Identifying possibly novel sources of antimicrobial resistance in uncultivated bacterial and archeal metagenome-assembled genomes

This manuscript (<u>permalink</u>) was automatically generated from <u>jackiepurdue/paper-2019-amr-uba@b65d6e9</u> on October 4, 2019.

Authors

- Jackie Purdue
 - © 0000-0003-3131-0688 · jackiepurdue · ➤ jackiepurdue

 Department of Computer Science, Dalhousie University · Funded by Grant XXXXXXXXX
- Jane Roe

Department of Something, University of Whatever; Department of Whatever, University of Something

Abstract

Antimicrobial resistance (AMR) is a global threat to human health. Efforts to prevent the spread of AMR rely on surveillance of possible AMR determinants. Using phylogenetics to highlight potential sources of AMR could guide researchers in chosing organisms for phenotypic resistance testing. A systematic way of classifying AMR gene variants is an important in comparing these phylogenetic relationshits, facilitating research in the feild. This study examines relationshipts within sets of previously classified canonical AMR family variants to demonstrate their usefulness in identifying AMR candidates. Canonical sequence data from the Comprehensive Antibiotic Resistance Database (CARD) was used to query CARD prevalence data, NCBI sequence data, and predicted resistomes in in a data set of 7903 draft quality metagenome assembled genomes 1 (MAGs) from various uncultivated bacterial and archeal sources (UBAs). Genes associated potentially associated with mobile colistin resistance (MCR) were found in the UBA sources. New Dheli beta-lactamases (NDM), Klebsiella pneumoniae carbapenemases (KPC), and OXA beta-lactamases were not found to be represented in the UBAs, but [TODO: "gradient of the AMR trees"] [TODO: conclusion]

Introduction

Antimicrobials are substances that can halt the growth of and/or kill microorganisms 2. An important part of the microbial adaptive process is the ability to adapt to pressure from antimicrobial agents within the environment. This ability allows microbes to continue to live and grow in the presence of antimicrobial compounds. This is known as antimicrobial resistance (AMR) https://www.cdc.gov/drugresistance/about.html. The abuse and misuse of antimicrobials has led to increasing levels of AMR within clinical settings 3.

The spread of AMR is a growing global health crisis. AMR within pathogenic microbes is a threat to our treatment of infectious diseases, reducing the efficacy of antibiotics, allowing for prolonged and more severe infections, and increased mortality 4. The European Centre for Disease Prevention and Control estimates that 25,000 people die per year due to infections related to AMR 5, with reports showing the annual cost to combat AMR infections in just the USA alone to be between \$21 billion and \$34 billion 6.

Microbes can be intrinsically resistant to AMR or acquire resistance via lateral gene transfer 7. These mechanisms allow for the constant change of the resistome, making determining the scope of AMR within both pathogenic and non-pathogenic bacteria difficult 3. To impede the spread of AMR the World Health Organization (WHO) created an action plan which emphasizes the need to strengthen the knowledge and surveillance of AMR. An important aspect of this research is to examine the spread of AMR between environmental and clinical samples url: https://www.who.int/antimicrobial-resistance/global-action-plan/en/. Characterization of many AMR genes is currently being done through curated databases. The Comprehensive Antibiotic Resistance Database (CARD) is an example of one of these curated databases. CARD contains reference data pertaining to the genes, proteins and mutations involved in AMR. CARD allows for the identification of resistance genes within a genome using the Resistance Gene Identifier (RGI). 3.

These curated databases contain a large amount of information on the resistomes that are most commonly studied, especially those pertaining to *Staphylococcus aureus*, *Escherichia coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa*. These pathogens are multidrug resistant and are the leading cause of hospital acquired infections §. Most antibiotics have no efficacy on these infections and are therefore treated with carbapenems or polymyxins §.

Carbapenems are broad spectrum beta-lactam antibiotics typically used to treat life threatening, high risk, and multidrug resistant bacterial infections. beta-lactamase producing bacteria have difficulty

degrading carbapenems when combined with beta-lactamase inhibitors making them effective treatments against pathogens with multi-drug resistance 5. However, there has been a recent emergence of resistance to these antimicrobials, presenting a new threat to public health. One way microbes can develop carbapenem resistance is through the acquisition of carbapenamase genes. 5. Three main carbapenamase genes of concern are Klebsiella pneumoniae carbapenemase (KPC), New Delhi metallo-beta-lactamase (NDM) and OXA beta-lactamases 5.

