ResFinder 4.0 for predictions of phenotypes from genotypes

Valeria Bortolaia†, Rolf S. Kaas †, Etienne Ruppe2, Marilyn C. Roberts3, Stefan Schwarz4, Vincent Cattoir5,6,7, Alain Philippin8, Rosa L. Allesoe1,9, Ana Rita Rebelo1, Alfred Ferrer Florensa1,10,11,12, Trinad Chakraborty10,11,13, Bernd Neumann13, Guido Werner13, Jennifer K. Bender13, Kerstin Stingl14, Minh Nguyen15, Jasmine Coppens15, Basil Britto Xavier15, Surbhi Malhotra-Kumar15, Henrik Westh16,17, Mette Pinholt16, Muna F. Anjum18, Nicholas A. Duggett18, Isabelle Kempf19, Suvi Nykäsena20, Satu Olkkola20, Kinga Wieczorek21, Ana Amaro22, Lurdes Clemente22, Joël Mossong23, Serge Losch24, Catherine Ragimbeau23, Ole Lund1 and Frank M. Aarestrup1,1*1

1 Technical University of Denmark, National Food Institute, European Union Reference Laboratory for Antimicrobial Resistance, WHO Collaborating Centre for Antimicrobial Resistance in Foodborne Pathogens and Genomics, FAO Reference Laboratory for Antimicrobial Resistance, Kgs. Lyngby, Denmark; 2 Université de Paris, IAME, INSERM, Paris, France; 3 Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, USA; 4 Institute of Microbiology and Epizootics, Centre for Infection Medicine, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany; 5 Rennes University Hospital, Department of Clinical Microbiology, Rennes, France; 6 National Reference Center for Antimicrobial Resistance (lab Enterococci), Rennes, France; 7 University of Rennes 1, INSERM U1230, Rennes, France; 8 Faculty of Medicine Paris Descartes, Bacteriology, Paris, France; 9 Institute of Medical Microbiology, Justus Liebig University Giessen, Giessen, Germany; 10 Institute for Infection Research, site Giessen-Marburg-Langen, Justus Liebig University Giessen, Giessen, Germany; 11 Institute of Hygiene and Environmental Medicine, Justus Liebig University Giessen, Giessen, Germany; 12 Robert Koch Institute, Wernigerode Branch, Department of Infectious Diseases, Division of Nosocomial Pathogens and Antibiotic Resistances, Wernigerode, Germany; 13 German Federal Institute for Risk Assessment, Department of Biological Safety, National Reference Laboratory for Campylobacter, Berlin, Germany; 14 Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp, Belgium; 15 Department of Clinical Microbiology, Hvidovre University Hospital, Hvidovre, Denmark; 16 Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; 17 Finnish Food Authority, Helsinki, Finland; 18 National Veterinary Research Institute, Pulawy, Poland; 19 National Institute of Agrarian and Veterinary Research (INIAV), National Reference Laboratory for Animal Health, Oeiras, Portugal; 20 Laboratoire de médecine vétérinaire de l’Etat, Veterinary Services Administration, Dudelange, Luxembourg; 21 Laboratory of Medical Microbiology, Justus Liebig University Giessen, Giessen, Germany; 22 Laboratory of Medical Microbiology, University of Copenhagen, Copenhagen, Denmark; 23 Laboratory of Medical Microbiology, Rennes University Hospital, Rennes, France; 24 Laboratoire de médecine vétérinaire de l’Etat, Veterinary Services Administration, Dudelange, Luxembourg

* Corresponding author. E-mail: fmaa@food.dtu.dk
† These authors contributed equally.

Received 17 April 2020; accepted 30 June 2020

Objectives: WGS-based antimicrobial susceptibility testing (AST) is as reliable as phenotypic AST for several antimicrobial/bacterial species combinations. However, routine use of WGS-based AST is hindered by the need for bioinformatics skills and knowledge of antimicrobial resistance (AMR) determinants to operate the vast majority of tools developed to date. By leveraging on ResFinder and PointFinder, two freely accessible tools that can also assist users without bioinformatics skills, we aimed at increasing their speed and providing an easily interpretable antibiogram as output.

Methods: The ResFinder code was re-written to process raw reads and use Kmer-based alignment. The existing ResFinder and PointFinder databases were revised and expanded. Additional databases were developed including a genotype-to-phenotype key associating each AMR determinant with a phenotype at the antimicrobial compound level, and species-specific panels for in silico antibiograms. ResFinder 4.0 was validated using Escherichia coli (n = 584), Salmonella spp. (n = 1081), Campylobacter jejuni (n = 239), Enterococcus faecium (n = 106), Enterococcus faecalis (n = 50) and Staphylococcus aureus (n = 163) exhibiting different AST profiles, and from different human and animal sources and geographical origins.
Results: Genotype-phenotype concordance was $\geq 95\%$ for 46/51 and 25/32 of the antimicrobial/species combinations evaluated for Gram-negative and Gram-positive bacteria, respectively. When genotype-phenotype concordance was $< 95\%$, discrepancies were mainly linked to criteria for interpretation of phenotypic tests and suboptimal sequence quality, and not to ResFinder 4.0 performance.

Conclusions: WGS-based AST using ResFinder 4.0 provides in silico antibiograms as reliable as those obtained by phenotypic AST at least for the bacterial species/antimicrobial agents of major public health relevance considered.

