, e00853-19, /mbio/10/3/mBio.00853-19.atom (2019).
35. Wang, C. *et al.* Identification of novel mobile colistin resistance gene *mcr-10*. *Emerg. Microbes Infect.* **9**, 508–516 (2020).

36. Wei, W. *et al.* Defining ICR-Mo, an intrinsic colistin resistance determinant from Moraxella osloensis. *PLOS Genet.* **14**, e1007389 (2018).

37. Economou, V. & Gousia, P. Agriculture and food animals as a source of antimicrobial-resistant bacteria. *Infect. Drug Resist.* 49 (2015) doi:10.2147/IDR.S55778.

38. Watts, J., Schreier, H., Lanska, L. & Hale, M. The Rising Tide of Antimicrobial Resistance in Aquaculture: Sources, Sinks and Solutions. *Mar. Drugs* **15**, 158 (2017).

39. Zhang, H. *et al.* Action and mechanism of the colistin resistance enzyme MCR-4. *Commun. Biol.* **2**, 36 (2019).

40. Almeida, A. *et al. A unified sequence catalogue of over 280,000 genomes obtained from the human gut microbiome*. http://biorxiv.org/lookup/doi/10.1101/762682 (2019) doi:10.1101/762682.

41. Arango-Argoty, G. *et al.* DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* **6**, 23 (2018).

42. Krawczyk, P. S., Lipinski, L. & Dziembowski, A. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* **46**, e35–e35 (2018).

43. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).

44. Lemoine, F. *et al.* NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res.* **47**, W260–W265 (2019).

45. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

46. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).

47. Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).

48. Lefort, V., Longueville, J.-E. & Gascuel, O. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* **34**, 2422–2424 (2017).

49. Hiippala, K., Kainulainen, V., Kalliomäki, M., Arkkila, P. & Satokari, R. Mucosal Prevalence and Interactions with the Epithelium Indicate Commensalism of Sutterella spp. *Front. Microbiol.* **7**, (2016).

50. Cheng, J. *et al.* Duodenal microbiota composition and mucosal homeostasis in pediatric celiac disease. *BMC Gastroenterol.* **13**, 113 (2013).

51. Mukhopadhya, I. *et al.* A Comprehensive Evaluation of Colonic Mucosal Isolates of Sutterella wadsworthensis from Inflammatory Bowel Disease. *PLoS ONE* **6**, e27076 (2011).

52. Ho, L. K. H. *et al.* Gut microbiota changes in children with autism spectrum disorder: a systematic review. *Gut Pathog.* **12**, 6 (2020).

53. the IMI-DIRECT consortium *et al.* Aberrant intestinal microbiota in individuals with prediabetes. *Diabetologia* **61**, 810–820 (2018).

54. Finegold, S. M. & Jousimies–Somer, H. Recently Described Clinically Important Anaerobic Bacteria: Medical Aspects. *Clin. Infect. Dis.* **25**, S88–S93 (1997).

55. Warren, Y. A., Citron, D. M., Merriam, C. V. & Goldstein, E. J. C. Biochemical Differentiation and Comparison of Desulfovibrio Species and Other Phenotypically Similar Genera. *J. Clin. Microbiol.* **43**, 4041–4045 (2005).

56. Ju, T., Kong, J. Y., Stothard, P. & Willing, B. P. Defining the role of Parasutterella, a previously uncharacterized member of the core gut microbiota. *ISME J.* **13**, 1520–1534 (2019).

57. Ye, Z. *et al.* Altered gut microbiome composition in patients with Vogt-Koyanagi-Harada disease. *Gut Microbes* **11**, 539–555 (2020).

58. Wang, R. *et al.* The global distribution and spread of the mobilized colistin resistance gene mcr-1. *Nat. Commun.* **9**, 1179 (2018).

