Consolidating and exploring antibiotic resistance gene data resources

Reference:

Xavier Basil Britto, Das Anupam, Cochrane Guy, De Ganck Sandra, Kumar-Singh Samir, Aerestrup Frank Møller, Goossens Herman, Malhotra Surbhi.- Consolidating and exploring antibiotic resistance gene data resources

Journal of clinical microbiology - ISSN 0095-1137 - (2016)

Full text (Publishers DOI): http://dx.doi.org/doi:10.1128/JCM.02717-15

Consolidating and exploring antibiotic resistance gene data resources

Basil Britto Xavier[a], Anupam J Das[a], Guy Cochrane[b], Sandra De Ganck[a], Samir Kumar-Singh[a,c], Frank Møller Aarestrup[d], Herman Goossens[a], Surbhi Malhotra-Kumar[a*]

Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, Universiteit Antwerpen, Antwerp, Belgium[a]; European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Cambridge, CB10 1SD, United Kingdom[b]; Laboratory of Molecular Pathology, Cell Biology and Histology, Faculty of Medicine and Health Sciences, Universiteit Antwerpen, Antwerp, Belgium[c], National Food Institute, Technical University of Denmark, Denmark[d]

Running Title: Antibiotic resistance gene data resources

Key words: Multi-drug resistant bacteria, ARDB, ResFinder, CARD, ResFams, CBMAR, beta-lactamase, carbapenemase, MDR, XDR, resistome

Text word count: 5049

Abstract word count: 109

*Corresponding author mailing address: Department of Medical Microbiology, Campus Drie Eiken, University of Antwerp, S6, Universiteitsplein 1, B-2610 Wilrijk, Belgium. Phone: 32-3-265-27-52. Fax: 32-3-265-26-63. E-mail: surbhi.malhotra@uantwerpen.be

ABSTRACT

The unrestricted use of antibiotics has resulted in rapid acquisition of antibiotic resistance (AR) and spread of multi-drug-resistant (MDR) bacterial pathogens. With the advent of next generation sequencing technologies and their application in understanding MDR pathogen dynamics, it has become imperative to unify AR gene data resources for easy accessibility for researchers. However, due to the absence of a centralized platform for AR gene resources, availability, consistency and accuracy of information vary considerably across different databases. In this article, we sought to explore existing AR gene data resources in order to make them more visible to the clinical microbiology community, to identify their limitations, and to propose potential solutions.

INTRODUCTION

Over the years, antibiotics have vastly benefitted human and animal heath in combatting bacterial infections. Apart from being widely used in clinical practice, antibiotics are also employed in agriculture, aquaculture and intensive animal farming either as prophylactic agents or for therapeutic purposes (1-3). The unrestrained use of antibiotics, however, has resulted in higher frequency of resistant human pathogens (4). Acquisition of antibiotic resistance can result from a variety of genomic alterations, for instance, single nucleotide mutations, large genomic changes such as insertions or deletions, chromosomal rearrangements, gene duplications, and importantly, factors that have facilitated their rampant spread i.e., carriage on plasmids and other mobile genetic elements (MGEs) including integrons and transposons (5, 6). A fitting example is the recently reported *mcr-1* gene that has been linked to colistin-resistance in humans and animals (7-10), and has been found to be associated with at least three different plasmids up till now (7, 8, 10).

In the last decade, emergence of multi drug-resistant (MDR) bacteria that harbor multiple antibiotic resistance mechanisms/genes has severely limited therapeutic options (4). Common examples are some of the most important Gram-negative human pathogens such as *Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Escherichia coli* that harbor MGEs carrying genes encoding enzymes like extended-spectrum beta-lactamases (ESBLs) that can hydrolyze penicillins, cephalosporins and monobactams, along with aminoglycoside-modifying enzymes, and the Qnr protection proteins that confer resistance to the fluoroquinolones (11, 12). Thus, a single conjugation event involving such MGEs is enough to transform an antibiotic sensitive pathogen to a MDR organism that can potentially cause infections that are non-treatable by the current antibiotic arsenal (13). Extremely worrisome are the rising rates of resistance to carbapenems, one of the most important last-line antibiotics available to us (14).

Carbapenemases such as *bla*$_{VIM}$, *bla*$_{IMP}$, *bla*$_{KPC}$, *bla*$_{NDM}$ are also primarily MGE-encoded and coupled with the high and varying antibiotic selection pressure in hospitals has led to a rapid evolution and spread of these enzyme with varying substrate specificities (15). Currently, more than 40 *bla*$_{VIM}$ and 10 *bla*$_{NDM}$ variants are known (16). On the other hand, in important Gram-positive pathogens such as *Staphylococcus aureus*, resistance to beta-lactam antibiotics is mediated by the 'staphylococcal cassette chromosome' (SCC*mec*) MGEs that integrate in a site-specific manner in the *Staphylococcus* genome (17). Interestingly, the marked differences in antibiotic resistance profiles of community- (CA) and hospital-associated (HA-) of methicillin-resistant S. aureus (MRSA) can be largely attributed to the kind of SCC*mec* element harboured by these strains. CA-MRSA harbor the smaller SCC*mec*IV, V, or VII elements, whereas HA-MRSA typically contain the larger SCC*mec*I, II, III, VI, or VIII elements that encode multiple resistance determinants in addition to the gene encoding beta-lactam resistance, *mecA* (17).While these examples highlight the complex combinations of emerging AR mechanisms and genes, they represent only the tip of the proverbial iceberg. The application of next-generation sequencing (NGS) technology to the study of pathogen genomes as well as to soil-, marine-, and human-associated metagenomes has given us unprecedented insights into unknown reservoirs and novel AR genes (18-21). Currently, a wealth of information with respect to AR genes is available online in AR gene databases (**Table 1**, **Figure 1**). As costs of sequencing are steadily decreasing and response times getting shorter, its utility as a tool for tracking MDR pathogens in real time for routine hospital epidemiology or as an early warning system for outbreak detection is steadily increasing. Application of NGS to routine clinical microbiology and diagnostics will be especially useful in simplifying the technical algorithms utilized for typing and for antibiotic resistance detection. Currently, majority of the routine microbiology laboratories still screen for MDR based on phenotypic susceptibility testing, which is not only subject to guidelines and breakpoints but is also time-consuming as it depends on pathogen growth. Previous NGS-based studies have

demonstrated high concordance between *in silico* predicted and phenotypic antimicrobial susceptibility (22, 23). Nonetheless, sequence-based predictions of phenotypic resistance for clinical purposes need to be made with caution. Firstly, in contrast to phenotypic testing, sequencing data yields information only on resistance to antibiotics but not on susceptibility to others. Also, absence of a resistance gene does not preclude sensitivity to that antibiotic as any new resistances that are not in the utilized AR gene database might have been missed. On the other hand, sequence-based predictions might potentially identify a gene that leads to resistance during treatment although is not expressed during sensitivity testing under specific growth conditions. Thus, while application of NGS to antimicrobial susceptibility testing can result in a more efficient workflow, the large datasets generated here will be heavily reliant on the available AR gene data resources for quality reference data and interpretation making it imperative that the latter are well curated, up-to-date and comprehensive.