Introduction

Antimicrobial susceptibility testing (AST) is the cornerstone of appropriate clinical use of antimicrobial agents and for surveillance programmes aiming at estimating the occurrence of antimicrobial resistance (AMR). Currently, in vitro phenotypic AST methods such as broth microdilution (BMD) and disc diffusion, which are based on principles laid down in the work by Alexander Fleming in the 1920s, are considered the gold standard for measuring AMR in bacteria. Despite being conceptually simple, phenotypic methods suffer from limitations that hinder reproducibility of results even when following international standards. These reproducibility issues range from the operator executing the test and reading the results, to parameters affecting bacterial growth and stability of reagents. In particular, the need for tightly controlled environmental conditions is a limiting factor for reliable AST in low- and middle-income countries where AMR claims the highest social and economic toll at present. An additional limitation of phenotypic AST is caused by incomplete agreements on panels of antimicrobial agents to test and on interpretive criteria by different organizations. It has also been argued that the laboratory conditions used for AST are poor predictors of the clinical efficacy for certain antimicrobial agent–bacterium combinations as a consequence of different bacterial growth dynamics in infection sites compared with in vitro conditions, inducible resistance genes, as well as silent genes that may revert during clinical antimicrobial treatment, among other factors. Furthermore, phenotypes can poorly inform conclusions on the epidemiology of resistance genes and bacteria, since phenotypic tests for AST have limited discriminatory power.

Genotypic approaches have been proposed as a valid alternative to phenotypic AST since the early 1990s. Although genotypic tests have been and are widely used for detection of AMR genes, their limited overall sensitivity hindered their application as reference AST methods. With increased accessibility and decreasing cost of next-generation sequencing (NGS), WGS has become an available methodology for routine characterization of bacteria. NGS-based methods outperform other genotypic approaches for detection of AMR since they allow detection of virtually any known AMR gene/mutation and identification of new variants of known AMR determinants. Furthermore, sequence data can be stored indefinitely and can be re-analysed when new AMR determinants are discovered phenotypically. The main caveat related to any genotypic AST method is represented by the fact that only known AMR mechanisms can be detected, whereas resistance caused by new mechanisms and/or modulation of gene expression (increased expression of efflux pumps, heteroresistance, etc.) might be overlooked.

In the last decade, at least 47 open-access bioinformatics pipelines for detecting AMR genes in NGS data have been published (reviewed in Hendriksen et al.) and continue to be developed, but they mostly lack easy and rapidly interpretable outputs including translation of genotypes into predicted phenotypes. In 2012, we published ResFinder, the first online bioinformatics tool aimed at users without specialized bioinformatic skills, which provides detection of AMR genes in WGS data submitted through a web server. The ResFinder tool has been widely used, and has so far executed more than 400,000 jobs from more than 32,000 IP addresses in over 100 countries. ResFinder was expanded with PointFinder, a tool that detects chromosomal point mutations mediating resistance to selected antimicrobial agents in a few selected bacterial species and recently extended to include Mycobacterium tuberculosis.

Here we describe the development and evaluation of a new and more advanced tool, ResFinder 4.0, which, in addition to the detection of AMR genes and chromosomal gene mutations, generates in silico antibiograms. ResFinder 4.0 was developed by improving and expanding the database of ResFinder and PointFinder and by rewriting the software. ResFinder 4.0 was validated using bacterial isolates of different species and genera from food, animal and human sources.

Methods

ResFinder 4.0 databases

ResFinder 4.0 contains four databases including AMR genes (ResFinder), chromosomal gene mutations mediating AMR (PointFinder), translation of genotypes into phenotypes and species-specific panels for in silico antibiograms. The databases of ResFinder and PointFinder were reviewed by experts and, when necessary, entries were removed or added. Furthermore, the PointFinder database was extended to include chromosomal gene mutations leading to ampicillin resistance in Enterococcus faecium, ciprofloxacin resistance in E. faecium and Enterococcus faecalis, and resistance to cefoxitin, chloramphenicol, ciprofloxacin, fusidic acid, linezolid, mupirocin, quinupristin–dalfopristin, rifampicin and trimethoprim in Staphylococcus aureus. The genotype-to-phenotype tables were created by experts, by using additional databases (www.bldb.eu for β-lactam resistance genes). The genotype-to-phenotype tables, the ResFinder and PointFinder entries have been associated with an AMR phenotype both at the antimicrobial class and at the antimicrobial compound level. A selection of antimicrobial compounds within each class was made to include antimicrobial agents important for clinical and surveillance purposes for the different bacterial species included (Table S1, available as Supplementary data at JAC Online).

The genotype-to-phenotype tables also include: (i) the PubMed ID of relevant literature describing the respective AMR determinants and...
phenotypes, when available; (ii) the mechanism of resistance by which each AMR determinant functions; and (iii) notes, which may contain different information such as warnings on variable expression levels (inducible resistance, cryptic genes in some species, etc.), structural and functional information, and alternative nomenclature.

Software and interface

ResFinder 4.0 was embedded using the same web interface as previous ResFinder versions and available at the link https://cge.cbs.dtu.dk/services/ResFinder-4.0/. Importantly, in the interface, the user is prompted to specify a bacterial species, which is needed to define the specific antimicrobial panel for the in silico antibiogram (Table S2). There is the option to include all antimicrobial agents from all panels (‘Other’ option). In this case, interpretation of results must be executed carefully and knowledge on intrinsic resistance is essential because, in the ‘Other’ option, isolates intrinsically resistant to an antimicrobial agent might appear predicted as susceptible since intrinsic resistance is often mediated by structural traits (e.g. reduced permeability of the outer membrane, among others) rather than by specific genes/mutations.19

Previous versions of ResFinder were written in Perl, whereas ResFinder 4.0 was rewritten in Python 3. The ResFinder software has not previously been able to process read data (FASTQ) directly but relied on an assembly step. ResFinder 4.0 has implemented KMA,20 which aligns reads directly to the databases without the need for assembly. Like all previous versions, ResFinder 4.0 is released as open source under the Apache 2.0 license and is available at: https://bitbucket.org/genomicepidemiology/resfinder/.