Carbapenamases are classified as one of four classes (A-D) based on their molecular structure 9. Class A, C, and D both use a serine residue to hydrolyze beta-lactam antibiotics 10. KPC genes are a part of class A. Many variants of KPC are spreading and being discovered, however carbapenamases within this class are more rare than the other three classes 5. Class D contains the OXA family of beta-lactamases. Unlike KPC and NDM, this family is characterized by phenotype rather than genotype. This family is characterized by specific hydrolases that can hydrolyze oxacillin and cloxacillin, with the enzyme being poorly inhibited by clavulanic acid https://card.mcmaster.ca/ontology/36026. This results in a low amount of sequence homology within the family. A subset of this family, OXA-48 possesses carbapenem hydrolyzing activity 11. The OXA-48 subfamily consists of five variants: OXA-48, OXA-162, OXA-163, OXA-181, OXA-204, and OXA-232 12. Class B beta-lactamases are metalloenzymes that use a zinc active site to hydrolyze beta-lactam antibiotics. The NDM gene is classified as a class B beta-lactamase 5.

When pathogens become resistant to carbapenems, the last line of defense is colistin. Colistin is a polymyxin antibiotic. Polymyxins act via disrupting membrane permeability 13. Colistin is limited to a last line defense against carbapenem resistant infections due to its neurotoxic and nephrotoxic effects 13. Despite its already limited use, mobilized colistin resistance (MCR) have been discovered 14. Due to the emergence of resistance to many last-line antibiotics, it is important to characterize the total diversity and spread of these resistance determinants.

The UBA dataset contains the data needed to expand AMR surveillance to a wide range of environmental samples. Coupling RGI predictions of the UBA metagenomes with phylogenetic analysis of the predicted AMR genes enables the possibility to discover previously unseen AMR diversity in genomes not yet analyzed for AMR, and provides insight into the diversity of AMR across non-clinical samples.

Methods

AMR Prediction

Materials

A data-set of 7903 draft quality MAGs which were recovered from the Sequence Read Archive by Parks $\underline{1}$ were used for this analysis. RGI version 5.0 with CARD database version 3.02 $\underline{3}$ was run on the contigs of the 7903 MAGS with the inclusion of loose, perfect, and strict hits.

CARD prevalence data version 3.0.4 ("based on sequence data acquired from NCBI on 28-Feb-2019, analyzed using RGI 4.2.2 (DIAMOND homolog detection) and CARD 3.0.1") was used to query the UBA data. [TODO: Other versions of software]

MCR

Obtaining reference MCR genes

24 canonical sequence variants (see Table S1) as indicated by CARD for the MCR phosphoethanolamine transferase [AMR Gene Family], were obtained from the CARD database.

Additional putative MCR phosphoethanolamine transferase gene family sequences were accumulated by querying the CARD prevalence data. The prevalence data was translated to a BLAST database. The 24 CARD canonical sequences were used to perform a multiple query BLASTP against the prevalence blast database with a e-value threshold of 1e-180 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. Many of the sequences are nearly identical, thus they were further processed by clustering with CD-HIT at a minimum sequence identity of 99%

The 24 CARD canonical sequences were used to perform a multiple query BLASTP against the NCBI non-redundant database blast database with a e-value threshold of 1e-180 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. There are many highly sampled taxa and genes in the non-redundant database. To balance the distribution, and reduce the size of the non-redundant sequence set, CD-HIT was used to cluster the data at 70%.

Obtaining candidate MCR genes from the UBA data

Sequences possibly containing AMR gene prediction data for the MCR phosphoethanolamine transferase gene family was produced by filtering RGI output based on the its association with 'MCR phosphoethanolamine'. The filtered data were translated to a blast database. The 24 CARD canonical sequences were used to perform a multiple query BLASTP against this UBA blast database with a evalue threshold of 1e-170 and a query coverage of 99%. For simplicity in identifying the taxonomic history of the non-redundant hits, MULTISPECIES sequences were removed from the analyses. Redundant results were filtered from these BLASTP results by retrieving only the longest sequence for each uniquely labeled result.

Phylogenetic analysis of the MCR genes

To compare the phylogenetic relationship of only the putative sequences, the sequences from the CARD prevalence data were concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the G+I+G4 model of substitution.

For the overall comparison of sequences, the filtered sequences from NCBI non-redundant data, CARD prevalence data, UBA data were all concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the G+I+G4 model of substitution.

NDM

Obtaining reference NDM genes

14 canonical sequence variants (see Table S1) as indicated by CARD for the NDM beta-lactamase [AMR Gene Family], were obtained from the CARD database.

Additional putative NDM beta-lactamase gene family sequences were accumulated by querying the CARD prevalence data. The prevalence data was translated to a BLAST database. The 14 CARD canonical sequences were used to perform a multiple query BLASTP against the prevalence blast database with a e-value threshold of 1e-180 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. Many of the sequences are nearly identical, thus they were further processed by clustering with CD-HIT at a minimum sequence identity of 99%

The 14 CARD canonical sequences were used to perform a multiple query BLASTP against the NCBI non-redundant database blast database with a e-value threshold of 1e-180 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. CD-HIT was used to cluster the data at 70%.