We reviewed the currently available AR gene data resources with the aim of making them more visible to the clinical microbiology community, particularly emphasizing on regular updates and easy accessibility to resources that include metadata from published literature. Additionally, we also demonstrated test runs on 4 available databases using in-house and publicly available data. This exercise revealed inconsistent search results, which we discuss in detail and propose two complementary approaches that call for a combined effort in addressing this issue.

ANTIBIOTIC RESISTANCE GENE DATABASES

AR gene data resources are online platforms that offer AR-related reference data in support of prediction of resistome and gene-based antibiograms along with online bioinformatics tools for sequence comparisons, alignment and annotation. These resources accept user nucleotide or protein sequences as queries and return predictions of their AR gene content, often with confidence-related statistics, annotation and onward links to external related resources. We consider first several generalist resources and then move on to AR-focused resources with an

anecdotal commentary and a tabulated summary of each open access data resource's key characteristics.

The NCBI non-redundant (NCBI-nr) data set (http://www.ncbi.nlm.nih.gov/nuccore) represents one of the largest publicly available generalist data resources that include (AR) genes and associated information. In some cases, however, search results obtained with NCBI-nr might not be specific in terms of gene sub-types resulting in multiple hits with similar level of identity and query coverage. It is important to note that results may vary depending on whether the query sequence is a part of the gene or also includes regions flanking the gene. Thus, additional manual verifications may be required for accurate predictions.

Popular further generalist options that relate to protein-level similarity are the UniProt Knowledgebase (UniProtB; http://www.uniprot.org/) and the Protein Families Database (Pfam; http://pfam.xfam.org/), which together provide information on protein sequences, functional annotations and conserved protein families (24, 25). UniProtKB offers an exhaustive collection of protein annotation, cross-references and literature-derived annotations, while Pfam offers conserved protein families. Pfam uses the profile HMM software, HMMER3 (http://hmmer.janelia.org/), in order to identify and build Hidden Markov Models (HMM) of protein families. These generalist resources are of value not only as they are comprehensive for publicly available data but also since they serve to feed data to more specialist AR-gene data resources (**Table 1**).

Antibiotic Resistance Genes Database (ARDB; http://ardb.cbcb.umd.edu/index.html), a manually curated specialist AR-gene database, appeared very promising at the time of its introduction, combining information from several existing resources, offering AR gene (sub)types and ontology information. At launch, it comprised 13293 genes, 377 types, 257 antibiotics, 632 genomes, 933 species and 124 genera to which were applied a two-step filtering of vector sequences, synthetic constructs and redundant genes and then removal of incomplete sequences, yielding 4545

antibiotic resistance gene sequences (26). The resource features various tools for annotation and comparison of genes and genomes. Furthermore, a tool for mutation detection is also provided (26). The site allows upload of data as a single gene or in a batch mode for multiple genes or protein sequences. Though the site is functional and user-friendly, the major concern is with the updates of the database as, according to the database statistics, the last update was in July 2009. Following the last update, the database reported 23137 genes, 380 types, 249 antibiotics, 632 genomes, 1737 species, 267 genera including information on 2881 vectors (vehicles for transmitting genetic material/ genes from one organism to another) including plasmids.

The Comprehensive Antibiotic Gene Database (CARD; http://arpcard.mcmaster.ca/) was first introduced with a β-lactamase ontology feature. Since its introduction in 2009, regular updates have been announced with the most recent one in April 2014. The database facilitates access to exhaustive knowledge resources regarding antibiotic resistance genes and their associated proteins that additionally include antibiotics and corresponding targets. CARD presents a well-developed AR Ontology (ARO) platform that has been expanded from the initial efforts of the ARDB. The ARO allows efficient investigation of molecular data by including classification of AR genes, functional ontology information, SNP details for resistance genes, extensive microarray targets, Gene Ontology (GO), Sequence Ontology (SO) and Infectious Disease Ontology (IDO) (24, 25, 27, 28). Additionally, CARD also features a graphical web tool called Resistance Gene Identifier (RGI), Version 2, in October 2011 for annotation of query sequences. As of the latest update, CARD provides 3008 genes tagged specifically for antibiotic resistance and 4120 genes with AR-related functions. It permits query sequence upload in both batch mode (limited to 20 Mb) and as single sequences. The graphical interface was found to be user friendly and highly descriptive with functional based classification of AR genes.

ResFinder version 2.1(https://cge.cbs.dtu.dk//services/ResFinder/), most recently updated in July 2015, is a database that provides exhaustive information on AR genes from sequenced or partially

sequenced bacterial isolates. Information of acquired resistance genes through horizontal gene transfer can be obtained here. ResFinder provides not only up-to-date information on AR genes but also offers enhanced flexibility in the user interface, which helps minimize unspecific hits. The current version of ResFinder allows a user to set the identity and length coverage thresholds as low as 30% and 20% respectively (29). One of the major advantages of Resfinder over other tools is that it accepts NGS raw reads including *de novo* assembled contigs without any limitations on size or sequence length. However, one of the limitations is that the information currently contained in the database is specific for acquired genes and therefore does not include AR mechanisms mediated by chromosomal mutations. Furthermore, the database accepts only nucleotide (and not protein) sequence queries for comparison.

The Lactamase Engineering Database (LacED; http://www.laced.uni-stuttgart.de/) provides systematic analysis and annotation of sequences that helps compare new entries to already existing ones. Furthermore, the database provides integrated tools for sequence comparison and multiple sequence alignments such as Basic Local Alignment Search Tool (BLAST) and ClustalW, respectively. LacED database, however, specializes in information related to mutations, sequences and structures of TEM and SHV, β –lactamases (30).