Datasets for validation

ResFinder 4.0 was validated with datasets consisting of MIC values (BMD or Etest, Table 1) and WGS data (Illumina sequencing) of Escherichia coli, Salmonella spp., Campylobacter jejuni, E. faecium, E. faecalis and S. aureus of different origins (Table 1). These datasets represent a convenience sample. Phenotypic AST results were interpreted using the EUCAST epidemiological cut-off values (ECOFFs) to categorize isolates as WT (MIC ≤ ECOFF) and non-WT (MIC > ECOFF) (www.eucast.org). Exceptions were: (i) one S. aureus dataset for which phenotypic AST was performed by disc diffusion and interpreted by EUCAST clinical breakpoints (Table 1); and (ii) one E. coli dataset that consisted of Illumina WGS data only and MIC values were available for the data provider but not for the ResFinder 4.0 developers, thus providing a blind test of the tool performance (Table 1). WGS data were obtained as raw reads and processed through a quality control (QC) pipeline as described here: https://bitbucket.org/genomicepidemiology/foadapp pipeline. In brief, reads were trimmed using bbduk2 (https://jgi.doe.gov/data-and-tools/bbtools/) to a phred score of 20, reads less than 50 bp were discarded, adapters were trimmed away and a draft de novo assembly was created using SPAdes.21 From the assemblies, contigs below 500 bp were discarded. The most important parameters that were used to assess quality of sequencing data were: number of bases left after trimming, N50, number of contigs and total size of assembly. QC parameters used as guidelines were: read depth of at least 25×, N50 of >30 000 bp and a limit on the number of contigs to <500.

WGS data (FASTQ) were used as input for ResFinder 4.0 using default parameters (≥80% identity over ≥50% of the length of the target gene) and also for SNP-based phylogenetic analysis as previously described23 to verify the genetic diversity of the validation datasets. SNP analysis was not performed for the Salmonella spp. dataset whose diversity was already described previously.23 The ResFinder 4.0 output was analysed to define AMR genotypes, i.e. patterns of resistance determinants observed for each antimicrobial, in each dataset.

Genotype–phenotype concordance was defined as presence or absence of a genetic determinant of resistance to a specific antimicrobial agent in non-WT (nWT) or WT isolates, respectively. Genotype–phenotype discordance was defined either as presence of a relevant AMR determinant in WT isolates or as absence of a relevant AMR determinant in nWT isolates. All discordances were individually analysed.

Sequence data that did not derive from previous studies (Table 1) have been deposited at NCBI (E. coli dataset from Germany: PRJNA616452; E. faecium dataset from Germany: PRJNA625631; E. faecium dataset from Belgium: PRJNA552025; S. aureus dataset from Belgium: PRJNA615176) and in the European Nucleotide Archive (S. aureus dataset from Denmark: PRJEB37586).

Table 1. Datasets for ResFinder 4.0 validation

| Species         | Isolates (n) | Observations (n) | Source          | Origin      | Country | Reference                  |
|-----------------|--------------|------------------|-----------------|-------------|---------|----------------------------|
| E. coli         | 95           | 1520             | animal, food    | surveillance| DK      | Hendriksen et al. 33, 44   |
| E. coli         | 99           | 890              | animal, food    | surveillance| UK      | Duggett et al. 44, 48      |
| E. coli<sup>2</sup> | 390          | 2559             | human, animal   | clinical, surveillance | DE | AbuOun et al. 23         |
| Salmonella enterica | 1081         | 7489             | animal, food    | surveillance| USA | McDermott et al. 23       |
| C. jejuni       | 239          | 1382             | animal, food    | surveillance| FIN, FR, DE, LU, PL, PT | Leekitchareonphon et al. 45 |
| E. faecium<sup>b</sup> | 50           | 363              | human           | clinical    | DE      | this study                 |
| E. faecium<sup>b</sup> | 56           | 159              | human           | clinical    | BE      | this study                 |
| E. faecalis     | 50           | 235              | human and animal| clinical, surveillance| DE | Neumann et al. 46, Bender et al. 47 |
| S. aureus<sup>c</sup> | 63           | 504              | human           | clinical, surveillance| BE | this study                 |
| S. aureus<sup>c</sup> | 100          | 598              | human           | clinical, surveillance| DK | this study                 |

If not otherwise specified, phenotypic AST results were obtained by BMD.

DK, Denmark; DE, Germany; FIN, Finland; FR, France; LU, Luxembourg; PL, Poland; PT, Portugal; BE, Belgium.

<sup>2</sup>Dataset for blind test of ResFinder 4.0 performance. Cefepime, chloramphenicol, ertapenem and nalidixic acid susceptibility testing were performed by Etest in a subset of isolates. All remaining AST were performed by BMD.

<sup>b</sup>Phenotypic AST results were obtained by Etest.

<sup>c</sup>Phenotypic AST results were obtained by disc diffusion.
Results

ResFinder 4.0 databases

The ResFinder 4.0 databases validated in this study include 2690 AMR genes (ResFinder) and 266 resistance-mediating mutations in 39 selected genes (PointFinder), a genotype-to-phenotype translation database that includes 57 antimicrobial compounds (Table S1) and in silico antimicrobial panels for eight species (Table S2). Of note, a few antimicrobial agents included in the genotype-to-phenotype translation database do not appear in the in silico antimicrobial panels. Thus the database can be used for any bacterial species and is not restricted to the species having assigned in silico antimicrobial panels. Furthermore, the in silico antimicrobial panels are limited to the antimicrobial agents for which the tool can actually provide an output. Hence, for example, antimicrobial agents for which the genetic bases of resistance have not been fully elucidated (e.g. daptomycin in Enterococcus spp., among others) are not included. The most updated versions of the databases are publicly available in BitBucket (https://bitbucket.org/genomicepidemiology/resfinder_db and https://bitbucket.org/genomicepidemiology/pointfinder_db). The databases are continuously updated relying both on users’ feedback and on curators'/developers’ own research. Importantly, the BitBucket repository allows access to outdated versions of the databases, which is particularly useful in case ResFinder 4.0 is used within quality-assured procedures.

