

Incorporation of metagenome-assembled genome data into antimicrobial resistance surveillance identifies inconsistencies in carbapenemase and colistin resistance gene nomenclature

This manuscript ([permalink](#)) was automatically generated from [jjackiepurdue/phylogenetic-amr-survey-manuscript@496508e](mailto:jackiepurdue/phylogenetic-amr-survey-manuscript@496508e) on September 30, 2020.

Authors

- **Jackie Purdue**
 [0000-0003-3131-0688](#) ·  [jackiepurdue](#) ·  [jackiepurdue](#)
Faculty of Computer Science, Dalhousie University
- **Jocelyn McDonald**
Faculty of Computer Science, Dalhousie University
- **Dayna Mikkelsen**
Faculty of Computer Science, Dalhousie University
- **Finlay Maguire**
 [0000-0002-1203-9514](#) ·  [fmaguire](#) ·  [finlaym](#)
Faculty of Computer Science, Dalhousie University · Funded by Donald Hill Family Fellowship in Computer Science
- **Robert G. Beiko**
Faculty of Computer Science, Dalhousie University

Abstract

The increase and spread of antimicrobial resistance (AMR) poses a global threat to human and animal health. The emergence of mobilised resistance to last-line antibiotics such as carbapenems and polymyxins are a particular cause for alarm. Given their ability to be laterally transferred between unrelated bacteria, surveillance of the evolution and transmission of these genes is vital to mitigating their impact. However, the poor characterisation of many important non-clinical environments and bacteria as well as failings in the existing classification and nomenclature of AMR genes pose a major barrier to this work. Metagenome assembled genome (MAG) methods, in which genomes are recovered from untargeted metagenomic sequencing, have been used to greatly increase the diversity of sampled bacterial genomes. Therefore, by profiling AMR genes detectable in MAG datasets we can greatly expand surveillance of these undersampled environments and organisms. Through this, we can also identify inconsistencies and potential refinements of the existing nomenclature.

In this study, the phylogenetic neighborhoods of several mobilised AMR gene families, of public-health importance, were examined for diversity, spread, and classification consistency. These gene families include: mobile colistin resistance (MCR), New Delhi beta-lactamase (NDM), *Klebsiella pneumoniae* carbapenemase (KPC), and the OXA-48-like beta-lactamase lactamase family. Canonical genes in the Comprehensive Antibiotic Resistance Database (CARD) for these families were used to identify putative homologues within CARD:Resistomes & Variants data, NCBI sequence databases, and draft quality MAGs from a large set of uncultivated bacterial and archeal sources (UBA). For each of these 4 AMR gene families, phylogenetic analyses were conducted for the expanded set of candidate resistance genes.

Although limited uncharacterised diversity in NDM, KPC, and OXA-48 carbapenemases was uncovered, several possibly novel clades of MCR genes were discovered. In general, the carbapenemases families were found to be more tightly defined, with most added diversity being either extremely close, or phylogenetically distant. The MCR family showed many new clades of diversity, even within well defined "canonical" clades. A large number of existing gene names within all 4 families were identified to be inconsistent with phylogeny or inconsequential in relation to the diversity they encompass. Additionally, there was a huge variability in the diversity encompassed by individual allele names across the AMR gene families analysed. These results highlight the need to incorporate MAG data into AMR surveillance work and to develop consistent, robust, phylogenetically defined nomenclature for AMR genes.

Introduction

Antimicrobials, substances which can kill or inhibit the growth of microbes [1], are key effectors in the ecology of microorganisms. The evolution of resistance to these substances, known as antimicrobial resistance (AMR), is an important and ancient process [2]. In the last century we have adopted and used antimicrobials to great effect in clinical medicine and agriculture. However, the use and abuse of antimicrobials has led to increasing levels of observed AMR [3]. AMR poses a growing global health risk by undermining our ability to treat infectious diseases and perform surgeries [4]. Infections by resistant microbes have greater associated mortality and morbidity [4], with the European Centre for Disease Prevention and Control estimating that 25,000 people per year die due to resistant infections [5].

To mitigate the spread of AMR, the World Health Organization (WHO) (and national government bodies e.g., [6]) have identified that strengthening surveillance of AMR, especially transmission of AMR between environmental and clinical samples [7], is vital. The WHO's priority list of antibiotic resistant bacteria [8] highlights the multi-drug resistant pathogens upon which these efforts should be focused. This includes the ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp. [9]. These pathogens have become resistant to many common classes of antimicrobials including penicillins, all generations of cephalosporins, glycopeptides, quinolones, and peptide antibiotics [9]. Concerningly, this also includes mobilised, i.e. laterally transferable between unrelated lineages, resistance to last-line carbapenem and colistin antibiotics [12].

Carbapenems are broad spectrum beta-lactam antibiotics that cannot be degraded by most beta-lactamases, as such they are used to treat life threatening, high risk, and multi-drug resistant bacterial infections [5]. The number of observed carbapenem resistant *K. pneumoniae*, *E. coli*, *E. cloacae*, and *A. baumanii* hospital-acquired infections have increased in Canada since 2013 [13], presenting a new threat to public-health. Resistance to carbapenems is largely mediated by acquisition of one of several families of carbapenemase genes [5]. Due to their mobility and host-range, 3 carbapenemase gene families of public-health concern are *Klebsiella pneumoniae* carbapenemase (KPC), New Delhi metallo-beta-lactamase (NDM), and OXA-48 type beta-lactamases [5].

Carbapenamases are assigned to one of four classes (A-D) based on their molecular structure [14]. Class A, C, and D use a serine residue to hydrolyze beta-lactam antibiotics [15]. KPC genes are a part of class A, and are the most prevalent in that class [5]. Many variants of KPC are spreading and being discovered, however carbapenamases within this class are more rare than those within classes B, C, and D [5]. Class B beta-lactamases are metallo-enzymes that use a zinc active site to hydrolyze beta-lactam antibiotics. The NDM gene is classified as a class B beta-lactamase [5], and is the most prevalent in that class [5]. Class D contains the OXA-48 beta-lactamases which are the most prevalent in that class [5]. Unlike KPC and NDM, the OXA 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 relatively few homology relationships within the family. A subset of this family, OXA-48 possesses carbapenem hydrolyzing activity [16]. The OXA-48 subfamily consists of five variants: OXA-48, OXA-162, OXA-163, OXA-181, OXA-204, and OXA-232 [17].

For carbapenem resistant infections, polymyxin E (colistin) and B antibiotics are frequently the only remaining option. Polymyxins act by disrupting membrane permeability [18] and has had limited use in humans due to its neurotoxic and nephrotoxic effects [18]. Despite this, mobilized colistin resistance (MCR) has emerged in clinical pathogens [19] likely due to transmission from agricultural usage of colistin [20]. This further underscores the vital importance of surveillance for evolution and potential transmission of AMR between environmental, agricultural, and clinical samples.

This kind of surveillance requires large-scale phylogenetic analyses which characterise the total diversity, context, and spread of these key resistance determinants. In order to perform these types of analyses we need two resources: a clear consolidated definition of what constitutes an AMR-related gene and a large amount of sequenced microbial genomes. However, naming and defining AMR genes is difficult to a complex legacy of historical nomenclature (prior to widespread DNA sequencing), and a huge quantity of active AMR research from a broad-range of stakeholders including academia, government organizations, human health, animal health, and agri-food industries [6].

AMR determinants are typically classified into families [21], some of which are based on phenotypic properties, and others on sequence variation. The nomenclature is usually an acronym with varying representations, such as the mechanism of resistance, the organism it is associated with, or the molecular structure. Each acronym is assigned a numerical value to distinguish variants. Each sequence, determined to be novel by some criterion, is assigned a new number. For example, the MCR phosphoethanolamine transferase gene family inhibits the binding of colistin to the cell membrane. The MCR-3 gene has a sequence identity which has been deemed to be significantly different from that of the other MCR genes in the family. MCR-3.1, has been deemed to be a variant of the MCR-3 gene, which has a slight variation in gene sequence from other MCR-3 genes [22]. Members of the NDM beta-lactamase gene family are named using a different set of criteria. In this case, NDM-1, and NDM-2 differ only by a single amino acid substitution. This system for classification has the potential to be misleading when conducting AMR research. The amount of diversity encompassed within different

names and alleles is hugely variable across AMR genes. For example, NDM has little diversity, while MCR has a large amount of sequenced microbial genomes. Some AMR families are not even defined on the basis of sequence space but phenotypically. An example of this is the OXA beta-lactamase family, which is a phenotypically characterized family, with several different homologous sets of genes. OXA-20 has less than 20% amino acid sequence identity when compared with several other OXA-48-like family beta-lactamases [23]. This is a problem when attempting to characterize the AMR determinants by sequence similarity.

High-quality manually curated AMR databases such as the Comprehensive Antibiotic Resistance Database (CARD) [21] provide a unified resource for the definition, nomenclature and classification of AMR genes. CARD organizes this information through the Antibiotic Resistance Ontology (ARO), a controlled vocabulary with defined relationships. There are two datasets within CARD, CARD:Canonical and CARD:Resistomes & Variants (R&V). Canonical is a conservative set of AMR genes and mutations that have been experimentally verified as being associated with resistance in a peer-reviewed publication. As a consequence, most of the entries in CARD:Canonical are derived from the most intensively studied organisms, such as the ESKAPE pathogens. On the other hand, CARD:R&V contains AMR sequences that have been detected by searching a set of 88 WHO priority organism genomes from central repositories of publicly available sequencing data such as National Center for Biotechnology Information (NCBI) for sequences similar to corresponding Canonical sequences. This *in silico* search is performed using CARD's BLAST/DIAMOND-based Resistance Gene Identifier (RGI) tool and greatly increases the sequence diversity in CARD [21].

However, this is still limited to a subset of all genomes currently available in central International Nucleotide Sequence Database Collaboration (INSDC) repositories like NCBI. If we want to thoroughly understand the evolution and spread of AMR genes we need to analyse as many genomes as possible. Unfortunately, the genomes in databases are largely sequenced from microbes that can be easily cultured. As only a subset of microbial diversity can be cultured, this means many of our existing genomes aren't necessarily representative of the environment from which they were sampled [24]. Recently, techniques have been developed that allow the recovery of genomes from metagenomic data e.g., [25]. As metagenomic sequencing, the direct sequencing of all DNA in a sample, doesn't require culturing these metagenome-assembled genomes (MAGs) represent a huge source of novel microbial diversity. Parks et. al. [24], generated new 7,903 bacterial and archaeal (Uncultured Bacteria and Archaea; UBA) MAGs representing >30% increase in the sampled phylogenetic diversity of the bacteria and archaea.

By combining this MAG dataset with CARD:R&V, and other NCBI non-redundant data, we are able to greatly increase the breadth of data available for surveillance beyond clinical samples and environments. Phylogenetic analyses of this data can then be used to characterise the evolution and previously unexplored diversity of these critical public-health priority carbapenemase (KPC, NDM, and OXA-48) and colistin AMR gene families. These analyses will also highlight inconsistencies and possible refinements in the classification and nomenclature of these genes and their variant alleles. Therefore, in this work we present an in-depth investigation of the classification and phylogeny of KPC, NDM, OXA-48, and MCR gene families across all currently available sequenced genomes and MAGs.

Methods

Data

We surveyed four separate databases to assess the diversity of known and candidate AMR genes. These sources include CARD:Canonical and CARD:R&V (version 3.0.5) sequences for the KPC, NDM, OXA-48, and MCR gene families, and associated homologs from NCBI non-redundant data and UBA MAG data. The 32 canonical MCR sequence variants (Table 1), 18 canonical KPC sequence variants (Table 2), 14 canonical NDM sequence variants (Table 3), and 6 canonical OXA-48-like sequence variants (Table 4), as labeled in the CARD database, were obtained as a reference. The NCBI non-redundant data consisted of all non-redundant GenBank, Protein Data Bank (PDB), SwissProt, Protein information resource (PIR), Protein research foundation (PRF) data from May 17, 2019. The MAG data was derived from a data-set of 7,903 draft quality MAGs generated from Sequence Read Archive data by Parks et. al [27]. The SRA metagenomic data from which these MAGs were derived were chosen to represent lineages which were under-sampled (e.g., environmental and non-human gastrointestinal samples).

Querying the data

The CARD canonical sequences from each family were used to perform a multiple query BLASTP (version 2.5.0 [28]) against the prevalence BLAST database with an e-value threshold and query coverages shown in Table 1. Many of the sequences are nearly identical, thus they were further processed by clustering with CD-HIT version 4.8.1 at a minimum sequence identity as per Table 1.

The reference CARD canonical sequences were also used to perform a multiple query BLASTP against the NCBI non-redundant database with an e-value and query coverage threshold indicated in Table 1. 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 version 4.8.1 was also used to cluster the data as per Table 1.

For the metagenomic data, RGI version 5.0 with CARD database version 3.02 [3] was run on the contigs of the 7903 MAGs with the inclusion of loose, perfect, and strict hits. The search strings in Table S1 were used to extract results from the RGI output. The CARD canonical sequences were used to perform a multiple query BLASTP against this UBA blast database with a e-value threshold and query coverage from Table 1. The minimum BLAST query coverages were chosen to produce meaningful alignments.

For all sequence variants obtained from each data source, redundant results for the prevalence, NCBI, and UBA queries were filtered from these BLASTP results by retrieving only the longest sequence for each unique Subject Taxonomy ID.

AMR Determinant	Maximum BLAST e-value			Minimum BLAST Query Coverage			CD-HIT Sequence Identity Threshold		
	CARD	NCBI	UBA	CARD	NCBI	UBA	CARD	NCBI	UBA
MCR	1e-160	1e-160	1e-160	98	98	98	100	100	100
KPC	1e-100	1e-40	1e-10	99	99	60	99	70	100
NDM	1e-160	1e-160	1e-10	98	98	60	100	100	100
OXA	1e-100	1e-100	1e-10	90	99	95	N/A	N/A	N/A

AMR Determinant	Maximum BLAST e-value			Minimum BLAST Query Coverage			CD-HIT Sequence Identity Threshold		
	CARD	NCBI	UBA	CARD	NCBI	UBA	CARD	NCBI	UBA
Table 1: Parameters used with BLAST and CD-HIT to obtain sequences for the phylogenies of MCR phosphoethanolamine transferase, and KPC, NDM, and OXA-48 beta-lactamases. For each determinant, the minimum BLAST e-values, minimum query coverages, and CD-HIT Sequence Identity threshold for clustering sequences are displayed for each database. CARD represents the RGI-CARD prevalence sequence data, NCBI represents the NCBI non-redundant sequence data, and UBA represents the MAG data source.									
After filtering, OXA did not have an overabundance of close sequences and was not clustered (marked as N/A)									

Table 1: Names of the canonical MCR phosphoethanolamine transferase variants from CARD database {#tbl:params}

Sequence Alignment

Two alignments were created for each AMR determinant under study. The first alignment was created to compare the phylogenetic relationship of the canonical and prevalence (putative) sequences. The second alignment was created for an comparison of sequences from all four databases.