ResFams (http://www.dantaslab.org/resfams) is a recently established resource for protein families, which are linked to their HMMs associated with AR function. It aims at providing accurate identification and annotation of AR genes. With the information provided, one can also get an overview of the ecology and evolution of the resistant pathogens. The ResFams platform is specifically targeted towards AR gene families and their HMMs, which are further associated with functional metagenomic datasets acquired from various sources such as soil and human feces as well as from 6000 sequenced microbial isolate genomes across diverse phylogenies and habitats. This data was then utilized to derive 166 HMM profiles comprising the major AR gene classes (31). The authors emphasize that for resistome analysis this HMM-based approach is superior to

that of BLAST-based pairwise sequence alignment to AR-specific databases that are biased towards human-associated organisms and vastly underestimate the potential impact of environmental resistance reservoirs on AR in pathogens (31). The authors demonstrated this by comparing ResFams HMMs with the BLAST-based ARDB and CARD databases for their ability to predict AR function, and showed that 64% of AR proteins identified using ResFams in both the soil and the human gut microbiota were not detected by BLAST. This increased sensitivity over other AR data resources is expected with HMM-based analysis. HMMs are specific models that are constructed based on observed sequence variation sampled across gene or protein families and capture nuanced positional variability for the family. Search of query sequence using these models returns resulting matches that can be distant, and not detectable using sequence-based matching such as in BLAST, but represent valid homologous genes or proteins. However, HMM approaches come at a computational cost, in particular where models are used in scans of high volume whole genome shotgun data. The implication of this cost for ResFams is that the user must provide local computational resources in order to run HMM-based searches. These HMMER tools need to be installed locally under LINUX/UNIX, and the results appear in a tabular form without a graphical interface.

Antibiotic Resistance Gene ANNOTation (ARG-ANNOT; http://en.mediterranee-infection.com/article.php?laref=283%26titre=arg-annot-) is a rapid bioinformatics tool that is used in identifying putative new AR genes in bacterial genomes. ARG-ANNOT also provides data relating to point mutations. In another study, the tool has been tested for its enhanced sensitivity and specificity for both complete and partial gene sequences (32).

The Pathosystems Resource Integrated System (Patric; https: //www. Patricbrc.org/portal/ portal/ patric/AntibioticResistance) provides a platform for genome assembly, protein family comparisons, genome annotations, meta-data information such as AR and pathway comparisons. Patric collects public genome data and currently provides AR data from ARDB and CARD (33).

The Human Microbiome Project (http://hmpdacc.org/HMGOI/), in its efforts to characterize the human microbiome, developed a reference set of 3000 microbial isolate genomes. The HMP also provides a large collection of AR genes (34).

Resistance Determinants Database (RED-DB; http://www.fibim.unisi.it/REDDB/) is a non-redundant collection of resistance genes from various nucleotide sequence databases. One can easily look up the database based on the cluster or reference gene names.

User-friendly Comprehensive Antibiotic Resistance Repository of *Escherichia coli* (U-CARE; http://www.e-bioinformatics.net/ucare/), is a manually-curated resource that provides *E.coli* related AR information, including information from 52 antibiotics, 107 resistance genes and associated information of transcription factors and SNPs (35).

β-lactamase Database (BLAD; http://www.blad.co.in) includes resistance patterns of all classes of β -lactamases collected from published data, NCBI and the crystal structure of proteins from the Protein Data Bank (PDB; http://www.rcsb.org/pdb/home/home.do). BLAD allows sequence comparison using BLAST search tool. Apart from facilitating information regarding the 3D structure and physiochemical properties of bound ligands, BLAD also provides links to the popular nucleotide and protein databases (36). The resource specializes in β-lactamase related information.

Comprehensive β-lactamase molecular annotation resource. (CBMAR; http://14.139.227.92/mkumar/lactamasedb) is a recently established AR gene resource, which provides a fully interactive environment for data access to exhaustive range of β-lactamase resources (37). It provides extensive metadata along with detailed molecular and biochemical information, which could reveal further insights into novel β –lactamases. CBMAR also features tools such as BLAST and searches for family-specific fingerprints employing MAST (Motif alignment and search tool; http://meme-suite.org/tools/mast). Information related to protein, nucleotide, protein structures, alignments, mutation profiles and phylogenetic trees can be

downloaded from the database. According to site statistics, the most recent update was performed on 9 September 2014.

Lahey (http://www.lahey.org/studies/) is a conventional database/repository for β -lactamase classification and amino acid sequences for *bla*$_{TEM}$, *bla*$_{SHV}$ and *bla*$_{OXA}$ extended-spectrum and other inhibitor-resistant enzymes. Information on plasmid-borne quinolone resistance genes *qnr* genes and *qnr* allele designations can also be accessed from http://www.lahey.org/qnrStudies/. At the time of this review, due to the transitioning of the database to a new location, it was not assessed.

Instituut Pasteur database (http://bigsdb.web.pasteur.fr/) provides MLST data for *Klebsiella pneumoniae*, including information particularly for specific β –lactamases, OKP, LEN and OXY.

Latest resources regarding tetracycline and macrolide-lincosamide-streptogramin (MLS) AR genes can be obtained from the Tetracycline and MLS Nomenclature database (faculty.washington.edu/marilynr/). The latest update for the tetracycline and MLS resources was in August 2015.

Antibiotic Resistance Gene Finder (ABRESfinder; http://www.bioindians.org/ABRES/) is an AR gene resource that includes information on the gene sub-families and on the mechanisms of resistance from various sources. ABRESfinder provides links to tools such as BLAST, ClustalW and Primer3Plus and is mainly focused on AR-related information from India.

Additionally, we would also like to shed light on some of the databases that provide information on integrons, AR-related gene cassettes and virulence factors. INTEGRALL, the integron database (http://integrall.bio.ua.pt) is a freely available web tool that provides information on integron sequences and their genetic arrangements with respect to AR genes (38). Furthermore, annotation and information of gene cassettes in mobile integrons can also be accessed using the Repository of Antibiotic Resistance Cassettes (RAC; http://rac.aihi.mq.edu.au/rac/) (39).

The MvirDB (http://mvirdb.llnl.gov/) is a database that targets genes for signature discovery, mainly in identification and characterization of both functional and genetic signatures. MvirDB has a collection of virulence factors and toxins, including AR gene related resources from several other databases. Apart from aiding medical researchers, the database also aims at centralizing information for bio-defense purposes concerning virulence factors and toxins, especially for tracking genetically engineered organisms. The web interface includes two tools – Virulence browser and Virulence BLAST Interface for sequence identification and comparison. MvirDB also compares entries to a high-throughput microbial annotation database (MannDB; http://manndb.llnl.gov/) (40). Although the web tools were found to be functional and easy to work with, we observed some broken links in the documentation section.