Datasets for ResFinder 4.0 validation

The SNP distances across strains used for ResFinder 4.0 validation were in the range 1–48 102 for E. coli from Denmark and the UK, 1–48 066 for E. coli from Germany, 1–20 134 for C. jejuni, 1–7 131 for E. faecium 15–17 067 for E. faecalis and 1–25 424 for S. aureus (Table S3). The proportion of WT and nWT isolates in each dataset is presented in Tables 2 and 3.

Concordance between genotypic and phenotypic AST

E. coli

For the dataset deriving from Danish AMR surveillance in animals and meat, the overall genotype–phenotype concordance resulting from 1520 observations encompassing 16 antimicrobial agents was 97%, ranging from 71.6% for cefepime to 100% for most antimicrobial agents (Table 2).

ResFinder 4.0 detected between 2 and 14 genotypes of resistance to most antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 99.6% (586/588) of cases in which AMR determinants were detected and in 95.3% (889/932) of cases in which no resistance determinant was detected. Discordances were caused by detection of sul2 in an isolate with the sulfamethoxazole MIC well below the ECOFF, and blaCTX-M-1 in an isolate with the ceftazidime MIC on the ECOFF (Figure 1). Both genes were detected with a reliable read depth (see the

Table 2. Antimicrobial resistance phenotypes and genotype–phenotype concordance for the Gram-negative bacteria datasets using ECOFFs

| Antimicrobial          | E. coli (animal surveillance; DK) | E. coli (animal surveillance; UK) | E. coli (human and animal clinical and surveillance; DE) | Salmonella sp. (human clinical, animal surveillance; USA) | C. jejuni (animal surveillance; EU) |
|------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
|                        | nWT    | concordance (%) | nWT    | concordance (%) | nWT    | concordance (%) | nWT    | concordance (%) | nWT    | concordance (%) |
| Ampicillin             | 95     | 0               | 81     | 18               | 98.9  | 202              | 0      | 100              | 249    | 822              | 98.7   | –                 | –     | –                 |
| Cefepime               | 78     | 17              | 71.6   | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |
| Cefotaxime             | 95     | 0               | 100    | 23               | 76    | 98.9             | 370    | 0                | 98.6   | –                | –      | –                 | –     | –                 |
| Cefoxitin              | 46     | 49              | 97.8   | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |
| Ceftazidime            | 94     | 1               | 98.9   | –                | –     | –                | 282    | 0                | 99.2   | –                | –      | –                 | –     | –                 |
| Chloramphenicol        | 8      | 87              | 100    | –                | –     | –                | 63     | 67               | 73.1   | 41                | 1030   | 99.7              | –     | –                 |
| Ciprofloxacin          | 29     | 66              | 87.3   | 64               | 35    | 95.9             | 275    | 0                | 99.2   | 22                | 1049   | 97                | 134   | 105               | 99.1  |
| Colistin               | 0      | 95              | 100    | –                | –     | –                | 11     | 88                | 100    | –                | –      | –                 | –     | –                 |
| Ertapenem              | 1      | 94              | 98.9   | –                | –     | –                | 60     | 70               | 54.6   | –                | –      | –                 | –     | –                 |
| Erythromycin           | –      | –               | –      | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |
| Gentamicin             | 15     | 80              | 100    | 34               | 65    | 100              | 129    | 258              | 97.6   | 126               | 945    | 98.9              | 0     | 239               | 100   |
| Imipenem               | 0      | 95              | 100    | 0                | 98    | 100              | 2      | 192              | 100    | –                | –      | –                 | –     | –                 |
| Meropenem              | 0      | 95              | 100    | 0                | 99    | 100              | 2      | 0                | 100    | –                | –      | –                 | –     | –                 |
| Nalidixic acid         | 25     | 70              | 98.9   | 39               | 60    | 90.9             | 99     | 28               | 99.2   | 10                | 1061   | 99.4              | 131   | 108               | 97.9  |
| Streptomycin           | –      | –               | –      | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |
| Sulfamethoxazole       | 61     | 34              | 98.9   | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |
| Tetracycline           | 53     | 42              | 100    | 79               | 20    | 97.9             | 138    | 33               | 98.8   | 656               | 415    | 98.6              | 130   | 109               | 99.1  |
| Tigecycline            | –      | –               | –      | –                | –     | –                | 5      | 147              | 96.7   | –                | –      | –                 | –     | –                 |
| Trimethoprim           | 29     | 66              | 100    | –                | –     | –                | –      | –                | –      | –                | –      | –                 | –     | –                 |