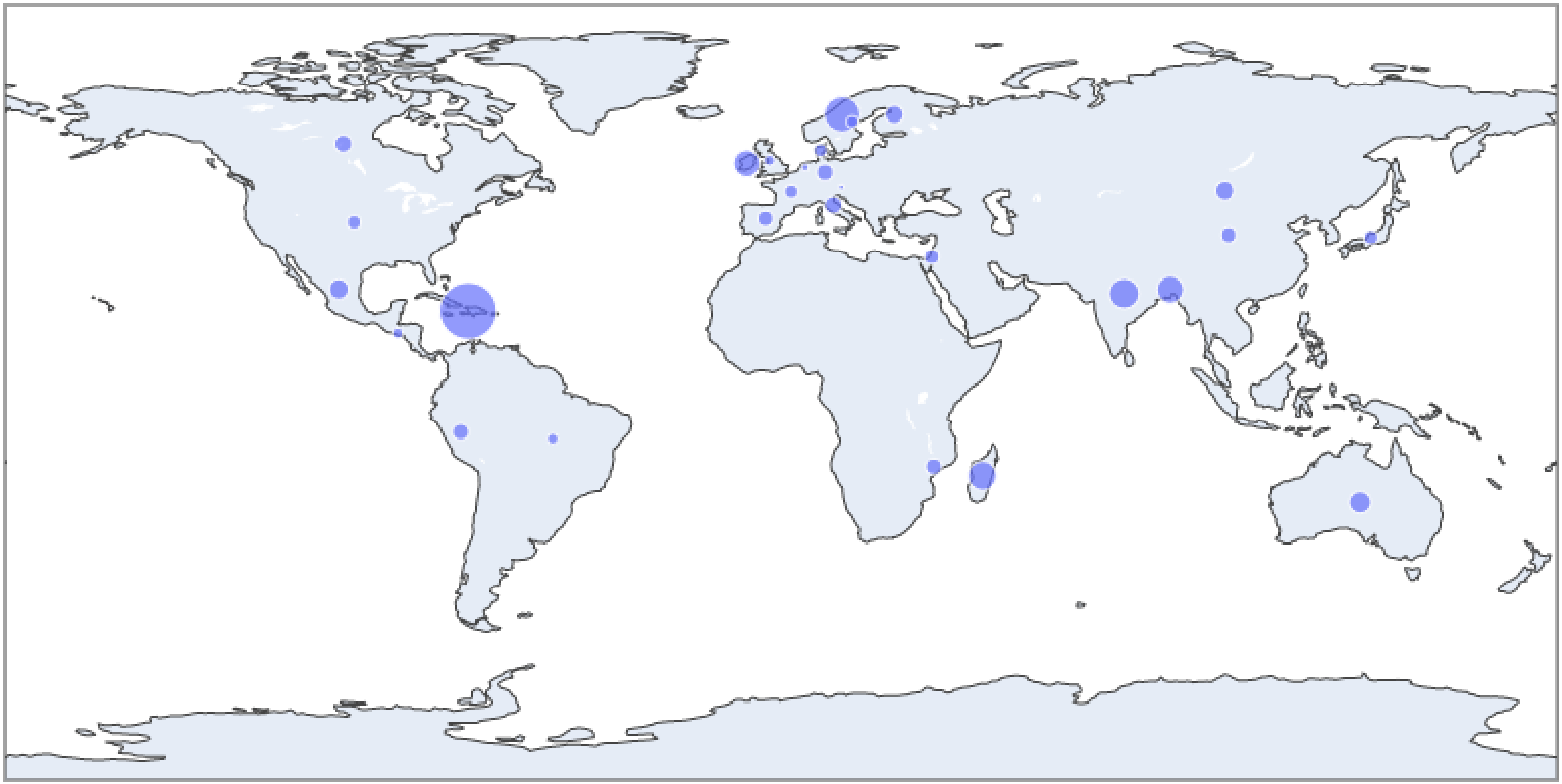
59. Ohnishi, M. *et al.* Spread of a Chromosomal Cefixime-Resistant penA Gene among Different Neisseria gonorrhoeae Lineages. *Antimicrob. Agents Chemother.* **54**, 1060–1067 (2010).

