

ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in Bacterial Genomes

Sushim Kumar Gupta,^a Babu Roshan Padmanabhan,^a Seydina M. Diene,^a Rafael Lopez-Rojas,^b Marie Kempf,^c Luce Landraud,^d Jean-Marc Rolain^a

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR CNRS 7278—IRD 198 IHU, Méditerranée Infection, Faculté de Médecine et de Pharmacie, Aix-Marseille Université, Marseille, France^a; Servicio de Enfermedades Infecciosas, Hospitales Universitarios Virgen del Rocío, Seville, Spain^b; Laboratoire de Bacteriologie, Institut de Biologie en santé-PHB, CHU, Angers, France^c; INSERM U895, Nice, France^d

ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) is a new bioinformatic tool that was created to detect existing and putative new antibiotic resistance (AR) genes in bacterial genomes. ARG-ANNOT uses a local BLAST program in Bio-Edit software that allows the user to analyze sequences without a Web interface. All AR genetic determinants were collected from published works and online resources; nucleotide and protein sequences were retrieved from the NCBI GenBank database. After building a database that includes 1,689 antibiotic resistance genes, the software was tested in a blind manner using 100 random sequences selected from the database to verify that the sensitivity and specificity were at 100% even when partial sequences were queried. Notably, BLAST analysis results obtained using the *rmtF* gene sequence (a new aminoglycoside-modifying enzyme gene sequence that is not included in the database) as a query revealed that the tool was able to link this sequence to short sequences (17 to 40 bp) found in other genes of the *rmt* family with significant E values. Finally, the analysis of 178 *Acinetobacter baumannii* and 20 *Staphylococcus aureus* genomes allowed the detection of a significantly higher number of AR genes than the Resfinder gene analyzer and 11 point mutations in target genes known to be associated with AR. The average time for the analysis of a genome was 3.35 ± 0.13 min. We have created a concise database for BLAST using a Bio-Edit interface that can detect AR genetic determinants in bacterial genomes and can rapidly and easily discover putative new AR genetic determinants.

Next-generation sequencing technologies have greatly reduced the cost for sequencing bacterial genomes and metagenomes and have increased the likelihood of rapid whole-bacterial-genome sequencing in clinical microbiology laboratories (1). An exponential increase in genome releases has occurred as more and more bacterial genomes are sequenced in a short time span. Although many of these genomes are not closed, they are released into the public domain without publication, and their annotation relies on automatic annotation pipelines (2–4). Rapid Annotation using Subsystem Technology (RAST) is one of the most widely used servers for bacterial genome annotation (2). It initially predicts the open reading frames (ORFs) followed by annotations. It also supports a comparative analysis with the annotated genomes maintained in the SEED environment. Although RAST is widely used, it annotates many novel proteins as hypothetical proteins or restricts the information to the domain function. RAST also provides little information about antibiotic resistance (AR) genes. Information on resistance genes can be found in the virulence section of an annotated genome or can be extracted manually from the generated Excel file using specific key words. This process is time-consuming and exhaustive. The largest barrier to the routine implementation of whole-genome sequencing is the lack of automated, user-friendly interpretation tools that translate the sequence data and rapidly provide clinically meaningful information that can be used by microbiologists. Moreover, because released sequences are not always complete sequences (for both bacterial genomes and metagenomes), sequence analysis and annotation should be performed on contigs or short sequences to detect putative functions, especially for AR genes. Recent evidence from environmental, human, and animal microbial metagenomic studies showed that AR has an ancient origin and evolution and that the pool of AR genes (the resistome) remaining to be discov-

ered in these reservoirs is immense (5, 6). As an increasing number of genomes and metagenomes are deposited in the NCBI GenBank database, a specific database needs to be implemented in a versatile tool to rapidly identify existing, putative, new, or emerging AR genes and point mutations in chromosomal target genes that are associated with AR. This bioinformatic tool should be permissive enough to predict the presence of any sequence in a given bacterial genome or metagenome that could be a new AR gene sequence when the given sequence is incomplete or has a low sequence similarity with AR genes already in the database. Such a tool should be easy to use locally with existing biological sequence analysis software, such as BioEdit. Several AR gene databases already exist, including Antibiotic Resistance Genes Online (ARGO), which was published in 2005 (7); the microbial database of protein toxins, virulence factors, and antibiotic resistance genes MvirDB, released in 2007 (8); Antibiotic Resistance Genes Database (ARDB), published in 2009 (9); Resfinder, recently released in 2012 (10), and the Comprehensive Antibiotic Resistance Database (CARD) (11). However, these databases are neither exhaustive nor regularly updated, with the exception of ResFinder and CARD (11). Although Res-

Received 20 June 2013 Returned for modification 18 August 2013

Accepted 14 October 2013

Published ahead of print 21 October 2013

Address correspondence to Jean-Marc Rolain, jean-marc.rolain@univ-amu.fr.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AAC.01310-13>.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AAC.01310-13

Finder and CARD are the most recently created databases, the tools are located in a website, focus only on acquired AR genes, and do not allow the detection of point mutations in chromosomal target genes known to be associated with AR. Additionally, for a sequence to be detected as an AR gene in ResFinder, it must cover at least two-fifths of the length of the AR gene in the database with identity of $\geq 50\%$ (10). In this study, the versatile ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) database was created and used in a local sequence alignment editor to detect existing and putative new AR genes in bacterial genomes and was then compared to the ResFinder database.

MATERIALS AND METHODS

Database generation. The gene sequences of AR genes were selected according to the existing comprehensive and official classification and nomenclature of AR genes (Lahey Clinic website, ARGO, MvirDB, ARDB, and Resfinder) and from previous publications from PubMed (7, 9, 12–20). The details regarding name of many beta-lactamase genes (available with references) were obtained from the Lahey Clinic website (<http://www.lahey.org/Studies/>). After information regarding AR genes was obtained from these resources, an exhaustive list of all AR genes was then created and each gene sequence was retrieved from the GenBank/EMBL sequence retrieval system/ARDB database (<http://www.ncbi.nlm.nih.gov>; <http://srs.ebi.ac.uk>; <http://ardb.cbcb.umd.edu>) and curated before being introduced into the database. The ARG-ANNOT database consists of a single file covering nucleotide sequences in FASTA format from all antibiotics classes (see Fig. S1 in the supplemental material). The nucleotide sequences included in this database from different antibiotics classes are abbreviated as follows: AGly, aminoglycosides; Bla, beta-lactamases; Fos, fosfomycin; Flq, fluoroquinolones; Gly, glycopeptides; MLS, macrolide-lincosamide-streptogramin; Phe, phenicols; Rif, rifampin; Sul, sulfonamides; Tet, tetracyclines; and Tmt, trimethoprim. A unified nomenclature system was followed in which the name contains all of the information regarding gene class, gene name, accession number, gene location in the sequence, and gene size. For example, (AGly)AadA1: M95287:3320–4111:792 tells the researcher that the class of antibiotics is AGly (the aminoglycosides), the gene product name is AadA1, the GenBank accession number is M95287, the gene location is nucleotides 3320 to 4111, and the gene size is 792 bp. The reference sequences of the chromosomal mutations were selected from wild-type *Escherichia coli* sp. strain K-12 substrain MG1655/*Mycobacterium tuberculosis* (H37Rv)/*Acinetobacter baumannii* (AB17978), and only flanked sequences from the mutational region were included in the database (Table 1) (12–17, 19, 21). The information regarding these small mutational regions reflects the position of reported mutations, base replacements, and mutation positions in the database sequence.