DK, Denmark; DE, Germany.
Table 3. Antimicrobial resistance phenotypes and genotype–phenotype concordance for the Gram-positive bacteria datasets using ECOFFs

| Antimicrobial      | E. faecium (human clinical; DE) | E. faecium (human clinical; BE) | E. faecalis (human clinical; DE) | S. aureus (human clinical and surveillance; BE) | S. aureus (human clinical and surveillance; DK) |
|--------------------|---------------------------------|---------------------------------|---------------------------------|-----------------------------------------------|-----------------------------------------------|
|                    | nWT | concordance (%) | nWT | concordance (%) | nWT | concordance (%) | nWT | concordance (%) | R   | S   | concordance (%) | R   | S   | concordance (%) |
| Ampicillin         | 50  | 0               | 100 |                | 55  | 1               | 100 |                | –   | –   | –               | –   | –   | –               |
| Cefoxitin          | –   | –               | –   |                | –   | –               | –   |                | –   | –   | –               | –   | –   | –               |
| Chloramphenicol    | 0   | 50              | 64  |                | –   | –               | –   |                | 63  | 0   | 100             | 99  | 1   | 100             |
| Ciprofloxacin      | 50  | 0               | 100 |                | 3   | 1               | 100 |                | –   | –   | –               | –   | –   | –               |
| Clindamycin        | –   | –               | –   |                | –   | –               | –   |                | –   | –   | –               | –   | –   | –               |
| Erythromycin       | 50  | 0               | 100 |                | –   | –               | –   |                | 39  | 11  | 96.8            | 36  | 64  | 97              |
| Gentamicin         | 13  | 0               | 100 |                | 6   | 2               | 75  |                | 31  | 4   | 97.1            | 63  | 0   | 100             |
| Linezolid          | 2   | 48              | 92  |                | 2   | 33              | 94.2|                | 16  | 34  | 96.2            | 0   | 63  | 100             |
| Tetracycline       | 22  | 28              | 92  |                | 2   | 33              | 94.2|                | 43  | 7   | 98              | 46  | 17  | 76.2            |
| Vancomycin         | 40  | 10              | 100 |                | 53  | 3               | 96.4|                | 16  | 34  | 98              | 0   | 63  | 100             |

For the S. aureus DK dataset only, R, resistant and S, susceptible interpretations according to EUCAST clinical breakpoints were available. DK, Denmark; DE, Germany; BE, Belgium.

Figure 1. Discordance between predicted (ResFinder 4.0) and observed phenotypes. The vertical dotted line divides the isolates having ‘WT phenotype with AMR determinant’ to the left and the isolates having ‘nWT phenotype without AMR determinant’ to the right. ‘Low depth’ refers to low read depth of the respective AMR determinant as explained in the Discussion. AMP, ampicillin; FEP, cefepime; CTX, cefotaxime; FOX, cefoxitin; CAZ, ceftazidime; CIP, ciprofloxacin; ETP, ertapenem; NAL, nalidixic acid; SMX, sulfamethoxazole; TET, tetracycline; ERY, erythromycin; GEN, gentamicin; LZD, linezolid; VAN, vancomycin; CHL, chloramphenicol; CLI, clindamycin.
Bortolaia et al.

Discussion. Additional discords were caused by lack of detection of AMR determinants in isolates with nWT phenotype for cefepime, cefoxitin, ciprofloxacin, ertapenem and nalidixic acid. Most of these isolates had the respective MIC one step dilution above the ECOFF (Figure 1 and Table S5).

For the dataset deriving from AMR surveillance in animals in the UK, the overall genotype–phenotype concordance for 890 observations encompassing nine antimicrobial agents was 98%, ranging from 90.9% for nalidixic acid to 100% for colistin, gentamicin, imipenem and meropenem (Table 2). ResFinder 4.0 detected between 2 and 17 genotypes of resistance to most antimicrobial agents, in addition to several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 96.7% (325/336) of cases in which AMR determinants were detected and in 98.9% (548/554) of cases in which no resistance determinant was detected. Discordances were caused by detection of: (i) (fluoro)quinolone resistance determinants (gyrA and parC mutations and qnrB1) in eight (fluoro)quinolone WT isolates for which only interpretation of the MIC was available (Table S5); (ii) blaTEM-106 and truncated tet(M) genes in isolates with WT cephalaxine and tetracycline phenotype, respectively, and with very low read depth (see the Discussion) of these AMR genes; and (iii) detection of the tet(C) gene with reliable read depth in a tetracycline WT isolate with the MIC on the breakpoint (Figure 1). Additional discordances were caused by lack of detection of AMR determinants in isolates with nWT phenotype for ampicillin, ciprofloxacin and nalidixic acid (Figure 1). Most (60%) of the (fluoro)quinolone nWT isolates had MICs one step dilution above the ECOFF (Figure 1 and Table S5).

For the dataset used for the blind test that comprised human and animal origin isolates from clinical and surveillance samples, the overall genotype–phenotype concordance resulting from 2559 observations encompassing 13 antimicrobial agents was 95.3%, ranging from 54.6% for ertapenem to 100% for ampicillin, cefepime, imipenem and meropenem (Table 2). ResFinder 4.0 detected between 2 and 17 genotypes of resistance to most antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 99.9% of cases in which AMR determinants were detected and in 98.9% (548/554) of cases in which no resistance determinant was detected. Discordances were mainly due to detection of AMR determinants in isolates with nWT phenotype for ampicillin, ciprofloxacin and nalidixic acid (Table S5). For C. jejuni, the overall genotype–phenotype concordance resulting from 1382 observations encompassing six antimicrobial agents was 99.2%, ranging from 97.9% for nalidixic acid to 100% for gentamicin and streptomycin (Table 2).

ResFinder 4.0 detected between one and eight genotypes for resistance to most antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 99.9% of cases in which AMR determinants were detected and in 99.3% of cases in which no resistance determinant was detected. Discordances were observed in isolates with (i) gyrA mutations known to mediate (fluoro)quinolone resistance and the nalidixic acid MIC close to the ECOFF, and (ii) tet(O) with <100% identity to the ResFinder 4.0 entries and reliable read depth in an isolate with the MIC close to the breakpoint. Additional discordances were caused by lack of detection of AMR determinants in isolates with nWT phenotype for ciprofloxacin, nalidixic acid, erythromycin and tetracycline with MICs notably higher than the ECOFF (Figure 1 and Table S5).