The putative AMR alignment was made up of the sequences from the CARD prevalence data were combined in one multi-FASTA format file with the canonical sequences and an outgroup chosen for each AMR family as per table ?? [TODO: make a table of outgroups in supplemental]. In the second, more diverse alignment, The filtered sequences from NCBI non-redundant data, CARD prevalence data, UBA data were all combined into one multi-FASTA format file with the canonical sequences with the same outgroup sequences (table ??).

This set of combined amino acid sequences were aligned with MAFFT-LINSI [29] version 7.40 and trimmed by trimal [30] version 1.4.rev22 using the automated1 option.

Creation of phylogenies

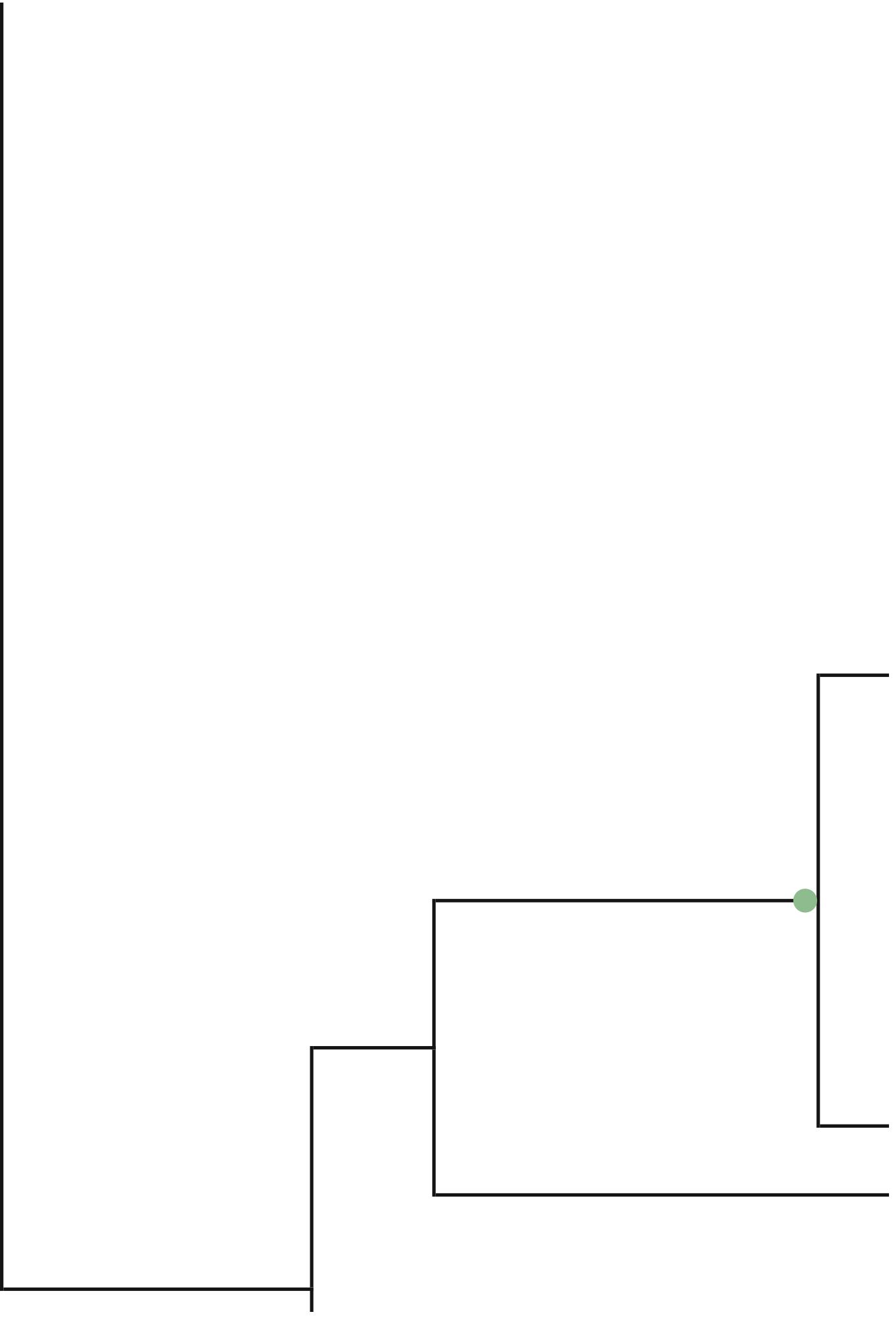
For each alignment under each AMR family under study, IQ-TREE multicore version 1.6.9 [31] was used to build a bootstrapped tree with -bb 1000 with the G+I+G4 model of substitution. Tree visualizations were created with ETE Toolkit [32] version 3. The trees were annotated with taxonomic information for each rank by using ETE's NCBI taxonomy module to translate the names result set from the various databases to taxonomic identifiers. The taxonomic identifier was used to query NCB's taxonomy database from May 17, 2019, for the taxonomy information. In cases where there were multiple hits for genes, the most recent common ancestor was used to query the taxonomic information, therefore some leaves in the trees were annotated with only partial taxonomic information.

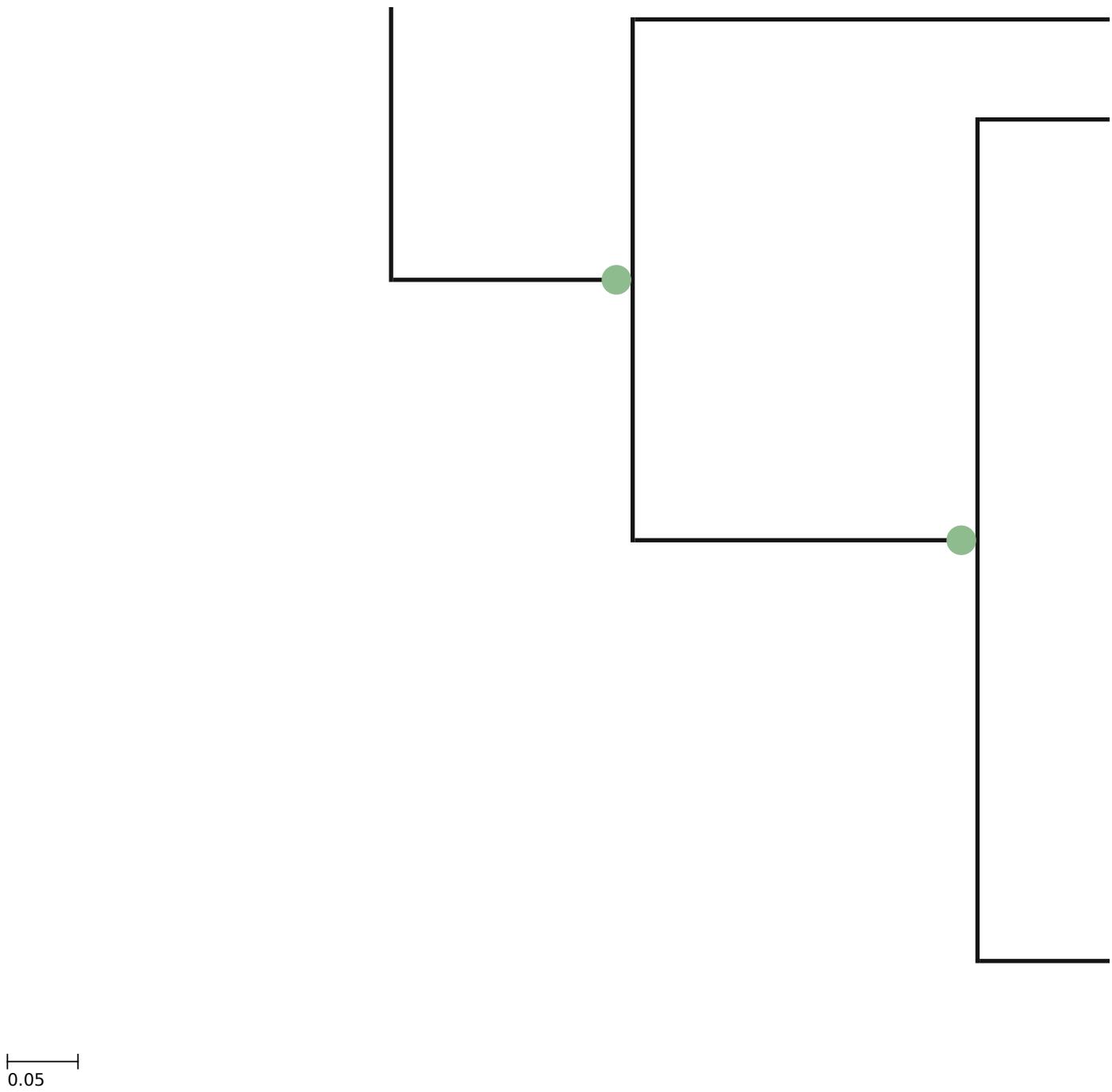
Results

Phylogenetic analysis of MCR sequences

A phylogeny with only CARD:Canonical and CARD:R&V MCR gene family sequences was created, showing the relationships without diversity from NCBI or the MAGs, including an out-group, 32 CARD:Canonical sequences, and 54 CARD:R&V sequences. The resulting tree in Figure {[??]} shows MCR family members 1 through 9 forming distinct sub-clades in the tree. MCR-1, MCR-2, and MCR-6 form a closely related clade. This clade is closely related to ICR-Mc [33], another phosphoethanolamine transferase which confers colistin resistance. MCR-3 and MCR-7, also form a closely related clade.







The tree visualisation was further condensed to only a single representative for each MCR allele (figure 1). Regions of the tree vary in their potential to accept additional diversity. Some clades are tightly related, while others have more distance between them. Newly discovered genes could some day be found to be sister to these more distantly related clades. Regions containing more evolutionary distance are accommodating to additional diversity, while closely related regions are not.

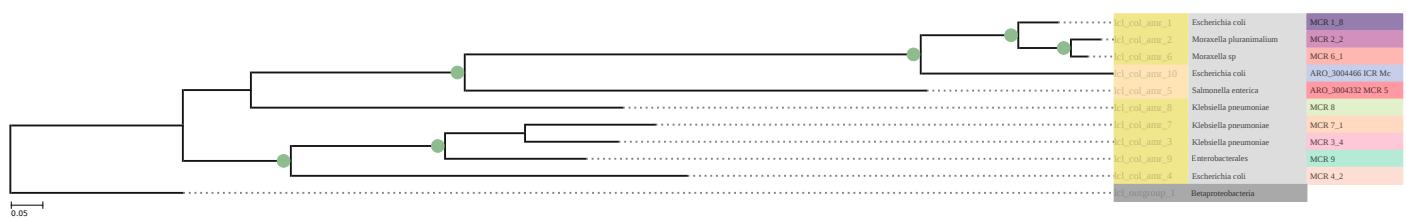


Figure 1: Phylogenetic relationship of 9 representative MCR family sequences, and an outgroup from Betaproteobacteria. (labels prefixed with 'lcl_canon' in yellow are canonical, labels prefixed with 'lcl_prev' in tan are CARD:R&V, and an outgroup from Betaproteobacteria is prefixed with 'lcl_outgroup' in grey)

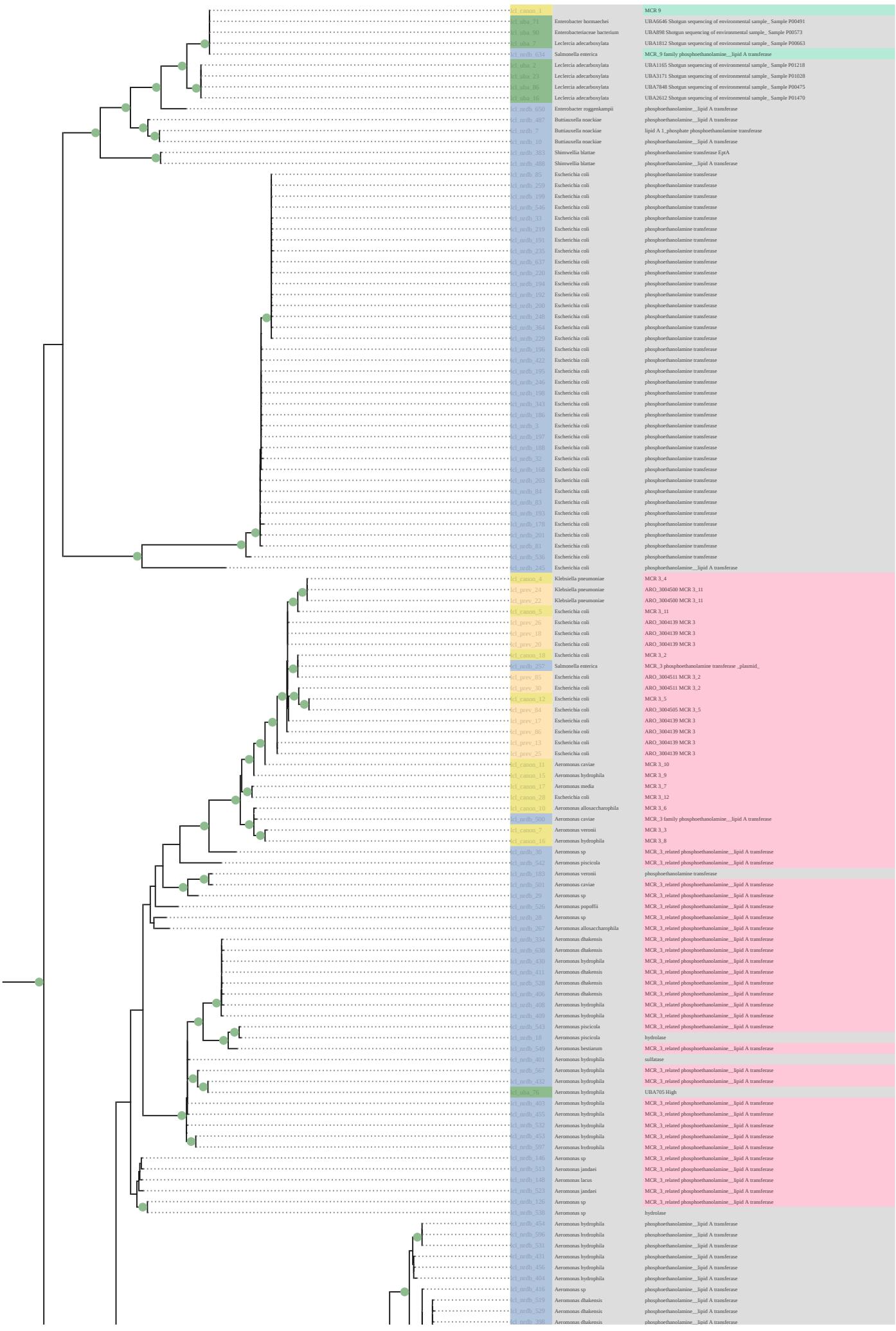
The phylogenies represented by Figures 2, 3, and 7 were created with additional diversity by querying the NCBI non-redundant database, and the MAGs from the Parks *et. al.* UBA data [27]. The filtering criteria in Table 1 were applied to these data sets.