Besides the specific characteristics discussed above, one of the common issues that arises with AR gene databases is of false positive predictions due to certain housekeeping genes that are ubiquitously present in bacterial and, sometimes, in mammalian genomes. For instance, dihydrofolate reductases (*dhfr*) are important enzymes catalyzing the folic acid pathway in bacteria, and are targeted by the antibiotic trimethoprim. Resistance arises either by overproduction of chromosomal DHFR due to a promoter mutation in *E. coli* or by an altered chromosomally encoded DHFR due to a single amino acid substitution in the *dhfr* gene in *S. aureus* (41). However, naturally insensitive enzymes have also been reported in some organisms (41, 42) and the mammalian *dhfr* genes are also highly similar to their bacterial counterparts. Expectedly, this gene is highly represented in metadata from various communities including fecal and soil sources (42), and may be challenging for most AR gene databases, which do not include information from soil and ecological microbiome studies, to single out as false positive predictions. ResFams is one database that includes soil resistome metadata, which aids accurate predictions. Furthermore, terminologies or gene names might vary, for example, *dhfr* is often referred to as *dfrA* in certain databases. This ambiguity can be counteracted by conducting a

parallel protein-domain based search using related databases such as Pfam. Thus, although a number of databases are available at our disposal, working with large amount of data still requires fine tuning of parameters, such as identity levels, e-value and bit score to predict the right AR genes and obtain better sensitivity.

ASSESSMENT OF THE PERFORMANCE OF AR GENE DATABASES USING GENE SEQUENCES, WHOLE GENOME SEQUENCES AND (FUNCTIONAL) METAGENOMICS DATA

Next, we carried out test runs on the selected databases namely the ARDB, CARD, Resfinder and CBMAR, using our in house data and those of others. We particularly selected these databases, as the ARDB, CARD and Resfinder are among the most popular AR related reference data sources. CBMAR is a recently established database that offers a comprehensive collection of data resources and tools related to AR genes. The query sequences used to assess the databases comprised of AR gene sequences, whole genome sequences and metagenomics data, including whole genome shotgun and functional metagenomics sequences.

To further verify the availability of latest resources and the accuracy of AR gene predictions from the 4 databases, we selected some of the known carbapenemase genes, *bla*$_{\text{VIM}}$ and *bla*$_{\text{NDM,}}$ and their variants as query sequences. The entire sequence of *bla*$_{\text{VIM-1}}$ (KT124311), *bla*$_{\text{VIM-2}}$ (KR337992.1), *bla*$_{\text{VIM-4}}$ (AJ585042.1), *bla*$_{\text{VIM-19}}$ (KT124310), *bla*$_{\text{VIM-35}}$ (JX982634.1) and *bla*$_{\text{NDM}}$ genes such as *bla*$_{\text{NDM-1}}$ (KP770030.1), *bla*$_{\text{NDM-2}}$ (JF703135.1), *bla*$_{\text{NDM-4}}$ (KP772213), *bla*$_{\text{NDM-6}}$ (KJ872581.1) and *bla*$_{\text{NDM-8}}$ (NG_036906.1) were downloaded from the NCBI database and used in our analysis. Runs were performed with the BLAST parameters set to the default for each of the databases used. Out of the 10 genes that we used for screening the 4 databases, 3 (*bla*$_{\text{VIM-1,}}$ *bla*$_{\text{VIM-2}}$, and *bla*$_{\text{VIM-4}}$) were predicted correctly by all of the 4 databases. *bla*$_{\text{VIM-19}}$ and *bla*$_{\text{VIM-35}}$ were incorrectly predicted by ARDB and CBMAR databases. ARDB returned several non-specific hits to *bla*$_{\text{VIM}}$ gene type with an average similarity percentage of 94.12%, 96.43% for *bla*$_{\text{VIM-19}}$ and *bla*$_{\text{VIM-35,}}$ respectively. The results shortlisted *bla*$_{\text{VIM}}$ genes but not the variants used as query. In

case of CBMAR, *bla*$_{\text{VIM-19}}$ and *bla*$_{\text{VIM-35}}$ yielded non-specific hits; the top 10 hits pointed to *bla*$_{\text{VIM-4,}}$ and *bla*$_{\text{VIM-1,}}$ *bla*$_{\text{VIM-4,}}$ *bla*$_{\text{VIM-5}}$ genes respectively. The CARD and Resfinder databases produced correct results. In case of the *bla*$_{\text{NDM}}$ genes, BLAST results from the ARDB and CBMAR databases returned no hits. Whereas, the CARD and Resfinder databases were found to consistently return correct hits (**Figure 2a**). We did not observe any differences in results upon use of the entire or partial sequence as query.

Next, we utilized whole genome sequences of 3 MRSA strains, UAS391, H-EMRSA-15 and JKD6008 with accession numbers CP007690, CP007659 (43) and CP002120 (44), respectively, in order to assess the databases' ability to predict AR genes and their variants from among whole genome sequence data. Results obtained using whole genome sequences showed that CARD detected the maximum number of AR genes – 6 for UAS391 and JKD6008, and 4 for the H-EMRSA-15 strains. Resfinder predicted 5 for JKD6008, and 1 each for UAS391 and H-EMRSA-15. CBMAR detected 1 each for JKD6008 and H-EMRSA-15 and no hits for UAS391 (**Figure 2b**). We were unable to receive results from ARDB as our sequence files of 2.8 Mb we not accepted as query.

Additionally, we also screened the databases with a publicly available whole genome shotgun metagenomics dataset with primary accession number PRJEB3977. The data was obtained from a study considering the effects of a decolonization strategy on the gut resistome (45). Utilizing this data as query, CARD database predicted the maximum number of resistance genes – a total of 11, including 2 aminoglycoside resistance genes, 7 β-lactamases and 2 either undefined or other genes. While the Resfinder detected a total of 2 including 1 aminoglycoside and 1 β -lactamase, the ARDB predicted a total of 4, 1 each from aminoglycoside, β-lactamase, tetracycline and others (**Figure 2c**). In this case, CBMAR detected no genes.

We also screened the databases using functional metagenomics data that came from a recently concluded study of the naso-oro-pharyngeal resistome from 150 healthy individuals across five

335 countries representing the Northern (Sweden), Southern (Spain), Eastern (Poland, Slovakia) and
336 Western (Belgium) parts of Europe (Vervoort J, Xavier B.B, Joossens M, Darzi Y, Versporten A,
337 Lammens C, Raes J, Goossens H and Malhotra-Kumar S**.** submitted for publication). Here, we
338 utilized a functional metagenomic approach in order to identify differences in presence of
339 antibiotic resistance genes harbored by healthy individuals and attempted to correlate it to
340 antibiotics consumed in that particular country. Samples were enriched overnight in presence of
341 different antibiotics, the DNA was isolated, sheared and cloned in *E. coli*. From the resistant
342 clones, plasmid DNA was sequenced by Illumina (HiSeq), followed by filtering out vector-
343 specific sequence reads and *denovo* assembly of remaining reads using Velvet v1.2.10 (46).
344 Derived contigs were used for BLAST search for AR genes against ResFinder, ARDB, CARD and
345 CBMAR. First, we utilized Resfinder and CARD –RGI for primary screening and results were
346 predicted by both tools. As Resfinder is restricted to acquired resistance genes, *norA*, a multi-drug
347 efflux transporter gene identified by CARD was not identified by Resfinder. Similarly,
348 trimethoprim resistance-conferring genes such as *dfrA/dhfr* gene were only predicted by Resfinder
349 and not by CARD (RGI), CBMAR and ARDB databases. Apart from these predictions, we
350 observed that HMM-based Resfams search on our resistome data gave us additional/novel *dfr*
351 variants and also predicted two (*dfrA8* and *dfrG*) resistance genes previously identified by
352 Resfinder. However, not all *dfr* genes identified by ResFams in our data were resistance-related,
353 which calls for caution while interpreting output from broader databases that include
354 soil/environmental microbiome data.