Development of ARG-ANNOT. (i) **BLAST search using Bio-Edit.** The generated database in FASTA format was used to create a local database for BLAST using BioEdit (22). It currently uses basic sequence annotation steps such as BLAST (23). BioEdit uses the same BLAST program as NCBI on the virtual interface. The required files for a local BLAST database were created in BioEdit, based on the sequences provided in the FASTA file, and are available in Fig. S1 in the supplemental material. A tutorial explaining how to set up a local BLAST database using BioEdit is provided in Fig. S2 in the supplemental material. The prerequisite to create a database for the BLAST is the sequence file in the FASTA. Bioedit uses the same program used by BLAST, i.e., *formatdb*, which converts FASTA files to binary files (index file, sequence file, and header file) with the extensions .nin, .nsq, and .nhr, respectively. All the generated files and the FASTA file used in *formatdb* are saved in the database folder of the BioEdit program. The results obtained using a local database in a BioEdit BLAST search can be compared directly with the results obtained from NCBI under similar conditions. There are several filter options for BLAST output on a BioEdit platform, including the number of hits to show, the

number of alignments to display, the use of tabular form with or without details, etc. The time taken to create a local BLAST database depends on the number of sequences in the reference set.

(ii) **Output information.** BLAST was used for the analyses, and the results were obtained on the basis of coverage and similarity. The database for automated analysis was arranged such that it could be used to annotate different classes independently or all AR genes simultaneously; the results were obtained either in tabular form or in the general NCBI BLAST format (for details, visit <http://www.mbio.ncsu.edu/bioedit/bioedit.html>) (Fig. 1).

Validation. The developed ARG-ANNOT system was validated using different sequences.

(i) **Specificity test with sequences of different sizes.** For the specificity test, 100 nucleotide sequences of different AR genes and 100 non-AR genes (housekeeping genes from *A. baumannii*, *E. coli*, and *Staphylococcus aureus*) were randomly split into fragments of different sizes (5%, 10%, 25%, 50%, 75%, and 100%), with 20% overlap. These fragments of different lengths, along with selected full-length genes, were subjected to BLAST analysis in ARG-ANNOT to evaluate the specificity and the sensitivity of the tool at different E values (1.0E-100 to 1.0E-05).

(ii) **Test with new sequences for putative new AR genes.** While generating this database, the system was further validated for the *rmtF* gene for the aminoglycoside resistance 16S rRNA methyltransferase family (24), which was recently added to GenBank, and BLAST was performed after curing the *rmtF* gene from the database under relaxed conditions (e-value 1.0E + 10).

(iii) **Analysis of full genomes.** The generated database was used to detect the ARG in the genomes of the 178 different strains of *A. baumannii* and 20 different strains of *S. aureus* from the Microbial Genomes Resources database (GenBank); different output results were obtained by altering the E value (10^{-5} to 10^0). The number of AR genes detected and the average time to obtain a full-genome sequence analysis with ARG-ANNOT were compared with those of Resfinder for first 25 genomes of different isolates of *A. baumannii* (10).

(iv) **Chromosomal mutations.** Mutations in genes known to be involved in AR include those in the 16S rRNA, 23S rRNA, *embB*, *gyrA*, *gyrB*, *katG*, *parC*, *pncA*, *rpoB*, *pmrA*, and *pmrB* genes (Table 1) (13–17, 19, 21, 25–28). The nucleotide sequences in a stretch of 250 to 600 bases (containing mutational regions) from the reference strain for a corresponding gene were used to build this database (see Fig. S6 in the supplemental material). The region of a gene picked for each reference strain was named to help scientists infer the mutation at a particular position. The mutations in the *parC* gene occur at positions 682 and 757, in which C replaces A and A replaces C, respectively; thus, this region is named A682(482)C: C757(557)A, where the numbers in parentheses indicate the positions of the mutation nucleotides in the database. Users can develop their own databases accordingly and look for mutations in different sequences.

RESULTS

Database generation. The sequence information for the AR genes was collected from different resources, and the information for 1,689 AR genes was included in the database (7, 9, 18, 20, 28). All of the sequences were formatted after being arranged in FASTA and used for BLAST analysis. The tool has the option to run single or multiple sequences in FASTA to perform BLAST analyses against the created database and provides the output on the basis of similarity. The ARG-ANNOT database can be downloaded with a tutorial at the following link: <http://www.mediterranean-infection.com/article.php?leref=282&titer=arg-annot>.

Specificity test analysis. The specificity and sensitivity test analysis for the selected genes was highly significant using full coverage and a 100% identity cutoff. Data from the sensitivity and specificity analysis are provided in Table 2. Specific hits were obtained with the corresponding gene in the database for all size

TABLE 1 Position of the mutation in the reference gene which leads to antibiotic resistance in different chromosomal genes