After filtering, 409 NCBI sequence hits, labeled as phosphoethanolamine lipid A transferase, and 104 NCBI sequence hits, labeled as MCR family, were selected for addition to the phylogeny. The 7903 draft quality MAGs were queried for AMR genes with RGI. RGI produced 1457246 loose hits, 7171 strict hits, and 310 perfect hits. After applying the filters from Table 1, only 63 MAG derived genes were retained for the phylogeny. Due to the large size, the phylogeny displayed in Figure 7 was subdivided into Figures 2, and 3. The subdivisions were chosen based on a pruning at the point of the most recent common ancestor of sets of canonical MCR labeled leaves. The set including MCR-3, MCR-7, and MCR-9, and the set including MCR-1, MCR-2, MCR-5,

MCR-6, and ICR-Mc, contained several clades sister to some of the canonical and CARD:R&V leaves. Other choices for leaf sets contained added diversity that was distant from canonical sequences, and were not chosen as a focused figures.

Sister to the canonical MCR-3 sequences, and the most recent common ancestor of MCR-3 and MCR-7 clades (figure 2), there is a clade of sequences from NCBI which have been reported as MCR-3. Present within this clade is a single result from the Parks *et. al.* UBA dataset [27], UBA705. This leaf is labelled as *Aeromonas hydrophila*, and is associated with rice paddy field-soil microbial fuel cells. This is present within the clade alone with several other canonical and non-redundant *Aeromonas*. *A. hydrophila* is a species which has been found to have an MCR-3 gene.

Sister to the clade containing MCR 7, and the most recent common ancestor of MCR 3 and MCR 7 clades, is a clade of NCBI sourced phosphoethanolamine lipid A transferase. This clade consists of the genus of Gram-negative bacteria, *Aeromonas*, [34] which is sometimes involved in human infection. This clade also includes an *A. veronii* hit from epidermal mucus of *Anguilla anguilla* in the UBA data. Sister to the MCR 9 clade and the most recent common ancestor of MCR 3 and 9, a clade of *Aeromonas* associated phosphoethanolamine lipid A transferases appear. The MCR 9 containing clade contains 8 UBAs and several NCBI non-redundant hits not reported as MCR family genes. 5 Loose hits for MCR in a *Leclercia adecarboxylata* branch within this clade. These *Leclercia* branch as a sister to 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. *Acinetobacter* is another opportunistic pathogen [35] which is becoming resistant to many antimicrobials.



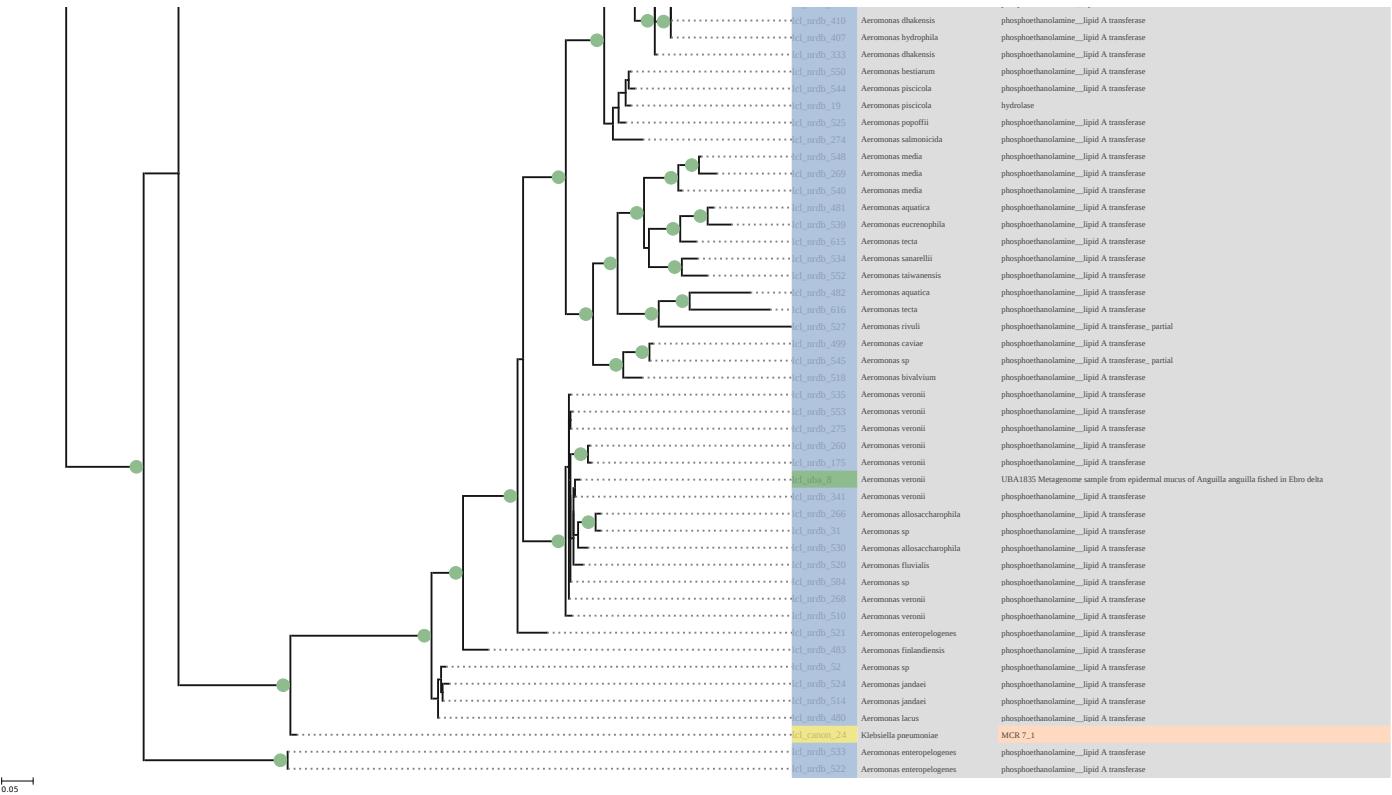
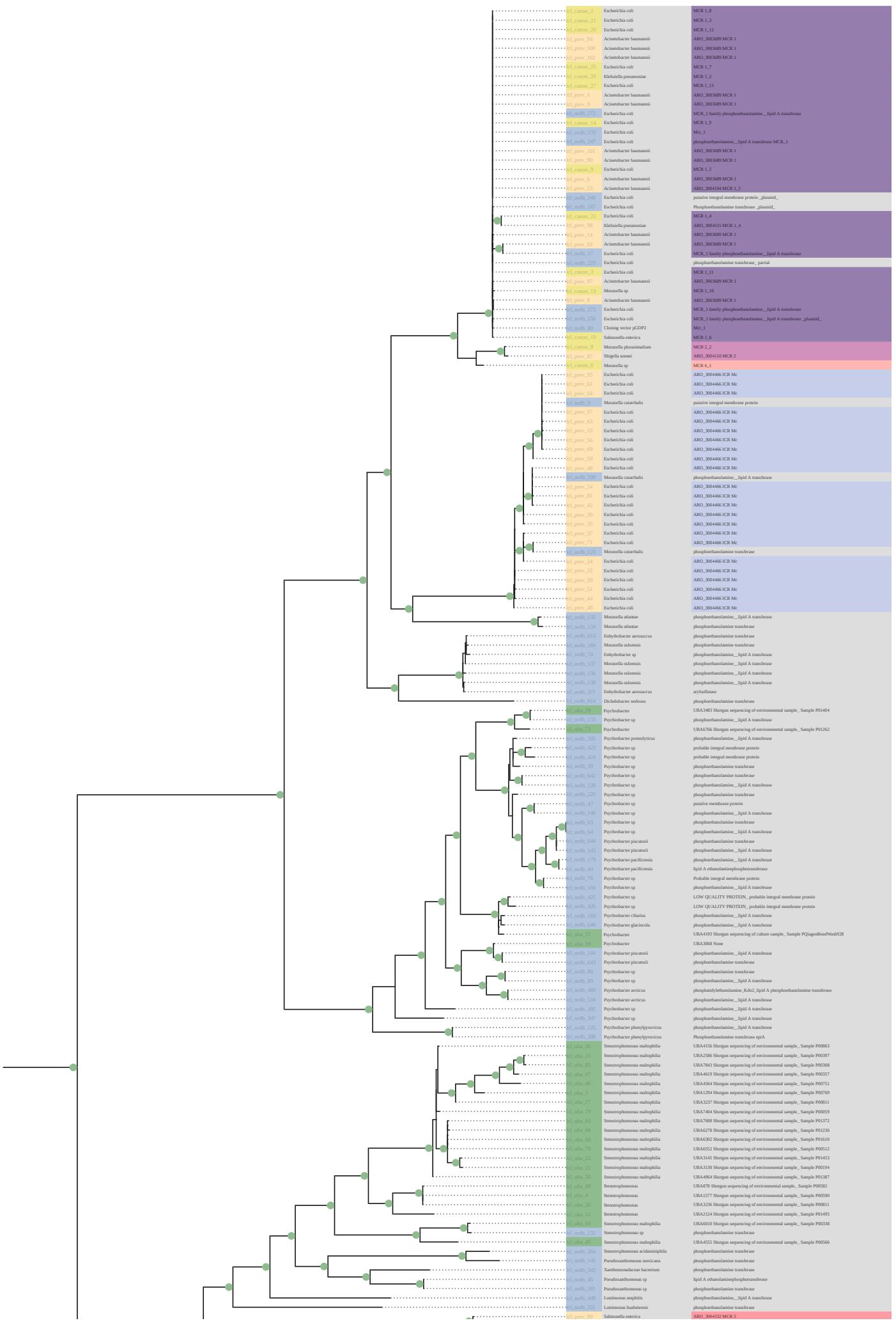


Figure 2: Clade containing putative MCR3 and MCR 9 clades pruned from phylogenetic relationship of 32 canonical (labels prefixed with 'cl_canon' in yellow), and 54 CARD:R&V (labels prefixed with 'cl_prev' in tan) MCR family sequences, 595 NCBI non-redundant sequences (labels prefixed with 'cl_nrdb' in blue), and 91 UBA sourced sequences (labels prefixed with 'cl_uba' in green). Each MCR variant is coloured based on its allele number. If the sequence is not reported to be MCR family it is coloured in grey.

Sister to the MCR 5 and ICR-Mc clades (figure {#fig:mcr-5mayb-icr}), there appear clades of diversity in the genus *Psychrobacter*. A descendant to this ancestor, is a clade of *Psychrobacter* species bacteria.

This clade includes several hits from the non-redundant database, and two hits from the UBA data. According to the UBA [27] data, the identity of these samples are, UBA3068, A *Psychrobacter* sp., sampled from Oil polluted marine microbial communities and, UBA4193, a *Psychrobacter* sp., sampled from the New York City subway. The quality information for these sequences, shown in table ?? [TODO: do we need a table with quality information], 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* [36] is a widespread genus and includes many cold adapted bacteria, it is an opportunistic pathogen, and has been found to sometimes be a cause of infections in humans, animals, and fish. Many new species of *Psychrobacter* have been discovered in cold climates [37]. Some of the species have been shown to be resistant to colistin, like *P. vallis* ps. nov. and *P. aquaticus* ps. nov. [38] and is sister to [TODO: display new tree such that *Acinetobacter* is shown] Another clade sister to these two variants contains phosphoethanolamine transferases from the genus *Stenotrophomonas*, including *S. maltophilia* and *S. acidaminiphila* which is a multidrug-resistant opportunistic pathogen [39]. Several UBA hits for *S. maltophilia* show up as shotgun sequencing of environmental samples. [TODO: quality information]



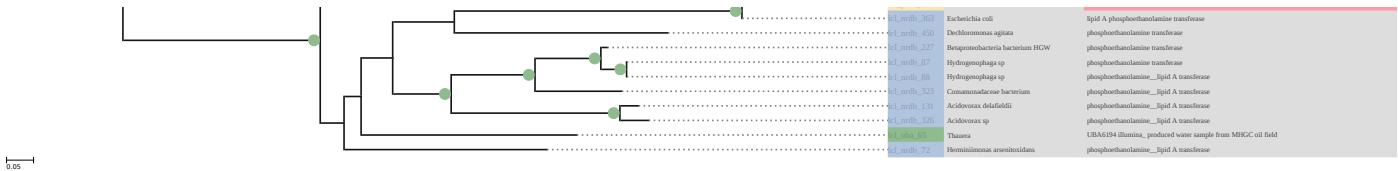


Figure 3: Clade containing putative MCR 5 and ICR-Mc clades pruned from phylogenetic relationship of 32 canonical labels (prefixed with 'cl_canon'_ in yellow), and 54 CARD:R&V labels prefixed with 'cl_prev' in tan) MCR family sequences, 595 NCBI non-redundant sequences (labels prefixed with 'cl_nrdb' in blue), and 91 UBA sourced sequences (labels prefixed with 'cl_uba' in green). Each MCR variant is coloured based on its allele number. If the sequence is not reported to be MCR family it is coloured in grey.

Phylogenetic analysis of KPC

Using the 18 canonical sequences from Table 2 as query sequences, 25 CARD:R&V sequences, 376 non-redundant sequences, and 6 UBA sequences were recovered from the various databases. The result was combined in the phylogeny displayed in figure 11. This phylogeny shows that many of the resulting sequences are distant from the canonical sequences. The main clade in figure 4 shows that many of the AMR determinants are named as distinct sequences, while not having much phylogenetic distance between them. The most closely related clade is a clade containing IMI and SME-type beta-lactamases. The remaining sequences are in a clade containing a large amount of class A beta-lactamases from the non-redundant data. All of the UBA sequences were present in this portion of the tree.