355 Finally, in order to check for availability of up-to-date reference information, we screened all of
356 the selected databases for the recently reported *mcr*-1 gene, which has been linked to colistin
357 resistance in bacteria. In our observation, as of 16 December 2015, Resfinder was the only
358 database that correctly detected this gene.

In summary, out of the 4 popular databases that we screened for latest information and accuracy, ARDB was found to provide information limited to the gene name, but not the actual variant. Records of the *bla*$_{NDM}$ genes were also missing in ARDB. Although ARDB is considered one of the most popular databases in identification of novel AR genes, lack of regular updates has limited its scope. CARD was found to accurately predict the query gene variant and provide several related hits in the BLAST results. We observed Resfinder to accurately predict all of the query genes. As for the CBMAR database, we found that predictions using nucleotide sequence of the variant gene (*bla*$_{VIM}$) provided correct hits to related *bla*$_{VIM}$ gene variants in 3 out of 5 searches. Whereas BLAST search with *bla*$_{NDM}$ variants returned no results, using a protein sequence query of these genes produced correct results.

Based on our results, we suggest that CARD and Resfinder are ideal while using single gene sequences as query. However, using whole genome sequences and metagenomic sequencing data, CARD performs better than the rest. The Resfinder database, which was found up to date and accurate, currently detects only acquired genes and ignores chromosomal mutations. The ARDB is limited in its scope due to lack of regular updates. As for the CBMAR database, on referring to the resources available for download, we found that the information to all the query variants were available and fully updated. This suggests that while the sequence repository of CBMAR database was found up-to-date, the search tools may need to be updated.

A FUTURE PATH

While we have noted the value of the data resources available to support AR-related work, we have also noted a number of limitations. These include gaps, inconsistent results of searches against different resources with the same query data and lack of up-to-date reference data. While it is beyond the scope in this minireview for us to formulate solutions to these issues, and it is certainly true that expertise beyond ours alone will be required for these solutions, we take the opportunity here to lay out some thinking that we hope will be useful in stimulating, and perhaps

steering, community discussions as to the solutions. We trust that we, and others, will be able to take advantage of existing initiatives, such as the Horizon 2020 COMPARE (COllaborative Management Platform for detection and Analyses of (Re-) emerging and foodborne outbreaks in Europe; http://www.compare-europe.eu/) project, to facilitate and energise these community discussions.

Our proposal is to rise to the challenge with two complementary approaches, the first simpler to lay out in practical terms, the second requiring significant conceptual planning before practical work. The first approach charges the community to develop and implement appropriate best practices and standards in the gathering of reference AR data, in the description, publication and dissemination of these data and in the presentation of methodologies and algorithms offered through the services of each data resource. Establishing best practice around the open sharing of richly and systematically described reference data (such as sequences, annotations, alignments and models) is a step that will reduce redundant effort in discovering source data for analysis and curation in specialist resources and will maximise opportunities to fill gaps. Systematic descriptions of computational methods and query services offered by specialist AR data resources will aid in users' selection of appropriate tools for their analyses and minimise risk of misinterpretation. In this first approach, we do not seek to fill gaps where they exist in AR data resource services, nor to benchmark precision and reliability, but rather seek to create a landscape of transparent and tractable elements that can contribute to many different current and future analytical infrastructures.

Our second approach calls for decisions as to how data resources, both generalist and AR specialist, should move forward to fill gaps in coverage, to provide consistency between query tools that are intended to serve the same function, to remove redundant data processing, curation and software development steps to maximise overall productivity and to guide consumers in making informed analyses using the most appropriate tools. Clearly, broad community

engagement will be required to tackle these issues. While the AR data resource provider and consumer community will no doubt present very specific needs, a number of successful initiatives in other domains will be informative. RNAcentral, for example, is the product of a broad collaboration between 38 specialist data resources covering different families of non-coding RNA genes (47) in a model that centralises database components for non-coding sequences and comprehensive search and discovery across, so far, 22 of the collaborating data resources, while maintaining expert curation and specialist web access at the expert site. A second example, which differs in its model, comes from the Generic Model Organism Database (GMOD) project (http://gmod.org/), which provides software tools for the maintenance and presentation of model organism data across many community projects, including FlyBase (http://flybase.org/), WormBase (http://www.wormbase.org/#01-23-6) and DictyBase (http://dictybase.org/), for example.

SUMMARY AND CONCLUSIONS

In this minireview, we have compiled information on the available data resources that relate to AR function. In the fight against the spread of MDR pathogens, a collective effort is being made in establishing these resources to share knowledge in free and accessible ways. While we find substantial value in what is already available, we note a number of limitations, including those that relate to frequency of updates, and functionality and comprehensiveness of the resources as a whole. Indeed, broader coverage, consistent gene terminologies, centralized (unifying soil, human and other microbiome/resistome data) up-to-date records will be crucial in identifying and tracking (novel) genomic alterations that are acquired by bacterial pathogens upon progression to antibiotic resistance.

At about a time when NGS has become affordable and relatively rapid, and MDR poses an ever-greater challenge to public and animal health, a greater need for comprehensive, up-to-date and interoperable AR gene data resources is created. We hope that this evaluation will initiate a

strategic response among data resource managers to come together to work out mutual solutions, which will make it easier for their operations to be sustained and kept more up-to-date. As coordination at the international level of pathogen genomics efforts grows, we urge that attention be paid to sustaining and extending AR gene data resources as a critical component of our response to MDR.

