Report from the EUCAST Subcommittee on the Role of Whole Genome Sequencing (WGS) in Antimicrobial Susceptibility Testing of Bacteria. Version 1, 2016.

Frank M. Aarestrup[1], Rafael Canton[2], Michel Doumith[3], Oskar Ekelund[4], Matthew J. Ellington[3], Christian Giske[5], Hajo Grundman[6], Henrik Hasman[7], Matthew Holden[8], Katie L. Hopkins[3], Jon Iredell[9], Gunnar Kahlmeter[4], Claudio U. Köser[10], Alasdair MacGowan[11], Dik Mevius[12], Mike Mulvey[13], Thierry Naas[14], Tim Peto[15], Jean-Marc Rolain[16], Ørjan Samuelsen[17], Neil Woodford[3*]

1) Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

2) Servicio de Microbiología. Hospital Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain

3) Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, National Infection Service, Public Health England, London, England

4) Department of Clinical Microbiology and the EUCAST development laboratory, Kronoberg Region. Central hospital, Växjö, Sweden

5) Department of Laboratory Medicine, Karolinska Institute and Karolinska University Hospital, Stockholm, Sweden

6) University Medical Center Freiburg, Infection Prevention and Hospital Hygiene, Faculty of Medicine, University of Freiburg, Freiburg, Germany

7) Statens Serum Institute, Department of Microbiology and Infection Control,

Copenhagen, Denmark

8) School of Medicine, Medical & Biological Sciences, North Haugh, University of St Andrews, United Kingdom

9) Westmead Institute for Medical Research, University of Sydney and Marie Bashir Institute, Sydney, NSW, Australia

10) Department of Medicine, University of Cambridge, Cambridge, UK

11) Department of Medical Microbiology, North Bristol NHS Trust, Southmead Hospital, Bristol

12) Central Veterinary Institute (CVI) part of Wageningen University and Research Centre (WUR), Lelystad, The Netherlands Department of Infectious Diseases and Immunology, Utrecht University, Utrecht, The Netherlands

13) National Microbiology Laboratory, Winnipeg, Manitoba, Canada

14) French National Reference Center for Antibiotic Resistance, Bacteriology-Hygiene unit, Hôpital Bicêtre, APHP, LabEx LERMIT, University Paris Sud, Le Kremlin-Bicêtre, France

15) Nuffield Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

16) PU-PH des Disciplines Pharmaceutiques, 1- URMITE CNRS IRD UMR 6236, IHU Méditerranée Infection, Valorization and Transfer, Aix Marseille Université, Faculté de Médecine et de Pharmacie, 27 Bd J Moulin, 13385 Marseille Cedex 05, France

17) Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, University Hospital of North Norway, Department of Microbiology and Infection Control, Tromsø, Norway

*To whom correspondence should be addressed:
Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit
National Infection Service
Public Health England
61 Colindale Avenue
London, NW9 5EQ

Tel. +44 (0)20 8327 7255
Email: neil.woodford@phe.gov.uk

1. *Abstract*

WGS offers the potential to predict antimicrobial susceptibility from a single assay. The European Committee on Antimicrobial Susceptibility Testing (EUCAST) established a sub-committee to review and report on the current development status of WGS for bacterial AST.

The available published evidence for using WGS as a tool to infer antimicrobial susceptibility accurately is currently either poor or non-existent and the evidence / knowledge base requires significant expansion. The primary comparators for assessing genotypic-phenotypic concordance from WGS data should be changed to epidemiological cut-off values (ECOFFs) in order to better differentiate wild-type from non-wild-type isolates (harbouring an acquired resistance). Clinical breakpoints should be a secondary comparator. This assessment will reveal if genetic predictions could also be used to guide clinical decision making. Internationally agreed principles and quality control (QC) metrics will facilitate early harmonization of analytical approaches and interpretative criteria for WGS-based predictive AST. Only datasets that pass agreed QC metrics should be used in AST predictions. Minimum performance standards should exist and comparative accuracies across different WGS laboratories and processes measured. To facilitate comparisons a single public database of all known resistance loci should be established, regularly updated and strictly curated using minimum standards for the inclusion of new resistance loci and control of resistance gene nomenclature. For most bacterial species the major limitations to widespread adoption for WGS based AST in clinical labs remain the current high-cost and limited speed of inferring antibiotic susceptibility from WGS data as well as the dependency on prior culture since analysis directly on specimens remains challenging.

Currently there is insufficient evidence to support the use of WGS-inferred AST to guide clinical decision making. WGS-AST should be a funding priority if it is to become a rival to phenotypic AST. We plan to update this report as the available evidence increases.

2. *Introduction / Foreword from the subcommittee chairman*

During the 2015 ECCMID meeting in Copenhagen, I was approached (separately) by Derek Brown and Gunnar Kahlmeter, who were both very keen for me to put together a group to consider how well whole genome sequencing (WGS) can predict antibiotic susceptibility patterns and how these game-changing technologies could impact on clinical microbiology, now and in the future. This coordinated and two-pronged attack clearly achieved its goal, and shortly afterwards this EUCAST subcommittee came into being.

Over the past year the team has 'met' virtually and this report marks its first output. We quickly agreed that there are too few data to present a definitive document on the topic, but it would be necessary to review the state-of-the–art as a first approach. This document is therefore presented as a baseline and as a discussion document, and should be considered as such. It marks where we are now, and we present it in the knowledge that it will require updating, probably regularly, as sequencing technologies become more affordable and more widely applied, and as the analysis of the WGS data becomes more rigorous and standardised, and the quantity and quality of evidence for phenotypic / genotypic concordance (or lack thereof) relating to antibiotic susceptibility improves.

Our remit was: (i) to perform a review of the literature describing the role of WGS in antimicrobial susceptibility testing (AST) of bacteria; (ii) to assess the sensitivity and specificity of WGS compared with standard phenotypic AST; (iii) to consider how WGS for AST may be applied in clinical microbiology laboratories and the likely implications for phenotypic and other genotypic methods in

use; (iv) to consider the epidemiological implications of using WGS; (v) to consider the clinical implications of WGS for the selection of antimicrobial therapy; (vi) to consider the principles of how the results of WGS for AST could be presented to clinical users; (vii) to describe the drivers and barriers to routine use of WGS; and, finally, (viii) to report at ECCMID 2016.

We chose to tackle this task on a 'by organism' basis, with particular focus on the use of technology for characterising cultured isolates of bacteria that have been identified as critical AMR threats by the World Health Organisation [1]. There are encouraging signs, but our report makes clear that more robust data are needed across these diverse bug/drug combinations. Furthermore, work is needed to overcome problems currently posed by particular species and/or certain drug classes. We highlight these gaps, make recommendations (summarised in Table 1), and encourage others to use these in order to generate the analyses that will move this important topic forwards.

I would like to thank all members of the Subcommittee for their efforts over the last year. It has been a pleasure working with them all. Special thanks are owed to Matt Ellington and Oskar Ekelund, who edited the report.

Neil Woodford
Chair, EUCAST WGS Subcommittee

April 2016

Table 1. Summary of the conclusions and recommendations of this report

1. **For most bacteria considered in this report, the available evidence for using WGS as a tool to infer antibiotic susceptibility (i.e. to rule-in as well as to rule-out resistance) accurately is either poor or non-existent. More focussed studies and additional funding resources are needed as a priority to improve knowledge.**

2. The primary comparator for WGS-based prediction of antibiotic susceptibility should, whenever possible, be the epidemiological cut-off value (ECOFF).

3. **Assessing genotypic data against clinical breakpoints represents a tougher challenge, but will be necessary if WGS-based testing is to guide clinical decision making. Clinical breakpoints should therefore be used as a secondary comparator, ideally using the same data sets as used for ECOFF-based assessments.**

4. Available published evidence does not currently support use of WGS-inferred susceptibility to guide clinical decision making (i.e. to replace routine phenotypic AST in most or all cases).

5. **There should be international agreement on the most appropriate and effective principles and quality control (QC) metrics to facilitate early standardisation and harmonization of analytical approaches and interpretative criteria for WGS-based predictive AST. Only datasets passing agreed QC metrics should be used in antimicrobial susceptibility predictions, since resistance genes or mutations might be missed in sequences of poor quality.**

6. Different bioinformatics tools for predicting AST should perform to minimum standards and should be calibrated and shown equivalent in terms of the results generated.

7. **A single database of all known resistance genes / mutations should be established to ensure that there is parity of analysis and to facilitate measurement of comparative accuracies across different systems and bioinformatics tools. This database should be updated regularly, and must have strictly curated minimum standards for the inclusion of new resistance genes and mutations. An important function of a centralised database would be to control resistance gene nomenclature.**

8. Expansion of the evidence base is a critical priority if WGS is to be considered seriously as a rival to phenotypic AST.

9. **For most bacterial species and in most countries the current cost and speed of inferring antibiotic susceptibility from WGS data remain prohibitive to wide adoption in routine clinical laboratories.**

10. This report should be considered as a baseline discussion document, which should be revisited and updated at regular intervals (likely every 18-24 months) as sequencing technologies become more affordable and more widely applied and the available evidence increases.

3. Defining Resistance
3.1 MIC distributions of wild-type bacteria, epidemiological cut-off values (ECOFFs) and their relationship to clinical breakpoints

In 2002, the European Committee on Antimicrobial Susceptibility Testing (EUCAST) introduced the concept of gathering large numbers of MIC values from many contributors to present on a web site as aggregated reference MIC distributions for every important combination of microorganisms and antimicrobial agents [2]. The original conditions of acceptance of individual MIC distributions were: (i) that each contribution of MIC values must consist of a minimum number of isolates; (ii) that the species was well defined; (iii) that MIC determinations were performed using standardized methodology (or a method calibrated to a standardized method); and (iv) that the concentrations tested were not truncated at the lower end of the concentration series. A EUCAST subcommittee is currently reviewing the rules for inclusion and exclusion of data sets and for how ECOFF values are defined.

Contributors of MIC data were not informed about whether or not their contributions were accepted. There are currently more than 26,000 MIC distributions in the EUCAST database, which amounts to many millions of MIC values. The distributions are from breakpoint committees, individual researchers in human and veterinary medicine, programs for the surveillance of antimicrobial resistance in humans and animals, EUCAST development projects, pharmaceutical companies as part of programs for the development of new agents, and more. The distributions are freely available on www.eucast.org.

A typical aggregated MIC distribution, in this case for *Escherichia coli* and cefotaxime, is shown in Figure 1.
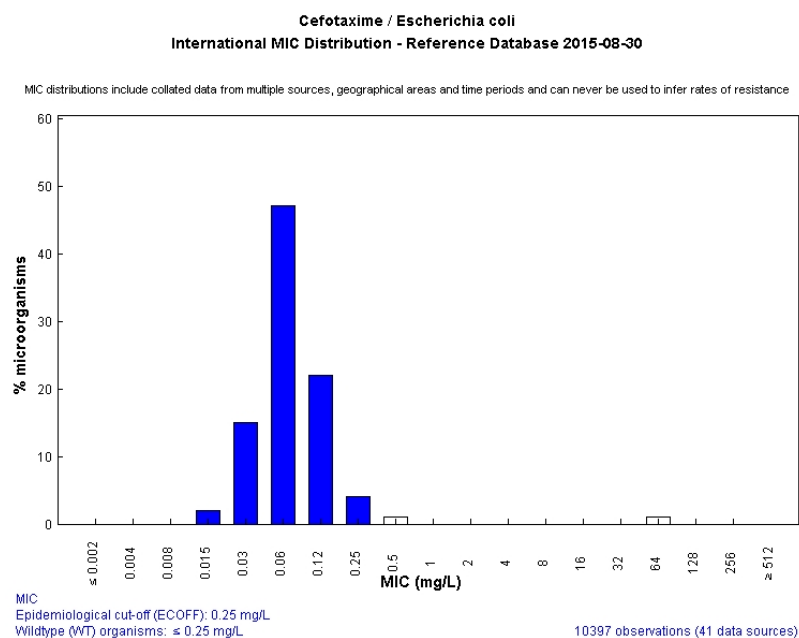

Figure 1. *Escherichia coli* (n=10,397) cefotaxime MIC distributions (n=41).

Of the 72 *E. coli* and cefotaxime distributions available in the database, 41 fulfilled the criteria for acceptance, were aggregated and are displayed on the EUCAST web site as Figure 1 above.

MIC distributions are uni- or multi-modal. The left-hand, most often dominating and Gaussian-shaped part of the distribution represents the isolates devoid of phenotypically detectable acquired resistance mechanisms, otherwise known as "the wild-type (WT) MIC distribution". Furthermore, MIC values for bacteria from humans and animals are distributed in the same way. It is not affected by the geographical location where the isolates are collected, the specimen source (i.e. humans or animals [3], or healthy or sick individuals) or the era of collection (as some of the distributions date from the 1950s whereas others are very recent). There are numerous ways (biological or statistical) to sample the upper end of the distribution to define the MIC-value best representing the end of the WT distribution [4-6]. Despite the fact that there is often no absolute and distinguishing threshold to mark the end of the WT or beginning of the non-wild-type (NWT), it has been useful to define the ECOFF as the "highest MIC of organisms devoid of phenotypically detectable acquired resistance mechanisms". It provides a means to distinguish between resistant and susceptible populations in a biological sense.

From a clinical point of view categorizing isolates into WT and NWT informs the clinician of whether or not the isolate causing infection is devoid of acquired resistance mechanisms or not, irrespective of its clinical susceptibility categorization as S, I or R.

There is no immediate relationship between the categorization of WT and NWT on one side and the clinical categorization "S", "I" and "R" on the other. A WT microorganism can be categorized as S, I or R to a particular antibiotic and a non-WT organism may still be categorized as S. This means that if one wants to encompass both WT and NWT on one hand and clinical S, I and R on the other, the possible susceptibility categories of an isolate to any antibiotic are S^{WT}, S^{NWT}, I^{WT}, I^{NWT}, R^{WT} and R^{NWT}.