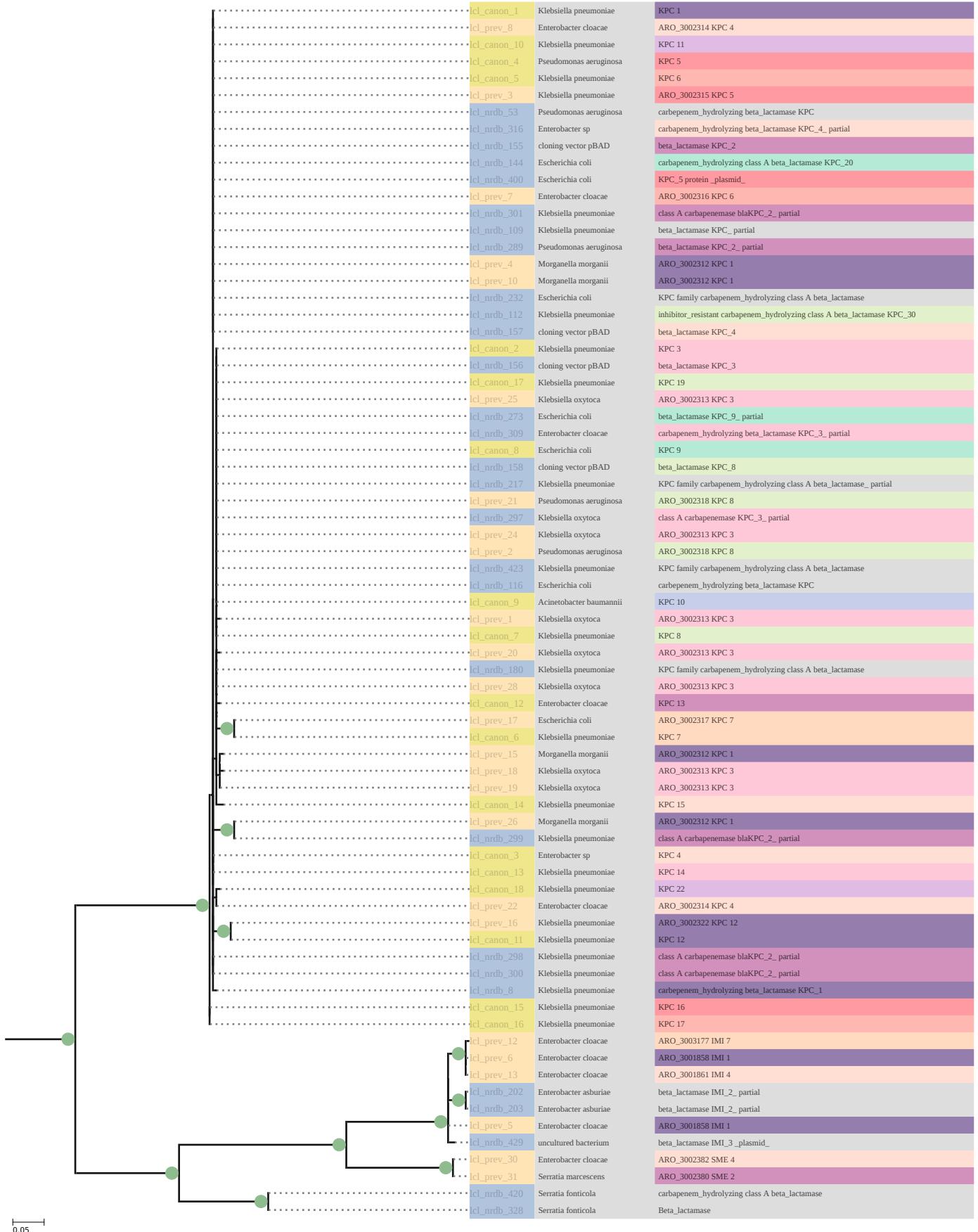


Figure 4: A clade pruned from the phylogeny in figure {#fig:kpc-tree} which represents the most closely related genes to the KPC family genes. 18 canonical labels (prefixed with 'cl_canon'_ in yellow), and 25 CARD:R&V (labels prefixed with 'cl_prev' in tan), and 28 NCBI non-redundant sequences (labels prefixed with 'cl_nrdb' in blue). Each KPC variant is coloured based on its allele number. If the sequence is not reported to be KPC family it is coloured in grey.

Phylogenetic analysis of NDM

14 canonical sequences were used to query the data for NDM beta-lactamase. The genes retained for the phylogeny were 14 canonical sequences, 8 CARD:R&V sequences, 12 non-redundant sequences, 1 UBA sequence, and an out-group. This resulted in the phylogenetic relationship in Figure 10.

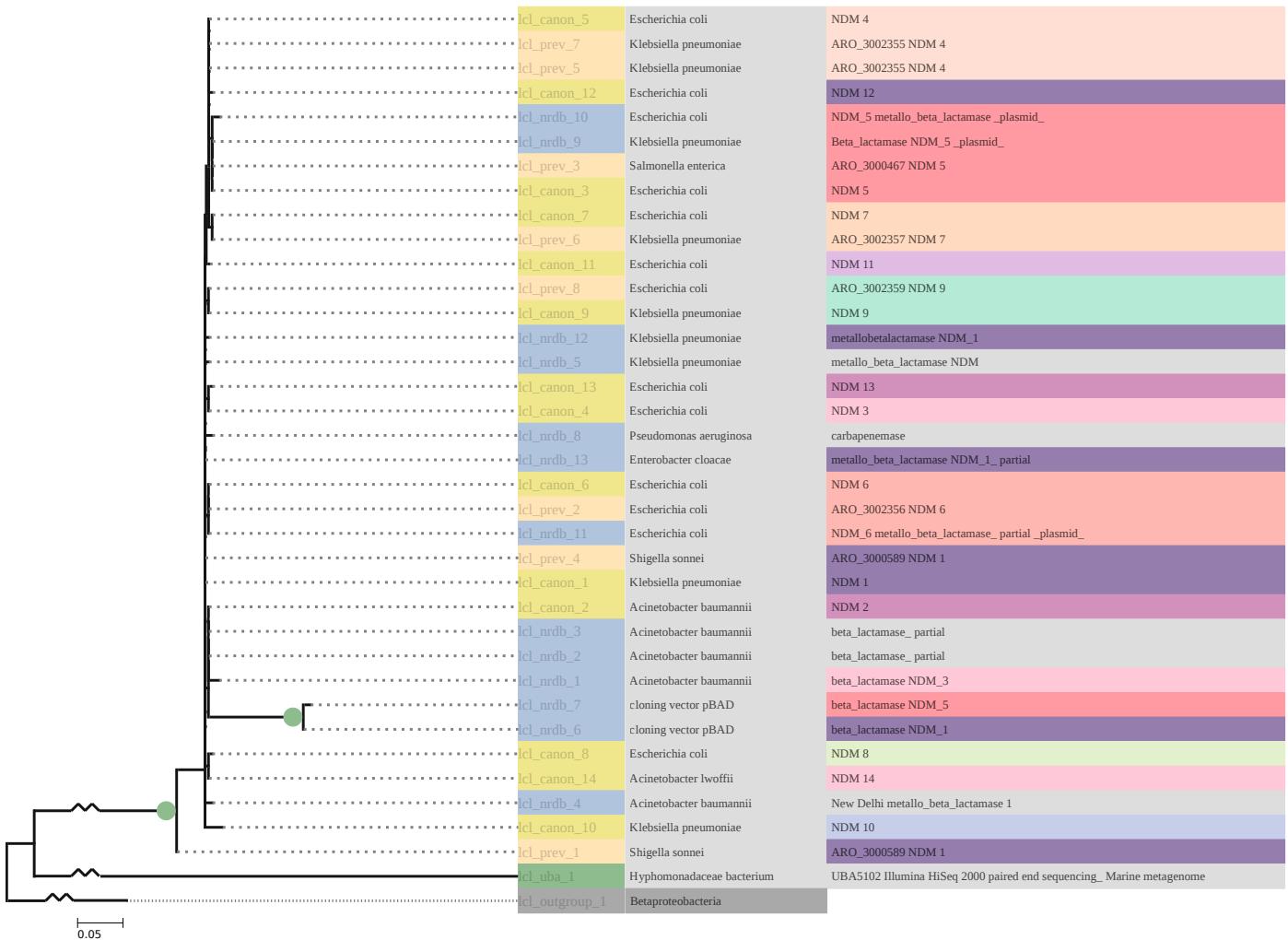


Figure 5: Phylogenetic relationship of The NDM genes with 14 canonical (labels prefixed with 'cl_canon' in yellow), 8 CARD:R&V (labels prefixed with 'cl_prev' in tan), 12 NCBI non-redundant database sequences (labeled as 'cl_nrdb' in blue), 1 UBA sequence (labeled as 'cl_uba' in green), and an outgroup from Betaproteobacteria (labeled as 'cl_outgroup' in grey). Each NDM variant is coloured based on its allele number.

The 12 sequences resulting from the NCBI data included 3 genes reported to be NDM-1, NDM-3, and NDM-5, while the other 4 genes were reported to be general beta-lactamase/carbenemase hits. As in the KPC phylogeny (figure 4), there is very little phylogenetic difference between the named variants of NDM. The only phylogenetically distinct genes in this clade were the CARD:R&V hit for NDM-1, and the NCBI hits labeled as NDM-1 and NDM-5. The naming of these sequences seem not to be congruent with the phylogenetic relationships present in the tree. Some of the NDM-1 and NDM-5 genes are more closely related to the other NDM variants than themselves. One BLAST result from the UBA data branches far from the clade containing the canonical sequences. This indicates 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% (see Table 1) is questionable. Reducing query coverage further could produce more sequences from the UBA data, but the resulting alignment would produce a meaningless phylogeny.

Phylogenetic analysis of OXA-48

In querying for the OXA-48 portion of the analysis, multiple BLAST results for UBA sequences were found, but the hits were too phylogenetically dissimilar to draw a conjecture about their relationship to the OXA family. 6 canonical sequences, and the 14 CARD:R&V sequences, and 20 non-redundant sequences are combined in the phylogeny in Figure 8

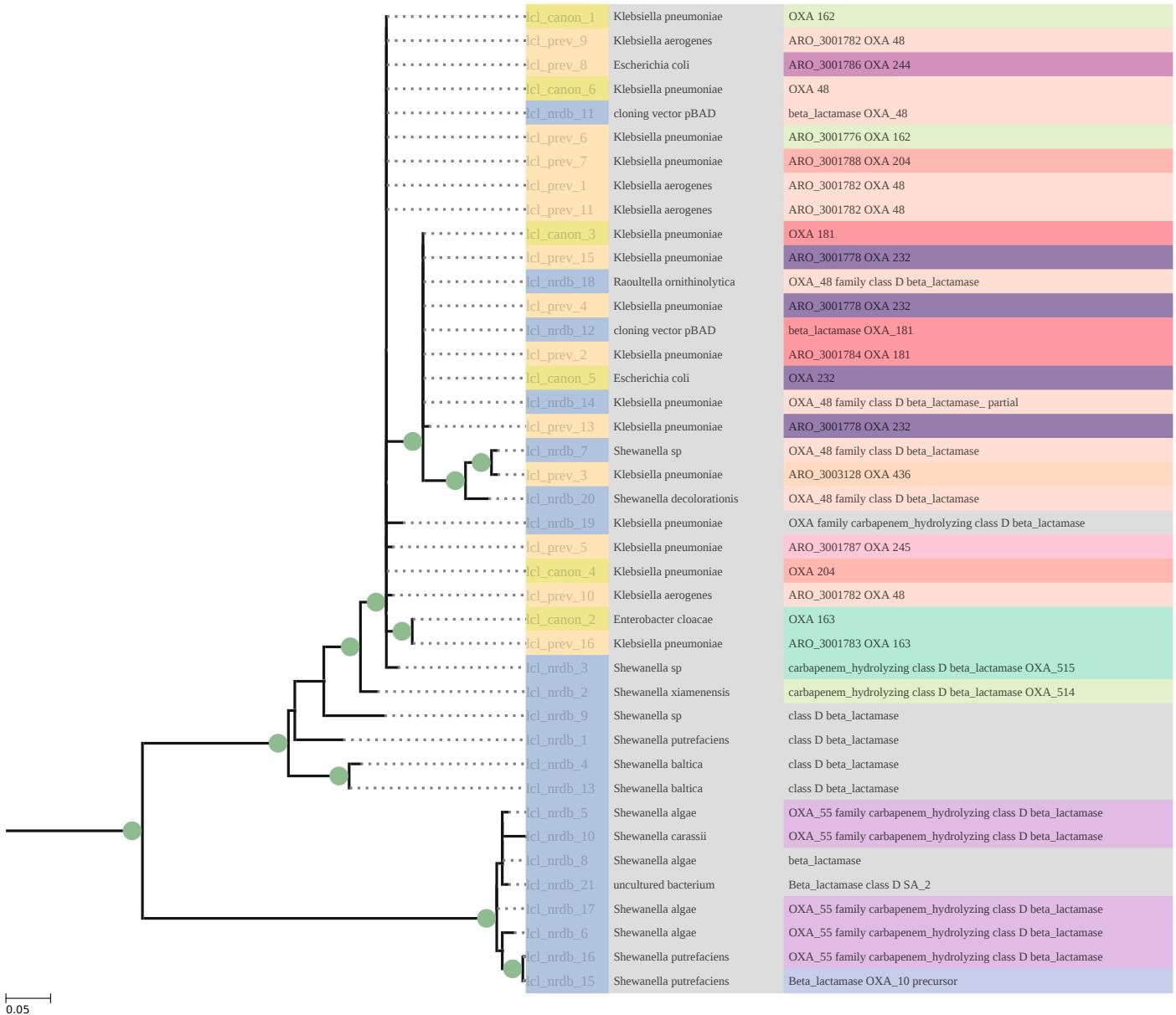


Figure 6: The main phylogenetic relationship pruned from the tree inferred from OXA-48-like sequences from figure (#fig:oxa-tree). 6 canonical (labels prefixed with 'cl_canon'_ in yellow), and 14 CARD:R&V (labels prefixed with 'cl_prev' in tan), 20 NCBI non-redundant sequences (labels prefixed with 'cl_nrdb' in blue). Each OXA-48 variant is coloured based on its allele number. If the sequence is not reported to be OXA-48 family it is coloured in grey.

The OXA-48 CARD:R&V hits added further diversity to the reference OXA-48-like sequences. OXA-436 [40] was found, clustered with no other gene, in the CARD:R&V data, and OXA-514 and OXA-515 were found in the non-redundant data. The named variants are too similar to produce significant phylogenetic resolution.

Discussion

An expanded set of candidate resistance genes

Homologous sequences from NCBI, CARD:R&V data, and a large set of MAGs, were queried, filtered, and added to phylogenies. The filtering criteria were chosen to ignore small changes between sequences (99-100% sequence identity threshold), and to produce meaningful alignments in the final phylogenies (90-99% query coverage).

NCBI and CARD:R&V added new homologues to the phylogenies, however, these conservative filtering criteria contributed obvious novel homologues from only the MCR MAG analysis, and not in the three beta lactamase MAG analyses. This difference could be explained by plasmid distributions. Plasmids could have been poorly recovered in the assembly phase of the SRA experiments, many of which produce short reads (Illumina HiSeq 2000 and 2500). TODO: Look at The number/percentages of completely sequenced genomes, completely sequenced plasmids analyzed for each pathogen, and see if there is any concrete evidence that the above is actually happening. This seems to be problematic with bias in sampling, so other ideas/sources that may be promising are welcome. Plasmids from short reads are known to be difficult to assemble[41] [41] [25].

Phylogenetic positioning of AMR genes from NCBI and MAGs

Another way in which the analyses of MCR and the beta lactamases differ is in the spread and diversity of the relationships. In addition to MAG sources, which are completely absent from the beta lactamase trees, the MCR tree contains several sequences from NCBI's non-redundant database which land close to canonical clades. Some of the added NCBI and MAG sourced sequences are not so close, but share ancestors with canonical sequences nonetheless. This suggests that there could be more diversity to be found within the MCR class. In contrast, the added diversity in phylogenies from OXA-48, KPC is distant from the canonical and CARD:R&V sequences. The matches which are close, are nearly identical to the canonical and CARD:R&V sequences. The beta-lactamase phylogenies are low in gene diversity, and are more tightly defined than the MCR phylogenies.