FUNDING

This work and AJD are financially supported by European Union Horizon 2020 Research and Innovation Programme: COMPARE (COllaborative Management Platform for detection and Analyses of (Re-) emerging and foodborne outbreaks in Europe: Grant Number 643476). B.B.X. is supported by University of Antwerp Research Funds (BOF-DOCPRO 2012-27450).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

448 **TABLE 1.** Characteristics of available AR gene data resources

No.	Database/Repository	AR Gene Spectrum	Functionality & Features	Last update*	References
1.	ARDB	All AR genes	Webtool, BLASTp, BLASTn	Jul-09	(26)
2.	CARD	All AR genes	Webtool, BLASTp, BLASTn, Gene ontology, Gene identifier V2, Annotation	Apr-14	(27)
3.	ResFinder	All AR genes (except chromosome-specific)	Webtool, BLASTn	Jun-15	(23)
4.	LacED	β-lactamases	Webtool, BLASTp, Clustalw	-	(30)
5.	ResFams	All AR genes	BLASTp, Local BLAST, HMM profile	Jan-15	(31)
6.	Patric	All AR genes	Webtool link to CARD and ARDB	Dec-15	(33)
7.	HMP	Human body site-specific study resources	Webtool, BLASTp, BLASTn	Nov-11	(34)
8.	RED-DB	All AR genes	BLASTn, BLASTp	-	-
9.	U-CARE	Organism-specific (*E.coli*)	BLASTp	-	(35)
10.	ARG-ANNOT	All AR genes	BLAST, BioEdit V7.25, Annotation	-	(32)
11.	BLAD	β-lactamases	Webtool	-	(36)
12.	CBMAR	β-lactamases	Webtool, BLASTn, BLASTp,Clustalw, MEME/MAST	Sep-14	(37)
13.	Lahey Clinic	β-lactamases	ß-lactamase classification and assigning allelic number**	Mar-15	-
14.	Instituut Pasteur	OKP, LEN, OXY	MLST database with additional information on specific β-lactamases	Aug-15	-
15.	Tetracycline & MLS Nomenclature	Tetacycline and macrolide AR genes	Information on resistance mechanisms and nomenclature	Jun-15	-
16.	ABRES Finder	All AR genes	Links to external databases	-	-
17.	INTEGRALL	Integron types and genetic context of AR genes	Webtool, BLASTn	Aug-15	(38)
18.	RAC	Genetic context of AR genes	Webtool, Resistance gene cassette annotation	-	(39)
19.	Mvirdb	Virulence and toxin factors	BLASTn, BLASTp, link to ARGODB for AR genes	Apr-14	(40)

449 *Based on information available on their respective websites and/or publications
450 **Moved to http://www.ncbi.nlm.nih.gov/pathogens/submit_beta_lactamase/

FIGURE 1: Non-exhaustive overview of available data resources in light of the functional classifications of resistance genes targeting different antibiotic classes. For instance, β-lactams refers to all beta-lactamase genes including ESBLs and carbapenemases. Colors indicate the subset of genes represented in the databases. Not drawn to scale.

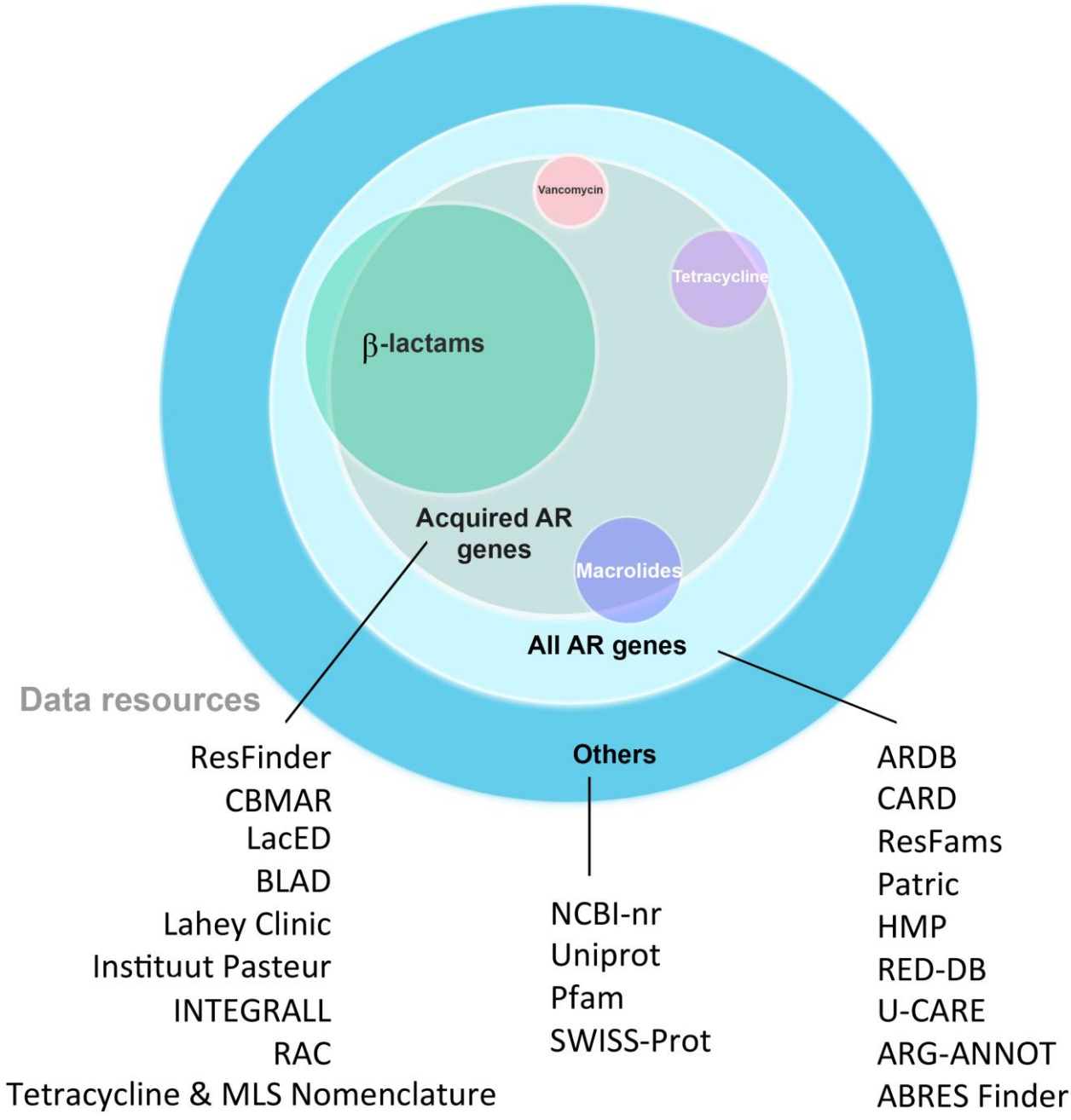
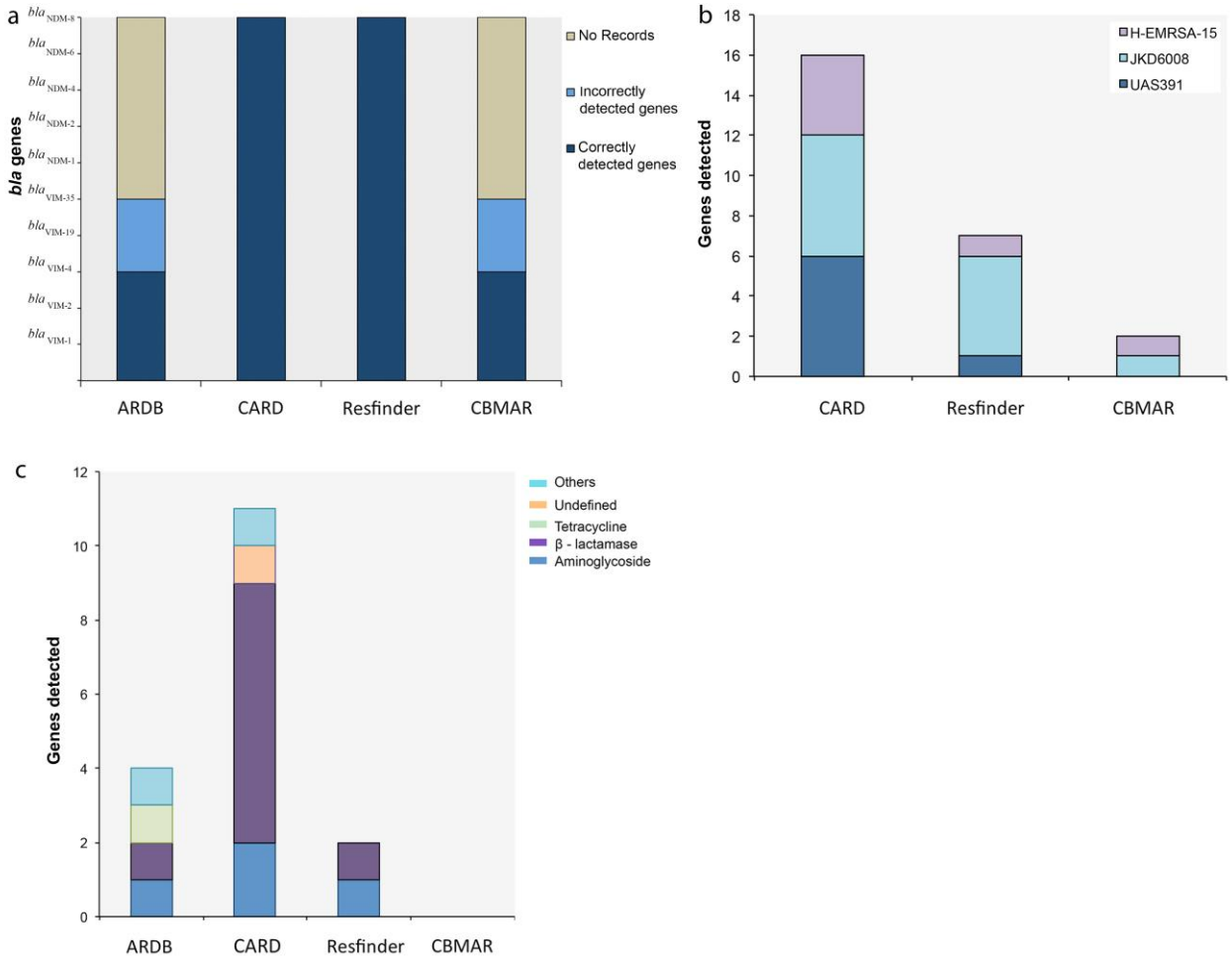