There are many examples of each of these in the breakpoint tables. To give one, an *E. coli* isolate with a ciprofloxacin MIC of 0.25 mg/L exemplifies the S^{NWT} category. Conversely, *Pseudomonas aeruginosa* and tigecycline or *Stenotrophomonas maltophilia* and carbapenems both exemplify the R^{WT} category. For ampicillin, TEM-1-producing *E. coli* represent R^{NWT}.

At the outset of reviewing all breakpoints for all antibiotic classes (2002), EUCAST decided that clinical breakpoints should not divide MIC WT distributions. If breakpoints are allowed to bisect WT MIC (or zone diameter) distributions, the methodological variation, would obviate reproducible S, I and R categorization [2]. This is despite the increased standardization of MIC determination [7]. We have not had reason to change this view. Following on from this, once it has been established that the species is a good clinical target for the antibiotic agent in question, the ECOFF is the lowest possible susceptible breakpoint.

The ECOFF is also the relevant 'cut-off' to screen for low-level resistance using phenotypic susceptibility testing. ECOFFs provide an opportunity to compare antimicrobial resistance and resistance development when clinical breakpoints: (i) differ between committees (e.g. EUCAST versus Clinical Laboratory Standards Institute - CLSI) or between committees and agencies (e.g. CLSI versus Food and Drug Administration - FDA) or between agencies (FDA vs. European Medicines Agency - EMA); (ii) change over time; or (iii) are different between humans and animals.

There is no principle difference between MIC distributions and ECOFFs in fast-growing non-fastidious and fastidious bacteria, and those exhibited by slow-growing bacteria such as *Mycobacterium tuberculosis* [8] or fungi such as *Candida* spp. and *Aspergillus* spp. [9].

More information on EUCAST in general and wild-type MIC distributions and ECOFFs in particular can be obtained from the EUCAST Website (www.eucast.org) and the recently published review of EUCAST activities since 2001 [10].

This subcommittee sought to assess whether available data are sufficient to test the hypothesis that the closest relationship between sequencing and phenotypic testing will be achieved by using the WT vs. NWT categories.

3.2 Molecular mechanisms of antimicrobial resistance (AMR)

For most of the clinically relevant bacterial pathogens, phenotypic analysis of bacterial susceptibility to antimicrobial agents is relatively straightforward and relies on well-proven methods, such as agar and broth micro-dilution (the latter being the gold standard) or disc diffusion, followed by interpretation according to agreed guidelines.

With the introduction of Sanger sequencing in the mid-1970s and PCR in the 1980s, it became possible to study some of the molecular mechanisms responsible for the observed non-susceptibility towards the various antimicrobial agents. Common examples of these molecular mechanisms are: 1) transferable AMR genes (e.g. extended-spectrum β-lactamases; ESBLs); 2) upregulation of AMR gene expression by point mutations (e.g. *ampC* in *E. coli* or regulatory mutations effecting efflux in many taxa); 3) porin modification or loss (e.g. by deletion events and/or lack of expression); 4) point mutations in essential single (e.g. *gyrA* and/or *parC* of Enterobacteriaceae) and multi-copy (e.g. mutations in one or more loci of the 23S rRNA gene) housekeeping genes.

To complicate the matter, some bacterial species (or higher taxonomic orders) can be intrinsically resistant to a given antimicrobial agent [11]. This intrinsic resistance can be caused by some of the mechanisms listed above (e.g. chromosomal AMR genes), but can also be a result of the lack or unavailability of targets for the antimicrobial agent.

Traditional Sanger sequencing and rapid molecular methods e.g. PCR etc, allow screening for a limited number of resistance genes, which are often selected because they confer resistance to key antibiotics (e.g. ESBLs or carbapenemases). The data offer very limited opportunities to compare genotype with phenotype. By contrast, WGS has potential to yield data about any resistance gene or mutation present and the data might therefore be analysed to create a genotypically-inferred antibiotic resistance profile (or antibiogram) or, perhaps, to infer susceptibility.

4. Using next generation sequencing data for in silico (genotypic) detection of AMR

Next generation sequencing data producing WGS information can originate from a variety of different sequencing platforms. A review of these technologies is beyond the scope of this report, but can be found elsewhere [12]. Short read (e.g. 100 to 500 bp) output with high accuracy may be complemented by that produce much longer reads. At the time of writing, these newer platforms come at significantly greater cost and higher error rates than the short-read technology. The dominant short-read technology produces single (raw) reads that are in most cases shorter than the gene(s) responsible for the reduced susceptibility to a given antimicrobial agent and either need to undergo *de novo* assembly to obtain larger fragments of the originally contiguous DNA ("contigs") or by reference ("mapping") to known genetic targets (in this case, AMR genetic determinants). Repeat regions in DNA fragments are particularly challenging and correct assembly may be problematic.

Using WGS data for detection of the many different molecular mechanisms leading to AMR yields far more information from a single physical test than other methods (e.g. PCR or microarrays) and, at its most fundamental level, does not require prior knowledge of the resistance phenotype of the isolate. Nevertheless, there is need to understand the potential 'added value' of WGS with regard to the clinical implications of AMR and so the validity of data generated by these novel technologies must be challenged against phenotypic methods to differentiate WT isolates from NWT or S isolates from R isolates. In this regard WGS is a genetic test that defines a genotype as WT or NWT and compares most directly with phenotypic criteria that do the same (ECOFFs).

Whilst more informative than conventional molecular techniques, WGS is no simple task, especially when the data have been generated by short-read ("second generation") technology. Detection of defined resistance genes can be achieved either by BLAST analysis of draft genomes towards a gene-based database or by mapping individual reads to the same type of database. Such solutions are already available as either downloadable tools, such as ARG-ANNOT [13], or web-based tools, such as ResFinder for BLAST analysis and MGmapper for mapping of reads [14, 15]. The gene-based solutions have the obvious requirement for full-length genes identical to already characterized (and preferably published) AMR genes. The bioinformatics solutions mentioned above are able to identify less-than-perfect hits (<100% nucleotide identity, truncated genes because of non-perfect *de novo* assembly), but such hits will always need to be subjected to some sort of assessment if they are to be translated into a predicted phenotype.

Accurate prediction of resistance by WGS can be complicated by insufficient knowledge about all genetic variation leading to reduced susceptibility for a given antimicrobial agent (such as colistin or daptomycin) as well the emergence of new mechanisms and when resistance arises due to altered expression of intrinsic genes (e.g. those encoding efflux pumps). Also, shortcomings of second generation sequencing technology may hamper accuracy. An example of the latter could be *Enterococcus faecium*, where a point mutation (G2576T) in 2-3 copies of the six 23S rRNA loci would lead to phenotypic resistance to linezolid [16]. De novo assembly of second generation sequence data from the same isolate would most likely lead to assembly into a wild-type version of the 23S gene due to the consensual nature of the assemblers, where only the most abundant base is reported in the draft genome data.

5. Quality metrics for WGS

Like any other test, the quality of WGS data can vary between individual test runs. Therefore prior to any actual bioinformatics analysis, quality control (QC) steps are essential to assess whether the WGS data have reached a suitable standard. **Only datasets passing these QC metrics should be used in antimicrobial susceptibility predictions, since resistance genes or mutations might be missed in sequences of poor quality.** These QC steps include assessing 1) the quality and quantity of the raw reads to ensure sufficient coverage (e.g. >30 times coverage) of the bacterial genome, 2) the quality of the *de novo* assembly (leading to a draft genome sequence) and 3) detecting possible contaminant DNA, originating either from upstream handling of the bacterial isolates and DNA purification or from the preparation and running of the DNA samples on the sequencer. This last key quality control step is predicated on comparison against the (wet-lab) species identification for the sample (isolate) that was sequenced.

Some of the different sequence QC parameters that have been used are listed in **Table 2**. The parameters most frequently used are highlighted in bold. **There are currently no international standards for QC-thresholds to use for assessing quality.** This seems to be individually decided by researchers and may also depend on the purpose of the study or the methods used for sequencing. A sequencing method with a high error rate can be compensated to some extent by greater depth of coverage. The necessary QC-threshold also depends on the species analyzed. Thus, before WGS can be routinely implemented into accredited clinical practice there is a need to establish necessary minimum QC-thresholds (e.g. by multiply sequencing reference isolates) for identification of resistance genes and their variants.

The Global Microbial Identifier initiative [15] is currently collaborating with the US-FDA and the COMPARE project [16] in proficiency testing of WGS data and isolates have been distributed to 50 laboratories worldwide. This and similar initiatives are important first steps towards setting objective QC-thresholds. There is, however, a need to expand this using more isolates as well as developing standard datasets of raw sequences to facilitate the assessment of performances across different laboratories.

Table 2. Selected quality control (QC) parameters used to evaluate WGS data (most commonly used in bold), Ref. = reference, No. = number.

QC-parameter	Explanation
No. of reads	No. of reads refers to sequence yield (the amount sequenced)
Average read length	The average length of all reads, measured in bp.
No. reads mapped to ref. sequence	The no. of reads that map to a closed (finished) genome (same strain).
Proportion of reads mapped to ref. sequence (%)	The proportion of reads that map to a closed genome (same strain).
No. reads mapped to ref. chromosome	The no. of reads that map to a closed chromosome (same strain).
Proportion of reads mapped to ref. chromosome (%)	The proportion of reads that map to a closed genome's chromosome (same strain). This cannot exceed more than 100%.
Reads mapped to ref. plasmids	The no. of reads that map to plasmids, if present.
Proportion of reads mapped to ref. plasmids (%)	The proportion of reads that map to plasmids (if present) of the closed genomes. This cannot exceed 100%.
Depth of coverage, total DNA sequence	Describes the no. of times the sequenced bps cover the reference DNA. No. of bps sequenced divided by the total size (both chromosome and plasmids) of the closed genome (same strain), often expressed with an "x" (e.g. 30x). A minimum depth of 30x is usually preferred.
Depth of coverage: chromosome	As above, but describes the no. of bps sequenced divided by the total size of the closed chromosome (same strain).
Depth of coverage: plasmid	As above, but describes the no. of bps sequenced divided by the total size of the closed plasmid (same strain).
Size of assembled genome	Often used to identify contamination. If the calculated size of all the contigs in bp exceeds that expected it could indicate >1 genome.
Size of assembled genome per total size of DNA sequence (%)	The proportion of contigs that map directly to the closed genome (same strain). This cannot exceed 100%.
Total no. of contigs	Generally, the total no. of contigs assembled, <1,000 contigs indicates good quality. For organisms with genomes 5-6 Mb in size then <100 contigs is (generally) realistic.
No. of contigs >500 bp	The total no. of contigs assembled that have a sequence length >500 bp. This should correspond well to the total no. of contigs.
Longest contig length	The length of the longest contig.
Shortest contig length	The length of the shortest contig
Mean, median and std. deviation	Mean, median and std. deviation of the contigs, used to evaluate quality.
N50	The length for which the collection of all contigs of that length or longer contains at least half of the sum of the lengths of all contigs, and for which the collection of all contigs of that length or shorter also contains at least half of the sum of the lengths of all contigs. N50 >15,000 normally indicates good quality, but minimum size of 30,000 bp is often preferred.
NG50	Helpful for comparisons between assemblies. As N50, except that 50% of the genome size must be of the NG50 length or longer. Where the assembly

	size ≤ the genome size then NG50 cannot exceed N50.

6. The need for a standardised, open-access database

Most of the genomes released now are not closed, so there is a need for better standardisation of annotation to facilitate detection of AMR genes because standard BLAST analysis will retrieve plenty of hits within annotated or raw sequences available in Genbank, and those hits will be inconsistently annotated even where the actual sequences are identical.

Currently several AMR databases exist, and they are either downloadable for use locally (e.g. ARG-ANNOT) [13] or are web-based solutions (ResFinder for BLAST analysis and MGmapper for mapping of reads [14, 15] and a Comprehensive Antibiotic Research Database, CARD [17]). In addition to a fully curated database of accurately annotated genes that seeks to avoid the pitfalls posed by erroneously annotated genes, there also exists a need for a single, standardised 'challenge database' solution that contains all validated AMR genes as well as those point mutations in chromosomal target genes that are known to be associated with antibiotic resistance. This can then be used as a standard reference dataset for different bioinformatics analysis tools. Any such solution should be flexible so that stringency of detection can be changed to allow detection from partial gene sequences (length less than 100%) and/or AMR genes with identities less than 100%. Looking at conserved sequence motifs in gene families (i.e. β-lactamase) should also help in assessing the validity of newly detected genes and in particular those exhibiting low sequence identity matches.
This single web solution should be iterative and enhanced by regular, validated updates of newly identified gene sequences and point mutations at a frequency that remains to be decided. Machine learning should be explored to iteratively and automatically improve the detection algorithms for this purpose.

However, in order to achieve this goal there must also be clear international consensus on the criteria used to define a gene as "new" (i.e. % of identity with existing genes) or as a variant of known genes. This is inextricably linked to issues of gene nomenclature. Currently, different criteria are used depending on the antibiotic class that a particular gene confers resistance to. For example, a "new" β-lactamase gene can be defined by as little as one amino acid difference from a known sequence and regardless of any impact that this change might have on the conferred resistance phenotype.

There should be minimum standards for inclusion of new resistance determinants in the standardised database, and these standards would probably differ from those currently required for publication (e.g. they may be more demanding). It seems reasonable that new genes should have a full gene sequence, which can be translated into a protein sequence, and that they should have been unequivocally linked to a predicted resistance phenotype, as was recently exemplified with *mcr-1* plasmid-mediated colistin resistance [18], before being added to the database.

7. Categories of systematic errors in WGS predictions of AMR

When comparing the concordance between phenotypic and genotypic AMR it is essential to consider the reasons that errors may occur. Three broad reasons for systematic errors are:

1) An inadequate limit of detection of WGS
2) Flaws with phenotypic AST
3) Incomplete understanding of genotypic basis of phenotypic resistance

Of these the limit of detection of WGS (1) applies to the detection of hetero-resistance, which is most applicable to TB, as for most other organisms WGS is usually performed from single bacterial colonies. 2) Flaws due to phenotypic detection issues become most apparent when the knowledge base of the genetic basis of resistance is relatively complete for a given organism and can point to such problems. For the purposes of this report will likely only apply to well progressed / characterised species (such as *Mycobacterium tuberculosis* and *Staphylococcus aureus*). 3) At this relatively early stage of development of WGS based genotype-phenotype comparisons it can be anticipated that there may be many gaps in the knowledge base and these will be explored and highlighted in the following evidence reports of this document.