Phylogenetic relationships vs. nomenclature

In assessing the value of the relationships present in these more tightly related phylogenies, it is surprising that the literature has produced so many named variants. In the tree for NDM, we have named variants numbered from NDM 1 to NDM 14, yet most of the tree would be more appropriately displayed with multifurcating nodes. There are fifteen named variants, but there are only three weakly distinct clades, and little bootstrap support (TODO: look at SNPs put this in results?). The most divergent NDM sequence is a CARD:R&V sequence which is labelled as NDM-1, yet is more distant from other homologs named NDM-1 than other NDM-1 genes. The criteria for naming these genes are seemingly arbitrary. It is also possible that incentive to find new names for producing publication contributes to many of the more redundant variants. Similarly, KPC is very tightly related and some of the same arguments could be made for classification of these sequences. In the tree for OXA-48 like genes, we see more diversity, with branches of more support. The tree forms few tightly related clades, and allows some diversity from NCBI to be added to the tree. The OXA group, as a whole, is very diverse, but only because they are classified phenotypically. If the OXA were more thoroughly separated into actual homologous groups, perhaps we would see more diversity added to the group, and perhaps even from metagenomic data. These groups of OXA could behave very differently and some could have data recovery similar to that of MCR, but we are missing out on this because of the phenotypic naming scheme over a rigorous scheme which takes into account the genetic similarity. In MCR, the naming situation is less arbitrary. The closely related variants are classified under the same primary number. Several versions of MCR-1 exist, for example, and all form a clade.

Should naming reflect phylogenetic relationships?

Finding an ideal nomenclature scheme which within, or generalizing across AMR families is problematic. Nomenclature which has been shown to be phylogenetically consistent, such as that of MCR, is nonetheless subject to ambiguity. Researchers have suggested, for example for MCR [22], using amino acid sequence identity differences rather than possibly trivial nucleotide differences, but this still subject to ambiguity in cutoff of percent identity. Fixed cutoffs of 2% have been suggested by some researchers [42], but depending on the mechanism of resistance, the level of diversity required to induce differences in substrate specificity will change between AMR genes. One could argue that the genomic and evolutionary context should be also considered in the nomenclature [43]. It has been suggested, for example that mobilized, and chromosomal genes differ in their evolutionary variance [43]. In addition, the current study demonstrates that the diversity and spread varies between gene families. This could be explained by the way MCR works by modifying the target of polymyxin, unlike the beta-lactamases, which work by destroying the beta-lactams. Justifying naming of a new NDM gene, when all genes available to the public are essentially identical, seems to be problematic, however, further investigations into certain patterns of changes in the amino acid sequences may further granularize our classification. Ultimately, many interesting questions about the nomenclature of these genes are rooted in ontology. Classifying these genes is of utmost importance for communication between scientists, and a naming scheme which seeks a balance between practicality, evolutionary relationships, function, and context, must be found.

Supplemental

Table 1: Names of the canonical MCR phosphoethanolamine transferase variants from CARD database

MCR-9	MCR-1.8	MCR-1.11	MCR-3.4	MCR-3.11	MCR-6.1	MCR-3.3	MCR-2.2	MCR-1.5	MCR-3.6	MCR-3.10	MCR-3.5	MCR-1.10	MCR-1.9	MCR-3.9	MCR-3.8	MCR-3.7	MCR-3.2	MCR-1.6
MCR-1.2	MCR-1.3	MCR-1.4	MCR-8	MCR-7.1	MCR-1.7	MCR-1.12	MCR-1.13	MCR-3.12	MCR-4.2	MCR-4.3	MCR-4.4	MCR-4.5	MCR-5.2					

Table 2: Names of the canonical beta-lactamase variants from CARD database

KPC-1	KPC-3	KPC-4	KPC-5	KPC-6	KPC-7	KPC-8	KPC-9	KPC-10	KPC-11	KPC-12	KPC-13	KPC-14	KPC-15	KPC-16	KPC-17	KPC-19	KPC-22	KPC-24
-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Table 3: Names of the canonical NDM beta-lactamase variants from CARD database

NDM-1	NDM-2	NDM-5	NDM-3	NDM-4	NDM-6	NDM-7	NDM-8	NDM-9	NDM-10	NDM-11	NDM-12	NDM-13	NDM-14	NDM-17
-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------

Table 4: Names of the canonical OXA beta-lactamase variants from CARD database

OXA-162	OXA-163	OXA-181	OXA-204	OXA-232	OXA-48
---------	---------	---------	---------	---------	--------

TODO: Quality/CHECKM information for UBAs

Table 5: Outgroups chosen in building the phylogenies of the 4 AMR families.

Outgroup to AMR Family	Gene	Species	Class
MCR	phosphoethanolamine-lipid A transferase	Oxalobacter formigenes	Betaproteobacteria
KPC	class A beta-lactamase	Mesorhizobium loti	Alpha Proteobacteria
NDM	VIM-4 metallo-beta-lactamase	Alcaligenes faecalis	Betaproteobacteria
OXA-48	OXA-10 family class D beta-lactamase OXA-454	Delftia acidovorans	Betaproteobacteria

MCR Alignment

[TODO: visualization with alignment/entropy]

KPC Alignment

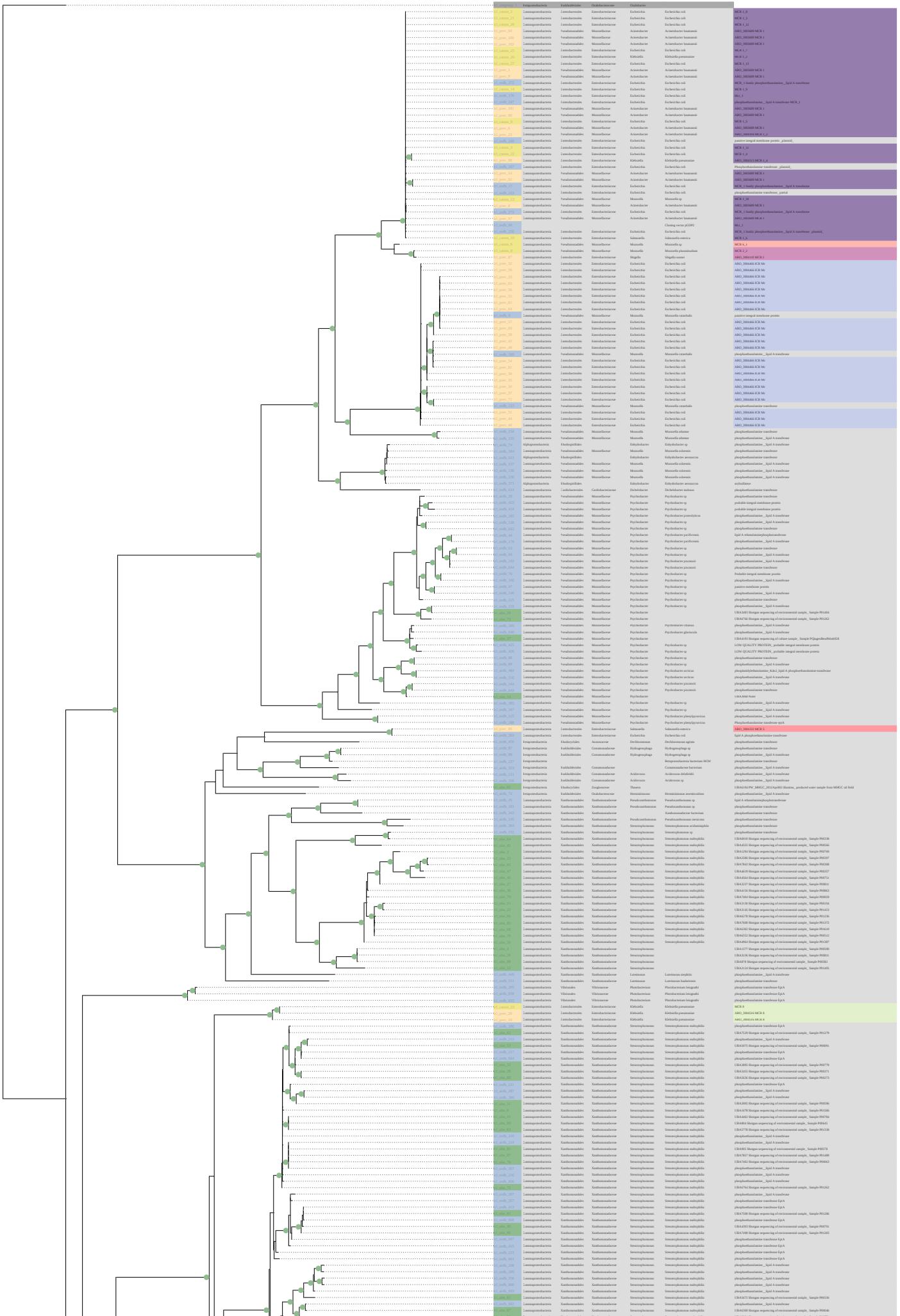
[TODO: visualization with alignment/entropy]

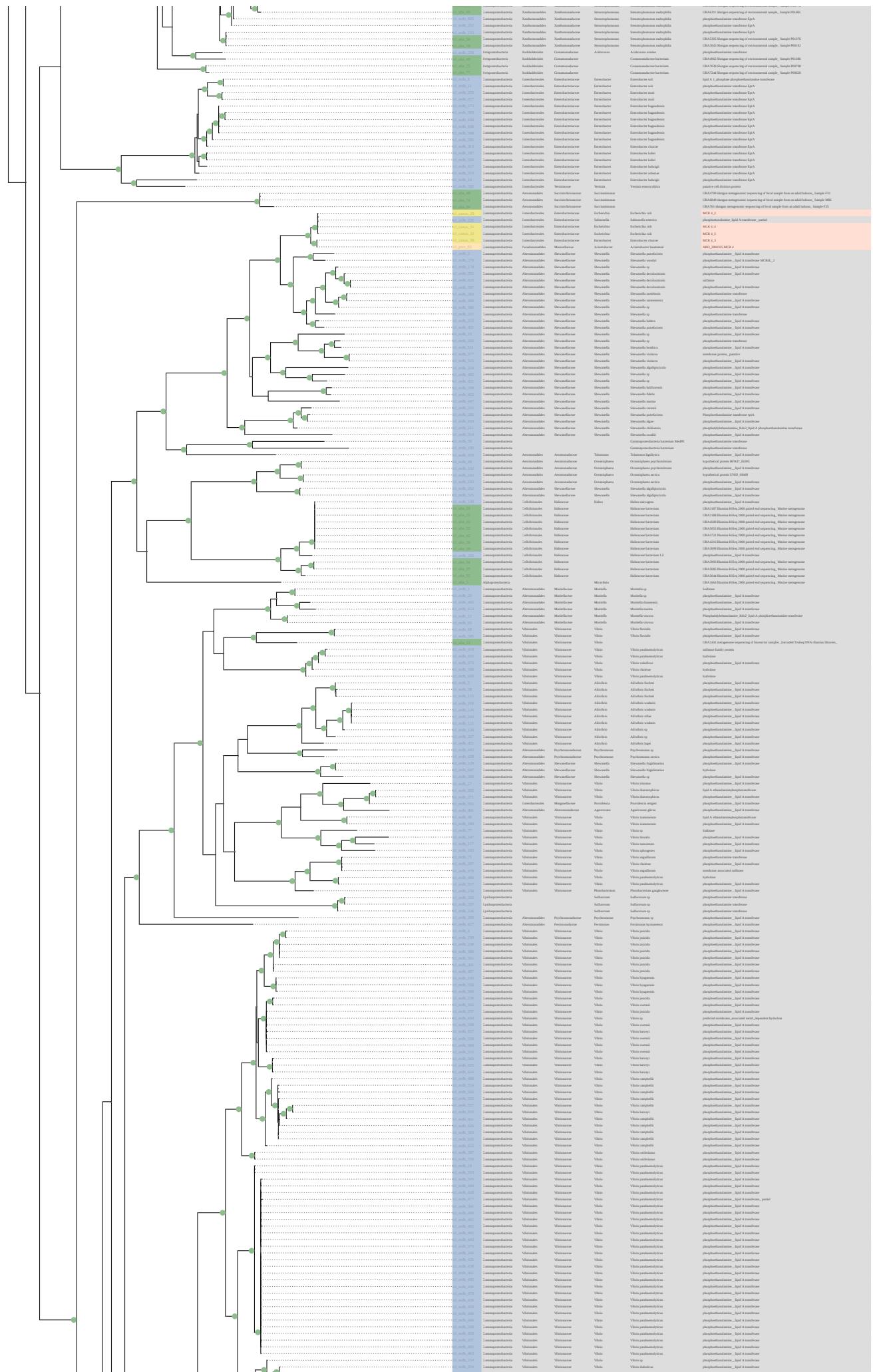
NDM Alignment

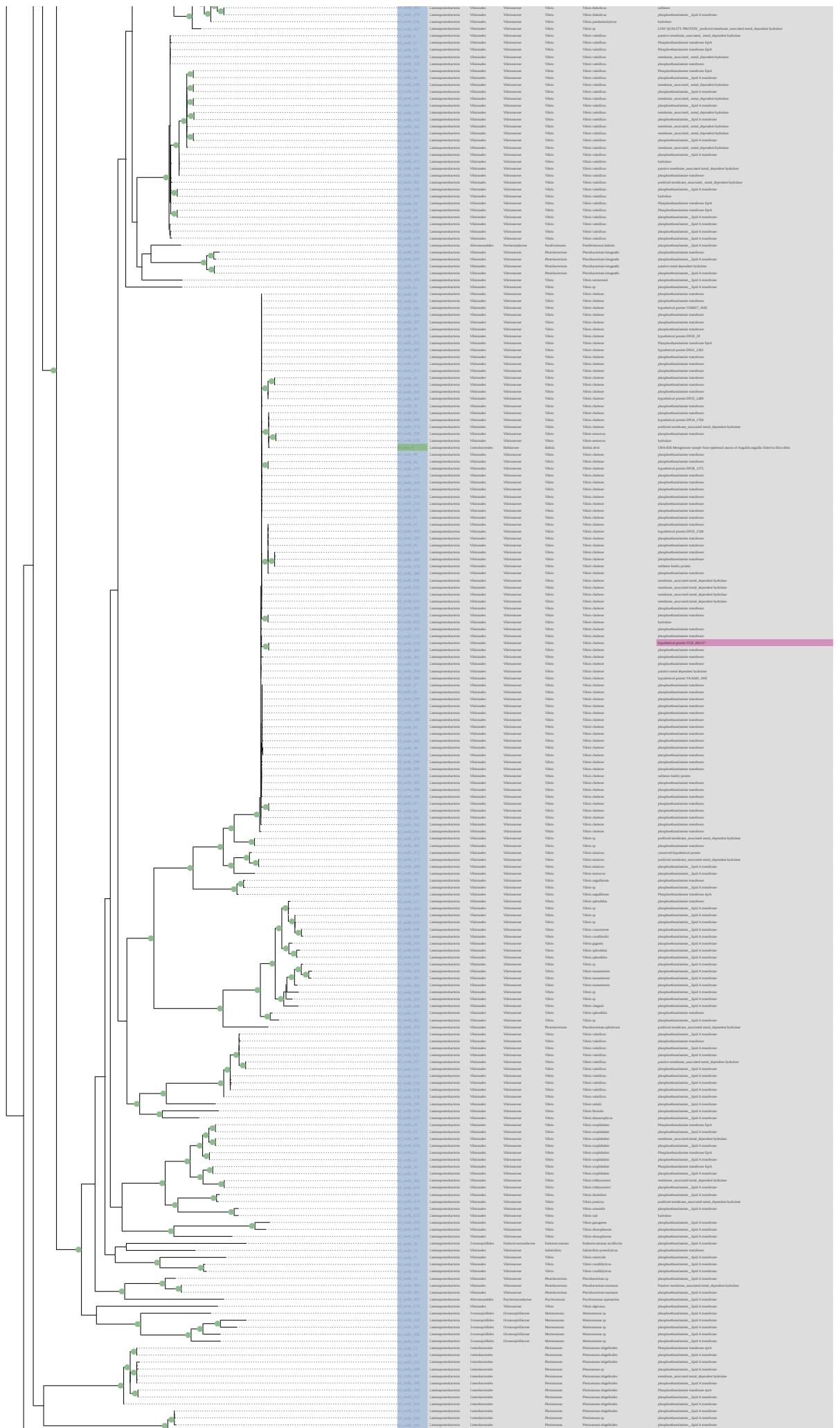
[TODO: visualization with alignment/entropy]