FIGURE 2. Comparison of AR gene data resources ARDB, CARD, Resfinder, and CBMAR using single gene sequences, whole genome sequences and metagenomics datasets as queries. Blast results obtained with *bla*$_{\text{VIM}}$ *and bla*$_{\text{NDM}}$ genes and their variants as query against the four databases (a). Results obtained using whole genome sequences (H-EMRSA-15, JKD6008, UAS391) (b) and metagenomic sequences (45) as query (c).



REFERENCES

1. **Cabello FC.** 2006. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. Environ Microbiol **8:**1137-1144.
2. **Burns A, Shore AC, Brennan GI, Coleman DC, Egan J, Fanning S, Galligan MC, Gibbons JF, Gutierrez M, Malhotra-Kumar S, Markey BK, Sabirova JS, Wang J, Leonard FC.** 2014. A longitudinal study of Staphylococcus aureus colonization in pigs in Ireland. Vet Microbiol **174:**504-513.
3. **Aarestrup FM.** 2015. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. Philos Trans R Soc Lond B Biol Sci **370:**20140085.
4. **Davies J, Davies D.** 2010. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev **74:**417-433.
5. **Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, Bruce D, Rubin E, Myers E, Siggia ED, Tomasz A.** 2007. Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing. Proc Natl Acad Sci U S A **104:**9451-9456.
6. **Ploy MC, Lambert T, Couty JP, Denis F.** 2000. Integrons: an antibiotic resistance gene capture and expression system. Clin Chem Lab Med **38:**483-487.
7. **Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu LF, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu JH, Shen J.** 2015. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis doi:10.1016/S1473-3099(15)00424-7.
8. **Hasman H, Hammerum AM, Hansen F, Hendriksen RS, Olesen B, Agerso Y, Zankari E, Leekitcharoenphon P, Stegger M, Kaas RS, Cavaco LM, Hansen DS, Aarestrup FM, Skov RL.** 2015. Detection of mcr-1 encoding plasmid-mediated colistin-resistant Escherichia coli isolates from human bloodstream infection and imported chicken meat, Denmark 2015. Euro Surveill **20**.
9. **Malhotra-Kumar S, Xavier BB, Das AJ, Lammens C, Hoang HTT, Pham NT, Goossens H.** 2016. Colistin-resistant Escherichia coli harbouring mcr-1 isolated from food animals in Hanoi, Vietnam. Lancet Infect Dis, in press.
10. **Malhotra-Kumar S, Xavier BB, Das AJ, Lammens C, Butaye P, Goossens H.** 2016. Colistin resistance conferring mcr-1 isolated from Belgian bovine and pig farms is harbored on a novel multi-drug resistant plasmid. Lancet Infect Dis, in press.
11. **Carattoli A.** 2013. Plasmids and the spread of resistance. Int J Med Microbiol **303:**298-304.
12. **Vervoort J, Gazin M, Kazma M, Kotlovsky T, Lammens C, Carmeli Y, Goossens H, Malhotra-Kumar S, Saturn WP, groups MWs.** 2014. High rates of intestinal colonisation with fluoroquinolone-resistant ESBL-harbouring Enterobacteriaceae in hospitalised patients with antibiotic-associated diarrhoea. Eur J Clin Microbiol Infect Dis **33:**2215-2221.
13. **Alekshun MN, Levy SB.** 2007. Molecular mechanisms of antibacterial multidrug resistance. Cell **128:**1037-1050.
14. **Pournaras S, Poulou A, Voulgari E, Vrioni G, Kristo I, Tsakris A.** 2010. Detection of the new metallo-beta-lactamase VIM-19 along with KPC-2, CMY-2 and CTX-M-15 in Klebsiella pneumoniae. J Antimicrob Chemother **65:**1604-1607.
15. **Queenan AM, Bush K.** 2007. Carbapenemases: the versatile beta-lactamases. Clin Microbiol Rev **20:**440-458, table of contents.