Obtaining candidate NDM genes from the UBA data

Sequences possibly containing AMR gene prediction data for the NDM beta-lactamase gene family was produced by filtering RGI output based on the its association with 'NDM beta-lactamase'. The filtered data were translated to a blast database. The 14 CARD canonical sequences were used to perform a multiple query BLASTP against this UBA blast database with a e-value threshold of 1e-10 and a query coverage of 60%. These values were loosened more for this analysis because at higher thresholds, there were no hits for UBA sequences in the results. For simplicity in identifying the taxonomic history of the non-redundant hits, MULTISPECIES sequences were removed from the analyses. Redundant results were filtered from these BLASTP results by retrieving only the longest sequence for each uniquely labeled result.

Phylogenetic analysis of the NDM genes

To compare the phylogenetic relationship of only the putative sequences, the sequences from the CARD prevalence data were concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

For the overall comparison of sequences, the filtered sequences from NCBI non-redundant data, CARD prevalence data, UBA data were all concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

OXA-48

Obtaining reference OXA-48-like genes

Because of the genetic diversity of the OXA gene family, only the OXA-48 like genes were used for analysis. 6 canonical sequence variants doi:10/db7w (see Table S1) OXA beta-lactamase [AMR Gene Family], were obtained from the CARD database.

Additional putative OXA beta-lactamase gene family sequences were accumulated by querying the CARD prevalence data. The prevalence data was translated to a BLAST database. The 6 CARD canonical sequences were used to perform a multiple query BLASTP against the prevalence blast database with a e-value threshold of 1e-100 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. Many of the sequences are nearly identical, thus they were further processed by clustering with CD-HIT at a minimum sequence identity of 99%

The 6 CARD canonical sequences were used to perform a multiple query BLASTP against the NCBI non-redundant database blast database with a e-value threshold of 1e-100 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. CD-HIT was used to cluster the data at 70%.

Obtaining candidate OXA-48-like genes from the UBA data

Sequences possibly containing AMR gene prediction data for the OXA beta-lactamase gene family was produced by filtering RGI output based on the its association with 'OXA beta-lactamase'. The filtered data were translated to a blast database. The 6 CARD canonical sequences were used to perform a multiple query BLASTP against this UBA blast database with a e-value threshold of 1e-10 and a query coverage of 95%. For simplicity in identifying the taxonomic history of the non-redundant hits, MULTISPECIES sequences were removed from the analyses. Redundant results were filtered from these BLASTP results by retrieving only the longest sequence for each uniquely labeled result.

Phylogenetic analysis of OXA-48-like genes

To compare the phylogenetic relationship of only the putative sequences, the sequences from the CARD prevalence data were concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

For the overall comparison of sequences, the filtered sequences from NCBI non-redundant data, CARD prevalence data, UBA data were all concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

KPC

Obtaining reference KPC genes

18 canonical sequence variants (see Table S1) KPC beta-lactamase [AMR Gene Family], as indicated by the card database, were obtained.

Additional putative KPC beta-lactamase gene family sequences were accumulated by querying the CARD prevalence data. The prevalence data was translated to a BLAST database. The 6 CARD canonical sequences were used to perform a multiple query BLASTP against the prevalence blast database with a e-value threshold of 1e-100 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. Many of the sequences are nearly identical, thus they were further processed by clustering with CD-HIT at a minimum sequence identity of 99%

The 6 CARD canonical sequences were used to perform a multiple query BLASTP against the NCBI non-redundant database blast database with a e-value threshold of 1e-40 and a query coverage of 99%. Redundant results were filtered from these BLAST results by retrieving only the longest sequence for each uniquely labeled result. CD-HIT was used to cluster the data at 70%.

Obtaining candidate KPC genes from the UBA data

Sequences possibly containing AMR gene prediction data for the KPC beta-lactamase gene family was produced by filtering RGI output based on the its association with 'KPC beta-lactamase'. The filtered data were translated to a blast database. The 6 CARD canonical sequences were used to perform a multiple query BLASTP against this UBA blast database with a e-value threshold of 1e-10 and a query coverage of 60%. For simplicity in identifying the taxonomic history of the non-redundant hits, MULTISPECIES sequences were removed from the analyses. Redundant results were filtered from these BLASTP results by retrieving only the longest sequence for each uniquely labeled result.