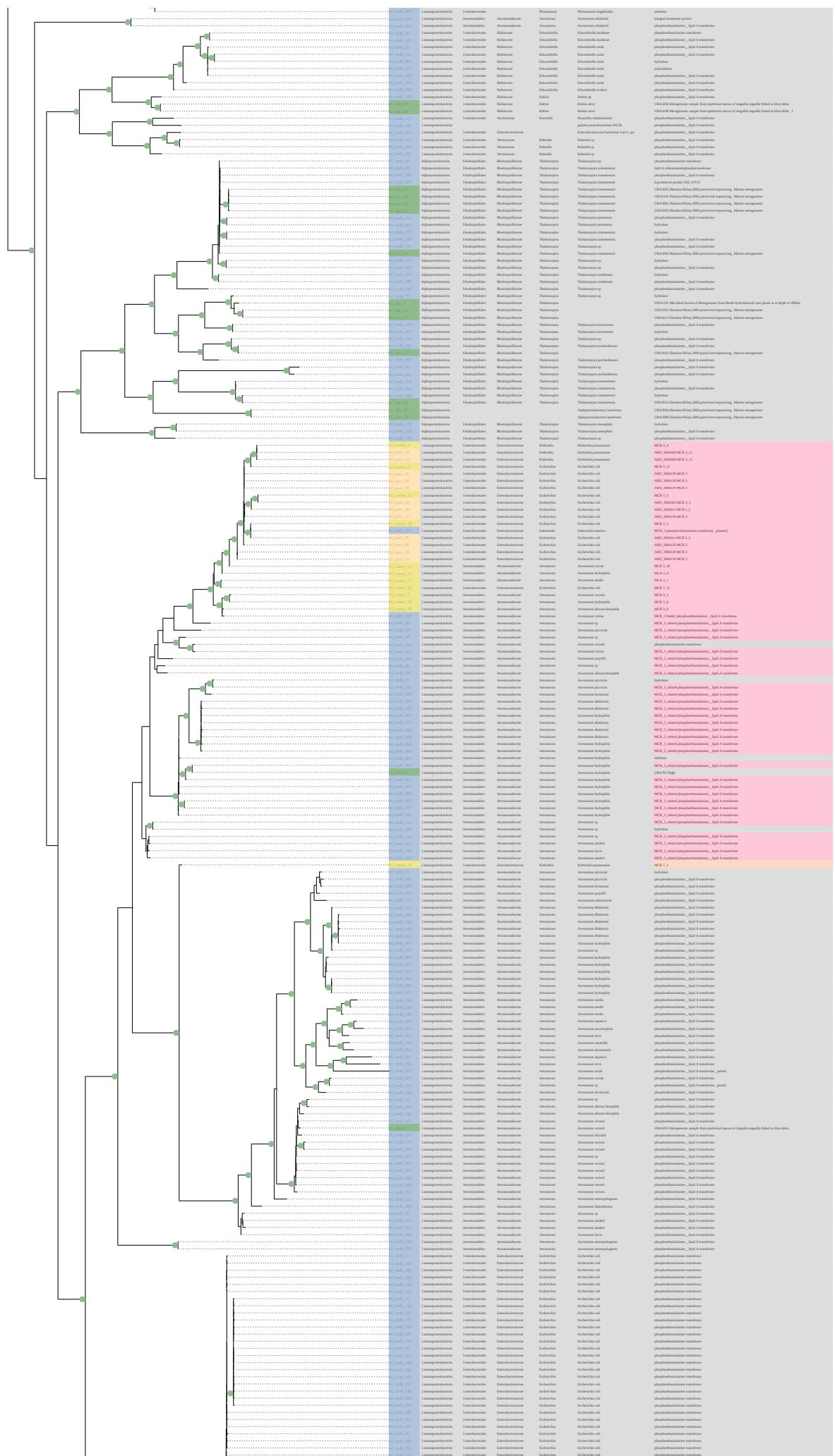
OXA-48 Alignment

[TODO: visualization with alignment/entropy]









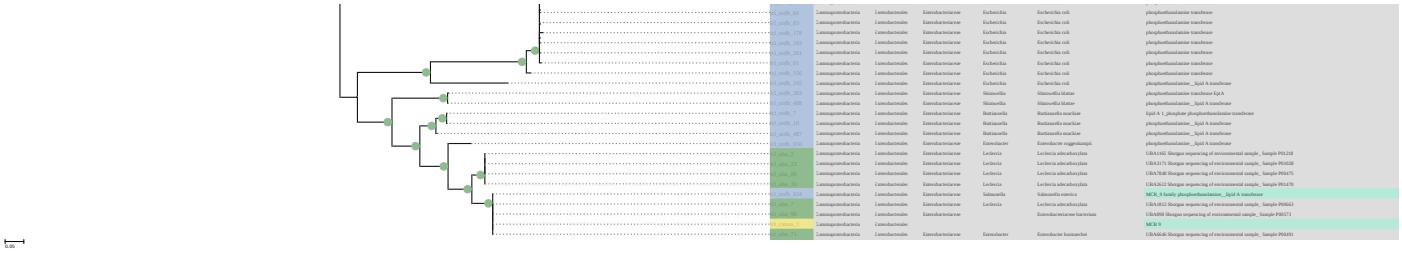


Figure 7: Phylogenetic relationship of 32 canonical (labels prefixed with lcl_canon_ in yellow), and 54 prevalence (labels prefixed with lcl_prev_ in tan) MCR family sequences, 595 NCBI non-redundant sequences (labels prefixed with lcl_prev_ in blue), and 91 UBA sourced sequences (labels prefixed with lcl_prev_ in green). Each MCR variant is coloured based on its primary numerical value. If the sequence is not reported to be MCR family it is coloured in grey.

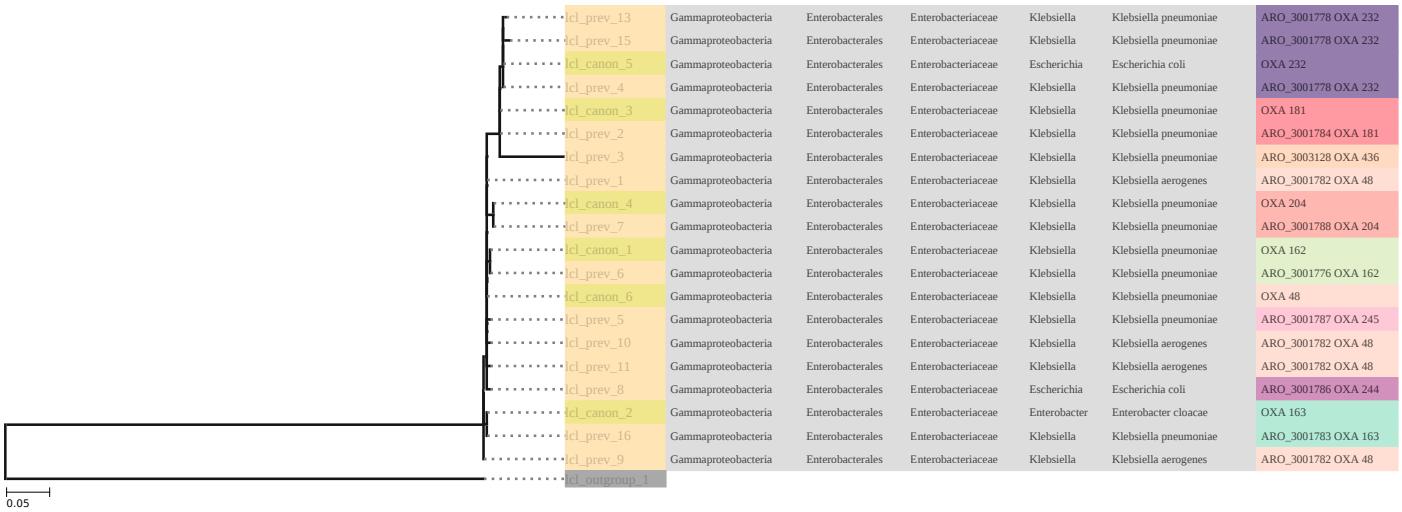
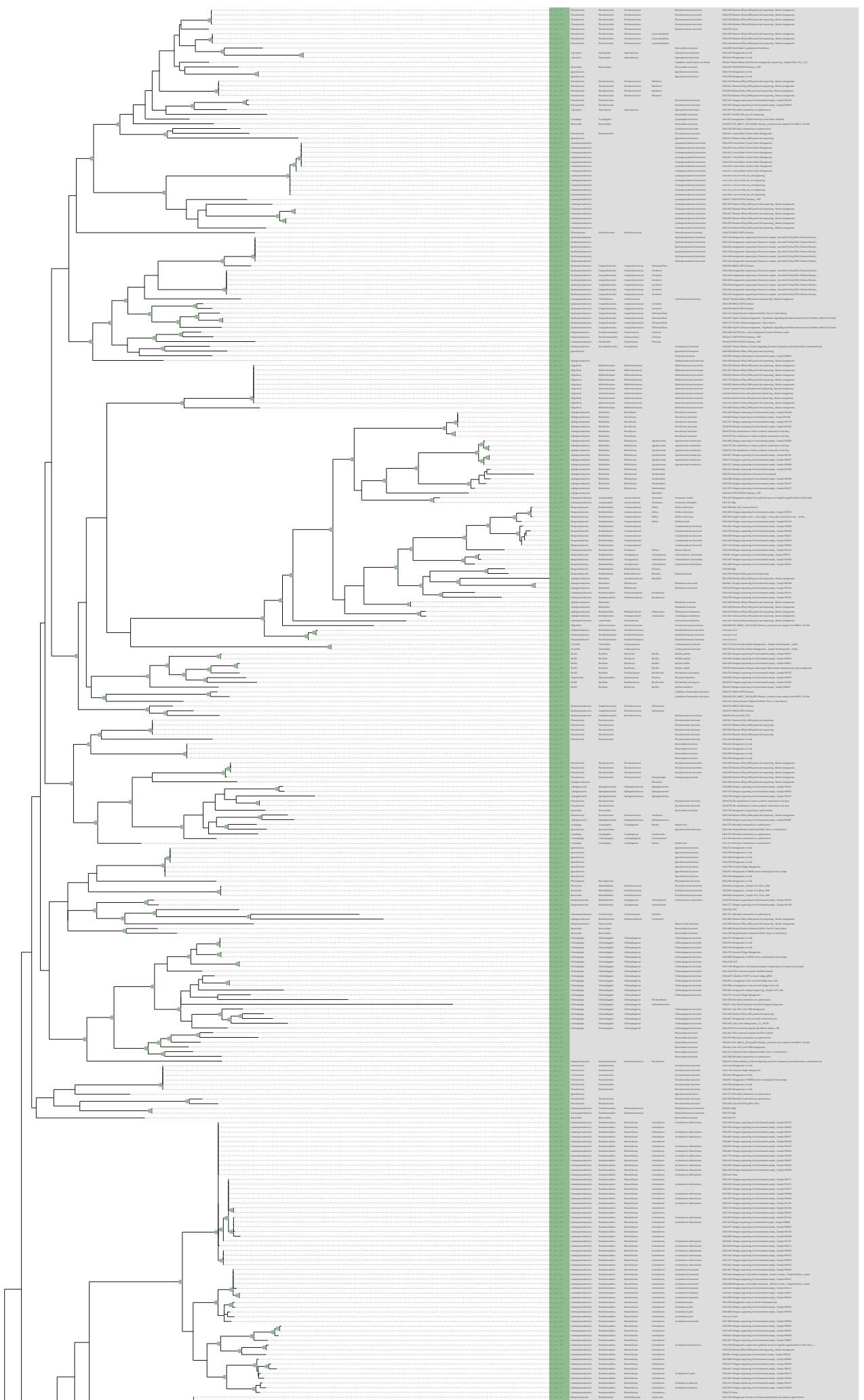


Figure 8: Phylogenetic relationship (lcl_canon) OXA-48 family sequences along with 14 prevalence (lcl_prev) sequences.



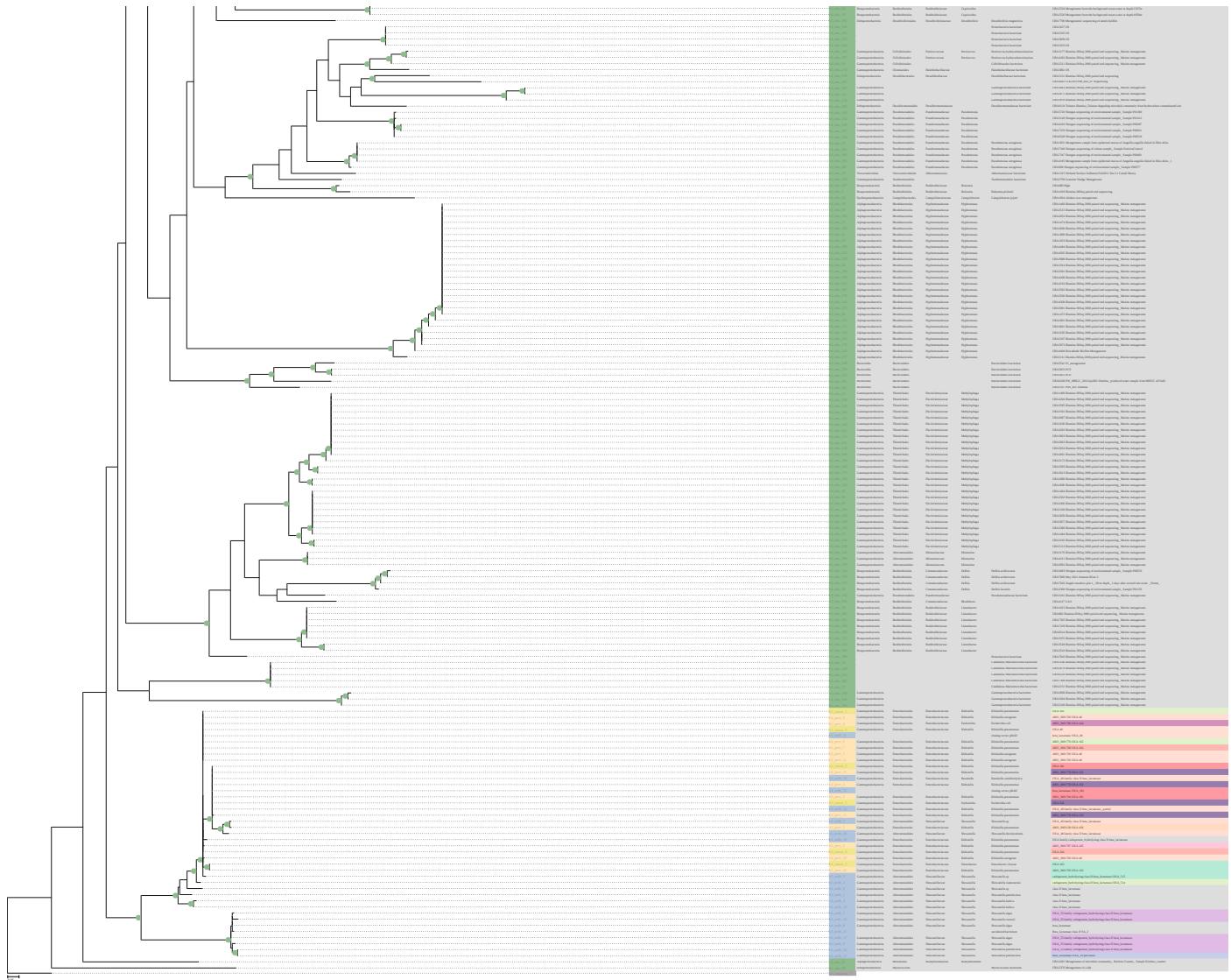


Figure 9: Phylogenetic relationship (lcl_canon) OXA-48 family sequences along with 14 prevalence (lcl_prev) sequences.