Gene ^a	Mutational point(s)	Reference sequence (5'-3') ^b
<i>rpoB</i>	C1564(64)A:A1565(65)T:A1565(65)G: G1573(73)T:A1574(74)T:C1603(103)G: C1603(103)T:A1604(104)T:C1619(119)A: C1619(119)T:C1696(196)T: A1741(241)T	1501(1):TTGATTAATGCAAACCAGTTGCTGCTGCAATCAAAGAATTCTTTGG TTCAAGCCAGTTATCTCAGTTTATGGACAAAACAACCCACTATCTGAGA TTACACATAAACGTCGTGTATCTGCGCTTGGTCTGGTGGTTTAAACACGT GAACGTGCGGGCTTCGAAGTACGTGACGTACACCAAACCTCACTATGGTTCG TGTTTGTCCAATTGAAACTCCTGAAGGTCCAAACATTGGTTTGTCAACT CGCTTCTGTATACGCAAAGCGAATGACTTCGGTTTCTTGGAAACTCC ATAC:1800(300)
<i>gyrA1</i>	T175(175)A:G224(224)C:C198(198)T: G259(259)T:G259(259)A	1(1):ATGAGCGACCTTGCAGAGAAAATTACACCGGTCAACATTGAGGAAGAG CTGAAGAGCTCCTATCTGGATTATGCGATGTCCGGTCACTTGGTGGCCGTGC GCTGCCAGATGTCCGAGATGGCCTGAAGCCGTACACCGTCCGGTACTT TACGCCATGAACGTACTAGGCAATGACTGGAACAAAGCCTATAAAAAATC TGCCCGTGTGTTGGTGACGTAATCGGTAATACCATCCCATGGTGACT CGGCGGTCTATGACACGATCGTCCGATGGCGCAGCCATTCTCGCTG CGTTAT:300(300)
<i>gyrA2</i>	G2599(299)T	2301(1):GACGAAAGGGTTATCTCCATCAAGGTTACCGAACGTAACGGTTTA GTTGTTGGCGCGGTACAGGTAGATGACTGCGACCAGATCATGATGATCA CCGATGCCGGTACGCTGGTACGTACTCGCGTTTCGGAAATCAGCATCGT GGGCCGTAACACCCAGGGCGTGATCCTCATCCGTAAGTGCAGGAAGATGAA AACGTAGTGGGTCTGCAACGTGTTGCTGAACCGGTTGACGAGGAAGATC TGGATACCATCGACGGCAGTCCCGCGGAAGGGGACGATGAAATCGCTC CGGAAGTGA:2600(300)
<i>parC</i>	G233(33)A:A238(38)C:G239(39)T: C240(40)A:G250(50)A: A682(482)G:G757(557)A	201(1):CGGTGACGTAAGGTAATACCATCCGCACGGCGATAGCGCCTGT TATGAAGCGATGGTCTGATGGCGCAACCGTTCTCTTACCCTTATCCGCT GGTTGATGGTCAGGGGAACTGGGGCGCGCCGACGATCCGAAATCGTTC GCGGCAATGCGTTACACCGAATCCCGTTGTGCGAAATATCCGAGCTGCT ATTGAGCGAGCTGGGGCAGGGGACGGTACTGGGTGCCAAACTTCGAC GGCACTTTGCAGGAGCCGAAAATGCTACCTGCCCGTCTGCCAAACATTT TGCTTAACGGCACCCCGGTTATGCGCTCGGCATGGCGACCGGATATCC ACCGCATAAACCTGCGTGAAGTGGCTCAGGCGGCAATCGCATTAAATCGAC CAGCCGAAAACACGCTCGATCAGCTGCTGGATATCGTGCAGGGGCCG GATTATCCGACTGAAGCGGAAATATCACTTCGCGCGCCGAGATCCGTA AAATCTACGAGAACGGACGTGGTTCAGTGCCTATGCGCGCGGTGTGGA AGAAAGAGATGGCGCGGTGGTTATCAGCGCATGCGCATCAGGTTT CAGGTGCGCGGCTACT:800(600)
16S rRNA	T1406(206)A:A1408(208)G: C1409(209)T:G1491(291)T	1201(1):ATCATGGCCCTTACGACCAGGGCTACACACGTGTACAATGGCGCA TACAAAGAGAAGCGACCTCGCGAGAGCAAGCGACCTCATAAAGTGCCT CGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCCGGAATCGCT AGTAATCGTGGATCAGAATGCCACGGTGAATACGTTCCCGGGCCTTGTAC ACACCGCCCGTACACCATGGGAGTGGGTTGCAAAAAGAAGTAGGTAGCT TAACCTTCGGGAGGGCGCTTACCCTTTGTGATTATGACTGGGGTGAA GTCGTAA:1500(300)
23S rRNA	A2058(8)G:A2059(9)G:C2611(561)T	2051(1):AAGACGGAAGACCCCGTGAACCTTACTATAGCTTGACACTGAAC ATTGAGCCTTGATGTGTAGGATAGGTGGGAGGCTTTGAAGTGTGGACGC CAGTCTGCATGGAGCCGACCTTGAATACCACCTTTAATGTTTGTATGTT CTAACGTTGACCCGTAATCCGGTTGCGGACAGTGTCTGGTGGGTAGTT GACTGGGGCGGTCTCCTCCTAAAGAGTAACGGAGGAGCACGAAGGTTG GCTAATCCTGGTCCGACATCAGGAGGTTAGTCAATGGCATAAGCCAGC TTGACTGCGAGCGTGACGGCGGAGCAGGTGCGAAAGCAGGTCATAGTG ATCCGGTGGTTCTGAATGGAAGGGCCATCGCTCAACCGGATAAAAGGTAC TCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCATATCGACGGCGGT GTTTGGCACCTCGATGTCCGCTCATCACATCCTGGGGTGAAGTAGGTC CCAAGGTATGGCTGTTCCGCAATTAAGTGGTACGCGAGCTGGGTTTA GAACGCTGTGAGACAGTTCGGTCCCTATCTGCCGTGGGCGCTGGAGAAC TGAGGGGGGCTGCT:2650(300)
<i>katG</i>	G871(171)A:G900(200)C:G944(244)C: G985(285)GG	701(1):GGCCGAACGGCAACCCGACCCCATGGCCGCGCGGTGACATTCCG CGAGACGTTTCGGCGCATGGCCATGAACGACGTGAAACAGCGGCGCTG ATCGTCCGGGTCACACTTTCGGTAAGACCCATGGCGCCGCGCCGCGCCG ATCTGGTCCGCCCCGAACCCGAGGCTGCTCCGCTGGAGCAGATGGGCTT GGGCTGGAAGAGCTCGTATGGCACCCGGAACCGGTAAGGACGCGATCAC CAGCGGCATCGAGGTCTGATGGACGAAACCCCCGACGAAATGGGACAA CAGTTTCCTCG:1000(300)

(Continued on following page)

TABLE 1 (Continued)