E. faecium

Of the two E. faecium datasets, the one from Germany included 363 observations encompassing eight antimicrobial agents and the one from Belgium included 159 observations encompassing five antimicrobial agents. In the first dataset, the overall genotype–phenotype concordance was 92.8%, ranging from 64% for chloramphenicol to 100% for ampicillin, ciprofloxacin, erythromycin, gentamicin and vancomycin (Table 3). In this dataset, ResFinder 4.0 detected between 2 and 17 genotypes of resistance to all antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 91.0% of cases in which AMR determinants were detected and in 96.6% of cases in which no resistance determinant was detected. Most discordances were due to detection of paxtA, cat and fexB with <100% identity to the database entries and often with low read depth (Figure 1). Furthermore, tet(M) was detected with reliable read depth in an isolate with the tetracycline MIC on the ECOFF (Figure 1). Additional discordances were due to lack of detection of tetracycline and linezolid resistance determinants in isolates with the respective MICs 1- and 3-fold dilution above the ECOFF, respectively (Figure 1 and Table S5).

In the second dataset, the overall genotype–phenotype concordance was 96.2%, ranging from 75% for gentamicin to 100% for ampicillin and ciprofloxacin (Table 3). In this dataset, ResFinder 4.0 detected between two and six genotypes for resistance to all
antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 96.6% (117/121) of cases in which AMR determinants were detected and in 94.7% (36/38) of cases in which no resistance determinant was detected. Discordances were due to detection of (i) gentamicin resistance genes in gentamicin WT isolates (n = 2), but further analysis revealed that such genes had low read depth, and (ii) vancomycin resistance genes with reliable coverage in isolates (n = 2) with different vancomycin MICs (Figure 1). Additional discordances were caused by lack of detection of AMR determinants in isolates with the linezolid MIC one step dilution above the ECOFF (Figure 1 and Table S5).

E. faecalis
For the E. faecalis dataset, the overall genotype–phenotype concordance resulting from 235 observations encompassing five antimicrobial agents was 97%, ranging from 96% for erythromycin and linezolid to 98% for tetracycline and vancomycin (Table 3).

ResFinder 4.0 detected between 1 and 17 genotypes of resistance to all antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 98.6% (140/142) of cases in which AMR determinants were detected and in 94.6% (88/93) of cases in which no resistance determinant was detected. Discordances were represented by isolates having erm(B) (n = 1) and tet(M) (n = 1) with 100% and <100% identity to the ResFinder 4.0 database, respectively, and an MIC distant from the ECOFF (Figure 1). Further analysis revealed that tet(M) had very low read depth. Additional discordances were represented by lack of detection of determinants of resistance to erythromycin, gentamicin, linezolid and vancomycin in a few isolates with clear nWT phenotype (Figure 1 and Table S5).

S. aureus
Of the two S. aureus datasets, the one from Belgium included 504 observations encompassing eight antimicrobial agents, whereas the one from Denmark included 598 observations encompassing six antimicrobial agents.

For the first dataset, the overall genotype–phenotype concordance was 96%, ranging from 76.2% for tetracycline to 100% for cefoxitin, ciprofloxacin, gentamicin, linezolid and vancomycin (Table 3). In this dataset, ResFinder 4.0 detected between one and two genotypes of resistance to most antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 95.1% (293/308) of cases in which AMR determinants were detected and in 97.4% (191/196) of cases in which no resistance determinant was detected. Most discordances were linked to detection of tet(M) with <100% identity to the database entry and in isolates with MICs on or below the ECOFF (Figure 1). Additional discordances were due to lack of detection of AMR determinants in isolates with clindamycin and erythromycin MICs notably higher than the ECOFF (Figure 1 and Table S5).

For the second dataset, the overall genotype–phenotype concordance was 97.3%, ranging from 93.9% for gentamicin to 100% for cefoxitin (Table 3). In this dataset, ResFinder 4.0 detected between one and nine genotypes of resistance to most antimicrobial agents besides several cases in which no AMR determinant was detected (Table S4). The predicted and observed AMR phenotypes matched in 94.4% (205/217) of cases in which AMR determinants were detected and in 98.9% (379/383) of cases in which no resistance determinant was detected. Discordances were due to detection of erm(A) (n = 2) and aac(6)-aph(2) with reliable read depth and to detection of tet(K) (n = 3) with low read depth (Table S4). Additional discordances were due to lack of detection of clindamycin, gentamicin, linezolid and tetracycline resistance determinants in isolates that were reported as resistant (Table S5). As phenotypic data were available only as interpretation (according to EUCAST clinical breakpoints) of disc diffusion results, we could not assess if the measured values were close to the breakpoint for resistance.

Discussion
We developed and implemented a freely accessible web server, a database and bioinformatics pipeline for detecting AMR genes and mutations and providing in silico prediction of expected phenotypes. ResFinder 4.0 was validated for six bacterial species of major public health relevance and has been evaluated for M. tuberculosis previously. ResFinder 4.0 can also be used to predict phenotypes for any additional bacterial species by users with profound knowledge of AMR.