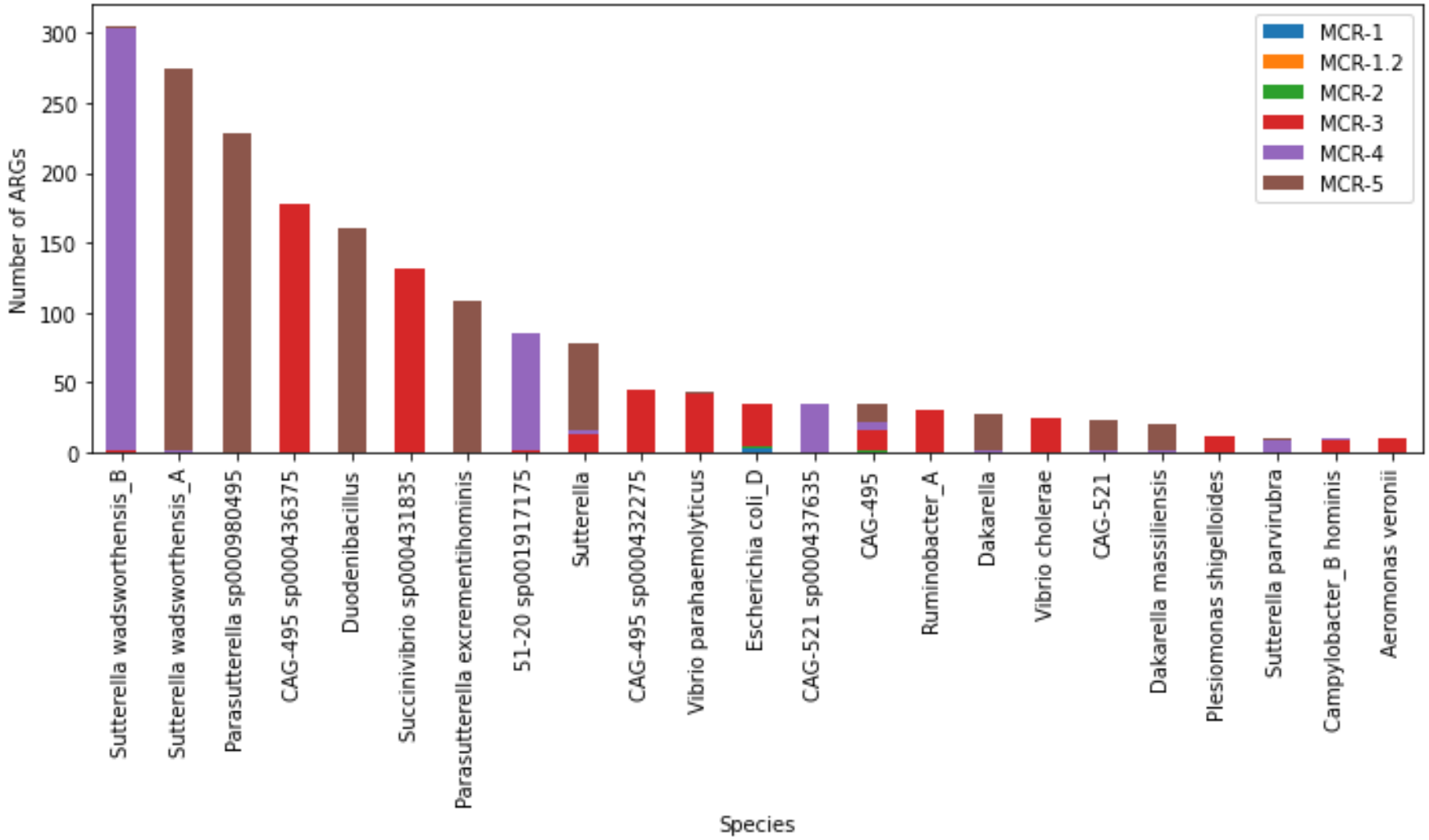
60. Munita, J. M. & Arias, C. A. Mechanisms of Antibiotic Resistance. *Microbiol. Spectr.* **4**, (2016).