Gene ^a	Mutational point(s)	Reference sequence (5'-3') ^b
<i>pncA</i>	A139(39)G:T192(92)TA	1(1):ATGCGGGCGTGTGATCATCGTTCGACGTGCAGAACGACTTCTGCGAGGGTGGCTCGCTGGCGGTAACCGGTGGCGCCGCGCTGGCCCCGCCATCAGCGACTACCTGGCCGAAGCGGGGACTACCATACGTCGTGGCAACCAAGGACTTCCACATCGACCCGGGTGACCACTTCTCCGGCACACCGGACTATTCCCTCGTCGTGGCCACCGCATTGCGTCAGCGGTAATCCCGCGCGGACTTCCA TCCCAGTCTGGACACGTCCGCAATCGAGGCGGTGTTCTACAAGGGTGCCTACACC:300(300)
<i>embB</i>	A916(16)G:A916(16)C:G918(18)T:G918(18)A:G918(18)C	901(1):GGTACATCTGGGCATGGCCCCGAGTCGCCGACCACGCCGGGTACA TGTCCAACTATTTCCGCTGGTTCGGCAGCCCGGAGGATCCCTTCGGCTGG TATTACAACCTGCTGGCGCTGATGACCCATGTACGCGACGCCAGTCTGTG GATGCGCCTGCCAGACCTGGCCGCCGGGCTAGTGTGCTGGCTGCTGCTG TCGCGTGAGGTGCTGCCCGCCTCGGGCCGGCGGTGGAGGCCAGCAAA CCGCCTACTGGGCGCGGCCATGGTCTTGCTGACCGCGTGGATGCCGT TCAACAAC:1200(300)
<i>rpsL</i>	A128(128)G:A263(263)G	1(1):ATGGCAACAGTTAACCAGCTGGTACGCAAACCACGTGCTCGCAAAGTT GCGAAAAGCAACGTGCCGTGCGCTGGAAGCATGCCCGCAAAAACGTGGCG TATGTACTCGTGTATATACTACCCTCTAAAAAACCGAATCCGCGCTG CGTAAAGTATGCCGTGTTCTGCTGACTAACGGTTTCGAAGTGACTTCTCA CATCGGTGGTGAAGGTCAACCTGCAGGAGCACTCCGTGATCCTGATC CGTGGCGGTGCTGTTAAAGACCTCCCGGGTGTTCGTTACCACACCGTAC GTGGT:300(300)
<i>pmrA</i>	G305(155)A	151(1):GAGTTGCTTGCCCGTATTCATGCCTTACTACGCCGTAGTGGAGTAGA AGTCAAACTTGCGAGTCAAGATCAACTATTAGAAAAGTGGTGTACTGGTTT TAAATGTTGAACAGCATATTGCGACGTTTAAAGGCCAACGCATTGATTTA TCAAATCGTGAATGGCAATCTTYKAATTCCACTTATGACTCACCCAAAT AAAATCTTTTCTAAAGCCAACTTAGAAGATAAGTTATATGATTTTGATAGT GATGTGACCAGTAATACTATTGAAGTATATGTTACCATTAAAGAGCG AAG:450(300)
<i>pmrB1</i>	C38(38)A:C41(41)T:G261(261)C:G261(261)T	1(1):GTGCATTATTCATTAAAAAACGACTGATTGGGGCACCTCAATTTTCA GTGTCACTTAGGTTGTATTTTAAATTTTTAGTGCTTATAAGGTTGCACTT CAAGAAGTCGATGAAATTCTAGATACTCAAATGAAGTATTTAGCGGAAAG AACAGCTGAGCACCTTTAAAAACTGTAAGCAGTAAGTTTGATTTTCAT AAAACTTACCAGGAAGAAGATCTGTTTATCGATATCTGGGCTTATAAGG ATCAGGCCCATTTGTCTCATCATTTACATTTGCTGGTCCACCTGTTGA GCAA:300(300)
<i>pmrB2</i>	T434(14)A:C680(260)T:C697(277)T: C697(277)A:G784(364)C	321(1):CTTGCGGGCAGTATGTTTATTCCGTATTTAATTATTTTACCTTTTGCA ATATTTGCCTTAGCAGCCATTATTCGTCGTGGTTTAAAACCAATAGATGA TTTTAAAAATGAGTTAAAAGAACGCGATTCCGAAGAACTCACACCAATTG AAGTACATGATTATCCTCAAGAGCTTTTACCTACTATTGACGAAATGAAC CGTCTTTTTGAGCGCATTCTAAAGCGCAAAATGAACAGAAGCAATTT ATTGCCGATGCTGCTCATGAATTACGAACACCTGTGACTGCATTGAACT TACAAACCAAGATTTTGCTAAGCCAGTTTCTGAGCATGAATCATTGCA AAACCTAAGCAAGGGTTTGGCGGTATTGAGCATT:700(380)
<i>pmrB3</i>	T1057(7)A:T1160(110)A:C1208(158)T	1051(1):GTTATTAATATTTTCAGTTTATAACCGATCCAGATCACTACGCATGTAT TCAAATTGAAGATAGCGGTGCAGGAATAGACCCTGAAAATTACGATAAAG TCCTTAAGCGTTTTTATCGCGTGCATCACCATCTTGAGGTGGGAAGTGGT CTAGGTTTATCTATTGTAGATCGTGCAACTCAAAGGCTTGGTGGGACTTT AACCTCGATAAGAGCTTAGAGCTTGGCGGTCTTTCTGTATTAGTGAAAT TAC:1300(250)

^a The nucleotide sequences for *gyrA*, *parC*, *rpsL*, 16S rRNA, and 23S rRNA were obtained from *Escherichia coli* sp. strain K-12 substrain MG1655; for *rpoB*, *pmrA*, and *pmrB* genes from *Acinetobacter baumannii* 17978; and for *katG*, *pncA*, and *embB* genes from *Mycobacterium tuberculosis* H37Rv.

^b The first number and the second number in parentheses represent the nucleotide position of the reference gene in NCBI GenBank and the nucleotide position in the table, respectively.

fragments; however, the hits were restricted to the group level for the smaller fragments at an elevated E value, which shows the suitability of the developed system. The tool was able to unambiguously identify all these genes with sensitivity and specificity at 100% for full-length sequences. For smaller fragments, we found sensitivity and specificity of 100% at the group level for AR gene sequences for at least 50% of the full sequence at an E value of

≤1.0E-5. Specificity corresponding to the specific name of the AR gene was 89% for at least 50% of the full sequence. Finally, specificity ranged from 74% to 86% for AR gene sequences covering at least 25% of the full sequence at an E value ≤ 1.0E-5 with 100% sensitivity.