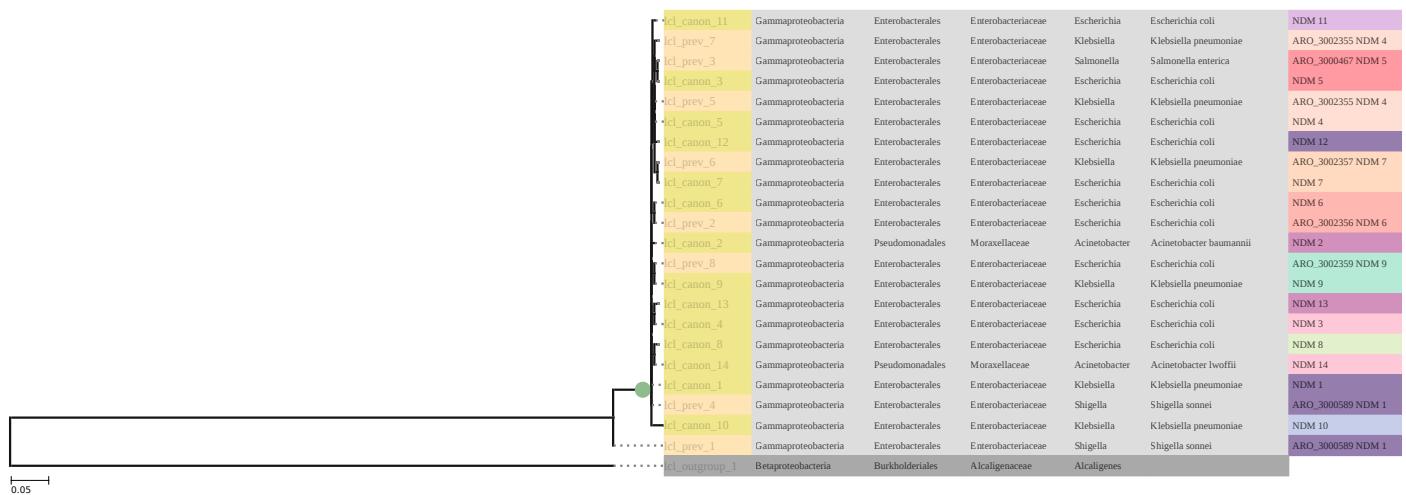


Figure 10: Phylogenetic relationship (lcl_canon) NDM family sequences along with 14 prevalence (lcl_prev) sequences.

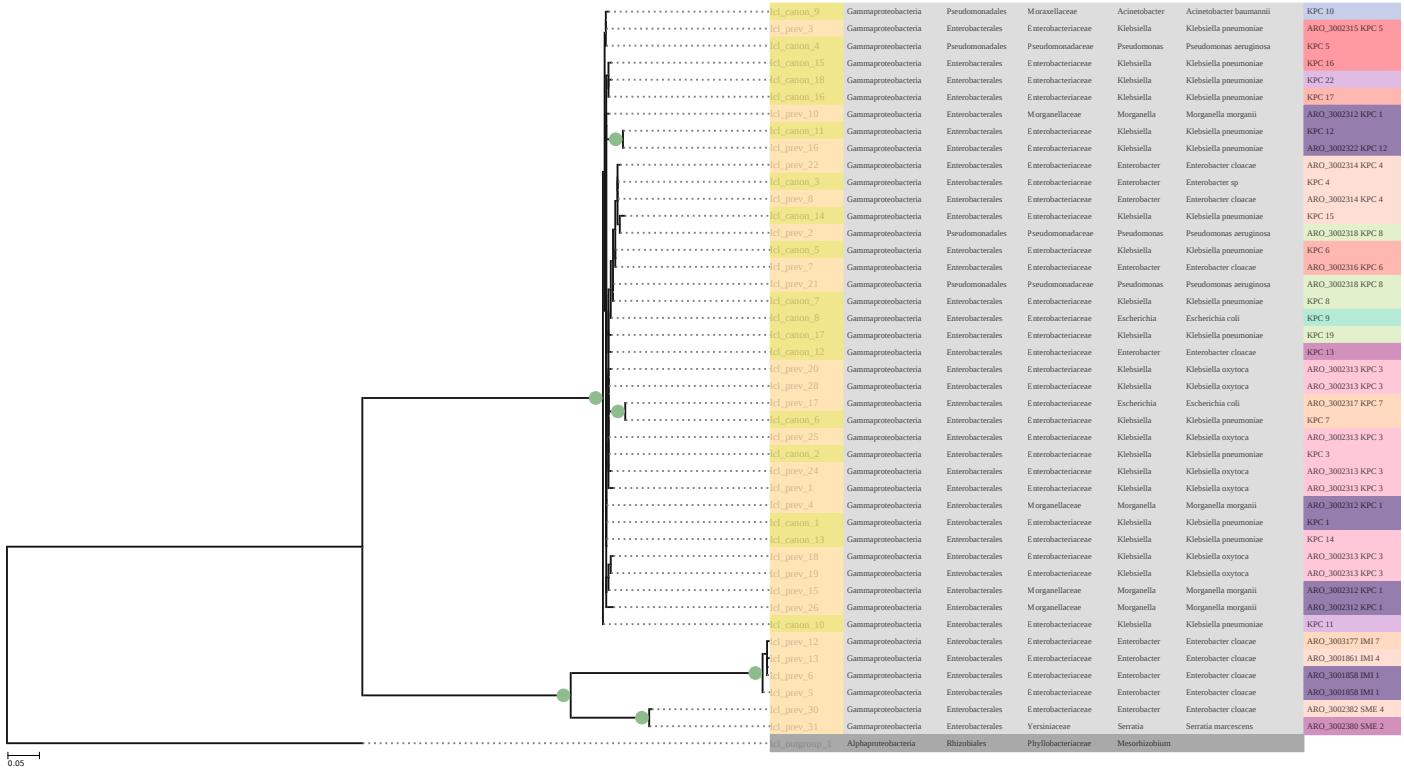
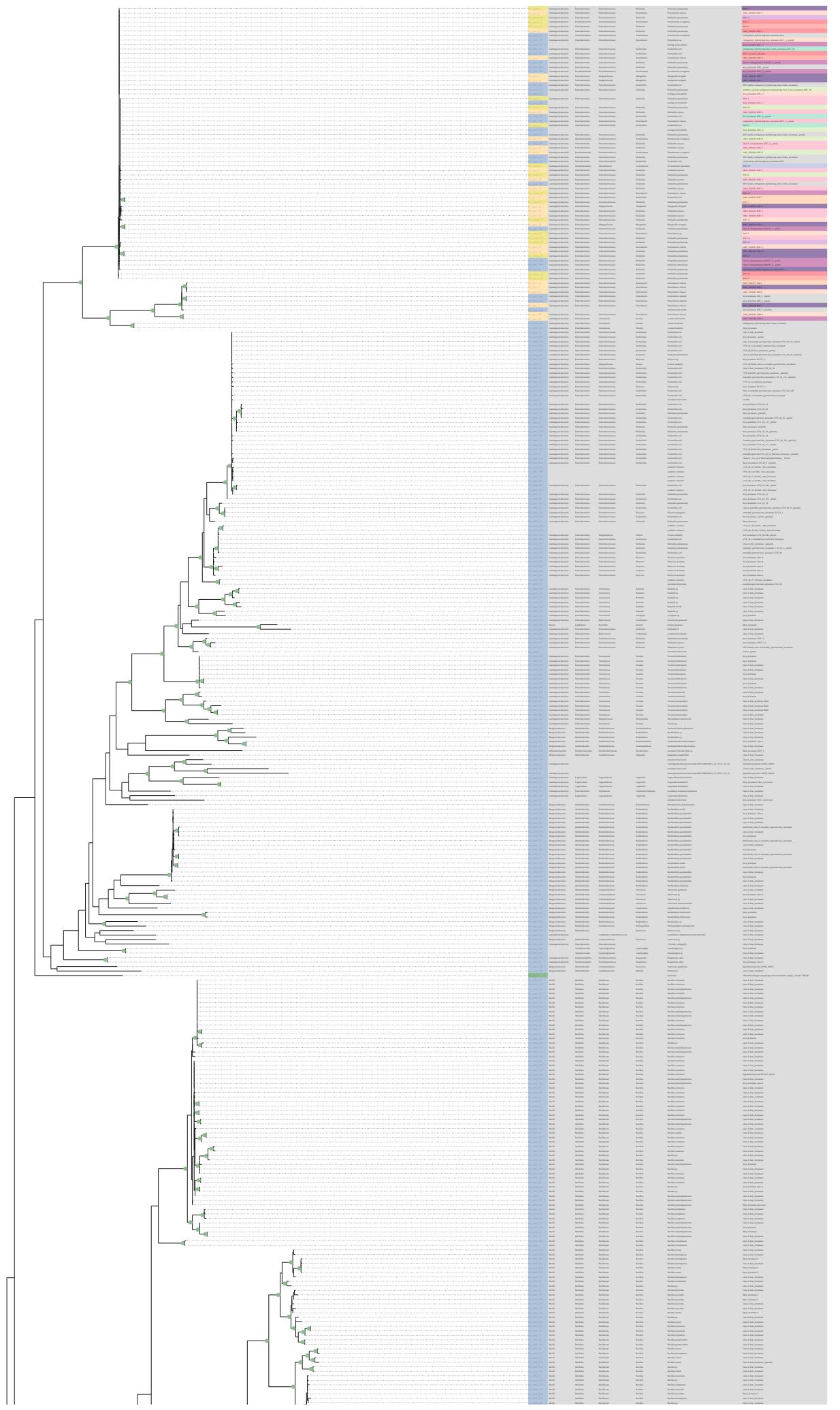


Figure 11: Phylogenetic relationship (lcl_canon) KPC family sequences along with 14 prevalence (lcl_prev) sequences.



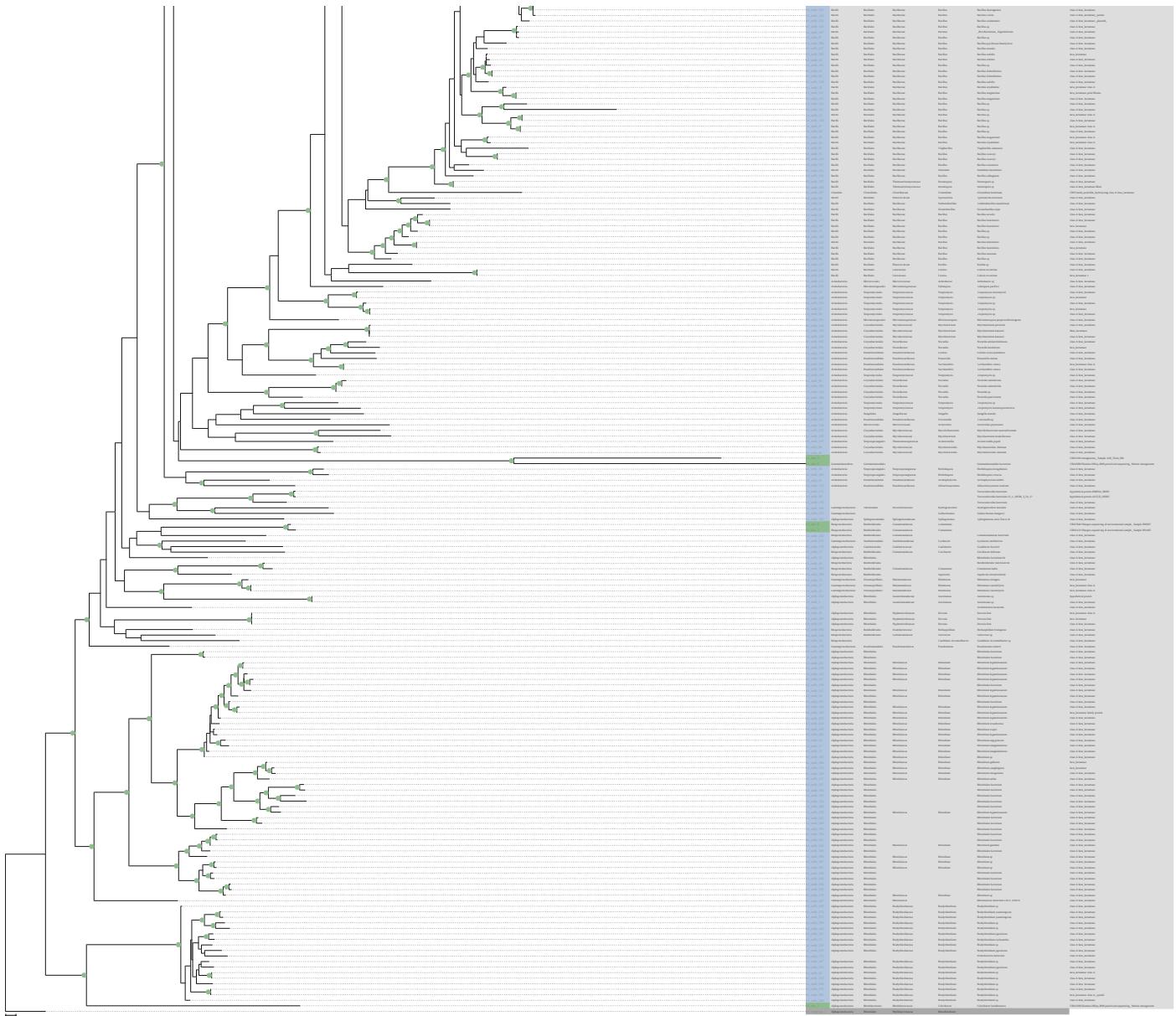


Figure 12: Phylogenetic relationship (lcl_canon) KPC family sequences along with 14 prevalence (lcl_prev) sequences.