16. **Canton R, Akova M, Carmeli Y, Giske CG, Glupczynski Y, Gniadkowski M, Livermore DM, Miriagou V, Naas T, Rossolini GM, Samuelsen O, Seifert H, Woodford N, Nordmann P, European Network on C.** 2012. Rapid evolution and spread of carbapenemases among Enterobacteriaceae in Europe. Clin Microbiol Infect **18:**413-431.
17. **Malachowa N, DeLeo FR.** 2010. Mobile genetic elements of Staphylococcus aureus. Cell Mol Life Sci **67:**3057-3071.
18. **Vervoort J, Xavier BB, Stewardson A, Coenen S, Godycki-Cwirko M, Adriaenssens N, Kowalczyk A, Lammens C, Harbarth S, Goossens H, Malhotra-Kumar S.** 2015. Metagenomic analysis of the impact of nitrofurantoin treatment on the human faecal microbiota. J Antimicrob Chemother **70:**1989-1992.
19. **Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Dore J, Meta HITC, Antolin M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C, Cultrone A, Delorme C, Denariaz G, Dervyn R, Foerstner KU, Friss C, van de Guchte M, Guedon E, Haimet F, Huber W, van Hylckama-Vlieg J, Jamet A, Juste C, Kaci G, Knol J, Lakhdari O, Layec S, Le Roux K, Maguin E, Merieux A, Melo Minardi R, M'Rini C, Muller J, Oozeer R, Parkhill J, Renault P, Rescigno M, Sanchez N, Sunagawa S, Torrejon A, Turner K, Vandemeulebrouck G, Varela E, Winogradsky Y, Zeller G, Weissenbach J, Ehrlich SD, Bork P.** 2011. Enterotypes of the human gut microbiome. Nature **473:**174-180.
20. **Canica M, Manageiro V, Jones-Dias D, Clemente L, Gomes-Neves E, Poeta P, Dias E, Ferreira E.** 2015. Current perspectives on the dynamics of antibiotic resistance in different reservoirs. Res Microbiol **166:**594-600.
21. **Didelot X, Bowden R, Wilson DJ, Peto TE, Crook DW.** 2012. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet **13:**601-612.
22. **Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agerso Y, Lund O, Larsen MV, Aarestrup FM.** 2013. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother **68:**771-777.
23. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV.** 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother **67:**2640-2644.
24. **Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G.** 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet **25:**25-29.
25. **Eilbeck K, Lewis SE, Mungall CJ, Yandell M, Stein L, Durbin R, Ashburner M.** 2005. The Sequence Ontology: a tool for the unification of genome annotations. Genome Biol **6:**R44.
26. **Liu B, Pop M.** 2009. ARDB--Antibiotic Resistance Genes Database. Nucleic Acids Res **37:**D443-447.
27. **McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJ, Spanogiannopoulos P, Sutherland AD,**

Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother **57:**3348-3357.

28. **Goldfain A, Smith B, Cowell LG.** 2010. Dispositions and the Infectious Disease Ontology. Formal Ontology in Information Systems (Fois 2010) **209:**400-413.
29. **Zankari E.** 2014. Comparison of the web tools ARG-ANNOT and ResFinder for detection of resistance genes in bacteria. Antimicrob Agents Chemother **58:**4986.
30. **Thai QK, Bos F, Pleiss J.** 2009. The Lactamase Engineering Database: a critical survey of TEM sequences in public databases. BMC Genomics **10:**390.
31. **Gibson MK, Forsberg KJ, Dantas G.** 2014. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. The ISME Journal **9:**207-216.
32. **Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM.** 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother **58:**212-220.
33. **Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW.** 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res **42:**D581-591.
34. **Human Microbiome Project C.** 2012. A framework for human microbiome research. Nature **486:**215-221.
35. **Saha SB, Uttam V, Verma V.** 2015. u-CARE: user-friendly Comprehensive Antibiotic resistance Repository ofEscherichia coli. Journal of Clinical Pathology doi:10.1136/jclinpath-2015-202927**:**jclinpath-2015-202927.
36. **Danishuddin M, Hassan Baig M, Kaushal L, Khan AU.** 2013. BLAD: a comprehensive database of widely circulated beta-lactamases. Bioinformatics **29:**2515-2516.
37. **Srivastava A, Singhal N, Goel M, Virdi JS, Kumar M.** 2014. CBMAR: a comprehensive beta-lactamase molecular annotation resource. Database (Oxford) **2014:**bau111.
38. **Moura A, Soares M, Pereira C, Leitao N, Henriques I, Correia A.** 2009. INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. Bioinformatics **25:**1096-1098.
39. **Tsafnat G, Copty J, Partridge SR.** 2011. RAC: Repository of Antibiotic resistance Cassettes. Database (Oxford) **2011:**bar054.
40. **Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T.** 2007. MvirDB--a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res **35:**D391-394.
41. **Eliopoulos GM HP.** 2001. Resistance to Trimethoprim-Sulfamethoxazole. Clinical Infectious Disease **32:**6.
42. **Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N, Dantas G.** 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature **509:**612-616.
43. **Sabirova JS, Xavier BB, Hernalsteens JP, De Greve H, Ieven M, Goossens H, Malhotra-Kumar S.** 2014. Complete Genome Sequences of Two Prolific Biofilm-Forming Staphylococcus aureus Isolates Belonging to USA300 and EMRSA-15 Clonal Lineages. Genome Announc **2**.
44. **Howden BP, Seemann T, Harrison PF, McEvoy CR, Stanton JA, Rand CJ, Mason CW, Jensen SO, Firth N, Davies JK, Johnson PD, Stinear TP.** 2010. Complete genome sequence of Staphylococcus aureus strain JKD6008, an ST239 clone of methicillin-

resistant Staphylococcus aureus with intermediate-level vancomycin resistance. J Bacteriol **192:**5848-5849.

45. **Buelow E, Gonzalez TB, Versluis D, Oostdijk EA, Ogilvie LA, van Mourik MS, Oosterink E, van Passel MW, Smidt H, D'Andrea MM, de Been M, Jones BV, Willems RJ, Bonten MJ, van Schaik W.** 2014. Effects of selective digestive decontamination (SDD) on the gut resistome. J Antimicrob Chemother **69:**2215-2223.

46. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18:**821-829.

47. **Consortium RN.** 2015. RNAcentral: an international database of ncRNA sequences. Nucleic Acids Res **43:**D123-129.