Analysis of a putative gene. The results obtained for the *rmtF* gene were highly significant: BLAST analysis revealed that the

NCBI Local BLAST

BLAST is government software obtained from the NCBI. For reference see:
 Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Program: **blastn** Nucleotide Database: **ARGANNOT1.txt**
 Protein Database: **aaARG.fas**

Query: **>AYE**
 CAATTGATTTTGAATTTAAATTTATTCCACAGAGATTTCTTTTTAGGCCAAAATCAAACGAAATCGTCATC
 TACCGTACAACAGGGCAATCAGCCTTTAACTTTAAATTTGCCTAAAGCAGTTTCTCAATATGATGAGAAC
 CTTTAGGTC AAGAGTTGGCTGCTAAACACTATTTCCGTCCTCGTAATCCTCAAGTTGCTGCTAAATATGC
 TGAACCATTAAATGACATATTCATCGCCATCAAGAACGGCTGTTGTACGAAACCGCTTGCAGTCTGAACCT
 CAACTTTTTCAGCTTCTGCTTTTCTCTATGCAAGTGTCTCCAACTGGGAGTTTTCGCCCAATTTATATC

Output file name: _____ Specify (default = file opened but not saved)

Open output
 Filter sequences for low-complexity regions
 Do Gapped BLAST (not available for tblastx)
 Show GI's in defines
 Tabular output

Expectation Value (E): **1**
 Matrix: **BLOSUM62**

Max number of hits to report: **500** Effective database size: **0** (0 = real size)
 Max number of alignments to show: **250**
 Threshold for extending hit: **0**

Additional parameters: **-r**

Warning! The complete combined command line (including file paths and auto-set parameters) cannot exceed 128 characters Under DOS. I have not yet found a way around this. If the program doesn't run, try saving the query file to C:\Temp first.

Usage

```
blastall arguments:
-p Program Name [String] (set internally with BioEdit)
-d Database [String] (set internally with BioEdit)
-i Query File [File In] (set internally with BioEdit)
```

Do Search Dismiss BLAST release notes NCBI BLAST help

A

```
programs", Nucleic Acids Res. 25:3389-3402.
Database: C:\BioEdit\database\ARGANNOT1.txt
1524 sequences: 1,315,037 total letters

Query= AYE
(3,936,291 letters)

Sequences producing significant alignments:

Score      E
(bits) Value

(Tet) TetA:JX424423:94438-95712:1275                2527  0.0
Bla(Beta-lactamase class A):NC_010410:2507740-2509014:1275 2527  0.0
PheCmlA5:HQ141279:907-2166:1260                    2498  0.0
(Phe) CmlA1:M64556:601-1860:1260                   2466  0.0
(Bla) ClassC:CP000863:c413167-411869:1299          2456  0.0
Bla(beta-lactamase class C):NC_010410:1184557-1185747:1191 2361  0.0
(Tet) TetG:AF133140:757-1932:1176                  2331  0.0
(Bla) ClassA:CP001921:c1809235-1808015:1221        2064  0.0
Bla(Beta-lactamase class A):NC_010410:1803480-1804499:1020 2022  0.0
Bla(Zn-dependent hydrolase, glyoxylase):NC_010410:882015-882977:963 1909  0.0
(Bla) VEB-1 DQ393569                                1784  0.0
(Bla) VEB-2 AY027870                                1776  0.0
(Bla) VEB-4 EF136375                                1768  0.0
(Bla) VEB-3 AY536519                                1768  0.0
(Bla) VEB-7 FJ825622                                1760  0.0
(Bla) VEB-6 EU259884                                1760  0.0
(Bla) VEB-5 EF420108                                1760  0.0
(Sul) SulI:AF071413:6700-7539:840                  1665  0.0
(Bla) OXA-69 AY750911                                1635  0.0
(AGly) AadA1-pm:JQ690540:7968-8798:831             1631  0.0
(Bla) OXA-112 EF660098                              1628  0.0
```

B

FIG 1 ARG-ANNOT database browser on Bio-Edit. (A) Paste the query sequences and select type of database, select the different parameters, and search. (B) BLAST output.

TABLE 2 BLAST analyses to validate the system under different stringent conditions using gene fragments of different percentages of length^a

% length e-value	Parameter	Value(s) for gene fragment of length:											
		100%		70%		50%		25%		10%		5%	
		X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
1E-100	Sp (%)	100	0	96 (4)	0	90 (10)	0	74 (26)	No	No	No	No	No
	Sn (%)	0	100	0	100	0	100	0	No	No	No	No	No
1E-50	Sp (%)	100	0	95 (5)	0	89 (11)	0	86 (14)	0	81 (19)	No	No	No
	Sn (%)	0	100	0	100	0	100	0	100	0	No	No	No
1E-25	Sp (%)	100	0	94 (6)	0	89 (11)	0	86 (14)	0	77 (23)	0	66 (34)	No
	Sn (%)	0	100	0	100	0	100	0	100	0	100	0	No
1E-10	Sp (%)	100	0	94 (6)	0	89 (11)	0	86 (14)	0	81 (19)	0	77 (23)	0
	Sn (%)	0	100	0	100	0	100	0	100	0	100	0	100
1E-5	Sp (%)	100	0	94 (6)	0	89 (11)	0	84 (16)	0	76 (24)	0	74 (26)	0
	Sn (%)	0	100	0	100	0	100	0	100	0	100	0	100

^a The table shows the specificity and sensitivity analysis of 100 randomly selected sequences (1.0E-100 to 1.0E-5). The e-values represent the different stringent conditions. X, randomly selected antibiotic resistance gene sequences from the database (see Fig. S8 in the supplemental material); Y, randomly selected gene sequences (no antibiotic resistance gene) from *Escherichia coli*, *Acinetobacter baumannii*, and *Staphylococcus aureus* genomes; Sp, specificity; Sn, sensitivity; No, no output. Values in parentheses show the numbers of specific hits at the group level.

total coverage was less than 35%, whereas similarity with other members of the *rmt* gene family (*rmtD*, *rmtD2*, and *rmtA*) was observed in small stretches of 17 to 40 bp (Fig. 2). The E values for a few of the hits (0.04, 0.19, and 0.74; Fig. 2) were not

statistically significant, although the database did help to predict the genes on the basis of similarity.