61. Bonnet, R. Growing Group of Extended-Spectrum β-Lactamases: the CTX-M Enzymes. *Antimicrob. Agents Chemother.* **48**, 1–14 (2004).

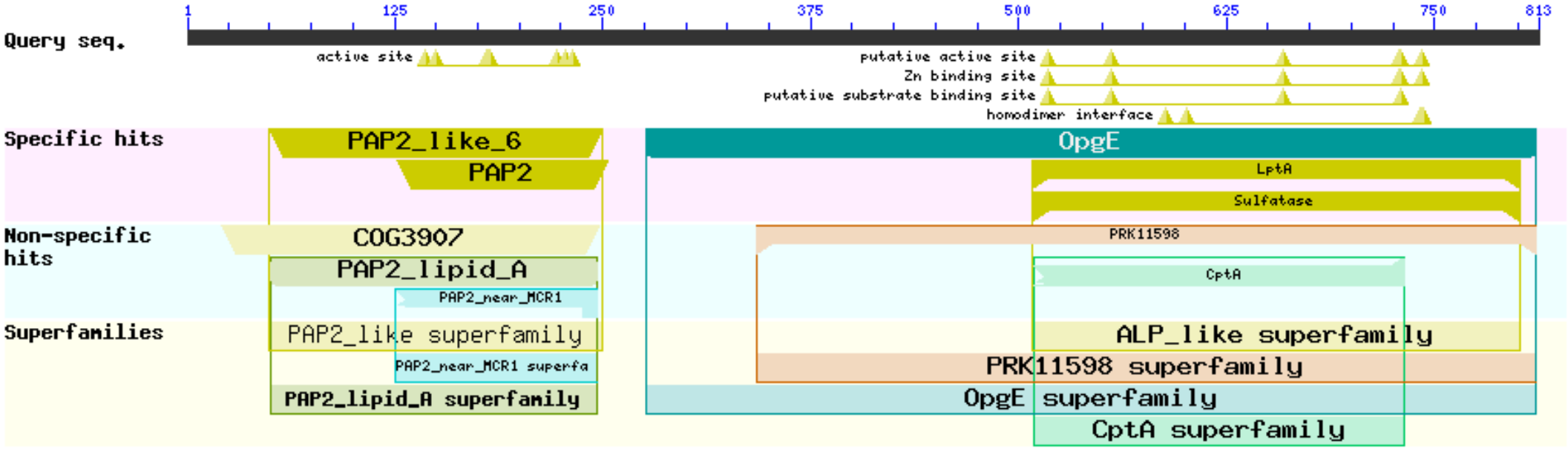
62. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, e00088-17, /cmr/31/4/e00088-17.atom (2018).