References

1. Antibiotic Resistance★

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

Reference Module in Biomedical Sciences (2015) <https://doi.org/dbhg>

DOI: [10.1016/b978-0-12-801238-3.02385-0](https://doi.org/b978-0-12-801238-3.02385-0)

2. 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/b3wbvz>

DOI: [10.1038/nature10388](https://doi.org/10.1038/nature10388) · PMID: [21881561](#)

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](https://doi.org/10.1093/nar/gkw1004) · PMID: [27789705](#) · PMCID: [PMC5210516](#)

4. Horizontal transfer of antibiotic resistance genes in clinical environments

Nicole A. Lermniaux, Andrew D. S. Cameron

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

DOI: [10.1139/cjm-2018-0275](https://doi.org/10.1139/cjm-2018-0275) · PMID: [30248271](#)

5. Carbapenem Resistance: A Review

Francis Codjoe, Eric Donkor

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

DOI: [10.3390/medsci6010001](https://doi.org/10.3390/medsci6010001) · PMID: [29267233](#) · PMCID: [PMC5872158](#)

6. Federal Action Plan on Antimicrobial Resistance

Public Health Agency of Canada

(2015) <https://www.canada.ca/en/health-canada/services/publications/drugs-health-products/federal-action-plan-antimicrobial-resistance-canada.html>

7. WHO Global Action Plan on Antimicrobial Resistance

World Health Organisation

(2015) <https://www.who.int/antimicrobial-resistance/global-action-plan/en/>

8. WHO Global Priority List of Antibiotic Resistant Bacteria

World Health Organisation

(2017) <https://www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en/>

9. Antimicrobial Resistance in ESKAPE Pathogens

David M. P. De Oliveira, Brian M. Forde, Timothy J. Kidd, Patrick N. A. Harris, Mark A. Schembri, Scott A. Beatson, David L. Paterson, Mark J. Walker
Clinical Microbiology Reviews (2020-05-13) <https://doi.org/dzzh>

DOI: [10.1128/cmr.00181-19](https://doi.org/10.1128/cmr.00181-19) · PMID: [32404435](#) · PMCID: [PMC7227449](#)

10. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens

Sirijan Santajit, Nitaya Indrawattana

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

DOI: [10.1155/2016/2475067](https://doi.org/10.1155/2016/2475067) · PMID: [27274985](#) · PMCID: [PMC4871955](#)

11. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE

Louis B. Rice

The Journal of Infectious Diseases (2008-04-15) <https://doi.org/frhgmw>

DOI: [10.1086/533452](https://doi.org/10.1086/533452) · PMID: [18419525](#)

12. Extended-Spectrum β-Lactamases: a Clinical Update

David L. Paterson, Robert A. Bonomo

Clinical Microbiology Reviews (2005-10) <https://doi.org/d595m9>

DOI: [10.1128/cmr.18.4.657-686.2005](https://doi.org/10.1128/cmr.18.4.657-686.2005) · PMID: [16223952](#) · PMCID: [PMC1265908](#)

13. Canadian Nosocomial Infection Surveillance Program

Public Health Agency of Canada

(2019) <https://www.canada.ca/en/public-health/services/publications/science-research-data/summary-report-healthcare-associated-infection-antimicrobial-resistance-antimicrobial-use-surveillance-data-2013-2017.html>

14. The structure of β-lactamases

Philosophical Transactions of the Royal Society of London. B, Biological Sciences

(1980-05-16) <https://doi.org/cdckks>

DOI: [10.1098/rstb.1980.0049](https://doi.org/10.1098/rstb.1980.0049) · PMID: [6109327](#)

15. Updated Functional Classification of β-Lactamases

Karen Bush, George A. Jacoby

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

DOI: [10.1128/aac.01009-09](https://doi.org/10.1128/aac.01009-09) · PMID: [19995920](#) · PMCID: [PMC2825993](#)

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

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

17. 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: [10.1128/jcm.01175-16](https://doi.org/10.1128/jcm.01175-16) · PMID: [27535687](https://pubmed.ncbi.nlm.nih.gov/27535687/) · PMCID: [PMC5078564](https://pubmed.ncbi.nlm.nih.gov/PMC5078564/)

18. Treatment Options for Carbapenem-Resistant Enterobacteriaceae Infections

Haley J. Morrill, Jason M. Pogue, Keith S. Kaye, Kerry L. LaPlante
Open Forum Infectious Diseases (2015) <https://doi.org/dbht>
DOI: [10.1093/ofid/ofv050](https://doi.org/10.1093/ofid/ofv050) · PMID: [26125030](https://pubmed.ncbi.nlm.nih.gov/26125030/) · PMCID: [PMC4462593](https://pubmed.ncbi.nlm.nih.gov/PMC4462593/)

19. 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: [10.1016/j.tim.2018.02.006](https://doi.org/10.1016/j.tim.2018.02.006) · PMID: [29525421](https://pubmed.ncbi.nlm.nih.gov/29525421/)

20. mcr-1 identified in Avian Pathogenic Escherichia coli (APEC).

Nicolle Lima Barbieri, Daniel W Nielsen, Yvonne Wannemuehler, Tia Cavender, Ashraf Hussein, Shi-Gan Yan, Lisa K Nolan, Catherine M Logue
PLoS one (2017-03-06) <https://www.ncbi.nlm.nih.gov/pubmed/28264015>
DOI: [10.1371/journal.pone.0172997](https://doi.org/10.1371/journal.pone.0172997) · PMID: [28264015](https://pubmed.ncbi.nlm.nih.gov/28264015/) · PMCID: [PMC5338808](https://pubmed.ncbi.nlm.nih.gov/PMC5338808/)

21. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database

Brian P Alcock, Amogelang R Raphanya, Tammy TY Lau, Kara K Tsang, Mégane Bouchard, Arman Edalatmand, William Huynh, Anna-Lisa V Nguyen, Annie A Cheng, Sihan Liu, ... Andrew G McArthur
Nucleic Acids Research (2019-10-29) <https://doi.org/ggckg6>
DOI: [10.1093/nar/gkz935](https://doi.org/10.1093/nar/gkz935) · PMID: [31665441](https://pubmed.ncbi.nlm.nih.gov/31665441/) · PMCID: [PMC7145624](https://pubmed.ncbi.nlm.nih.gov/PMC7145624/)

22. Proposal for assignment of allele numbers for mobile colistin resistance (mcr) genes

Sally R Partridge, Vincenzo Di Pilato, Yohei Doi, Michael Feldgarden, Daniel H Haft, William Klimke, Samir Kumar-Singh, Jian-Hua Liu, Surbhi Malhotra-Kumar, Arjun Prasad, ... Basil Britto Xavier
Journal of Antimicrobial Chemotherapy (2018-07-19) <https://doi.org/dw7v>
DOI: [10.1093/jac/dky262](https://doi.org/10.1093/jac/dky262) · PMID: [30053115](https://pubmed.ncbi.nlm.nih.gov/30053115/) · PMCID: [PMC6148208](https://pubmed.ncbi.nlm.nih.gov/PMC6148208/)

23. Molecular Characterization of OXA-20, a Novel Class D β -Lactamase, and Its Integron from *Pseudomonas aeruginosa*

Thierry Naas, Wladimir Sougakoff, Anne Casetta, Patrice Nordmann
Antimicrobial Agents and Chemotherapy (1998-08-01) <https://doi.org/dwsr>
DOI: [10.1128/aac.42.8.2074](https://doi.org/10.1128/aac.42.8.2074) · PMID: [9687410](https://pubmed.ncbi.nlm.nih.gov/9687410/)

24. Author Correction: 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 (2017-12-12) <https://doi.org/c8rq>
DOI: [10.1038/s41564-017-0083-5](https://doi.org/10.1038/s41564-017-0083-5) · PMID: [29234139](https://pubmed.ncbi.nlm.nih.gov/29234139/)

25. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities

Dongwan D. Kang, Jeff Froula, Rob Egan, Zhong Wang
PeerJ (2015-08-27) <https://doi.org/gdf329>
DOI: [10.7717/peerj.1165](https://doi.org/10.7717/peerj.1165) · PMID: [26336640](https://pubmed.ncbi.nlm.nih.gov/26336640/) · PMCID: [PMC4556158](https://pubmed.ncbi.nlm.nih.gov/PMC4556158/)

26. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy

Christian M. K. Sieber, Alexander J. Probst, Allison Sharrar, Brian C. Thomas, Matthias Hess, Susannah G. Tringe, Jillian F. Banfield
Nature Microbiology (2018-05-28) <https://doi.org/gfwvfg>
DOI: [10.1038/s41564-018-0171-1](https://doi.org/10.1038/s41564-018-0171-1) · PMID: [29807988](https://pubmed.ncbi.nlm.nih.gov/29807988/) · PMCID: [PMC6786971](https://pubmed.ncbi.nlm.nih.gov/PMC6786971/)

27. 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](https://doi.org/10.1038/s41564-017-0012-7) · PMID: [28894102](https://pubmed.ncbi.nlm.nih.gov/28894102/)

28. Basic local alignment search tool

Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, David J. Lipman
Journal of Molecular Biology (1990-10) <https://doi.org/cnsjsz>
DOI: [10.1016/s0022-2836\(05\)80360-2](https://doi.org/10.1016/s0022-2836(05)80360-2)

29. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform

K. Katoh
Nucleic Acids Research (2002-07-15) <https://doi.org/dbpmq3>
DOI: [10.1093/nar/gkf436](https://doi.org/10.1093/nar/gkf436) · PMID: [12136088](https://pubmed.ncbi.nlm.nih.gov/12136088/) · PMCID: [PMC135756](https://pubmed.ncbi.nlm.nih.gov/PMC135756/)

30. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses

S. Capella-Gutierrez, J. M. Silla-Martinez, T. Gabaldon
Bioinformatics (2009-06-08) <https://doi.org/bjhdh7>
DOI: [10.1093/bioinformatics/btp348](https://doi.org/10.1093/bioinformatics/btp348) · PMID: [19505945](https://pubmed.ncbi.nlm.nih.gov/19505945/) · PMCID: [PMC2712344](https://pubmed.ncbi.nlm.nih.gov/PMC2712344/)

31. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies

Lam-Tung Nguyen, Heiko A. Schmidt, Arndt von Haeseler, Bui Quang Minh

Molecular Biology and Evolution (2014-11-03) <https://doi.org/f3srt>
DOI: [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300) · PMID: [25371430](https://pubmed.ncbi.nlm.nih.gov/25371430/) · PMCID: [PMC4271533](https://pubmed.ncbi.nlm.nih.gov/PMC4271533/)

32. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data

Jaime Huerta-Cepas, François Serra, Peer Bork

Molecular Biology and Evolution (2016-02-26) <https://doi.org/gfzpph>

DOI: [10.1093/molbev/msw046](https://doi.org/10.1093/molbev/msw046) · PMID: [26921390](https://pubmed.ncbi.nlm.nih.gov/26921390/) · PMCID: [PMC4868116](https://pubmed.ncbi.nlm.nih.gov/PMC4868116/)

33. Substrate Recognition by a Colistin Resistance Enzyme from *Moraxella catarrhalis*

Peter J. Stogios, Georgina Cox, Haley L. Zubyk, Elena Evdokimova, Zdzislaw Wawrzak, Gerard D. Wright, Alexei Savchenko
ACS Chemical Biology (2018-04-09) <https://doi.org/gdnrmn>

DOI: [10.1021/acscchembio.8b00116](https://doi.org/10.1021/acscchembio.8b00116) · PMID: [29631403](https://pubmed.ncbi.nlm.nih.gov/29631403/) · PMCID: [PMC6197822](https://pubmed.ncbi.nlm.nih.gov/PMC6197822/)

34. The Genus *Aeromonas*: Taxonomy, Pathogenicity, and Infection

J. Michael Janda, Sharon L. Abbott

Clinical Microbiology Reviews (2010-01) <https://doi.org/cbnfww>

DOI: [10.1128/cmr.00039-09](https://doi.org/10.1128/cmr.00039-09) · PMID: [20065325](https://pubmed.ncbi.nlm.nih.gov/20065325/) · PMCID: [PMC2806660](https://pubmed.ncbi.nlm.nih.gov/PMC2806660/)

35. 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: [10.1111/j.1469-0691.2004.00892.x](https://doi.org/10.1111/j.1469-0691.2004.00892.x) · PMID: [15301675](https://pubmed.ncbi.nlm.nih.gov/15301675/)

36. The Genus *Psychrobacter*

Elliot Juni

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

DOI: [10.1007/978-1-4757-2191-1_12](https://doi.org/10.1007/978-1-4757-2191-1_12)

37. 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](https://doi.org/10.1099/00207713-46-4-841) · PMID: [8863407](https://pubmed.ncbi.nlm.nih.gov/8863407/)

38. *Psychrobacter vallis* sp. nov. and *Psychrobacter aquaticus* sp. nov., from Antarctica

Sisinthy Shivaji, Gundlapalli S. N. Reddy, Korbole Suresh, Pratima Gupta, Suresh Chintalapati, Peter Schumann, Erko Stackebrandt, Genki I. Matsumoto
International Journal of Systematic and Evolutionary Microbiology (2005-03-01) <https://doi.org/cdq9w4>

DOI: [10.1099/ijs.0.03030-0](https://doi.org/10.1099/ijs.0.03030-0) · PMID: [15774658](https://pubmed.ncbi.nlm.nih.gov/15774658/)

39. *Stenotrophomonas maltophilia*: an Emerging Global Opportunistic Pathogen

J. S. Brooke

Clinical Microbiology Reviews (2012-01-01) <https://doi.org/fxx2tt>

DOI: [10.1128/cmr.00019-11](https://doi.org/10.1128/cmr.00019-11) · PMID: [22232370](https://pubmed.ncbi.nlm.nih.gov/22232370/) · PMCID: [PMC3255966](https://pubmed.ncbi.nlm.nih.gov/PMC3255966/)

40. 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: [10.1128/aac.01260-17](https://doi.org/10.1128/aac.01260-17) · PMID: [29061750](https://pubmed.ncbi.nlm.nih.gov/29061750/) · PMCID: [PMC5740310](https://pubmed.ncbi.nlm.nih.gov/PMC5740310/)

41. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data

Sergio Arredondo-Alonso, Rob J. Willems, Willem van Schaik, Anita C. Schürch

Microbial Genomics (2017-10-01) <https://doi.org/gf6b63>

DOI: [10.1099/mgen.0.000128](https://doi.org/10.1099/mgen.0.000128) · PMID: [29177087](https://pubmed.ncbi.nlm.nih.gov/29177087/) · PMCID: [PMC5695206](https://pubmed.ncbi.nlm.nih.gov/PMC5695206/)

42. Resistance gene naming and numbering: is it a new gene or not?

Ruth M. Hall, Stefan Schwarz

Journal of Antimicrobial Chemotherapy (2015-10-28) <https://doi.org/f8gn4r>

DOI: [10.1093/jac/dkv351](https://doi.org/10.1093/jac/dkv351) · PMID: [26510717](https://pubmed.ncbi.nlm.nih.gov/26510717/)

43. Comment on: Resistance gene naming and numbering: is it a new gene or not?

Benjamin A. Evans

Journal of Antimicrobial Chemotherapy (2016-03-06) <https://doi.org/dw7r>

DOI: [10.1093/jac/dkw037](https://doi.org/10.1093/jac/dkw037) · PMID: [26953334](https://pubmed.ncbi.nlm.nih.gov/26953334/)