Analysis of genomic sequences. Table S3 and Table S5 in the supplemental material show the results of comparative analysis

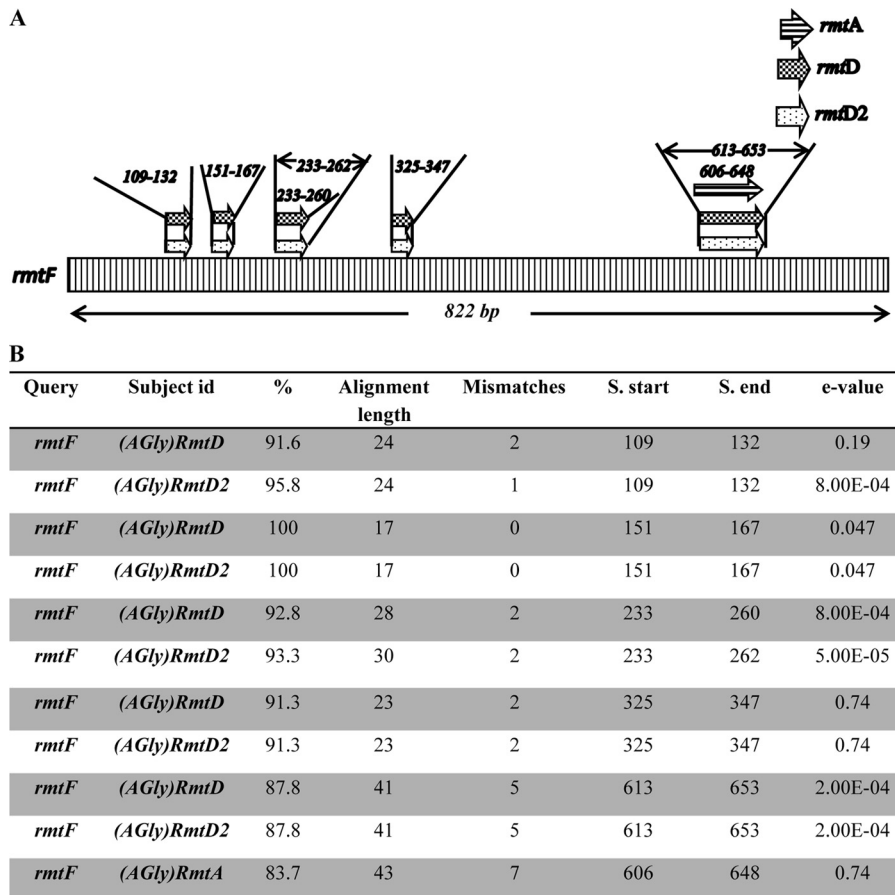


FIG 2 (A) Topographical representation of BLAST results for *rmtF* after curing it from the database. (B) BLAST analysis of *rmtF* gene using ARG-ANNOT.

of *A. baumannii* and *S. aureus* genomes for ARG using ARG-ANNOT and Resfinder (10). We have analyzed 178 genomes (closed and draft) of different *A. baumannii* strains and compared ARG-ANNOT with the existing Web-based ARG analyzer Resfinder (10). ARG-ANNOT identified a significantly higher number of AR genes (2,011 genes) than Resfinder (1,004 genes). The average number of genes detected by ARG-ANNOT was 11.29 ± 4.87 ($P = 1.0 \times 10^{-6}$), whereas the number of genes detected by Resfinder was 5.64 ± 4.31 ($P = 1.0 \times 10^{-6}$). The average times taken for the analysis by ARG-ANNOT and Resfinder were 3.35 ± 0.13 and 2.29 ± 0.06 min, respectively. The real-time analysis was done for first 25 randomly picked genomes in a stretch (Resfinder and ARG-ANNOT), as mentioned in Table S3 in the supplemental material. The ARG-ANNOT system recognized 50% more genes than Resfinder. Approximately 5.65 additional genes were detected by ARG-ANNOT. Some of the genes detected by ARG-ANNOT were not detected by Resfinder; those genes either showed a low (<50%) level of sequence similarity or were not present in the database (Resfinder could not detect the *tetR* gene in *A. baumannii* strains AB1583-8, AB4857, AB-NAVAL-13, AB-NAVAL-81, AB-BC-1, AB-BC-5, AB-OIFCO-98, AB-OIFCO-109, AB-OIFCO-137, AB-WCA694, and AB-TG27339 because of the absence of these genes in the database, and it detected the *catB8* gene as the *catB3* gene with 84.84% similarity). The genes which were present in multiple copies or belong to the same group were detected once by Resfinder. Only one *sull* gene was detected by Resfinder in the ABAYE strain of *A. baumannii*; however, four copies of the *sull* gene were detected by ARG-ANNOT (see Table S4 in the supplemental material). Two *bla*_{OXA} genes (*bla*_{OXA-23} and *bla*_{OXA-69}) were detected in *A. baumannii* strain AB0057 using ARG-ANNOT, whereas only one *bla*_{OXA} gene was detected by the Resfinder. Thus, ARG-ANNOT may have the upper hand with respect to detecting the exact number of ARGs in the genome. We have analyzed 20 closed genomes of different *S. aureus* strains and compared the results with Resfinder (10). As with the *A. baumannii* results, ARG-ANNOT identified more than 50% additional genes (203 genes) compared to Resfinder (98 genes) in this Gram-positive bacterium (see Fig. S5 in the supplemental material) with a significant P value $\leq 10^{-3}$. The detected ARGs were compared to a gold-standard manual annotation of well-known reference genomes, including two *A. baumannii* (AYE and SDF) and two *S. aureus* (08BA02176 and MW2) reference genomes, and sensitivity of 100% was found. It was observed that a few genes were annotated either as hypothetical proteins or by other names; however, ARG-ANNOT detected them as ARGs which showed >98% similarity to the reported ARGs. For, e.g., two beta-lactamase genes in the AYE genome (CU459141; positions 882015 to 882977 and positions 1996441 to 1997136) and SDF (CU468230; positions 764865 to 765827 and positions 1797960 to 1798655), the gene products are annotated as hypothetical proteins in NCBI, and the products of the aminoglycoside gene *aph* and of the *aac*(6′)-I_k gene are annotated as hypothetical proteins in the annotated genome of *S. aureus* strains 08BA02176 (NC_018608) and MW2 (NC_003923) at positions 1844600 to 1845400 and positions 722585 to 723028, respectively.

Analysis of mutations. The mutation analysis was conducted for the *rpoB* gene in different *A. baumannii* genomes (178 genomes). Single-point mutations were detected in *A. baumannii* strains AB210 and TG2018 at positions 1604 and 1565 of the reference sequence (*A. baumannii* 17978), respectively. A single-

point mutation was detected in the draft genome of strain AB210 at position 104 for the *rpoB* gene in the database, where A is replaced by T (as shown in Fig. S7 in the supplemental material), which corresponds to nucleotide position 20420 in contig 37 (AEOX01000037). However, another point mutation was detected in draft genomes of strain AB-TG2018, where A is replaced by T at position 65 of the database, which corresponds to nucleotide position 5589 in contig 34 (AMIE01000034). No reported *rpoB* mutation was detected in the rest of the analyzed genomes.