Phylogenetic analysis of KPC genes

To compare the phylogenetic relationship of only the putative sequences, the sequences from the CARD prevalence data were concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

For the overall comparison of sequences, the filtered sequences from NCBI non-redundant data, CARD prevalence data, UBA data were all concatenated to one multi-FASTA format file with the canonical sequences. These the concatenated amino acid sequences were aligned with MAFFT-LINSI. The aligned sequences were trimmed by trimal using the automated1 option. IQ-TREE was then used to build a bootstrapped tree with -bb 1000 with the LG+I+G4 model of substitution.

Labeling of phylogenies with metadata

All leaves of the trees were labeled with as much taxonomic information as possible for each rank, based on information from metadata as per table S2.

Results

Phylogenetic analysis of putative MCR sequences

The phylogenetic relationships of the CARD canonical sequences, and the CARD prevalence sequences involving the MCR family were investigated. A total of 30 genes, an out-group, the 23 canonical sequences, and the 6 prevalence sequences (clustered as per table S3), were selected for analyses. The tree in Figure 1 shows several distinct clades. Each MCR variant forms its own clade. MCR-1, MCR-2, and MCR-6 form a clade, appearing to more closely related to one another than with the other MCR family members.

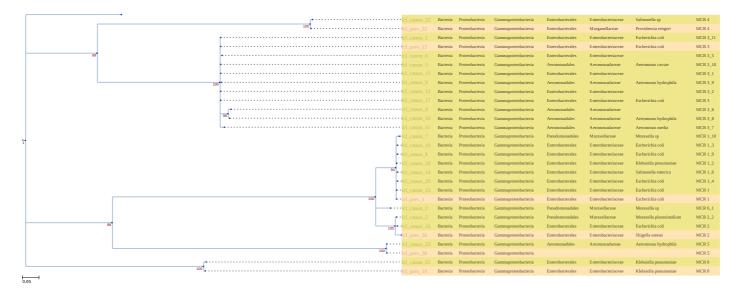


Figure 1: Phylogenetic relationship for 23 canonical (lcl_canon) MCR family sequences along with 6 prevalence (lcl_prev) sequences.

Phylogenetic analysis of NCBI data

Sequences from the NCBI non-redundant data were added to the analysis to be compared with the canonical and prevalence sequences. This resulted in a total of 409 sequences for subsequent analysis, all labeled as phosphoethanolamine lipid A transferase genes, where 104 hits were labeled as MCR family genes.

One of the BLAST results was a variant which was not included in the CARD canonical or prevalence sequences. The variant, MCR-9, was reported in the literature to be a novel MCR homologue with similarity to MCR-3 15. The gene was shown to confer resistance to colistin when expressed in *Escherichia coli*. The minimum inhibitory concentration (MIC) was inferred by broth microdilution as per the European Committee on Antimicrobial Susceptibility Testing (EUCAST; http://www.eucast.org)

The phylogeny in Figure 2 shows that the, the MCR 3 clade branches with MCR-9. This MCR-3 and MCR-9 clade is separated by a large amount diversity from the MCR 1,2,4,5, and 8 clades. This diversity is represented by many bacteria containing phosphoethanolamine lipid A transferase with none reported to be in the MCR family.