8. Evidence reports for in silico prediction of antimicrobial resistance

8.1 Enterobacteriaceae (other than Salmonella spp.)

8.1.1 Background

Multidrug-resistant Enterobacteriaceae are emerging as a serious infectious disease challenge; they can accumulate many antibiotic resistance genes through horizontal transfer of genetic elements, those coding for β-lactamases (e.g. ESBLs and carbapenemases), fluoroquinolones and aminoglycosides being of particular concern.

8.1.2 Published studies

A small number of studies have assessed the feasibility of using WGS to infer AMR in *E. coli* and *Klebsiella pneumoniae* genomes; they are largely based on screening for known acquired AMR genes and a small number of known resistance-conferring mutations, such as those associated with ciprofloxacin resistance. In one study, Stoesser *et al.* reported 95% concordance between phenotypic and WGS-predicted susceptibility for seven commonly used antibiotics (amoxicillin, amoxicillin/clavulanic acid, ciprofloxacin, gentamicin, ceftriaxone, ceftazidime and meropenem) by querying 143 assembled genomes from *E. coli* and *K. pneumoniae* with a compiled database of acquired AMR sequences and mutations in the quinolone resistance determining regions (QRDRs) of *gyrA* and *parC* [19]. An even higher level of concordance (99.74%) between phenotypic susceptibility testing and the WGS-predicted resistance to five classes of antibiotics (β-lactams, chloramphenicol, sulphonamides, tetracycline and trimethoprim) in *E. coli* genomes was reported in an earlier study using the same approach with a database of 1,411 different AMR sequences, confirmed using simple blotting and PCR approaches to the expected genes [14, 20]. A recent study investigating 76 *E. coli* from farm cattle also showed good phenotype-genotype correlation (97.8%), with the majority of discordant results attributed to the prediction of streptomycin resistance [21].

There are important limitations with identifying the mode of transmission of these acquired genes in short-read sequences due to exclusion of repeat regions during the 'cleaning' stages while initial contigs are assembled. This typically results in inaccuracies in annotation. These are particularly marked in highly recombinant plasmids, which unfortunately carry most of the AMR genes that are relevant to β-lactam and aminoglycoside resistance in the Enterobacteriaceae. Nevertheless, using sequencing technologies with longer reads (and greater cost), such as PacBio and MinION, improved their detection [22-25]. These bioinformatics challenges also include the development of tools that can detect signature sequences of AMR determinants (e.g. β-lactamase motifs) to identify potentially new variants conferring acquired AMR, which can be explored in more detail.

8.1.3 Problems found or anticipated - Gaps in the knowledge base

Chromosomal mutations that alter the cell membrane permeability due to modification in the structure or the levels of expression of outer membrane proteins (OMPs), antibiotic efflux due to efflux pumps such as resistance-nodulation-division (RND) and major facilitator superfamily (MDF) pumps, or changes in the lipopolysaccharide structure have been linked to decreased susceptibility and resistance to β-lactams, quinolones, chloramphenicol, tetracyclines, tigecycline and colistin in Enterobacteriaceae, but have yet to be fully elucidated. This makes comprehensive phenotypic-genotypic comparisons difficult [26-32] by limiting the sensitivity of the WGS-based data. In particular, the relationships between chromosomal mutations and the related phenotypic changes responsible for resistance are not always well characterised and screening genome sequences for

insertion sequences interrupting or modifying the expression of resistance-associated genes, including intrinsic β-lactamases, could be problematic due to constraints inherent in using short reads. Therefore, and in contrast to horizontally-acquired resistance genes, the ability of WGS to predict resistance due to, or modulated by, chromosomal alterations is likely to be restricted by existing knowledge, as in the case of carbapenem non-susceptibility resulting from the combination of decreased permeability and AmpC or ESBL enzymes and also for antimicrobial agents for which the underlying genetic backgrounds of resistance are yet to be fully characterised (e.g. amoxicillin-clavulanic acid, nitrofurantoin, temocillin, colistin and tigecycline). Screening for the loss-of-function mutations via nonsense mutations, frameshifts or insertion elements that are meaningfully less complex than substitutions affecting the structure, dynamics and substrate specificity of resistance-conferring proteins is realistically achievable [33-35]. The effect of amino acid changes in the transmembrane β-strand loop 3 that constitutes the porin channel eyelet that was associated with diminished carbapenem uptake in the endemic KPC-producing *K. pneumoniae* ST258 clone illustrates the complexity of interpreting amino acid substitutions identified by WGS in the absence of experimental evidence [36, 37]. Recent studies have shown that the genetic basis of resistance to colistin in *K. pneumoniae* clinical isolates can be attributed in the majority of cases to alterations in the *mgrB* regulator or the two component systems *pmrAB* or *phoEQ* that regulate the expression of the biosynthesis pathway of lipid A [29, 38]. However, incorrectly inferring susceptibility remains the risk if resistance is mediated by genuinely novel, undiscovered genetic factors, as evidenced by the description of *mcr-1*, the first known transferable colistin resistance determinant [18].

The gaps in the existing knowledge of genotype-phenotype relationships could be augmented in some cases by directly detecting the levels of gene expression by sequencing RNA extracts. This approach was successfully used as a proof-of-concept for the detection of *ompF* down-regulation associated with cephalosporin resistance and over-expression of the RND pump component *acrB* leading to decreased susceptibility to quinolones, tetracycline and chloramphenicol in an *E. coli* laboratory selected mutant [39]. Whilst a similar approach should also be feasible for the detection of hyperproduction of intrinsic chromosomal β-lactamases (e.g. AmpC in *Enterobacter* spp.) the use of such methods which accuracy is highly dependent on bacterial growth conditions, is likely to occur in a sub-set of laboratories in the foreseeable future and will not be considered further for the purposes of this document, which seeks to examine widely used techniques only.

8.1.4 Summary
The relatively limited number of acquired resistance genes and resistance-associated mutations that dominate epidemiologically in the Enterobacteriaceae (compared with the large number of those that have been reported in the resistome) could explain the high levels of accuracy of genotype-phenotype correlation in published studies and means that well-informed screening approaches can be very accurate. However, susceptibility to some drugs will be harder to predict than for others and understanding the full range of mechanisms and their interplay will require more study if improved levels of accuracy across large genetically diverse datasets are to be achieved.

8.2 Salmonella *spp.*

8.2.1 Background

Molecular mechanisms conferring reduced susceptibility to antimicrobial agents are relatively well characterized in *Salmonella* spp. The majority of these are encoded by horizontally-transferable genes, which potentially makes genotypic detection a reliable alternative to phenotypic testing, as these genes generally assemble into full-length genes when using short-read sequencing data, as long as the quality and quantity of these are adequate to produce good assemblies. In addition to acquired gene-based AMR, mutational acquired AMR also exists in *Salmonella*. The most clinically important examples currently are single or double mutations in the *gyrA* DNA gyrase and *parC* topoisomerase genes leading to reduced susceptibility to quinolones and fluoroquinolones. Also clinically important to consider is resistance to third-generation cephalosporins due to acquired extended-spectrum and AmpC β-lactamases.

8.2.2 Published studies

Very few comprehensive studies have been published where phenotypic susceptibility data have been compared with the underlying molecular mechanisms identified in WGS datasets from a collection of *Salmonella* isolates. Zankari et al. used a set of 50 *Salmonella enterica* serovar Typhimurium originating from pigs and previously tested phenotypically towards 17 different antimicrobial agents as part of the DANMAP surveillance program [20, 40]. WGS was performed on these 50 isolates of which 49 produced sufficient WGS data to create draft genomes for analysis with the ResFinder web-tool. Here, complete agreement (100% sensitivity and specificity) between tested and predicted susceptibility/resistance phenotype (S/R) was observed [20]. However, this perfect correlation between phenotypes and genotypes generated by ResFinder was to some extent biased by 1) the fact that none of the isolates showed phenotypic resistance to quinolones or fluoroquinolones which would, in most cases, have been unnoticed by ResFinder as it currently does not detect chromosomally-acquired mutations leading to AMR and 2) there was a relatively low level of diversity amongst the resistance phenotypes and hence resistance genes found.

In addition, a few studies exist where only a small number of isolates have been analyzed using both phenotypic and genotypic methods. In a study of extremely drug-resistant *Salmonella enterica* serovar Senftenberg, two isolates from Zambia were analyzed both by conventional phenotypic methods and by WGS analysis [41]. Here, genes conferring reduced susceptibility to nine drug classes including fluoroquinolones and extended-spectrum cephalosporins, were identified, again with the use of the ResFinder web-tool. These strains also demonstrated high-level resistance to fluoroquinolones caused by mutations in GyrA (S83Y and D87G) and ParC (S80I), which were identified manually. An underlying molecular mechanism was identified for all the AMR phenotypes displayed by the two isolates. In a similar study, two ESBL-producing *Salmonella enterica* serovar Typhi isolates were tested phenotypically towards 25 different antimicrobial agents belonging to 10 different classes [42]. Here, ResFinder found seven different AMR genes, which in combination could explain the observed AMR phenotypes of the two isolates.

8.2.3 Problems found or anticipated - Gaps in the knowledge base

The overall degree of resistance within *Salmonella* spp. varies depending on the serovar and phage type; some may be associated with resistance to particular antimicrobials, whilst others may have an increased propensity for multidrug resistance. Diversity within the strain panel tested will therefore likely impact on the conclusions drawn regarding the utility of WGS for AMR prediction in *Salmonella* spp. WGS should therefore be applied to further strain panels reflecting the diversity of

Salmonella serovars (and their common resistance phenotypes) associated with clinical and veterinary infections. Included in these panels should be representatives of some of the particularly multiresistant clones that are currently circulating in human and animal populations, e.g. serovars Kentucky, Infantis and monophasic Typhimurium. Nevertheless, recent data from the European Centre for Disease Control ECDC data indicate that nearly 55% of *Salmonella* spp. from humans are susceptible to all antimicrobial classes tested, suggesting that resistance prediction in *Salmonella* spp. may be more straightforward than in other organisms where multidrug resistance is the norm [43].

As with the other Enterobacteriaceae, detection of chromosomal mutations leading to acquired antimicrobial resistance is a challenge that still needs to be addressed in order to be able to predict resistance phenotypes fully based on WGS data. Priority should be afforded to the detection of mutations leading to fluoroquinolone resistance, as this drug class has high clinical relevance and as phenotypic resistance is commonly detected in salmonellae. Development of fluoroquinolone resistance can be a multifactorial process involving acquisition of mutations leading to amino acid substitutions within the topoisomerase genes and altered expression of outer membrane proteins and/or multidrug efflux pumps. Fortunately, the most common chromosomal mutations leading to acquired fluoroquinolone resistance in salmonellae are well characterized; decision rules to translate mutations into a predicted phenotype are therefore available and can, in principle, be incorporated into existing tools such as ResFinder. This, however, requires detection of amino acid variation rather than nucleotide variation, which is currently used to detect (transferable) resistance genes. Other possible candidates for detection of chromosomal mutations are *pmrA* and *pmrB* leading to reduced susceptibility towards colistin [44].

8.2.4 Summary
In conclusion, the relatively limited examples that are available on the feasability of using WGS data to predict antimicrobial resistance in salmonellae are showing promising results, but the impact of the sample sets tested on genetic diversity needs to be explored in detail before further conclusions are drawn about the use of WGS data for AMR prediction in *Salmonella* spp.

8.3 Staphylococcus aureus

8.3.1 Background

Staphylococcus aureus exhibits intrinsic sensitivity to commonly used antibiotics. Resistance is associated with mutations in core genes or the acquisition of specific antibiotic resistance genes. Generally, the history of antibiotic resistance in this species is associated with the evolution of resistance shortly after the first introductions of new antibiotic into clinical practice, for example resistance to meticillin was detected in clinical isolates of *S. aureus* within a year of introduction into the UK [45]. The problem of resistance in this species has driven extensive studies identifying the genetic basis of resistance, and as such there is a large body of literature documenting resistance mechanisms for most of the clinically relevant antibiotics used. This has revealed a well-characterized spectrum of mechanisms bestowing resistance in *S. aureus*, and in the cases of some AMR determinants, in other staphylococcal species and other genera too. This has proved to be a valuable resource for the *in silico* prediction of antibiotic resistance and has contributed to the overall efficacy of the results.

8.3.2 Published studies

To date, several studies have been published that demonstrate the ability to predict antibiotic resistance from genome sequence data [46-50]. Initial studies examined the concordance between phenotype and genotype [46-48], and more latterly blinded studies have made predictions [49]. In addition, bioinformatics resources have also been produced, which will generate predictions of antibiotic resistance from genome sequence data [51].

The initial demonstration of the potential of WGS data to predict AMR as a clinical tool came from proof-of-concept studies using the benchtop Illumina MiSeq platform to investigate suspected meticillin-resistant *S. aureus* (MRSA) outbreaks. Köser *et al.* sequenced 14 isolates belonging to four different clonal complexes of *S. aureus*, and demonstrated 100% concordance of the *in silico* resistance prediction with the phenotypic results for 13 different antibiotics (cefoxitin, erythromycin, ciprofloxacin, gentamicin, tetracycline, rifampin, fusidic acid, mupirocin, clindamycin, kanamycin, tobramycin, trimethoprim, and linezolid) [46]. The authors used an in-house database of resistance determinants derived from literature mining, and mapped sequence reads to a resistome pseudo-molecule (concatenated resistance genes in a single DNA sequence), followed by manual inspection to predict the resistance profile of each isolate. Investigating an MRSA cluster in an intensive care unit, Eyre *et al.* sequenced 10 isolates belonging to the same *spa*-type (t5973) and conducted *in silico* predictions for two antibiotics (penicillin and tetracycline) [47]. The authors took a different bioinformatic path to investigate the antibiotic resistance, using the *de novo* assemblies to look for the presence and absence of two gene, *tetK* and *blaZ*. In all cases the presence of these genes correlated with the phenotypic resistances for respective antibiotics.