Excellent concordance between genotypically predicted and phenotypically detected phenotypes was found for most antimicrobial–species combinations (Tables 2 and 3). In the vast majority of cases in which the automatically evaluated genotype–phenotype concordance was <95%, we could identify issues related to ECOFFs and/or read depth of WGS data (Figure 1), indicating that phenotypic tests and suboptimal sequence quality, respectively, were likely sources of discrepancies. Misclassifications by the phenotypic tests could have happened in the cases in which the MIC was either on or bordered the ECOFF (i.e. ± one 2-fold dilution from the ECOFF). This is due to limitation in the reproducibility of MIC results, whereby one 2-fold dilution difference across repeated tests is considered acceptable. Furthermore, for certain antimicrobial–bacteria combinations including chloramphenicol–enterococci, MIC reading is difficult due to trailing growth, and we cannot exclude that this might have also contributed to the genotype–phenotype discordances. In fact, in a previous study, we found that most discrepancies between genotype and phenotype were resolved when repeating the phenotypic test. Unfortunately, the isolates used in this study were not available for repeating MIC determination. Also the ECOFFs themselves might be a source of misclassification of phenotypic tests. It has been shown that the number of isolates showing genotype–phenotype concordance notably increased when adjusting the ECOFF of certain antimicrobial agents. Low read depth was also a likely source of some of the genotype–phenotype discordances observed (Figure 1). There is no unequivocally accepted threshold for acceptable read depth for detection of AMR genes and mutations, and our definition of ‘low read depth’ derived from a case-by-case examination. Specifically, the genes scored as ‘low read depth’ had either ‘depth <10’ or ‘depth <1/10 compared with that of additional AMR genes in the same isolate and query coverage <100%’ (which indicates an imperfect match between the query and the subject). This shows the importance of visualizing
the ‘read depth’ parameter in the output. At present, such a parameter appears only in the output obtained when running ResFinder 4.0 in Unix but will also be included in the web tool output in the future. Of note, issues related to read depth depend on sequencing procedures and are not indicative of ResFinder 4.0 performance and/or feasibility of WGS-based AST. ResFinder 4.0 performed poorly in prediction of linezolid resistance. This was expected since the tool only detects linezolid resistance mediated by acquired genes, whereas resistance mediated by mutations in chromosomal genes is currently detected by LREFinder.27 This underlines the importance of becoming familiar with features of the tools used for AMR detection in WGS data to choose the most appropriate according to the user’s aims.3 Genotype–phenotype discordance might have also occurred due to inducible genes. For example, the S. aureus dataset from Denmark included erm(A)-positive isolates phenotypically classified as clindamycin susceptible (Tables S4 and S5). erm(A) can be inducible, and mutations in the translational attenuator can occur very rapidly, leading to phenotypic resistance to all lincosamides and ketolides.28 Also, occurrence of genes with <100% identity and/or coverage (in length) compared with the ResFinder 4.0 database entries might have caused some genotype–phenotype discordances since genes that underwent mutations, insertions or deletions might have become silent. In addition, there are genes for which an nWT phenotype is linked to specific sequences only and will not be expressed if a mutation appears. For example, only specific mutations in aac(6’)-Ib mediate fluoroquinolone resistance in E. coli and Salmonella,29 which explains why aac(6’)-Ib-cr with <100% identity to the database entry is detected in ciprofloxacin WT Salmonella spp. isolates (Table S5). Nonetheless, in our datasets, the prediction of a resistance phenotype matched the measured nWT phenotype even in most of the cases presenting an AMR gene with <100% identity and/or coverage (in length) compared with the database entries (Tables S4 and S5). From a surveillance perspective, detection of potentially silent genes is still more informative than the reporting of a WT phenotype. As silencing mutations can be transient,30 also from a clinical perspective the finding of AMR genes may be even more relevant than the observed in vitro phenotype as it may be used as an early warning to avoid initiating inappropriate therapies.

In conclusion, we implemented and evaluated ResFinder 4.0, a bioinformatics tool whose unique feature is to provide a phenotype for each antimicrobial resistance gene detected and easily interpretable in silico antibiograms for eight bacterial species including both Gram-positive and Gram-negative bacteria of major public
health relevance. The tool will be extended to allow for input of Nanopore sequence data and to include antibiograms of additional bacterial species. ResFinder 4.0 allows for rapid updates of all databases to promptly respond to users’ feedback and keep pace with improvements in knowledge on AMR. We would encourage users around the world to input additional curated databases for better and faster improvement.

Acknowledgements

We thank Katrinne Valentin Jensen for her skilful assistance in formatting the databases. We also thank Carola Fleige and Christine Gunther from the Robert Koch Institute (Germany), G. Mourand from Anses Ploufragan (France), and Edyta Denis, Iwona Kania and Magdalena Staniek from the National Veterinary Research Institute (Poland) for their excellent technical assistance in laboratory analyses.

Funding

This work was supported by the European Union Horizon 2020 research and innovation programme under grant agreement 643476 to the COMPARE project (http://www.compare-europe.eu), and The Novo Nordisk Foundation (NNF16OC0021856) - Global Surveillance of Antimicrobial Resistance. The work by L. Falgenhauer and T. Chakraborty was supported by grants to the German Center of Infection Research (DZIF), and the Zoonoses Network ‘ESBL and fluoroquinolone resistance in Enterobacteriaceae (RESET)’ Consortium through the German Federal Ministry of Education and Research (BMBF; grant numbers 8000 701-3 [HZI], 01KI1313G, T106.001, 8032808811, 8032808818 and 8032808820 to T.C.). Data for the E. faecium (dataset from Germany) and E. faecalis isolates were generated as part of a research project called InfectControl2020-IRMRESS funded by the German BMBF (grant no. 03Z20815A).

Transparency declarations

None to declare.

Supplementary data

Tables S1–S7 are available as Supplementary data at JAC Online.