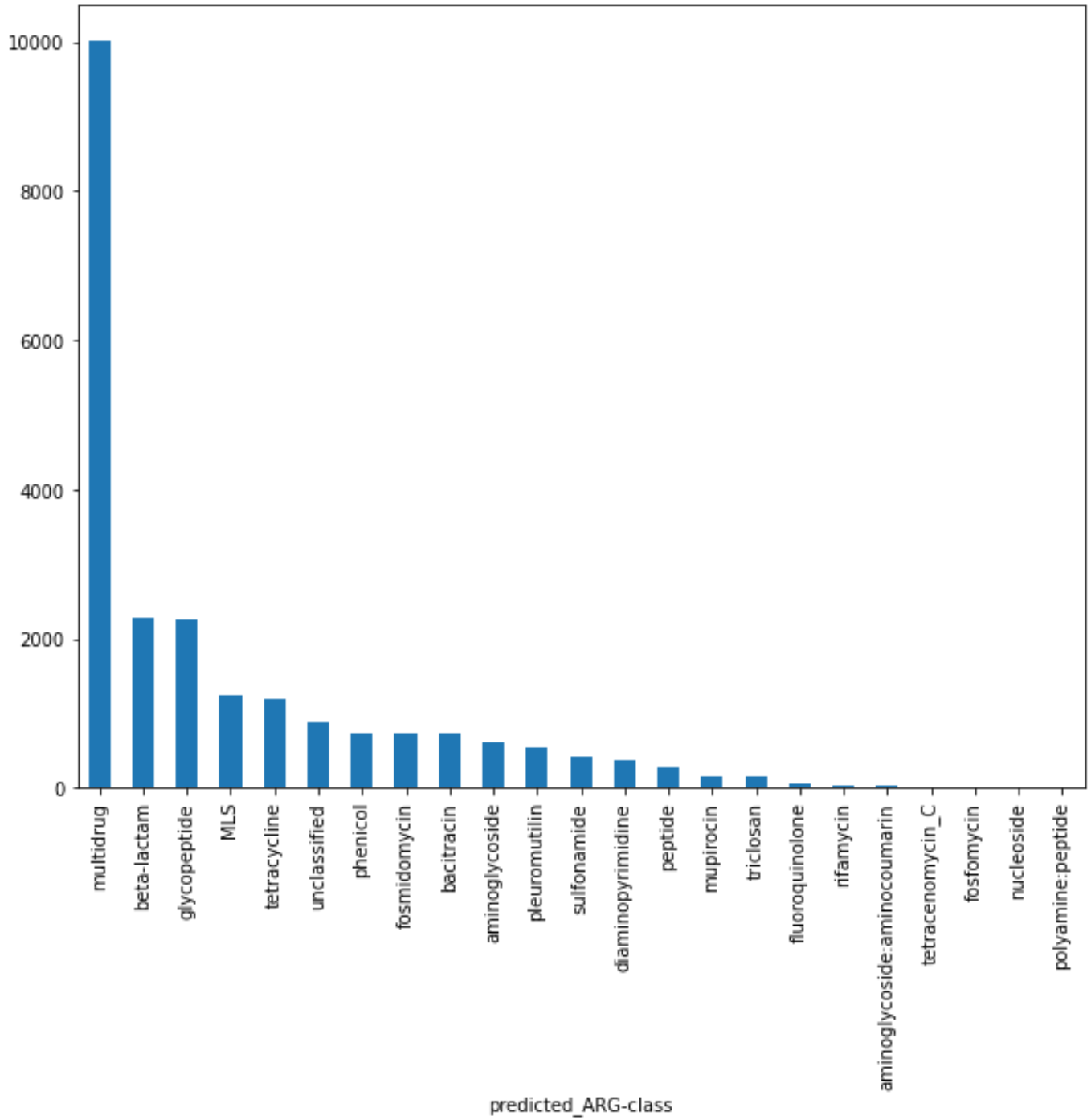
63. Cantón, R., González-Alba, J. M. & Galán, J. C. CTX-M Enzymes: Origin and Diffusion. *Front. Microbiol.* **3**, (2012).