Examining 13 isolates belonging to USA300 clone, Lee et al. predicted the antibiotic resistance profiles for nine antibiotics (ciprofloxacin, clindamycin, doxycycline, erythromycin, gentamicin, oxacillin, tetracycline, trimethoprim-sulfamethoxazole, and vancomycin) in complete concordance with the phenotypic results [50]. The authors generated *in silico* predictions from short read data using SRST2, a mapping-based tool for fast and accurate detection of genes and alleles from WGS data from a user defined database [52]. In the case of this study details of the antibiotic resistant database used were not provided.

The effectiveness of *in silico* prediction for antibiotic resistance in *S. aureus* has been further demonstrated in larger studies, both in terms of the number of isolates and also the antibiotics investigated. Using the genome data of 193 isolates belonging to a global collection of ST22, Holden *et al.* used a mapping based approach, coupled with manual inspection, to identify molecular determinants that explained 99.8% of the measured phenotypic resistance traits [48]. In total 847 resistance traits were tested for 18 antibiotics (penicillin, oxacillin gentamicin, linezolid, erythromycin, clindamycin, ciprofloxacin, fusidic acid, mupirocin, moxifloxacin, co-trimoxazole, tetracyline, vancomycin, teicoplanin, rifampicin, fosfomycin, tigecycline, and daptomycin), utilizing an enhanced version of the library used by Köser *et al.* in their earlier study.

Using the WGS data and phenotype data for 501 *S. aureus* isolates as a derivation set to optimise predictions for 12 antibiotics (penicillin, methicillin, erythromycin, clindamycin, tetracycline, ciprofloxacin, vancomycin, trimethoprim, gentamicin, fusidic acid, rifampin, and mupirocin), Gordon *et al.* then conducted a blind validation of their refined method on a query set of 491 isolates and demonstrated sensitivity of 0.97 and specificity of 0.99 [49]. As with their previous study, the authors used *de novo* assemblies and BLASTn for their *in silico* prediction [47]. In their resistance database 18 acquired genes were included: *blaZ*, *mecA*, *msr*(A), *erm*(A), *erm*(B), *erm*(C), *erm*(T), *tet*(K), *tet*(L), *tet*(M), *vanA*, *fusB*, *far*, *dfrA*, *dfrG*, *aacA-aphD*, *mupA*, and *mupB*, and variation in six core genes: *gyrA* (n=6), *grlA* (n= 13) *grlB* (n=6), *fusA* (n=59), *rpoB* (n=28) and *dfrB* (n=8).

In a significant departure from previous studies in *S. aureus*, Bradley *et al.* described a stand-alone tool, Mykrobe predictor (http://www.mykrobe.com/), for antibiotic resistance prediction directly from fastq files and which does not rely on mapping or assembly-based approaches [51]. This tool utilizes a de Bruijn graph-based approach to compare sequence reads to a reference graph representation. The method has the advantage of being quicker than the mapping and assembly-based methods, and it can also identify minority variants in sequencing data and therefore identify potential contamination issues. In their study, Bradley *et al.* used 495 isolates as a training set, and then validated the tool with a collection of WGS data from a further 471 isolates. The tool utilizes the Gordon *et al.* database [49] with some additional refinements and makes predictions for the same 12 antibiotics. Using the tool Bradley *et al.* were able to demonstrate sensitivity of 99.1% and specificity of 99.6% for the genotypic predictions in comparison to the phenotypes.

8.3.3 Problems found or anticipated

Genetic instability: The evolution of antibiotic resistance in *S. aureus* occurs by point mutation in core genes, and also by horizontal acquisition of resistance genes via mobile genetic elements (MGEs). In the studies conducted thus far, it is apparent that in some cases the relative genetic instability of MGEs carrying resistance genes can be a cause of discrepancy in the genotypic and phenotypic comparisons. One the most prone antibiotics is erythromycin. In *S. aureus* genes encoding erythromycin resistance are often found on plasmids, such as in the case of *erm*(C). The instability of the *erm*(C)-carrying plasmid has been well-documented, and it can be lost during passage of strains in the laboratory. In the study by Holden *et al.* discrepancies in erythromycin resistance prediction were thought to be due to the loss of the *erm*(C) during propagation and transfer between testing- and genomics- laboratories. Similar observations about the loss of the SCC*mec* element from the chromosome have also been made, which can account for discrepancies in cefoxitin resistance, albeit at a far lower frequency. In this case, evidence of the deletion of the whole SCC*mec* element carrying the *mecA* gene encoding cefoxitin resistance can be observed. Genetic stability of core components can also affect the observed resistance levels for some antibiotics in *S. aureus*. Hetero-resistance has

been observed whereby a subpopulation of cells in a cultured population exhibits a higher MIC than their 'siblings'. WGS of sub-populations has uncovered genetic variation associated with hetero-resistance to vancomycin, daptomycin and oxacillin [53, 54].

Gaps in the knowledge base: Whilst the studies that have been published thus far have generally demonstrated the effectiveness of *in silico* resistance prediction, there is evidence emerging that the performance for some antibiotics will be less accurate than for others. Aanensen *et al*. [55] recently conducted a blinded study of 308 isolates and used a range of antibiotics phenotypically tested by EUCAST (16 core antibiotics tested for all isolates: penicillin, cefoxitin, ciprofloxacin, moxifloxacin, amikacin, gentamicin, tobramycin, erythromycin, clindamycin, tetracycline, tigecycline, fusidic acid, linezolid, mupirocin, rifampicin, trimethoprim; and additionally teicoplanin, vancomycin, daptomycin tested for MRSA isolates). Overall the total performance of the *in silico* prediction was in line with previous studies, with 98.6% concordance, although for some antibiotics such as amikacin (92.5% concordance) and teicoplanin (97.5% concordance) the *in silico* prediction proved less effective.

For some antibiotics there are clearly gaps in the knowledge base for the genetic basis of resistance that require further investigation. For example, in the case of glycopeptides, such as vancomycin, the multiplicity and diversity of mutational changes linked with increased MICs confounds accurate prediction, where gene acquisition and mutations in an array of core genes have been characterized [56-58].

Laboratory variation: Discrepancies between the genotype and phenotype in some studies have been revealed to be laboratory artefacts and errors, where phenotype re-testing led to concordance. Technical variation in some of the tests is also a possible contributing factor in mis-matches. For example, in the study by Aanensen *et al.* it was noted that in the case of five isolates that had incorrect *in silico* predictions for mupirocin, the inhibition zones (29 mm) for these isolates were close to the breakpoint (S ≥30 mm), suggesting that growth conditions were potentially influencing the results.

8.3.4 Summary

The *in silico* prediction of AMR susceptibility for *S. aureus* is effective for most clinically relevant antibiotics. There are, however, some antibiotics that are more challenging to make predictions for, and further investigation is required to characterize the genetic and phenotypic basis of resistance.

8.4 Streptococcus pneumoniae

8.4.1 Background

Streptococcus pneumoniae is a highly clinically important community pathogen, in which genetic detection of acquired resistance is particularly challenging because most resistance results from mosaicism or mutations in chromosomally-encoded genes [59]. The most clinically important groups of antimicrobials with activity against pneumococci are the β-lactams, macrolides, tetracyclines and newer fluoroquinolones (for example, moxifloxacin). WGS-based approaches have been used for characterizing resistance mechanisms for several of these antimicrobial groups. However, no specific user-friendly database has been developed thus far, and WGS has mostly been used to study new mechanisms, and not in the context of predicting phenotypic resistance from whole genome data.

8.4.2 Published studies

For β-lactams, resistance is mostly mediated through the development of mosaic genes encoding altered penicillin-binding proteins (PBPs), as a result of intra- and interspecies DNA transfer by natural transformation [60]. Variants of PBP2x, PBP2b and PBP1a are considered most relevant to penicillin resistance in pneumococci. Still, there have also been reports of non-PBP-mediated resistance, such as enrichment in branched-chain muropeptides and mutations in genes encoding other enzymes involved in the peptidoglycan synthesis [61]. Fani *et al* used WGS to study mutants selected for penicillin resistance by step-wise penicillin increments until reaching a final MIC of 2 mg/L [61]. Sequencing was done by the 454 platform, generating a genome assembly of 28x coverage, with 97% of the reads assembled into 78 large contigs. Comparative sequence analysis identified mutations that were confirmed by Sanger sequencing. PBP2x mutations were shown to be important, but the relationship between genotypic and phenotypic resistance was complex, with a mutated iron transport system found in several of the resistant mutants.

Later work by the same research group proposed phenotypic reconstruction by whole genome transformation of penicillin-susceptible *S. pneumoniae* of known genetic backgrounds with genomic DNA from resistant clinical isolates [62]. This procedure would then be followed by WGS of the antimicrobial-resistant transformants. Selection of transformants was done by gradual increments of penicillin concentrations. The genome sequences of the fully-resistant and intermediate-step transformants were compared with the reference genome of the wild-type *S. pneumoniae* strains used in the experiments. The study confirmed the importance of mosaic PBP2x, -2b and -1a, but also suggested a role of PBP2a in some isolates. In another study, analysis of cefotaxime-resistant mutants revealed mosaic PBPs as well as mutations in other genes important for peptidoglycan synthesis [63]. Although these data suggest that predicting phenotypic β-lactam resistance based on WGS could be feasible in *S. pneumoniae*, there are so far no studies with clinical isolates to confirm this.

Macrolides also have clinically important activity against *S. pneumoniae*. Resistance is often mediated through RNA methylase (*erm*) or macrolide efflux (*mef*) genes, both of which are coupled to mobile genetic elements. One study was conducted in the USA with 147 pneumococcal isolates collected over an 18-year period both before and after the introduction of conjugate vaccines [64]. Genomes were then compared and mapping of macrolide resistance genes and their genetic environment was carried out. Resistance genes were detected in all strains, but the study was in no way investigator-blinded, as all included isolates were macrolide-resistant with phenotypic methodology.

For tetracyclines, Lupien *et al.* investigated mutants selected for resistance to tetracyclines [65]. Resistance to tetracycline in bacteria occurs through enzymatic inactivation or, more often, by active

efflux (via intrinsic or acquired pumps) or by ribosome protection. Resistance to tetracyclines in pneumococci is very common, and most often mediated by *tet* genes, which are found on mobile genetic elements. Lupien *et al*. used WGS to investigate not only genomic DNA, but also RNA sequencing libraries depleted of rRNA. RNA expression was compared in parent strains and mutants, identifying differentially expressed genes. Quantitative RT-PCR was used to confirm overexpression of some of the genes identified by comparison of sequenced mRNA in mutants and wild-type strains. Gene ontology classification of genes whose expression is significantly altered in *S. pneumoniae* thus seems like a feasible way of studying new chromosomal resistance mechanisms, although this approach has not been used in clinical isolates. Finally, whole genomic DNA transformation combined with WGS has also been used to study isolates with resistance to ciprofloxacin [66]. In addition to identifying efflux (using qRT-PCR) and QRDR-mutations, the methodology could also point to the potential role of mutations in drug transporters and redox enzymes in ciprofloxacin resistance.

8.4.3 Summary

A number of mechanistic studies have been carried out with laboratory mutants including whole genome transformation. These studies have shed light on a number of putative new mechanisms, but there is at present a lack of studies of the utility of WGS for predicting phenotypic resistance to antimicrobials used in the treatment of *S. pneumoniae*.

8.5 Neisseria gonorrhoeae

8.5.1 Background

Gonorrhoea is a global public health concern with the World Health Organization estimating 106 million cases every year [67]. *Neisseria gonorrhoeae* has developed resistance to every class of antimicrobial used to treat infections caused by the organism [68]. The emergence of multidrug resistant *N. gonorrhoeae* has led the Centers for Disease Control and Prevention (CDC) to classify it at an urgent threat level requiring serious public health attention [69]. Empiric combination therapy with intramuscular ceftriaxone and oral azithromycin (AZI) is recommended for first-line treatment in North America and Europe [70, 71]. Hence, this report focuses on the mechanisms of resistance to these antibiotics and the use of WGS to predict cephalosporin and AZI resistance.

CLSI has set the clinical breakpoint for susceptible isolates for both ceftriaxone and cefixime at ≤0.25 mg/L, although no breakpoint for resistance has been established [72]. EUCAST established a clinical resistance breakpoint for ceftriaxone and cefixime at >0.125 mg/L [73]. However, treatment failures with cephalosporin monotherapy have recently been observed in a number of countries in cases where the cefixime MICs of the infecting gonococci were as low as 0.032 mg/L [74-78]. Although there are no AZI breakpoints established by the CLSI, susceptibility and resistance breakpoints have been set by EUCAST at ≤0.25 mg/L and >0.5 mg/L, respectively [73]. In addition, the CDC has used a resistance breakpoint to AZI of ≥2 mg/L [79].

Numerous genetic mechanisms exist in *N. gonorrhoeae* for the development of elevated MICs to the extended-spectrum cephalosporins (ESCs; ceftriaxone and cefixime). Alterations in *penA*, which encodes penicillin-binding protein 2 (PBP2), have been described either through amino acid alterations (A501, G542, P551) or via the acquisition of a *penA* mosaic allele, which contains segments of *penA* from non-gonococcal *Neisseria* species [80-82]. These alterations confer reduced susceptibility to ESCs mediated by reduced binding to PBP2 [83]. Up-regulation of the MtrCDE efflux pump via a deletion in the promoter at -35 (A-del) or alterations of the MtrR repressor protein at positions G45D and A39T have also been associated with decreased susceptibility to the ESCs [80, 84]. A third mechanism for increased MICs to ESCs involves alterations in the PorB1b porin at amino acid positions G120 and A121. These permeability changes may reduce entry of ESCs into the cell leading to reduced susceptibility [80, 84].