DISCUSSION

In this study, we generated a recursive database of all known AR genes to date and a simple and rapid bioinformatic tool, ARG-ANNOT, for the detection of AR genes and point mutations known to be associated with AR in bacterial genome sequences. Compared to other existing AR gene databases, ARG-ANNOT shows excellent sensitivity and specificity for the identification of known AR genes not only for complete gene sequences but also for partial sequences and/or sequences with low levels of similarity with existing sequences in the database. For example, the ARDB database has a plethora of information concerning AR genes and is a robust database (9). However, this database has not been upgraded since 2009 and has a restricted sequence entry (≤ 5 kb) for BLAST that is not compatible with whole-genome sequence analysis (9). The ARGO database cannot be informative because the data are restricted to the β -lactamase, tetracycline, and vancomycin genes, which results in limited sequence information (7). In addition, some of the virtual links between ARGO and GenBank are not correct (e.g., the link directs users to the genome instead of to the target gene in NCBI). MvirDB is a broad repository of virulence-associated genes, including toxins, virulence factors, and AR genes (8). The Lahey Clinic website has a comprehensive collection of AR genes and attempts to standardize the nomenclature for these genes but does not provide the sequence information. However, the CARD database enables efficient analysis of genomic sequences for AR genes but lacks a means to provide information about numbers of genes and their locations in the genome or in the contigs. Moreover, it accepts only nucleotide sequences for BLAST analysis (11). The efficient Resfinder gene analyzer was introduced recently and has been used successfully for the detection of known AR genes, including genes from bacterial genome sequences that show good concordance between genetic detection and phenotypic antibiotic susceptibilities at a threshold of 98% of sequence identity (10, 29, 30). However, because that system detects only existing AR genes on the basis of greater sequence similarities ($\geq 50\%$) and high sequence coverage (at least two-fifths of an existing sequence), it can predict only known AR genes, which is not helpful for the detection of putative new AR genes with low levels of similarity and/or short sequences, as demonstrated in this study. The comparative analysis for the detection of AR genes from the genomes of different *A. baumannii* and *S. aureus* strains performed using ARG-ANNOT and Resfinder (10) revealed a significantly higher number of AR genes detected by ARG-ANNOT. Hence, the probability of finding existing genes increased with better inputs and relaxed searched criteria. This analysis further suggests that this tool may be useful to analyze a wide range of bacterial genomes to detect ARG. The ARG-ANNOT database is not exhaustive but possesses a significantly higher number of AR gene sequences than Resfinder, which may be a reason for the detection of extra genes (see Fig. S4 in the

supplemental material). Moreover, the relaxed search criteria (high E value) increased the processing time for ARG-ANNOT compared to Resfinder (10); consequently, the system became more sensitive (wider search criteria enabled the system to recognize diverse genes with a low level of similarity). Some of the genes with low percentages of similarity were not detected by Resfinder (10) because of the restricted search criteria ($\geq 50\%$). In contrast, these genes were detected by ARG-ANNOT. Multiple copies of a few genes or closely related genes were detected once by Resfinder (10), whereas ARG-ANNOT detected these correctly. Thus, ARG-ANNOT enables the detection of the correct number of genes present in a genome. In a report from Tan et al., the threshold used for the identification of AR genes in two multidrug-resistant *A. baumannii* isolates was also 98%, and they found only 15 AR genes (29), whereas we detected 18 AR genes with ARG-ANNOT (see Fig. S3 in the supplemental material). Similarly, ARG-ANNOT detects specific genes on the basis of homology and helps in predicting putative new AR genes with low sequence similarities. This property is exemplified by the prediction that the recently described *rmtF* sequence, which was initially absent from our database, belongs to the *rmt* aminoglycoside resistance 16S rRNA methyltransferase family, even though the deduced protein sequence showed only 25% to 46% identity with other aminoglycoside resistance 16S rRNA methyltransferases (24). We detected similar results with CARD (11). Moreover, the Resfinder and CARD tools are unable to detect point mutations in target genes that are known to be associated with AR. Moreover, Resfinder was unable to detect AR genes present in multicopies. The tool can be used at two levels: first, to identify all known existing genes in a given genome within a few minutes; second, using less-stringent conditions, to look for the presence of putative new or unknown AR genes with lower similarities. One of the main limitations is that such putative sequences with low similarities could be false-positive results from the BLAST output. To circumvent this problem, each new or putative AR gene should be further studied biologically to verify its activity, i.e., by cloning and expression in a vector. If activity is confirmed, then the new AR gene would be described and published before being added in the database. Because there is a tremendous amount of redundancy in the BLAST output, reflecting many variants of each gene, we have posted in the Web interface an XL sheet for a post-BLAST cleanup script that produces a usable resistance gene annotation of a genome. Although the ARG-ANNOT system currently does not detect mutations automatically, its quick processing and the output file in the form of small stretches (250 to 600 bases) enable the rapid detection of mutations. Thus, we believe that many AR genes in bacterial genome sequences that are currently available remain unknown; ARG-ANNOT will be helpful to reannotate any genome from the public domain and can be used to perform comparative analyses to discover new AR gene sequences from previously annotated genomes. Finally, our report demonstrates that ARG-ANNOT is a promising tool to search for putative new AR genes in a given sequence, including the metagenomes from human microbiota or the environment that represent a huge reservoir of AR genes yet to be discovered (5, 6).

Conclusion. We have generated a concise database of all currently known AR genes and a simple and rapid bioinformatic tool for the detection of known and putative new AR genes using local BLAST in Bioedit. The ARG-ANNOT database and tutorial are available as Fig. S1 and S2 in the supplemental material. We plan

to update the ARG-ANNOT database every month. A Web interface (<http://www.mediterranee-infection.com/article.php?laref=282&titer=arg-annot>) has been developed in which different links are available for access to Bio-Edit via download, tutorials to create a local database and perform BLAST analysis, a screen shot to help to create databases, BLAST and post-BLAST analyses, and updated database sequences. We welcome submissions of any new AR gene or genes not included in the database from any source; thus, a submission form is provided as Fig. S8 in the supplemental material.

ACKNOWLEDGMENTS

We thank Infectiopole Sud and IHU Méditerranée Infection for financial support. We thank American Journal Experts for English corrections.

We declare that we have no conflicts of interest.