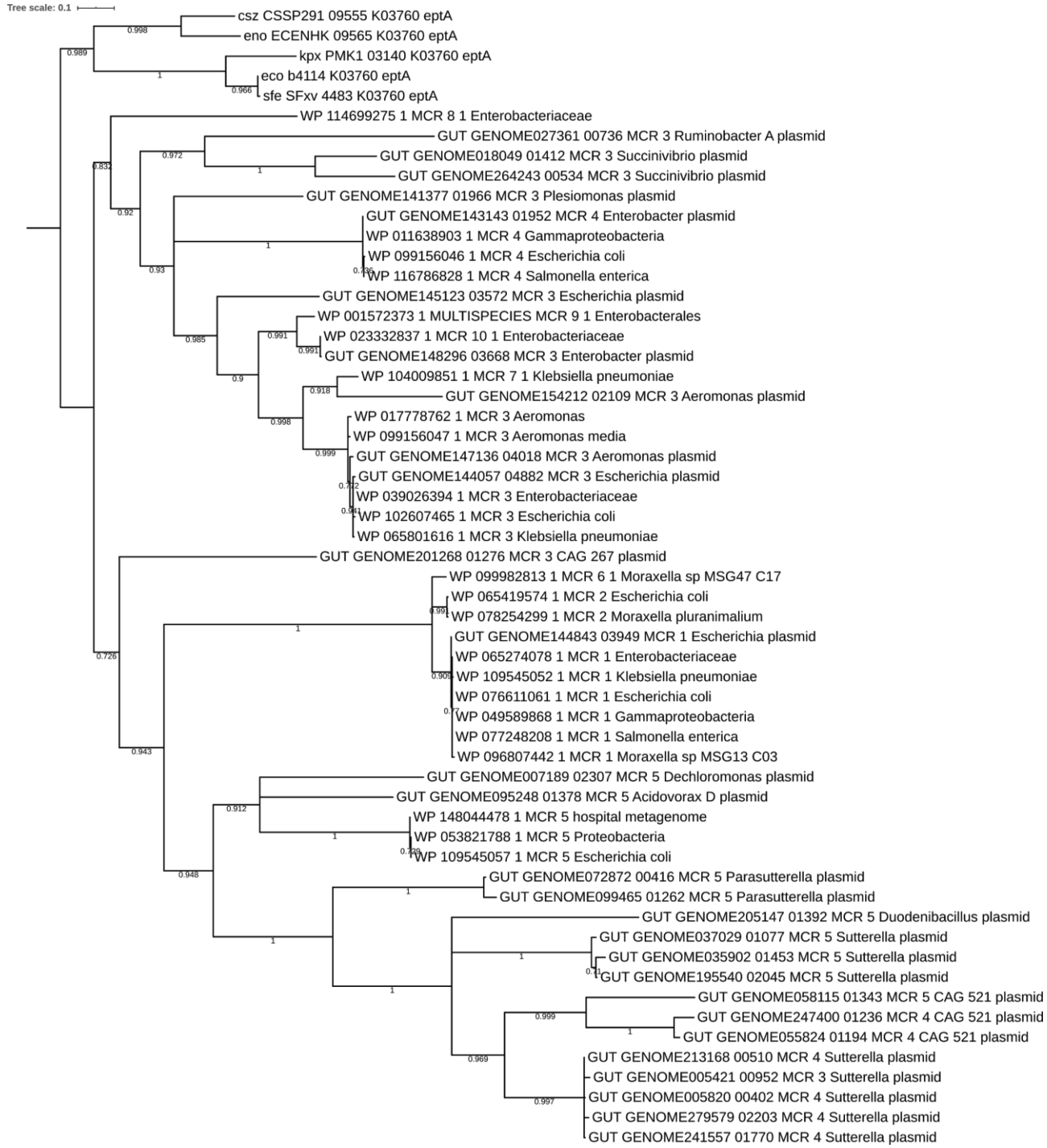



MCR-1
MCR-1.2
MCR-2
MCR-3
MCR-4
MCR-5
Number of ARGs
0
50
100
150
200
250
300
Sutterella wadsworthensis_B
Sutterella wadsworthensis_A
Parasutterella sp000980495
CAG-495 sp000436375
Duodenibacillus
Succinivibrio sp000431835
Parasutterella excrementihominis
51-20 sp001917175
Sutterella
CAG-495 sp000432275
Vibrio parahaemolyticus
Escherichia coli_D
CAG-521 sp000437635
CAG-495
Ruminobacter_A
Dakarella
Vibrio cholerae
CAG-521
Dakarella massiliensis
Plesiomonas shigelloides
Sutterella parvirubra
Campylobacter_B hominis
Aeromonas veronii
Species

## Conserved domains on [GUT_GENOME241557_01770|MCR-4|Sutterella|plasmid]




10000
8000
6000
4000
2000
0
multidrug
beta-lactam
glycopeptide
MLS
tetracycline
unclassified
phenicol
fosmidomycin
bacitracin
aminoglycoside
pleuromutilin
sulfonamide
diaminopyrimidine
peptide
mupirocin
triclosan
fluoroquinolone
rifamycin
aminoglycoside:aminocoumarin
tetracenomycin_C
fosfomycin
nucleoside
polyamine:peptide
predicted_ARG-class