8.5.2 Published studies

Only a few published studies have used WGS to examine the phenotypic and genotypic antimicrobial resistance patterns observed in *N. gonorrhoeae*. The first such study examined the genomes of 236 isolates of *N. gonorrhoeae* collected in the USA from 2009-10, and included 118 isolates with decreased susceptibility to cefixime (MIC ≥0.25 mg/L) [85]. The mosaic *penA* XXXIV allele was present in 115 of 118 isolates with decreased cefixime susceptibility and one other isolate carried a novel mosaic allele termed XLI [86]. No other alleles were clearly associated with reduced susceptibility to cefixime. A second study examined the genomes of 76 *N. gonorrhoeae* from numerous countries [87]. The mosaic *penA* XXXIV allele once again had the best positive predictive value, with this locus detected in six of seven cefixime-resistant isolates. Mutations in the *mtrR/mtrCDE* operon promoter region, and *penB* gene did not have such a strong predictive value, being found in only 2 of 7 and 3 of 7 cefixime-resistant isolates, respectively. A third study applied WGS to 169 Canadian isolates of *N. gonorrhoeae* with various antimicrobial resistance patterns [88]. There were 67 isolates with ceftriaxone MICs ranging from 0.125 to 2 mg/L. Of these, 40 (59.7%) harboured the *penA* mosaic, and all but one isolate had either *porB* mutations or the *mtrR* promoter

mutations. Of the remaining 27 isolates without the *penA* mosaic, all had *porB* mutations and only a single isolate did not contain a mutation in the *mtrR* promoter. However, when isolates with low MICs to ceftriaxone (<0.032 mg/L; n=65) were examined, one isolate (1.5%) was found to have the *penA* mosaic, 33 (50.7%) isolates contained *mtrR* promoter mutations, and 5 isolates also contained mutations in *porB*.

A second highly clinically significant phenotype that has been examined for *N. gonorrhoeae* is azithromycin resistance. The rates of AZI resistance in *N. gonorrhoeae* have been increasing in recent years in many countries, and the emergence of high-level resistance (MIC ≥256 mg/L) has been reported from Scotland, the United Kingdom, Argentina, Canada, and the USA [89-93]. The genetic mechanisms of resistance to AZI in *N. gonorrhoeae* include: accumulated changes in the four different alleles of the 23S rRNA genes; the presence of an 23S rRNA methylase encoded by *erm*(A)*, *erm*(B), *erm*(C) or *erm*(F); mutations in *rplD* and *rplV*; as well as the *penB* and *mtr* operon genes described above for cephalosporin resistance [94-96].

Although reports using WGS-based detection of these mechanisms are limited, there have been attempts to compare the AZI resistance phenotype with the WGS genotype and these are discussed below.

8.5.3 Problems found or anticipated - Gaps in the knowledge base

Ezewudo *et al.* examined two isolates that were resistant to AZI and found that one isolate contained the 23S rRNA mutation whereas the other did not contain any of the known mutations examined [87]. In a second study, involving the WGS analysis of 213 Canadian AZI-resistant isolates, 23S rRNA mutations A2045G and/or C2597T, disruptions in the *mtrR* promoter, or the presence of *erm*(C) were strongly associated with phenotypic resistance [88]. Seventy *N. gonorrhoeae* only contained the *mtrR* -35 deletion, and of these 21 were susceptible to AZI suggesting other potential unknown mechanisms conferring resistance may exist.

Although there is strong association with *penA* mosaic alleles and reduced susceptibility to ESCs, caution should be taken with this dataset as it relates to predicting the ESC MIC phenotype. The study by Demczuk suggested reduced susceptibility to ceftriaxone remains complex, involving additional genetic markers [88]. In addition, there has been a report of the mosaic *penA* allele XXXVIII displaying a susceptible phenotype to ESCs [97].

8.5.4 Summary

In summary, inferring ESC and AZI resistance in *N. gonorrhoeae* is possible with a high probability if certain genetic markers are detected in WGS data. However, the elevated MICs for some isolates result from combinations of multiple genetic changes, and further mechanisms for resistance have yet to be elucidated. Hence, predicting resistance to these antimicrobial agents can be problematic. Additional studies are required before the use of WGS can be advocated for use on a routine basis to predict resistance for these antimicrobial agents.

8.6 Mycobacterium tuberculosis *complex*

8.6.1 Background
Tuberculosis (TB) is caused by members of the *Mycobacterium tuberculosis* complex (MTBC) and, more rarely, by *Mycobacterium canettii* [98]. MTBC is monomorphic and strictly clonal and antibiotic resistance is therefore only caused by chromosomal changes. These are single nucleotide polymorphisms (SNPs) in the vast majority of cases, but small in-frame insertions/deletions (indels) in essential resistance genes and large indels in non-essential genes are also possible [99, 100]. By contrast, *M. canettii* is genetically diverse and displays evidence of recombination and lateral gene transfer [98]. The resistance mechanisms in this species are unclear [101-103].

Phenotypic susceptibility testing for TB is expensive, technically challenging, and time-consuming owing to the slow growth rate of MTBC [104]. Current molecular AST assays only interrogate the most frequent mutations conferring resistance to a limited number of drugs. In theory, this limitation can be overcome by WGS. In practice, however, routine WGS of all TB cases is unlikely to be cost-effective if performed to predict antimicrobial susceptibility alone, despite the decreasing cost of WGS [105, 106]. This is primarily because rates of drug resistance are low (typically below 5%) in countries that can afford WGS [107]. Instead, the main driver for the introduction of WGS for all TB cases will be the desire to replace traditional typing techniques by the ultimate resolution provided by WGS to improve outbreak investigations [108-110]. Moreover, WGS from the initial liquid culture can replace current techniques for pathogen identification (WGS directly from a clinical sample is technically possible, but less reliable and prohibitively expensive for clinical practice at the moment [111, 112]). In this scenario, the sequence data can also be used to rule-in resistance at no additional cost [106].

8.6.2 Published studies
Several large-scale studies reporting WGS of MTBC have focused on elucidating the genetic basis of drug resistance, which complement more focused efforts [106, 113-122]. Nevertheless, major gaps in this area remain, as discussed below. Several tools have been developed to automate WGS data analysis, although most do not meet clinical standards as they do not provide the necessary record keeping capabilities, have not been evaluated extensively, and often there are no plans to accredit them [51, 106, 123-127]. One exception is the analysis infrastructure that is currently being evaluated by Public Health England for clinical accreditation and routine use of WGS for TB in the UK [106].

8.6.3 Problems found/anticipated – Gaps in the knowledge base
As discussed briefly in section 9 above, three main challenges limit the utility of genotypic AST compared with phenotypic alternatives.

1) Systematic errors due to inadequate limit of detection of WGS: AST for TB is usually done on a significant fraction of the primary culture, as opposed to just 1-3 colonies from a primary agar plate, which is the approach taken for the vast majority of other clinically relevant bacterial pathogens. Resistance is deemed clinically significant if resistant organisms are present at or above a critical proportion, which is set at 10% for pyrazinamide and at 1% for the remaining drugs, and gold standard phenotypic AST (i.e. the proportion method) is calibrated to detect resistance at this limit [128, 129]. The limit of detection of traditional genotypic AST methods is poorer and depends on the assay and specific mutation, which can result in systematic false-negative results for strains with low-level hetero-resistance [130-132]. The magnitude of this source of error depends on a several factors, including the level of mixed infections with unrelated strains that have different susceptibilities, and

the proportion of resistance that is transmitted vs. resistance that evolves during treatment [133]. In practice, these factors vary between patient groups, geographic settings, and antibiotics. Moreover, the precise mechanism of resistance is relevant in this context. Low-level hetero-resistance SNPs can be identified by increasing the sequencing coverage, although this makes WGS prohibitively expensive in a clinical context at the moment [106, 134]. This strategy is not an option for hetero-resistance indels, particularly large ones, because of the limited read lengths of the most commonly used platforms for clinical sequencing, coupled with the fact that most analysis algorithms are not optimised to identify indels [135].

2) Systematic errors due to poorly defined breakpoints for phenotypic AST used as the gold standard for the validation of WGS-based AST: The clinical breakpoints (known as critical concentrations (CCs) in the tuberculosis field) are currently defined by the CLSI and the WHO [136, 137]. Clinical breakpoints should be defined by committees based on representative MIC distributions, pharmacokinetic/pharmacodynamic (PK/PD) data and, ideally, clinical outcome data, which, for a variety of reasons, are difficult to obtain for TB drugs [10, 138]. In practice, however, the evidence used to set the current CCs is not clear and emerging data from systematic MIC testing and PK/PD studies indicate that the CCs for some agents need to be revised [137-142]. In addition, no CCs exist for some second-line drugs, such as clofazimine [137]. Finally, the reproducibility of some phenotypic AST assays is poor, particularly for second-line drugs [137]. A new EUCAST sub-group has been launched to set breakpoints for TB using the same rigorous and transparent methodologies used for other pathogens [10, 143].

3) Incomplete understanding of the genotypic basis of phenotypic resistance: The Bill & Melinda Gates Foundation has funded the Foundation for Innovative New Diagnostics (FIND) and the Critical Path to TB Drug Regimens (CPTR) to create a clinical grade database, akin to the HIV Stanford resistance database, to enable the interpretation of TB WGS data for AST [144]. As part of this effort, FIND and CPTR will together with the World Health Organization, New Diagnostic Working Group of the Stop TB Partnership, the US Centres for Disease Control and Prevention and the National Institute of Allergy and Infectious Diseases collect and analyse WGS with associated phenotypic AST results for tens of thousands of strains to gain sufficient confidence in the association between particular mutations and resistance, as even the largest WGS studies published to date have been underpowered and were not designed to achieve this goal [118, 122].

The complexity of this task depends on whether a resistance gene is essential or non-essential. In the former case, only a limited spectrum of resistance mutations is possible. Consequently, the correlation between genotype and phenotype should be relatively easy to resolve, provided that methodological problems such as poorly defined CCs are addressed (e.g. there is a near perfect correlation between genotype and phenotype for rifampicin resistance and *rpoB* mutations, although this can depend on the medium used for phenotypic AST) [145-147]. The situation with non-essential genes is more complicated. For genes that are non-essential *in vitro* as well as *in vivo*, it is impossible to study the genetic basis of resistance comprehensively given that there are too many possible resistance mutations. The best example of this type of resistance gene is *pncA*, which is responsible for the activation of the pro-drug pyrazinamide [148]. Any loss-of-function mutation in this gene can confer resistance and a wide variety of mutations are found clinically (e.g. more than 4,000 single codon changes are possible in *pncA* (excluding start codon changes and nonsense mutations), not all of which will cause resistance). By combining large datasets it is possible to distinguish resistance mutations from neutral polymorphisms, but novel mutations will continue to be discovered, albeit at a

lower rate over time [122, 149]. Similar considerations apply to the catalase-peroxidase KatG, which is required for the activation of the pro-drug isoniazid [150]. The gene encoding KatG has been found to be non-essential *in vitro*, but clinically most resistance mutations are selected against [151]. By contrast, the KatG S315T change only confers a low fitness cost [152]. Consequently, this mutation accounts for the majority of isoniazid resistance clinically [122]. The remaining isoniazid resistance is due to a large number of rare mutations that are impossible to study in their entirety.

8.6.4 Summary

Some of the aforementioned challenges to introducing and validating WGS for AST of TB can be overcome over time. For example, the ability of WGS to detect hetero-resistance will improve as the cost of sequencing decreases and read lengths improve. Similarly, the ongoing re-evaluation of CCs will likely resolve some of the current systematic differences between genotype and phenotype. Moreover, the pooling of large datasets will clarify the role of rare resistance mechanisms and the level of resistance conferred by different resistance mutations or mechanisms. For example, some low-level isoniazid-resistant strains due to *inhA* mutations remain treatable with higher doses of the drug and the same may apply for some strains with low-level resistance to new-generation fluoroquinolones (codon 90 mutations of *gyrA*) [136, 153, 154]. Yet, it is impossible to study the genetic basis of antibiotic resistance to all clinically relevant drugs comprehensively because of the large number of possible resistance mutations for some drugs. This means that WGS can mainly be used to rule-in resistance, as opposed to rule-out resistance. Nevertheless, this constitutes a significant improvement to current clinical practice, since WGS directly from the first positive culture would allow for established resistance mutations in key drugs to be identified rapidly, thereby allowing for regimens to be adjusted within days as opposed to weeks or even months for phenotypic AST [155]. Based on these results, reference laboratories could also immediately commence phenotypic AST for all remaining relevant drugs, including second-line drugs (which are usually only tested if resistance to first-line drugs is found, which introduces long delays). Consequently, WGS is unlikely to completely replace phenotypic AST for TB in the near future, but will result in less phenotypic testing over time and in more rapid identification of resistant isolates in many cases. However, it is likely that different countries will adopt their own policies in terms of how much phenotypic confirmation of genotypic results is required based on the resources available and the local rates of resistance.

8.7 Clostridium difficile

8.7.1 Background

Clostridium difficile is the leading cause of health-care associated diarrhoea, the severity of which may vary from mild and self-limiting symptoms to fulminant disease, including pseudomembranous colitis. Hospital outbreaks are occurring with an increasing frequency, and the most severe ones have been caused by hypervirulent *C. difficile* strains 027/NAP1, although other ribotypes (such as 078/NAP7&8) also seem to have the ability both to cause outbreaks and severe disease in affected individuals.

Acquired phenotypic resistance to tetracyclines, clindamycin, fluoroquinolones and rifampicin and corresponding resistance genes have frequently been reported in *C. difficile*, and moxifloxacin resistance is used as an epidemiological marker for hypervirulent strains, and for ribotype 027 (NAP1) in particular. However, resistance to the compounds that are used as primary therapeutics for *C. difficile* infection (i.e. vancomycin, metronidazole and fidaxomicin) is less common.

Phenotypic AST of *C. difficile* suffers from some drawbacks. Since anaerobic conditions are required, it is costly and time-consuming and therefore often not routinely performed in the clinical laboratory and the correlation between *in vitro* susceptibility and clinical outcome in the individual patient has not been thoroughly studied. In light of this, genotypic AST using WGS appears an attractive alternative. While resistance towards fidaxomicin has been associated with mutations in genes encoding RNA polymerase (*rpoB* and *rpoC*) or in the *marR* homologue CD22120, the mechanisms underlying resistance to vancomycin and metronizadole are less well defined.