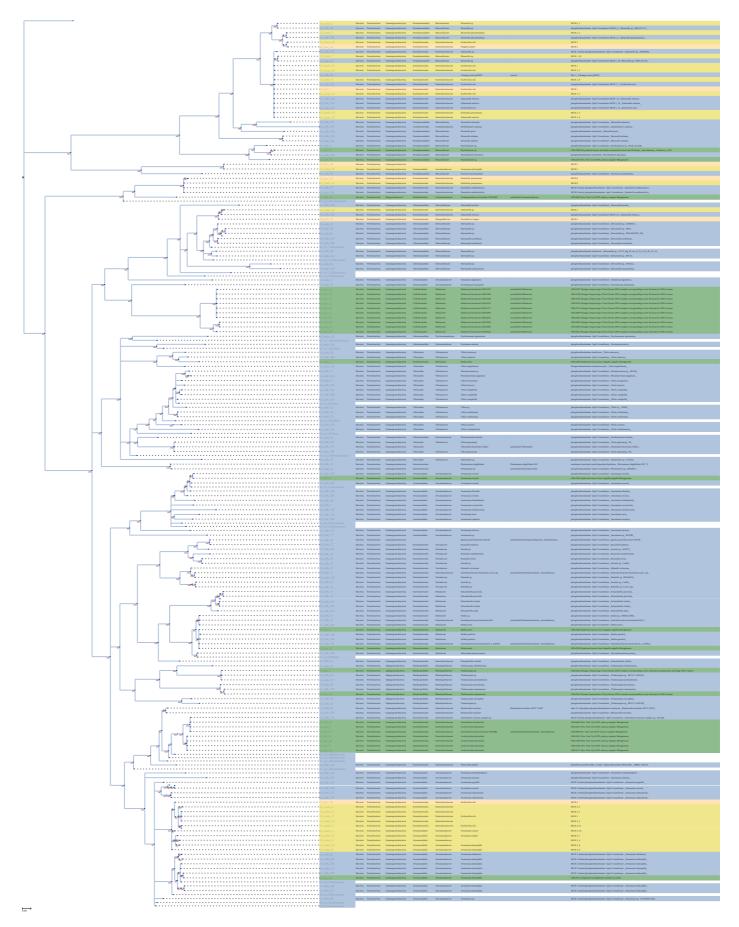


Figure 2: Phylogenetic relationship (lcl_canon) MCR family sequences along with 6 prevalence (lcl_prev) sequences.

Phylogenetic analysis of UBA sequences with RGI-CARD

RGI produced 1457246 results AMR determinants under the loose cutoff from the UBA data, 7171 for the strict cutoff,and 310 for the loose cutoff. It was hoped that pyhlogenetic analysis could find AMR

determinants would be found in under the loose criterion that may have been missed by RGI-CARD. The UBA BLAST results were included in the phylogeny in Figure 2 for the analysis. The diversity between clades, swathes of which are not reported to be in the MCR family, contain multiple sources from UBAs. It is unclear if these relationships could be involved with the MCR family, but there are more enticing relationships present in the tree. The remainder of the analysis deals with relationships which are deemed to be interesting based on the locations of the UBAs between MCR family clades.

MCR-5, MCR-1,2,6 clades

The MCR-5, MCR1,2,6clades shown in figure ?? share, with 100 bootstrap support, a common ancestor. Branching between these two clades, as a descendant to this ancestor, is a clade of Psycrobacter species bacteria. This clade includes two hits from the non-redundant database, and two hits from the UBA data. According to the Parks 1 data, the identity of these samples are, UBA3068, A Psychrobacter sp., sampled from Oil polluted marine microbial communities from Coal and Oil in Point Santa Barbara, California, USA and, UBA4193, a Psychrobacter sp, sampled from the New York City MTA subway samples. The quality information for these sequences, shown in Table ??, shows that UBA3068 is near complete, while UBA4193 is only partial. It is encouraging to see that even with the quality difference, these two sequences branch in the expected clade. Psychrobacter 16 is a Genus is widespread and includes many cold adapted bacteria, it is an opportunistic pathogen, and has been found sometimes be a cause of infections in humans, animals, and fish. Many new species of Psychrobacter have been discovered in cold climates 17. Some of the species have been shown to be resistant to colistin, like Psychrobacter vallis ps. nov. and Psychrobacter aquaticus ps. nov. 18.

MCR-3 clade

A single result, UBA705, which the Parks data 1 reports as a Comparative metagenome analyses of anode-associated microbial communities developed in rice paddy field-soil microbial fuel cells, is reported to be Aeromonas hydrophilia. This present within the clade alone with several other canonical and non-redundant Aeromonadalacea. Aeromonas hydrophilia is a species which has been found to have an MCR-3 gene. Again, even though qualities vary (Table ??), the sequences branch in the expected location.

MCR 3,9 clade

This clade contains 8 UBAs from the loose CARD hits under the given constraints, and several NCBI non-redundant hits not reported as MCR family genes. 7 Loose hits for MCR in a /Leclercia/ adecarboxylata branch within this clade, branching closely with the NCBI non-redundant MCR9. These /Leclercia/ branch below the common ancestor of MCR-9 and MCR-3 which are well supported. The /Leclercia/ UBAs were all sampled from New York City MTA subway samples Metagenome. Acinetobacter is another opportunistic pathogen 19 which is becoming resistant to many antimicrobials.

Phylogenetic analysis of NDM

The phylogenetic relationships of the NDM family were investigated. The 14 canonical sequences were used to query the same databases as in the phylogenetic analyses of MCR.

Table 1: Example table(TODO: supplemental?).

lcl prev 2	lcl prev 3
lcl prev 4 at 99.63%	lcl prev 6 at 99.26%
lcl prevd 5 at 99.26%	lcl prev 7 at 99.25%

lcl prev 2	lcl prev 3
Icl prev 8 at 99.26%	-
lcl prev 9 at 99.63%	-

The genes retained for the phylogeny were the 14 canonical sequences, 3 prevalence sequences ((clustered as per table S3)), 7 non-redundant sequences, 1 UBA sequence, and 1 out-group.

This total of 27 resulted in the phylogeny in Figure 3.