References

1 Wheat PF. History and development of antimicrobial susceptibility testing methodology. J Antimicrob Chemother 2001; 48 Suppl 1: 1–4.
2 Pedersen SK, Wagenaar JA, Vige H et al. Proficiency of WHO global foodborne infections network external quality assurance system participants in identification and susceptibility testing of thermostolerant Campylobacter spp. from 2003 to 2012. J Clin Microbiol 2018; 56: e01066–18.
3 Hendriksen RS, Seyforth AM, Jensen AB et al. Results of use of WHO Global Salm-Surv external quality assurance system for antimicrobial susceptibility testing of Salmonella isolates from 2000 to 2007. J Clin Microbiol 2009; 47: 79–85.
4 World Health Organization. Antimicrobial Resistance: Global Report on Surveillance. 2014. https://apps.who.int/iris/bitstream/handle/10665/112642/9789241564748_eng.pdf;jsessionid=845ECA593D3087F909CF083DC999B B7?sequence=1.
5 World Bank Group. Resistant Infections. A Threat to Our Economic Future. 2017. http://documents1.worldbankgroup.org/curated/en/32331149336993758/pdf/final-report.pdf.
6 Thaker MN, Kalan L, Wagelechner N et al. Vancorycin-variable enterococci can give rise to constitutive resistance during antibiotic therapy. Antimicrob Agents Chemother 2015; 59: 1405–10.
7 Acar JF, Courvalin P, Chabbert YA. Methicillin-resistant staphylococci: bacteriological failure of treatment with cephalosporins. Antimicrob Agents Chemother 1970; 10: 280–5.
8 Mouton JW. Soup with or without meatballs: impact of nutritional factors on the MIC, kill-rates and growth-rates. Eur J Pharm Sci 2018; 125: 23–7.
9 Courvalin P. Genotypic approach to the study of bacterial resistance to antibiotics. Antimicrob Agents Chemother 1991; 35: 1019–23.
10 Pfaller MA. Molecular approaches to diagnosing and managing infectious diseases: practicality and costs. Emerg Infect Dis 2001; 7: 312–18.
11 Anjum MF, Zankari E, Hasman H. Molecular methods for detection of antimicrobial resistance genes. In: Schwarz S, Cavaco LM, Shen J, eds. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. American Society for Microbiology, 2018; 33–50.
12 Hendriksen RS, Bortolaia V, Tate H et al. Using genomics to track global antimicrobial resistance. Front Public Health 2019; 7: 242.
13 Rägden DJ, Fernandez XM. The 27th annual Nucleic Acids Research database issue and molecular biology database collection. Nucleic Acids Res 2020; 48: D1–8.
14 Panuzzi LG. sraX: a novel comprehensive resistome analysis tool. Front Microbiol 2020; 11: 52.
15 Zankari E, Hasman H, Cosentino S et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 2012; 67: 2640–4.
16 Zankari E, Allesæ R, Jaensen KG et al. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J Antimicrob Chemother 2017; 72: 2764–8.
17 Johnsen CH, Clausen P, Aarestrup FM et al. Improved resistance prediction in Mycobacterium tuberculosis by better handling of insertions and deletions, premature stop codons, and filtering of non-informative sites. Front Microbiol 2019; 10: 2464.
18 Naas T, Oueslati S, Bonnin RA et al. β-lactamase database (BLDB)—structure and function. J Enzyme Inhib Med Chem 2017; 32: 917–19.
19 EUCAST. Intrinsic Resistance and Exceptional Phenotypes, Expert Rules. Version 3.1. 2016. https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Expert_Rules/Expert_rules_intrinsic_exceptional_V3.1.pdf.
20 Clausen P, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinformatics 2018; 19: 307.
21 Bankovich A, Nurk S, Antipov D et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012; 19: 455–77.
22 Kaas RS, Leekitcharoenphon P, Aarestrup FM et al. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS One 2014; 9: e104984.
23 McDermott PF, Tyson GH, Kabera C et al. Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal Salmonella. Antimicrob Agents Chemother 2016; 60: 5515–20.
24 CLSI Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically—Eleventh Edition: M07. 2018.
25 Zankari E, Hasman H, Kaas RS et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother 2013; 68: 771–7.

3499
plasmid-mediated quinolone resistance.

Lancet Infect Dis

Robicsek A, Jacoby G, Hooper DC. The worldwide emergence of

Chemother

in vitro

of the 23S rRNA mutations and the

Schmitz FJ, Petridou J, Jagusch H

ing linezolid resistance in enterococci from whole-genome sequences.

Hasman H, Clausen P, Kaya H

ment precision—authors’ response.

ments: the contribution of strain and laboratory variability to measure-

mBio

irreconcilable? Genotypic and phe notypic amoxicillin–clavulanate

Davies TJ, Stoesser N, Sheppard AE

et al.

fables.

Bader JC, Lakota EA, Andes DR

and fables.

Mouton JW, Meletiadis J, Voss A et al.

variation of MIC measure-

Wallmann J, Bo¨ttner A, Goossens L

et al.

Use of whole genome sequenc-

Eur J Clin Microbiol Infect

Microbes Infect

Aytan-Aktug D, Clausen P, Bortolaia V et al. Prediction of acquired anti-

mSystems

Rappe E, Schrenzel J. Messages from the third International Conference

Microbes Infect

38 Ellington MJ, Ekelund O, Aarestrup FM et al. The role of whole genome

J Antimicrob Chemother

et al.

Int J Antimicrob Agents

Leekitcharoenphon P, Garcia-Graells C, Botteldoorn N

Neumann B, Prior K, Bender JK et al. Comparative Genomics of Quinolone-resistant and Susceptible Campylobacter jejuni of

J Antimicrob Agents

73: 3025–33.

45 Leekitcharoenphon P, Garcia-Graells C, Botteldoorn N et al. Comparative

J Clin Microbiol

57: e01686–18.

47 Bender J, Fleige C, Lange D et al. Rapid emergence of highly variable and

Enterococcus spp. clinical isolates. Int J Antimicrob Agents 2018; 52: 819–27.

48 AbuOun M, O’Connor HM, Stubberfield LJ et al. Characterizing antimicro-

Front Microbiol

3500