Some of the main lineages of *C. difficile* contain a *vanG* locus which is expressed; however, this does not appear to play a role in resistance to vancomycin in *C. difficile* [156]. To date, no clinical isolates have been identified that are resistant to vancomycin. Two laboratory-derived vancomycin-resistant isolates have been described, one had a substitution mutation in the *rpoC* gene and the second had two mutations, one in *murG* (CD2725) and the second in a locus named CD3659 [157].

Nitroimidazole genes (*nimA-E*) associated with metronidazole resistance in other anaerobic species, including several species of the *Clostridium* genus, have not been described in *C. difficile*. The exact mechanism(s) behind reduced susceptibility to metronidazole in *C. difficile* still remains to be determined, although there have been several reports of strains exhibiting elevated MICs. Such reports of metronidazole resistance have all observed loss of the resistant phenotype after passaging or low temperature storage [158-160]. There has been one reported clinical isolate 027/NAP1 from Canada that initially had an unstable resistance phenotype, but after serial passage in the presence of metronidazole, the phenotype became stable [161]. Whole genome sequence comparisons between the resistant and reverted susceptible isolate revealed many SNPs between the two isolates [161]. Proteomic analysis of these isolates suggested a multifactorial response maybe associated with the high-level metronidazole resistance observed [162]. Thus, the genetic mechanism for metronidazole in *C. difficile* remains elusive.

8.7.2 Gaps in the knowledge base

To date there have not been any publications comparing large scale phenotypic to WGS-based AST for *C. difficile*.

8.7.3 Summary

AST of *C. difficile* using WGS could be a useful tool, both for guiding the choice of treatment of the individual patient and for epidemiological purposes. However, the knowledge gaps regarding the mechanisms underlying resistance to several of the first-line treatment options pose a great challenge. Studies comparing WGS-based approaches with phenotypic testing are needed and future work on resistance mechanisms to frontline antimicrobials are required.

8.8 Acinetobacter baumannii *and* Pseudomonas aeruginosa

8.8.1 Background

Among non-fermentative Gram-negative bacteria *Pseudomonas aeruginosa* and *Acinetobacter baumannii* are important pathogens due to their ability to cause a variety of opportunistic infections, persist in the hospital environment and acquire antimicrobial resistance [163-165]. Genomic studies have shown that both *P. aeruginosa* and *A. baumannii* are associated with high genomic diversity and gene content due to frequent transfer/acquisition of mobile genetic elements, mobilization of IS-elements, IS-mediated deletions and genome-wide homologous recombination [166-172]. In addition to being intrinsically resistant to many antimicrobials [11], the increase in multidrug-resistant and in particular carbapenem-resistant *P. aeruginosa* and *A. baumannii* has resulted in infections caused by extensively drug-resistant (XDR) and even pandrug resistant (PDR) isolates with limited or no validated therapeutic options [166, 173-175]. High-risk clones of both species have been identified to be responsible for the spread of relevant resistance genes, such as carbapenemases [175, 176].

In both *P. aeruginosa* and *A. baumannii* acquired resistance genes are associated with various horizontally-acquired resistance elements, although the majority of acquired resistance genes exist as gene cassettes in integron structures [175, 177]. In *A. baumannii* these resistance elements are frequently clustered in antibiotic resistance islands (AbRs), which vary in structure, size and genomic location [167] and plasmid-borne resistance genes also contribute to the evolution of antibiotic resistance in both species [175, 178]. In addition, both species and in particular *P. aeruginosa* have an extraordinary capacity for modification of endogenous genes affecting functions such as membrane permeability, efflux, expression of intrinsic β-lactamases, antibiotic targets, and regulatory genes contributing to multidrug resistance [164, 175, 179].

8.8.2 Published studies

To date, few comprehensive studies have investigated the concordance between phenotypic AST and WGS-based resistance prediction for *P. aeruginosa* or *A. baumannii*. Kos *et al.* related phenotypic susceptibility data for meropenem, levofloxacin and amikacin to the genome sequences of ~390 clinical *P. aeruginosa* isolates [180]. The results showed that the sensitivity and specificity for genotypic inference of meropenem and levofloxacin resistance was 91% and 94%, respectively. In contrast, a genotypic marker for amikacin resistance was only identified for 60% of the amikacin non-susceptible isolates. In addition, 30 of 283 amikacin-susceptible isolates were found to harbor genes associated with amikacin resistance. This is in contrast to a study by Wright *et al.*, where a strong association between amikacin resistance and the presence of *aphA6* and *armA* genes was observed in a collection of 75 clinical *A. baumannii* isolates [181].

Although there is a lack of phenotypic/genotypic comparison studies with respect to prediction of clinical resistance, several genomic studies have been performed for epidemiological purposes and to decipher mechanisms of resistance to various antibiotics in selected resistant isolates [167, 182, 183]. These studies are important to identify both intrinsic and acquired genotypic resistance determinants associated with resistance to various antibiotics. For instance, recent investigation of isogenic colistin-susceptible and -resistant isolates of both *P. aeruginosa* and *A. baumannii* from single patients revealed novel determinants associated with colistin resistance [184, 185]. Further, the use of WGS as a tool to predict antibiotic resistance has been recently studied using 178 *A. baumannii* bacterial genomes to evaluate the antibiotic resistance gene database ARG-ANNOT showing that such approach could be used as a routine test [13].

8.8.3 Problems found or anticipated - Gaps in the knowledge base

Although prediction of antimicrobial susceptibility/resistance based on the presence of a relatively limited number of acquired resistance genes and chromosomal resistance-associated mutations might give high sensitivity and specificity, the major challenge with respect to both *P. aeruginosa* and *A. baumannii* lies in the identification or prediction of resistance due to chromosomal alterations caused by modification of expression levels, particularly with respect to efflux pumps, outer membrane proteins or intrinsic β-lactamases.

For instance, resistance to β-lactams in *A. baumannii* can occur due to the insertion of elements such as IS*Aba1* and IS*Aba125* upstream of the intrinsic β-lactamase genes bla_{ADC} and bla_{OXA-51}, increasing the expression of these genes and consequently resistance to cephalosporins and carbapenems, respectively [186, 187]. Further, IS elements have been implicated in disruption of genes encoding outer membrane proteins that contribute to resistance and in particular to β-lactams [164]. Screening of genomes for IS elements in close association with resistance-associated genes as well as for gene loss will pose a significant challenge.

For *P. aeruginosa* the challenge is expected to be even greater due to the plethora of genes associated with the species' intrinsic resistome that when altered or alterations of regulatory genes can confer resistance to several antibiotics, even from different antibiotic classes [179]. Further, alterations to one or mostly several of these mechanisms might be required to achieve clinical resistance (e.g. combination of decreased porin expression, increased efflux and/or increased β-lactamase expression) [188]. Altered expression of genes could be overcome by investigation of gene-expression analysis by RNA sequencing, however specific studies on *P. aeruginosa* indicate that the correlation between expression of genes due to exposure of sub-MIC concentration of antibiotics and genes implicated in intrinsic resistance is not always clearly observed [189].

8.8.4 Summary

In general these studies showed that prediction of resistance based on the detection of known acquired resistance genes and resistance-conferring mutations in drug targets can be used to investigate the phenotype-genotype relationship. However, additional comparative studies between phenotypic and genotypic methods using representative strain collections of *P. aeruginosa* and *A. baumannii* are required to evaluate the possibility of confidently predicting antimicrobial susceptibility/resistance by WGS. Further, for both species a greater understanding of the contribution with respect to alterations of the intrinsic resistome in terms of clinical resistance is required. This will require not only WGS but also knock-out and complementation studies of deleted/mutated determinants in a comprehensive framework.

9. The epidemiological implications of using WGS

The epidemiology of AMR is determined by the spread of the host organisms harbouring resistance genes, and the spread of the resistance genes by different routes of horizontal gene transfer.

Classical methodologies to study the epidemiology include strain genotyping with a variety of methods with a large variation in reproducibility and discriminatory power. This includes techniques such as MLST, PFGE, VNTR, MLVA, AFLP, Eric-PCR etc. [12]. Resistance genes can be identified by micro-array approaches, PCR and sequence analysis. A variety of molecular techniques are needed to identify and characterize epidemiologically-relevant mobile genetic elements involved in the horizontal gene transfer of AMR genes, such as plasmids, conjugative transposons or genomic islands. Dedicated PCRs and sequencing are needed to identify the genetic environment of the AMR genes such as integrons and/or transposons and insertion sequences, which is crucial to understand the epidemiological behaviour of specific AMR genes [190].

PCR-based replicon typing is most commonly applied for plasmid characterization in Enterobacteriaceae. Relaxase typing is more comprehensive and phylogenetically more informative, but is less discriminatory within the Enterobacteriaceae, where the major concerns are at present. Plasmid MLST and similar techniques such as Double Locus Sequence Typing or RFLP are used to subtype plasmids. In addition toxin/anti-toxin systems present in plasmids in Enterobacteriaceae may be key epidemiological determinants. Because of its complexity, plasmid epidemiology is currently beyond the capabilities of standard microbiology laboratories and is labour-intensive even for the reference laboratory.

WGS opens a world of opportunities for enhanced (molecular) epidemiology of AMR because, in principle, all essential information needed to study the epidemiology of AMR will be available in the sequences obtained. WGS is particularly effective for identifying and characterizing clonal distribution of monomorphic species such as *S. aureus* [55] and *M. tuberculosis* [191] and high-risk clones such as CTX-M-15-producing *E. coli* O25:H4-ST131 [192] or KPC-producing *Klebsiella pneumoniae* (e.g. ST11, 258 and 512) [193]. Importantly, WGS provides high-resolution typing information making most if not all of the traditional molecular typing approaches redundant. With its potential for an objective assessment of the gene content such as presence of absence of resistance genes of particular public health importance, multicentre surveillance approaches would greatly benefit from the reporting of genomic resistance markers obviating the need to rely on phenotypic AST profiles of doubtful inter-laboratory reproducibility.

However, WGS also has its weaknesses. As an example, WGS is weak in managing direct repeats and insertions in plasmids and current bioinformatic cleaning often omits those from contigs. As a result, short-read WGS data can be misleading if studying plasmid-mediated outbreaks in which a broad host-range plasmid is moving freely between different species, in all of which it has a different phenotype.
To be able to analyse sequence output rapidly and identify all information needed for epidemiology as listed above, the AMR genes and plasmid types need to be determined in these sequences using genomic databases such as ResFinder or PlasmidFinder [194]. For pMLST-typing, identification of genomic islands, insertion sequences and transposons the genomic database PLACnet is available [195]. *In silico* arrays or PCRs are also commonly applied but short-read sequencing (e.g. as obtained with Illumina) is generally not sufficient to study the genetic environment of AMR genes.

Transfer (by transformation or conjugation) of a plasmid of interest into a 'workhorse' bacterium with a known genetic background, with or without subsequent plasmid enrichment during DNA extraction, will facilitate complete and correct plasmid sequence assembly. Long-read or single cell sequencing (e.g. by PacBio or Oxford Nanopore) may be necessary [196, 197], either alone or as a 'scaffold' for high-coverage short-read data (e.g. Illumina). However, both are beyond the capabilities of most standard microbiology laboratories, and plasmid handling is relatively labour intensive. This may still leave specialised annotation problems, although direct annotation grammars can be helpful with these [198, 199].

Since different users may have different demands for WGS data, a tiered approach can be applied:

Diagnostic information
- Rapid identification of a targeted set of AMR genes may provide important information at clinical level. The output could vary from answering a specific dichotomous question (e.g. does a sample contain an ESBL-producer or an MRSA) or a more complete resistance/susceptibility profile [20, 200, 201]. The output should be based on the bacterial species and the information required for the clinicians.
- The question of the positive and negative predictive value of WGS will be important, although their usefulness will depend on the targets to be identified, their diversity and prevalence in the gene pools [202, 203].

Epidemiological information
- Identification of genes and subtypes (e.g. $bla_{CTX-M-15}$ versus $bla_{CTX-M-1,}$ or $bla_{CTX-M-3}$, or *mecA* versus *mecC)* may be important for outbreak management and infection control.
- *In silico* strain typing in cultured organisms based on:
    a. MLST (7 or more gene targets) for evolutionary relatedness
    b. Genomic Islands such as SCC*mec*, SGI1, SXT
    c. Single nucleotide polymorphisms (SNPs), insertions/deletions (INDELS), and large structural DNA rearrangements (e.g. for tracking outbreaks/ mapping transmission chains)
- *In silico* plasmid typing by mapping to a reference database (*in silico* microarray)
    a. Subtyping of plasmids by in silico PCR (pMLST, DLST, RST)
- Phylogenetic analysis of the total sequence output

WGS approaches can be used to track markers from the chromosome (e.g. MLST), from plasmids (e.g. incompatibility markers, or post-segregational killing / toxin/antitoxin markers) and individual genes – barcoding to tie them together in individual isolates by WGS using third generation approaches such as PacBio or Nanopore.

The ideal method will provide sufficient depth and coverage to answer all these questions, but will vary with the starting material: metagenomic approaches to DNA extracted directly from clinical samples will require a considerably higher number of sequencing reads than for analysis of a microorganism in pure culture.

At the time of writing the availability of reference databases for epidemiological questions remains limited both in the number of species and typing methods that are represented. Further development in this area will be crucial.

10. Clinical & wider impacts

Routine use of WGS in diagnostic and public health laboratories holds the promise of a revolution in the identification, typing, susceptibility testing and determination of pathogenicity of potential pathogens [104].

At present, at the initiation of antimicrobial chemotherapy or thereafter when definitive therapy is selected based on phenotypic AST, clinicians have no routine data provided on the likely pathogenic potential of any strains of pathogens isolated. Future data from WGS linking pathogenicity determinants to adverse clinical outcomes for certain highly pathogenic strains may have significant impacts on chemotherapy - perhaps by identifying those at higher risk of infection-related complications, those who may require more aggressive or combination chemotherapy or prolonged intravenous courses of antibiotics. Conversely reassurance that some potential pathogens are of low pathogenicity may allow for shorter duration therapy, oral therapy, less intensive patient monitoring, fewer investigations and perhaps earlier discharge. Such approaches are, at present, almost entirely speculative but may have a greater clinical impact than the work done so far on the value of WGS in predicting phenotypic susceptibility or resistance when tested by conventional methodologies.