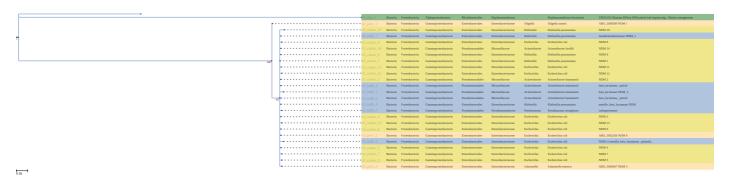


Figure 3: Phylogenetic relationship (lcl_canon) NDM family sequences along with 14 prevalence (lcl_prev) sequences.

The 7 sequences resulting from the NCBI data included 3 genes reported to be NDM-1, NDM-3, and NDM-5. The other 4 genes were reported to be general beta lactamase/carbapenemase hits. These sequences form a closely related multifurcation (due to poor support values) with the prevalence and canonical data. The only phylogenetically resolved gene in this clade was a prevalence hit for NDM-1 found in Shigella sonnei representing only itself in the cd-hit cluster. Interestingly other representatives from other purported NDM families are more closely related to each other, including the canonical NDM-1 sequence from Klebsiella pneumoniae. The one UBA BLAST result branches far from the clade containing the canonical indicating that under the coverage queried, there are no reasonably detectable NDM homologues in the UBA data. The alignment of this UBA under this relaxed query coverage of 60% is already pushing the limits of a "good alignment", and reducing this further would produce meaningless results.

Phylogenetic analysis of OXA-48

In investigating the phylogenetic relationship of OXA-48, the result was similar to that of NDM. There were multiple BLAST results for UBA sequences, but the hits were too phylogenetically dissimilar to draw a conjecture about their relationship to the OXA family. 6 canonical sequences, and the 5 prevalence sequences (clustered according to Table S3), 11 non-redundant sequences, and 112 UBA sequences were combined in the phylogeny in Figure 4

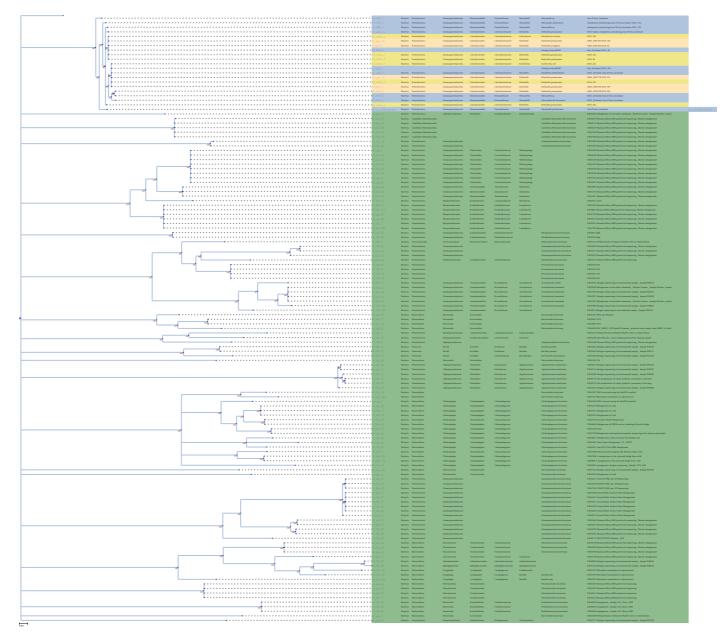


Figure 4: Phylogenetic relationship (lcl_canon) OXA-48 family sequences along with 14 prevalence (lcl_prev) sequences.

The OXA-48 prevalence hits added further diversity to the referecne OXA-48-like sequences. OXA-436 20 was found, clustered with no other gene, in the prevalence data, and OXA-514 and OXA-515 were found in the non-redundant data.

Phylogenetic analysis of KPC

18 canonical sequences, and the 8 prevalence sequences (clustered according to Table S3), non-redundant sequences, and 7 UBA sequences were combined in the phylogeny in Figure 4 In investigating the phylogenetic relationship of KPC also resulted in UBA sequences which were too distant from the canonical sequences to make a conjecture regarding their relationship to KPC family. In Figure 5 the 7 resulting UBA sequences branch with class A beta-lactamases as reported by the non-redundant BLAST.