REFERENCES

- Didelot X, Bowden R, Wilson DJ, Peto TE, Crook DW. 2012. Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* 13:601–612. <http://dx.doi.org/10.1038/nrg3226>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27:4636–4641. <http://dx.doi.org/10.1093/nar/27.23.4636>.
- Markowitz VM, Szeto E, Palaniappan K, Grechkin Y, Chu K, Chen IM, Dubchak I, Anderson I, Lykidis A, Mavromatis K, Ivanova NN, Kyrpides NC. 2008. The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucleic Acids Res.* 36(Database issue):D528–D533. <http://dx.doi.org/10.1093/nar/gkm846>.
- Bhullar K, Waglechner N, Pawlowski A, Koteva K, Banks ED, Johnston MD, Barton HA, Wright GD. 2012. Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* 7:e34953. <http://dx.doi.org/10.1371/journal.pone.0034953>.
- D'Costa VM, King CE, Kalan L, Morar M, Sung WW, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD. 2011. Antibiotic resistance is ancient. *Nature* 477:457–461. <http://dx.doi.org/10.1038/nature10388>.
- Scaria J, Chandramouli U, Verma SK. 2005. Antibiotic Resistance Genes Online (ARGO): a database on vancomycin and beta-lactam resistance genes. *Bioinformatics* 15:5–7. <http://dx.doi.org/10.6026/97320630001005>.
- 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(Database issue):D391–D394. <http://dx.doi.org/10.1093/nar/gkl791>.
- Liu B, Pop M. 2009. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res.* 37(Database issue):D443–D447. <http://dx.doi.org/10.1093/nar/gkn656>.
- 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. <http://dx.doi.org/10.1093/jac/dks261>.
- McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De PG, 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. <http://dx.doi.org/10.1128/AAC.00419-13>.
- Biswas S, Raoult D, Rolain JM. 2007. Molecular mechanisms of resistance to antibiotics in *Bartonella bacilliformis*. *J. Antimicrob. Chemother.* 59:1065–1070. <http://dx.doi.org/10.1093/jac/dkm105>.
- Doktor SZ, Shortridge VD, Beyer JM, Flamm RK. 2004. Epidemiology of macrolide and/or lincosamide resistant *Streptococcus pneumoniae* clin-

- ical isolates with ribosomal mutations. *Diagn. Microbiol. Infect. Dis.* 49: 47–52. [http://dx.doi.org/10.1016/S0732-8893\(03\)00130-5](http://dx.doi.org/10.1016/S0732-8893(03)00130-5).
14. Feuerriegel S, Oberhauser B, George AG, Dafaie F, Richter E, Rusch-Gerdes S, Niemann S. 2012. Sequence analysis for detection of first-line drug resistance in *Mycobacterium tuberculosis* strains from a high-incidence setting. *BMC Microbiol.* 12:90. <http://dx.doi.org/10.1186/1471-2180-12-90>.
 15. Froelich JM, Tran K, Wall D. 2006. A *pmrA* constitutive mutant sensitizes *Escherichia coli* to deoxycholic acid. *J. Bacteriol.* 188:1180–1183. <http://dx.doi.org/10.1128/JB.188.3.1180-1183.2006>.
 16. Fujimoto-Nakamura M, Ito H, Oyamada Y, Nishino T, Yamagishi J. 2005. Accumulation of mutations in both *gyrB* and *parE* genes is associated with high-level resistance to novobiocin in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 49:3810–3815. <http://dx.doi.org/10.1128/AAC.49.9.3810-3815.2005>.
 17. Nessar R, Reytrat JM, Murray A, Gicquel B. 2011. Genetic analysis of new 16S rRNA mutations conferring aminoglycoside resistance in *Mycobacterium abscessus*. *J. Antimicrob. Chemother.* 66:1719–1724. <http://dx.doi.org/10.1093/jac/dkr209>.
 18. Ramirez MS, Tolmasky ME. 2010. Aminoglycoside modifying enzymes. *Drug Resist. Updat.* 13:151–171. <http://dx.doi.org/10.1016/j.drup.2010.08.003>.
 19. Shi D, Li L, Zhao Y, Jia Q, Li H, Coulter C, Jin Q, Zhu G. 2011. Characteristics of *embB* mutations in multidrug-resistant *Mycobacterium tuberculosis* isolates in Henan, China. *J. Antimicrob. Chemother.* 66: 2240–2247. <http://dx.doi.org/10.1093/jac/dkr284>.
 20. Zou LK, Wang HN, Zeng B, Zhang AY, Li JN, Li XT, Tian GB, Wei K, Zhou YS, Xu CW, Yang ZR. 2011. Phenotypic and genotypic characterization of beta-lactam resistance in *Klebsiella pneumoniae* isolated from swine. *Vet. Microbiol.* 149:139–146. <http://dx.doi.org/10.1016/j.vetmic.2010.09.030>.
 21. Ruiz J, Moreno A, Jimenez de Anta MT, Vila J. 2005. A double mutation in the *gyrA* gene is necessary to produce high levels of resistance to moxifloxacin in *Campylobacter* spp. clinical isolates. *Int. J. Antimicrob. Agents* 25:542–545. <http://dx.doi.org/10.1016/j.ijantimicag.2004.10.016>.
 22. Hall T. 24 September 2013. BioEdit: biological sequence alignment editor for Win95/98/NT/2K/XP, p 95–98. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>.
 23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
 24. Galimand M, Courvalin P, Lambert T. 2012. *rmtF*, a new member of the aminoglycoside resistance 16S rRNA N7 G1405 methyltransferase family. *Antimicrob. Agents Chemother.* 56:3960–3962. <http://dx.doi.org/10.1128/AAC.00660-12>.
 25. Bilgin N, Richter AA, Ehrenberg M, Dahlberg AE, Kurland CG. 1990. Ribosomal RNA and protein mutants resistant to spectinomycin. *EMBO J.* 9:735–739.
 26. Herrera CM, Hankins JV, Trent MS. 2010. Activation of PmrA inhibits *lpxT*-dependent phosphorylation of lipid A promoting resistance to antimicrobial peptides. *Mol. Microbiol.* 76:1444–1460. <http://dx.doi.org/10.1111/j.1365-2958.2010.07150.x>.
 27. Moskowitz SM, Ernst RK, Miller SI. 2004. *PmrAB*, a two-component regulatory system of *Pseudomonas aeruginosa* that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A. *J. Bacteriol.* 186:575–579. <http://dx.doi.org/10.1128/JB.186.2.575-579.2004>.
 28. van Hoek AH, Mevius D, Guerra B, Mullany P, Roberts AP, Aarts HJ. 2011. Acquired antibiotic resistance genes: an overview. *Front. Microbiol.* 2:203. <http://dx.doi.org/10.3389/fmicb.2011.00203>.
 29. Tan SY, Chua SL, Liu Y, Hoiby N, Andersen LP, Givskov M, Song Z, Yang L. 2013. Comparative genomic analysis of rapid evolution of an extreme-drug-resistant *Acinetobacter baumannii* clone. *Genome Biol. Evol.* 5:807–818. <http://dx.doi.org/10.1093/gbe/evt047>.
 30. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agero 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. <http://dx.doi.org/10.1093/jac/dks496>.