At present, proof-of-principle studies have been completed for WGS on common pathogens already isolated in pure cultures and hence most data related to WGS for predicting antibiotic susceptibility assume an initial culture step. This is an obvious limitation in terms of speed of diagnosis, compared with direct testing of specimen material. To date common pathogens such as *E. coli* and *K. pneumoniae* [14, 19, 21], *S. aureus* [46-49]; *S. pneumoniae* [64, 65] and *P. aeruginosa* [180] have had WGS related to conventional susceptibility with some success. In addition, there is also work to show potential in *Salmonella* spp. [20], *Acinetobacter* spp. [13, 181], *N. gonorrhoeae* [88] and *M. tuberculosis* [122, 155]. Whilst for other organisms including *C. difficile* little has been demonstrated to date. The most work remains where there are significant gaps in the resistance mechanisms knowledge base.

However, at present, we lack a clear understanding of how antimicrobial susceptibility data can be generated from WGS in a timely way for incorporation into clinical care pathways and what the likely clinical impacts will be. In particular, we do not fully understand the barriers or facilitators to increased clinical use assuming technical problems can be overcome. The costs of routine delivery of WGS data to predict AST have not been balanced against potential financial savings across the patient care pathway or the clinical impacts. At present, even feasibility studies to start answering these questions have not been reported.

One obvious potential advantage of WGS in AST is the increased speed of information flow even if at present WGS would depend on an initial culture step. Increased speed of diagnosis has been identified as a way of improving antimicrobial stewardship and patient outcomes. If WGS could be made to deliver pathogen identification and predict susceptibility for common pathogens within 8 hours of initial culture it may offer enough to impact on measures of patient outcome and antibiotic drug use to justify higher costs within the laboratory. However, the longer it takes for data to become available then the poorer the potential clinical impacts.

Thought needs to be given as to how WGS data will be presented to end-users. It is possible we will move from dichotomous reporting of susceptible or resistant, to reporting the probability of an isolate being susceptible or resistant based on pre-test probabilities (perhaps different in different hospitals or areas within a hospital), the completeness of our genetic database for a particular pathogen and the

presence or absence of resistance determination as determined by WGS. We may even be able to give measures of confidence to these predictions. Such approaches will require significant staff education and evaluation as it is not clear how prescribers would respond to such data.

As >95% of pathogen identification and susceptibility testing in present clinical microbiology laboratories is based on 20 bacterial species and a limited number of antibiotics it may not be necessary to cover all possibilities to provide useful data rapidly, but rather we might focus on a limited number of antimicrobials for each pathogen initially and let more detailed data become available later.

At present the use of WGS outside reference or research laboratories to determine antimicrobial susceptibility has not been tested. Preliminary data are promising and feasibility studies need to be conducted in a more clinical environment. It is likely that WGS will first be used as a tool to predict antimicrobial susceptibility in public health microbiology laboratories in the coming years with subsequent use closer to the patient in order to predict susceptibility in pathogens such as *M. tuberculosis* before its wider application in diagnostic laboratories.

11. Conclusions & Recommendations

This EUCAST Subcommittee report on the role of WGS in AST of bacteria has reviewed the state-of-the–art as a first approach, it refers to almost 200 published works and describes where we are at the time of writing (late 2015 to early 2016). Despite the volume of published literature already available we concluded that, at present, there are still insufficient data to present a definitive document on the topic. Instead, **this report is intended to form a baseline discussion document, which can be revisited and updated at regular intervals** (likely every 18-24 months). This will be important as sequencing technologies become more affordable and more widely applied. This first version will provide the first point against which to compare and assess future progress in the area.

Our Subcommittee is aware of many ongoing, as yet unpublished studies of phenotypic/genotypic AST concordance and it is certain that the amount of available data will increase in the near future. However, the quality of those data needs to improve and to be assured via more rigorous and 'standardised' approaches to data analysis. Bacterial AST is a fundamental activity undertaken in any microbiology laboratory, but it is important to appreciate that the MIC or zone size diameter measured reflects more than gene presence / absence; these values reflect multiple and complex interplays between different systems including cellular permeability, influx/efflux, target availability and binding as well as enzymatic expression levels and activities. So **there are many challenges in gathering and assessing evidence to consider whether AST can be replaced by a genotypic method such as WGS, which does not assess bacterial growth in the presence of antibiotics**.

At the present time, **WGS-based analyses cannot yield an inferred MIC or zone diameter**. Hence the potential utility of WGS-based approaches for AST must be considered at the level of detecting gene presence or absence. We will need more powerful bioinformatics tools in future if we seek to make inferences about antibiotic susceptibility based on combinations of multiple different genes or contributory mutations. Furthermore, WGS does not directly inform gene expression levels. Although other technologies can do so, e.g. RNASeq, it seems unlikely that these will find a place in a clinical laboratory before WGS.

It is our recommendation that **the primary AST comparator for WGS-based prediction should be the ECOFF**, wherever possible, in order to assess WGS-inferred 'antibiograms' (based on gene positivity) against phenotypically-defined categories of wild-type or non-wild-type. Adoption of ECOFFs as the primary comparator would make meta-analysis across different publications simpler, since comparison of data would not be subject to confounders such as differences in breakpoints adopted. Nevertheless, demonstrating concordance with **interpretation against clinical breakpoints will ultimately be necessary for the use of WGS-based testing to guide clinical decision making**, but this will likely be more difficult to demonstrate for all bugs and drugs. For this reason, assessing WGS-derived data against clinical breakpoints represents a tougher challenge, but should be encouraged as a secondary comparator and should ideally be done using the same data sets as are used for ECOFF-based assessments.

The challenges of harmonising antimicrobial susceptibility/resistance breakpoints across multiple parallel and independent national and international systems have been ongoing for >50 years, and we still lack a globally harmonised system. When considering the introduction of WGS-based approaches, we need to balance the needs of clinical laboratories, where standardised and validated procedures are needed in order to meet accreditation standards, with the need for intellectual and

innovative academic challenge, which drives many of those who generate bioinformatics tools. We recommend that **there should be international agreement on the most appropriate and effective principles to facilitate early standardisation and harmonization of analytical approaches and interpretative criteria for WGS-based predictive AST**. However we also recommend at the present time that **we need to be pragmatic and must accept that bioinformatics algorithms will vary**, it is unrealistic to suggest a single analytical approach. We recommend that **different bioinformatics tools should perform to minimum standards and should be calibrated and equivalent in terms of the results generated.**

In order to facilitate such comparisons, we recommend that **performance of different bioinformatics tools should be calibrated against a single database of all known resistance genes / mutations**. There have been efforts and investments in this direction, but multiple solutions exist and are used, thereby confounding comparisons. Establishing a single database will ensure that there is parity of analysis and will facilitate measurement of comparative accuracies across different systems. Such a global reference database would need to be updated regularly, and must have strictly curated minimum standards for the inclusion of new resistance genes and mutations. An important function of a centralised database would be to control resistance gene nomenclature (since poor annotation can confound current analyses, where multiple 'hits' from searches may reflect inconsistent annotation of the same gene). The inclusion criteria for any new determinant would probably need to be set higher than those accepted for publication because strong evidence of causal association would maximize the predictive values of inferring AST phenotype from genotype.

The organisms considered in this report can be divided into three main groups in terms of the available evidence for predicting AMR using WGS. **Firstly**, at present most is known for *S. aureus* and *M. tuberculosis* and it is apparent that there is now momentum behind their deeper investigation. For a **second** group of organisms, including the Enterobacteriaceae (including Salmonella), initial studies have shown promise, but serve to highlight through poor concordance where gaps exist in the knowledge base about resistance mechanisms either in some genera or species or for some antibiotics. For a **third** group of organisms, including *S. pneumoniae, N. gonorrhoeae, P. aeruginosa, A. baumannii* and *C. difficile* it is apparent that more studies are required before we can even define the extent of the gaps in the knowledge base accurately. More focussed study and additional funding resources are needed as a priority to improve knowledge for the second and third of these groups.

Expansion of the knowledge base is a critical priority if WGS is to be considered seriously as a rival to phenotypic AST, better defining resistance determinants across all organisms. It seems likely that WGS may replace phenotypic testing 'soon' for surveillance purposes, where the low error rate has low impact. This would need to be phased to reflect the evidence base for the bug/drug being reported, and would require surveillance schemes to expand their inclusion criteria to accept WGS-inferred data. In reference laboratories, WGS-based AST may also be adopted 'soon', unless the reason for investigation relates to individual patient management, is for antibiotics or species shown to have poor genotypic/phenotypic concordance, or is to assess the activity of novel antibiotics.

Available published evidence does not currently support use of WGS-inferred susceptibility to guide clinical decision making. Such a paradigm shift would require large-scale education and behavioural change amongst microbiologists and prescribers. Since gene (or mutation) absence cannot always reliably predict susceptibility, robust evidence will be needed to show that the potential of genotypic tests for very major errors does not adversely impact on treatment outcomes. It seems likely that this

may first be considered for *M. tuberculosis*, where the speed of WGS-generated results offers advantage over traditional AST methods. However, even if the evidence can be generated and expectations changed, for most bacteria and in most countries **the current cost and speed of inferring antibiotic susceptibility from WGS data remain prohibitive to wide adoption in routine clinical laboratories** (in comparison with AST using antibiotic discs, for example). Nevertheless, as advances in the knowledge of polymorphisms associated with drug resistance, technology, data sharing and training become more widely available in high burden countries, sequencing technologies will be more attractive and cost effective as the cost of goods come down.

Finally, there may even be scope for WGS-based approaches to be used to better understand and improve some areas of phenotypic AST. For some antibiotics, there are technical challenges in measuring susceptibility in any way that meaningfully correlates with outcome. If WGS data could be correlated directly with outcome, then this revolutionary tool might aid development of improved criteria for interpreting phenotypic data.

12. Acknowledgements

13. Funding & conflicts of interest

14. References

1   WHO. *Antimicrobial resistance* http://www.who.int/mediacentre/factsheets/fs194/en/: World Health Organisation. 2013. http://www.who.int/mediacentre/factsheets/fs194/en/

2   Kahlmeter G, Brown DF, Goldstein FW, et al. European harmonization of mic breakpoints for antimicrobial susceptibility testing of bacteria. *J Antimicrob Chemother*. 2003; **52**: 145-148.

3   Sjolund M, Bengtsson S, Bonnedahl J, Hernandez J, Olsen B, Kahlmeter G. Antimicrobial susceptibility in escherichia coli of human and avian origin--a comparison of wild-type distributions. *Clin Microbiol Infect*. 2009; **15**: 461-465.

4   Kronvall G. Normalized resistance interpretation as a tool for establishing epidemiological mic susceptibility breakpoints. *J Clin Microbiol*. 2010; **48**: 4445-4452.

5   Kronvall G, Kahlmeter G, Myhre E, Galas MF. A new method for normalized interpretation of antimicrobial resistance from disk test results for comparative purposes. *Clin Microbiol Infect*. 2003; **9**: 120-132.

6   Turnidge J, Kahlmeter G, Kronvall G. Statistical characterisation of bacterial wild-type mic value distributions and the determination of epidemiological cut-off values. *Clin Microbiol Infect*. 2006; **12**: 418-425.

7   (ISO). IOfS. Clinical laboratory testing and in vitro diagnostic test systems. Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices. . *Part 1: Reference Method for Testing The in Vitro Activity of Antimicrobial Agents Against Rapidly Growing Aerobic Bacteria Involved In Infectious Diseases International Standard 20776-1*. Geneva, Switzerland: ISO 2006.

8   Jureen P, Angeby K, Sturegard E, et al. Wild-type mic distributions for aminoglycoside and cyclic polypeptide antibiotics used for treatment of mycobacterium tuberculosis infections. *J Clin Microbiol*. 2010; **48**: 1853-1858.

9   Arendrup MC, Kahlmeter G, Rodriguez-Tudela JL, Donnelly JP. Breakpoints for susceptibility testing should not divide wild-type distributions of important target species. *Antimicrob Agents Chemother*. 2009; **53**: 1628-1629.

10   Kahlmeter G. The 2014 garrod lecture: Eucast - are we heading towards international agreement? *J Antimicrob Chemother*. 2015; **70**: 2427-2439.

11   Leclercq R, Canton R, Brown DF, et al. Eucast expert rules in antimicrobial susceptibility testing. *Clin Microbiol Infect*. 2013; **19**: 141-160.

12   Sabat AJ, Budimir A, Nashev D, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill*. 2013; **18**: 20380.

13   Gupta SK, Padmanabhan BR, Diene SM, et al. Arg-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother*. 2014; **58**: 212-220.

14   Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012; **67**: 2640-2644.

15   Nordahl Petersen T, Rasmussen S, Hasman H, et al. Meta-genomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance. *Sci Rep*. 2015; **5**: 11444.

16   Sinclair A, Arnold C, Woodford N. Rapid detection and estimation by pyrosequencing of 23s rrna genes with a single nucleotide polymorphism conferring linezolid resistance in enterococci. *Antimicrob Agents Chemother*. 2003; **47**: 3620-3622.

17   McArthur AG, Waglechner N, Nizam F, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother*. 2013; **57**: 3348-3357.

18   Liu YY, Wang Y, Walsh TR, et al. Emergence of plasmid-mediated colistin resistance mechanism mcr-1 in animals and human beings in china: A microbiological and molecular biological study. *Lancet Infect Dis*. 2016; **16**: 161-168.

19   Stoesser N, Batty EM, Eyre DW, et al. Predicting antimicrobial susceptibilities for escherichia coli and klebsiella pneumoniae isolates using whole genomic sequence data. *J Antimicrob Chemother*. 2013; **68**: 2234-2244.

20   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 Chemoth*. 2013; **68**: 771-777.

21   Tyson GH, McDermott PF, Li C, et al. Wgs accurately predicts antimicrobial resistance in escherichia coli. *J Antimicrob Chemother*. 2015; **70**: 2763-2769.