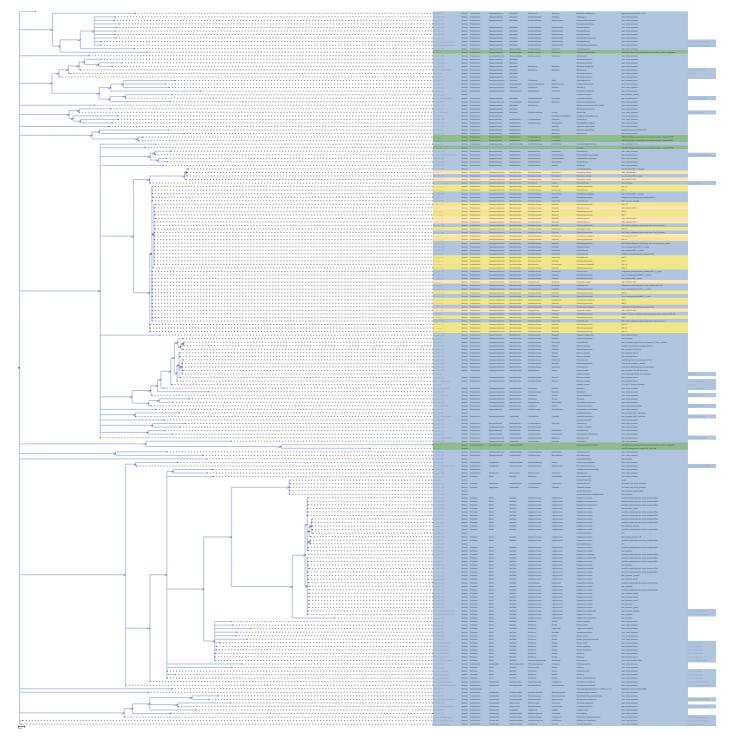


Figure 5: Phylogenetic relationship (lcl_canon) KPC family sequences along with 14 prevalence (lcl_prev) sequences.

Discussion

Conclusion

S1: MCR 3.11, 6.1, 2.2, 3.6, 3.10, 3.5, 1.10, 1.9, 3.9, 3.8, 3.7, 3.2, 3.1, 1.6, 1, 2, 3, 1.2, 1.3, 1.4, 8, 4, 5, 7.1

References

1. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life

Donovan H. Parks, Christian Rinke, Maria Chuvochina, Pierre-Alain Chaumeil, Ben J. Woodcroft, Paul N. Evans, Philip Hugenholtz, Gene W. Tyson

Nature Microbiology (2017-09-11) https://doi.org/cczd

DOI: 10.1038/s41564-017-0012-7 · PMID: 28894102

2. Antibiotic Resistance☆

B. Périchon, P. Courvalin, C. W. Stratton

Reference Module in Biomedical Sciences (2015) https://doi.org/dbhq

DOI: <u>10.1016/b978-0-12-801238-3.02385-0</u>

3. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database

Baofeng Jia, Amogelang R. Raphenya, Brian Alcock, Nicholas Waglechner, Peiyao Guo, Kara K. Tsang, Briony A. Lago, Biren M. Dave, Sheldon Pereira, Arjun N. Sharma, ... Andrew G. McArthur Nucleic Acids Research (2016-10-26) https://doi.org/f9wbjs

DOI: 10.1093/nar/gkw1004 · PMID: 27789705 · PMCID: PMC5210516

4. Horizontal transfer of antibiotic resistance genes in clinical environments

Nicole A. Lerminiaux, Andrew D. S. Cameron

Canadian Journal of Microbiology (2019-01) https://doi.org/gfnrkj

DOI: <u>10.1139/cjm-2018-0275</u> · PMID: <u>30248271</u>

5. Carbapenem Resistance: A Review

Francis Codjoe, Eric Donkor

Medical Sciences (2017-12-21) https://doi.org/c9zd

DOI: 10.3390/medsci6010001 · PMID: 29267233 · PMCID: PMC5872158

6. Combating Antimicrobial Resistance: Policy Recommendations to Save Lives Clinical Infectious

Diseases (2011-05-01) https://doi.org/cgt3qk

DOI: 10.1093/cid/cir153 · PMID: 21474585 · PMCID: PMC3738230

7. Antibiotic resistance is ancient

Vanessa M. D'Costa, Christine E. King, Lindsay Kalan, Mariya Morar, Wilson W. L. Sung, Carsten Schwarz, Duane Froese, Grant Zazula, Fabrice Calmels, Regis Debruyne, ... Gerard D. Wright Nature (2011-08-31) https://doi.org/b3wbvx