Tree scale: 0.1
csz CSSP291 09555 K03760 eptA
eno ECENHK 09565 K03760 eptA
kpx PMK1 03140 K03760 eptA
eco b4114 K03760 eptA
sfe SFxv 4483 K03760 eptA
WP 114699275 1 MCR 8 1 Enterobacteriaceae
GUT GENOME027361 00736 MCR 3 Ruminobacter A plasmid
GUT GENOME018049 01412 MCR 3 Succinivibrio plasmid
GUT GENOME264243 00534 MCR 3 Succinivibrio plasmid
GUT GENOME141377 01966 MCR 3 Plesiomonas plasmid
GUT GENOME143143 01952 MCR 4 Enterobacter plasmid
WP 011638903 1 MCR 4 Gammaproteobacteria
WP 099156046 1 MCR 4 Escherichia coli
WP 116786828 1 MCR 4 Salmonella enterica
GUT GENOME145123 03572 MCR 3 Escherichia plasmid
WP 001572373 1 MULTISPECIES MCR 9 1 Enterobacterales
WP 023332837 1 MCR 10 1 Enterobacteriaceae
GUT GENOME148296 03668 MCR 3 Enterobacter plasmid
WP 104009851 1 MCR 7 1 Klebsiella pneumoniae
GUT GENOME154212 02109 MCR 3 Aeromonas plasmid
WP 017778762 1 MCR 3 Aeromonas
WP 099156047 1 MCR 3 Aeromonas media
GUT GENOME147136 04018 MCR 3 Aeromonas plasmid
GUT GENOME144057 04882 MCR 3 Escherichia plasmid
WP 039026394 1 MCR 3 Enterobacteriaceae
WP 102607465 1 MCR 3 Escherichia coli
WP 065801616 1 MCR 3 Klebsiella pneumoniae
GUT GENOME201268 01276 MCR 3 CAG 267 plasmid
WP 099982813 1 MCR 6 1 Moraxella sp MSG47 C17
WP 065419574 1 MCR 2 Escherichia coli
WP 078254299 1 MCR 2 Moraxella pluranimalium
GUT GENOME144843 03949 MCR 1 Escherichia plasmid
WP 065274078 1 MCR 1 Enterobacteriaceae
WP 109545052 1 MCR 1 Klebsiella pneumoniae
WP 076611061 1 MCR 1 Escherichia coli
WP 049589868 1 MCR 1 Gammaproteobacteria
WP 077248208 1 MCR 1 Salmonella enterica
WP 096807442 1 MCR 1 Moraxella sp MSG13 C03
GUT GENOME007189 02307 MCR 5 Dechloromonas plasmid
GUT GENOME095248 01378 MCR 5 Acidovorax D plasmid
WP 148044478 1 MCR 5 hospital metagenome
WP 053821788 1 MCR 5 Proteobacteria
WP 109545057 1 MCR 5 Escherichia coli
GUT GENOME072872 00416 MCR 5 Parasutterella plasmid
GUT GENOME099465 01262 MCR 5 Parasutterella plasmid
GUT GENOME205147 01392 MCR 5 Duodenibacillus plasmid
GUT GENOME037029 01077 MCR 5 Sutterella plasmid
GUT GENOME035902 01453 MCR 5 Sutterella plasmid
GUT GENOME195540 02045 MCR 5 Sutterella plasmid
GUT GENOME058115 01343 MCR 5 CAG 521 plasmid
GUT GENOME247400 01236 MCR 4 CAG 521 plasmid
GUT GENOME055824 01194 MCR 4 CAG 521 plasmid
GUT GENOME213168 00510 MCR 4 Sutterella plasmid
GUT GENOME005421 00952 MCR 3 Sutterella plasmid
GUT GENOME005820 00402 MCR 4 Sutterella plasmid
GUT GENOME279579 02203 MCR 4 Sutterella plasmid
GUT GENOME241557 01770 MCR 4 Sutterella plasmid