DOI: 10.1038/nature10388 · PMID: 21881561

8. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens

Sirijan Santajit, Nitaya Indrawattana

BioMed Research International (2016) https://doi.org/f9h4q3

DOI: 10.1155/2016/2475067 · PMID: 27274985 · PMCID: PMC4871955

9. The Structure of \$\beta\$-Lactamases

R. P. Ambler

Philosophical Transactions of the Royal Society B: Biological Sciences (1980-05-16)

https://doi.org/cdckks

DOI: 10.1098/rstb.1980.0049 · PMID: 6109327

10. Updated Functional Classification of -Lactamases

K. Bush, G. A. Jacoby

Antimicrobial Agents and Chemotherapy (2009-12-07) https://doi.org/bp6dp2

DOI: 10.1128/aac.01009-09 · PMID: 19995920 · PMCID: PMC2825993

11. Class D OXA-48 Carbapenemase in Multidrug-Resistant Enterobacteria, Senegal

Olivier Moquet, Coralie Bouchiat, Alfred Kinana, Abdoulaye Seck, Omar Arouna, Raymond Bercion, Sebastien Breurec, Benoit Garin

Emerging Infectious Diseases (2011-01) https://doi.org/cn235v

DOI: <u>10.3201/eid1701.100244</u> · PMID: <u>21192883</u> · PMCID: <u>PMC3204621</u>

12. Rapid Identification of OXA-48 and OXA-163 Subfamilies in Carbapenem-Resistant Gram-Negative Bacilli with a Novel Immunochromatographic Lateral Flow Assay

Fernando Pasteran, Laurence Denorme, Isabelle Ote, Sonia Gomez, Denise De Belder, Youri Glupczynski, Pierre Bogaerts, Barbara Ghiglione, Pablo Power, Pascal Mertens, Alejandra Corso *Journal of Clinical Microbiology* (2016-08-17) https://doi.org/dbhs

DOI: <u>10.1128/jcm.01175-16</u> · PMID: <u>27535687</u> · PMCID: <u>PMC5078564</u>

13. Treatment Options for Carbapenem-Resistant Enterobacteriaceae Infections

H. J. Morrill, J. M. Pogue, K. S. Kaye, K. L. LaPlante

Open Forum Infectious Diseases (2015-05-05) https://doi.org/dbht DOI: 10.1093/ofid/ofv050 · PMID: 26125030 · PMCID: PMC4462593

14. Towards Understanding MCR-like Colistin Resistance

Jian Sun, Huimin Zhang, Ya-Hong Liu, Youjun Feng

Trends in Microbiology (2018-09) https://doi.org/gdqcfq

DOI: <u>10.1016/j.tim.2018.02.006</u> · PMID: <u>29525421</u>

15. Identification of Novel Mobilized Colistin Resistance Gene mcr-9 in a Multidrug-Resistant, Colistin-Susceptible Salmonella enterica Serotype Typhimurium Isolate

Laura M. Carroll, Ahmed Gaballa, Claudia Guldimann, Genevieve Sullivan, Lory O. Henderson, Martin Wiedmann

mBio (2019-05-07) https://doi.org/gf7r4q

DOI: <u>10.1128/mbio.00853-19</u> · PMID: <u>31064835</u> · PMCID: <u>PMC6509194</u>

16. The Genus Psychrobacter

Elliot Juni

The Prokaryotes (1992) https://doi.org/dbpr

DOI: <u>10.1007/978-1-4757-2191-1 12</u>

17. Novel Psychrobacter Species from Antarctic Ornithogenic Soils

J. P. BOWMAN, J. CAVANAGH, J. J. AUSTIN, K. SANDERSON

International Journal of Systematic Bacteriology (1996-10-01) https://doi.org/fd9rb9

DOI: 10.1099/00207713-46-4-841 · PMID: 8863407

18. Psychrobacter vallis sp. nov. and Psychrobacter aquaticus sp. nov., from Antarctica

S. Shivaji

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY (2005-03-01)

https://doi.org/cdq9w4

DOI: <u>10.1099/ijs.0.03030-0</u> · PMID: <u>15774658</u>

19. Natural antimicrobial susceptibility patterns and biochemical profiles of Leclercia adecarboxylata strains

I Stock, S Burak, B Wiedemann

Clinical Microbiology and Infection (2004-08) https://doi.org/b2zr9k

DOI: <u>10.1111/j.1469-0691.2004.00892.x</u> · PMID: <u>15301675</u>

20. Dissemination and Characteristics of a Novel Plasmid-Encoded Carbapenem-Hydrolyzing Class D β -Lactamase, OXA-436, Found in Isolates from Four Patients at Six Different Hospitals in Denmark

Ørjan Samuelsen, Frank Hansen, Bettina Aasnæs, Henrik Hasman, Bjarte Aarmo Lund, Hanna-Kirsti S. Leiros, Berit Lilje, Jessin Janice, Lotte Jakobsen, Pia Littauer, ... Anette M. Hammerum *Antimicrobial Agents and Chemotherapy* (2017-10-23) https://doi.org/db7x

DOI: <u>10.1128/aac.01260-17</u> · PMID: <u>29061750</u> · PMCID: <u>PMC5740310</u>