22   Pecora ND, Li N, Allard M, et al. Genomically informed surveillance for carbapenem-resistant enterobacteriaceae in a health care system. *MBio*. 2015; **6**.

23   Hudson CM, Bent ZW, Meagher RJ, Williams KP. Resistance determinants and mobile genetic elements of an ndm-1-encoding klebsiella pneumoniae strain. *PLoS One*. 2014; **9**: e99209.

24   Doi Y, Hazen TH, Boitano M, et al. Whole-genome assembly of klebsiella pneumoniae coproducing ndm-1 and oxa-232 carbapenemases using single-molecule, real-time sequencing. *Antimicrob Agents Chemother*. 2014; **58**: 5947-5953.

25   Conlan S, Thomas PJ, Deming C, et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing enterobacteriaceae. *Sci Transl Med*. 2014; **6**: 254ra126.

26   Ruzin A, Visalli MA, Keeney D, Bradford PA. Influence of transcriptional activator rama on expression of multidrug efflux pump acrab and tigecycline susceptibility in klebsiella pneumoniae. *Antimicrob Agents Chemother*. 2005; **49**: 1017-1022.

27   Keeney D, Ruzin A, McAleese F, Murphy E, Bradford PA. Mara-mediated overexpression of the acrab efflux pump results in decreased susceptibility to tigecycline in escherichia coli. *J Antimicrob Chemother*. 2008; **61**: 46-53.

28   Keeney D, Ruzin A, Bradford PA. Rama, a transcriptional regulator, and acrab, an rnd-type efflux pump, are associated with decreased susceptibility to tigecycline in enterobacter cloacae. *Microb Drug Resist*. 2007; **13**: 1-6.

29   Cannatelli A, Giani T, D'Andrea MM, et al. Mgrb inactivation is a common mechanism of colistin resistance in kpc-producing klebsiella pneumoniae of clinical origin. *Antimicrob Agents Chemother*. 2014; **58**: 5696-5703.

30   Beabout K, Hammerstrom TG, Perez AM, et al. The ribosomal s10 protein is a general target for decreased tigecycline susceptibility. *Antimicrob Agents Chemother*. 2015; **59**: 5561-5566.

31   Doumith M, Ellington MJ, Livermore DM, Woodford N. Molecular mechanisms disrupting porin expression in ertapenem-resistant klebsiella and enterobacter spp. Clinical isolates from the uk. *J Antimicrob Chemother*. 2009; **63**: 659-667.

32   Nielsen LE, Snesrud EC, Onmus-Leone F, et al. Is5 element integration, a novel mechanism for rapid in vivo emergence of tigecycline nonsusceptibility in klebsiella pneumoniae. *Antimicrob Agents Chemother*. 2014; **58**: 6151-6156.

33   Weeks JW, Celaya-Kolb T, Pecora S, Misra R. Acra suppressor alterations reverse the drug hypersensitivity phenotype of a tolc mutant by inducing tolc aperture opening. *Mol Microbiol*. 2010; **75**: 1468-1483.

34   Bredin J, Saint N, Mallea M, et al. Alteration of pore properties of escherichia coli ompf induced by mutation of key residues in anti-loop 3 region. *Biochem J*. 2002; **363**: 521-528.

35   Blair JM, Bavro VN, Ricci V, et al. Acrb drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity. *Proc Natl Acad Sci U S A*. 2015; **112**: 3511-3516.

36   Woodford N, Turton JF, Livermore DM. Multiresistant gram-negative bacteria: The role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiol Rev*. 2011; **35**: 736-755.

37   Clancy CJ, Chen L, Hong JH, et al. Mutations of the ompk36 porin gene and promoter impact responses of sequence type 258, kpc-2-producing klebsiella pneumoniae strains to doripenem and doripenem-colistin. *Antimicrob Agents Chemother*. 2013; **57**: 5258-5265.

38   Olaitan AO, Morand S, Rolain JM. Mechanisms of polymyxin resistance: Acquired and intrinsic resistance in bacteria. *Front Microbiol*. 2014; **5**: 643.

39   Suzuki S, Horinouchi T, Furusawa C. Prediction of antibiotic resistance by gene expression profiles. *Nat Commun*. 2014; **5**: 5792.

40   *Danmap 2011. Use of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from food animals, foods and humans in denmark.* 2012. www.danmap.org

41   Hendriksen RS, Joensen KG, Lukwesa-Musyani C, et al. Extremely drug-resistant salmonella enterica serovar senftenberg infections in patients in zambia. *J Clin Microbiol*. 2013; **51**: 284-286.

42   Hendriksen RS, Leekitcharoenphon P, Mikoleit M, et al. Genomic dissection of travel-associated extended-spectrum-beta-lactamase-producing salmonella enterica serovar typhi isolates originating from the philippines: A one-off occurrence or a threat to effective treatment of typhoid fever? *J Clin Microbiol*. 2015; **53**: 677-680.

43   *The european union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2014.* http://ecdc.europa.eu/en/publications/Publications/antimicrobial-resistance-zoonotic-bacteria-humans-animals-food-EU-summary-report-2014.pdf

44   Sun S, Negrea A, Rhen M, Andersson DI. Genetic analysis of colistin resistance in salmonella enterica serovar typhimurium. *Antimicrob Agents Chemother*. 2009; **53**: 2298-2305.

45   Jevons MP. "Celbenin"-resistant staphylococci [letter]. *Br Med J*. 1961; **1**: 124-125.

46   Koser CU, Holden MT, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal mrsa outbreak. *N Engl J Med*. 2012; **366**: 2267-2275.

47   Eyre DW, Golubchik T, Gordon NC, et al. A pilot study of rapid benchtop sequencing of staphylococcus aureus and clostridium difficile for outbreak detection and surveillance. *BMJ Open*. 2012; **2**.

48   Holden MT, Hsu LY, Kurt K, et al. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant staphylococcus aureus pandemic. *Genome Res*. 2013; **23**: 653-664.

49   Gordon NC, Price JR, Cole K, et al. Prediction of staphylococcus aureus antimicrobial resistance by whole-genome sequencing. *J Clin Microbiol*. 2014; **52**: 1182-1191.

50   Lee GC, Long SW, Musser JM, et al. Comparative whole genome sequencing of community-associated methicillin-resistant staphylococcus aureus sequence type 8 from primary care clinics in a texas community. *Pharmacotherapy*. 2015; **35**: 220-228.

51   Bradley P, Gordon NC, Walker TM, et al. Rapid antibiotic-resistance predictions from genome sequence data for staphylococcus aureus and mycobacterium tuberculosis. *Nat Commun*. 2015; **6**: 10063.

52   Inouye M, Dashnow H, Raven LA, et al. Srst2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med*. 2014; **6**: 90.

53   Cui L, Isii T, Fukuda M, et al. An rpob mutation confers dual heteroresistance to daptomycin and vancomycin in staphylococcus aureus. *Antimicrob Agents Chemother*. 2010; **54**: 5222-5233.

54   Dordel J, Kim C, Chung M, et al. Novel determinants of antibiotic resistance: Identification of mutated loci in highly methicillin-resistant subpopulations of methicillin-resistant staphylococcus aureus. *MBio*. 2014; **5**: e01000.

55   Aanensen DM, Feil EJ, Holden MTG, et al. Whole-genome sequencing for routine pathogen surveillance in public health: A population snapshot of invasive staphylococcus aureus in europe. *MBio*. 2016; **7**.

56   Weigel LM, Clewell DB, Gill SR, et al. Genetic analysis of a high-level vancomycin-resistant isolate of staphylococcus aureus. *Science*. 2003; **302**: 1569-1571.

57   Chen CJ, Huang YC, Chiu CH. Multiple pathways of cross-resistance to glycopeptides and daptomycin in persistent mrsa bacteraemia. *J Antimicrob Chemother*. 2015; **70**: 2965-2972.

58   McEvoy CR, Tsuji B, Gao W, et al. Decreased vancomycin susceptibility in staphylococcus aureus caused by is256 tempering of walkr expression. *Antimicrob Agents Chemother*. 2013; **57**: 3240-3249.

59   Hakenbeck R. Discovery of beta-lactam-resistant variants in diverse pneumococcal populations. *Genome Med*. 2014; **6**: 72.

60   Chambers HF. Penicillin-binding protein-mediated resistance in pneumococci and staphylococci. *J Infect Dis*. 1999; **179 Suppl 2**: S353-359.

61   Fani F, Leprohon P, Legare D, Ouellette M. Whole genome sequencing of penicillin-resistant streptococcus pneumoniae reveals mutations in penicillin-binding proteins and in a putative iron permease. *Genome Biol*. 2011; **12**: R115.

62   Fani F, Leprohon P, Zhanel GG, Bergeron MG, Ouellette M. Genomic analyses of DNA transformation and penicillin resistance in streptococcus pneumoniae clinical isolates. *Antimicrob Agents Chemother*. 2014; **58**: 1397-1403.

63   Fani F, Brotherton MC, Leprohon P, Ouellette M. Genomic analysis and reconstruction of cefotaxime resistance in streptococcus pneumoniae. *J Antimicrob Chemother*. 2013; **68**: 1718-1727.

64   Chancey ST, Agrawal S, Schroeder MR, Farley MM, Tettelin H, Stephens DS. Composite mobile genetic elements disseminating macrolide resistance in streptococcus pneumoniae. *Front Microbiol*. 2015; **6**: 26.

65   Lupien A, Gingras H, Bergeron MG, Leprohon P, Ouellette M. Multiple mutations and increased rna expression in tetracycline-resistant streptococcus pneumoniae as determined by genome-wide DNA and mrna sequencing. *J Antimicrob Chemother*. 2015; **70**: 1946-1959.

66   Lupien A, Billal DS, Fani F, et al. Genomic characterization of ciprofloxacin resistance in a laboratory-derived mutant and a clinical isolate of streptococcus pneumoniae. *Antimicrob Agents Chemother*. 2013; **57**: 4911-4919.

67   World Health Organization G, Switzerland. Global action plan to control the spread and impact of antimicrobial resistance in neisseria gonorrhoeae. (accessed 10.12.2015). 2012.

68   Unemo M, Shafer WM. Antimicrobial resistance in neisseria gonorrhoeae in the 21st century: Past, evolution, and future. *Clin Microbiol Rev*. 2014; **27**: 587-613.

69   Prevention. CfDCa. *Antibiotic resistance threats in the united states, 2013.* http://www.cdc.gov/drugresistance/threat-report-2013/

70   Public Health Agency of Canada. 2013. Canadian guidelines on sexually transmitted infections giRJCAID, Centre for Communicable Diseases and Infection Control, Public Health Agency of Canada, Ottawa, Ontario, Canada. (http://www.phac-aspc.gc.ca/std-mts/sti-its/index-eng.php) (Accessed 22 July 2014).

71   Bignell C, Unemo M, European STIGEB. 2012 european guideline on the diagnosis and treatment of gonorrhoea in adults. *Int J STD AIDS*. 2013; **24**: 85-92.

72   Clinical and Laboratory Standards Institute. 2015. Performance standards for antimicrobial susceptibility testing. Twenty-fourth Informational Supplement. CLSI document ASM-SCaLSI, Wayne, PA.

73   (www.eucast.org) ECoASt.

74   Allen VG, Mitterni L, Seah C, et al. Neisseria gonorrhoeae treatment failure and susceptibility to cefixime in toronto, canada. *JAMA*. 2013; **309**: 163-170.

75   Unemo M, Golparian D, Nicholas R, Ohnishi M, Gallay A, Sednaoui P. High-level cefixime- and ceftriaxone-resistant neisseria gonorrhoeae in france: Novel pena mosaic allele in a successful international clone causes treatment failure. *Antimicrob Agents Chemother*. 2012; **56**: 1273-1280.

76   Unemo M, Golparian D, Syversen G, Vestrheim DF, Moi H. Two cases of verified clinical failures using internationally recommended first-line cefixime for gonorrhoea treatment, norway, 2010. *Euro Surveill*. 2010; **15**.

77   Ison CA, Hussey J, Sankar KN, Evans J, Alexander S. Gonorrhoea treatment failures to cefixime and azithromycin in england, 2010. *Euro Surveill*. 2011; **16**.

78   Unemo M, Golparian D, Stary A, Eigentler A. First neisseria gonorrhoeae strain with resistance to cefixime causing gonorrhoea treatment failure in austria, 2011. *Euro Surveill*. 2011; **16**.

79   Center for Disease Control and Prevention. Sexually Transmitted Disease Surveillance 2007 Supplement gISPGArA, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, March 2009. Available: http//www.cdc.gov/std/GISP2007/.

80   Barry PM, Klausner JD. The use of cephalosporins for gonorrhea: The impending problem of resistance. *Expert Opin Pharmacother*. 2009; **10**: 555-577.

81   Osaka K, Takakura T, Narukawa K, et al. Analysis of amino acid sequences of penicillin-binding protein 2 in clinical isolates of neisseria gonorrhoeae with reduced susceptibility to cefixime and ceftriaxone. *J Infect Chemother*. 2008; **14**: 195-203.

82   Whiley DM, Limnios EA, Ray S, Sloots TP, Tapsall JW. Diversity of pena alterations and subtypes in neisseria gonorrhoeae strains from sydney, australia, that are less susceptible to ceftriaxone. *Antimicrob Agents Chemother*. 2007; **51**: 3111-3116.

83   Ito M, Deguchi T, Mizutani KS, et al. Emergence and spread of neisseria gonorrhoeae clinical isolates harboring mosaic-like structure of penicillin-binding protein 2 in central japan. *Antimicrob Agents Chemother*. 2005; **49**: 137-143.

84   Zhao S, Duncan M, Tomberg J, Davies C, Unemo M, Nicholas RA. Genetics of chromosomally mediated intermediate resistance to ceftriaxone and cefixime in neisseria gonorrhoeae. *Antimicrob Agents Chemother*. 2009; **53**: 3744-3751.

85   Grad YH, Kirkcaldy RD, Trees D, et al. Genomic epidemiology of neisseria gonorrhoeae with reduced susceptibility to cefixime in the USA: A retrospective observational study. *Lancet Infect Dis*. 2014; **14**: 220-226.

86   Ohnishi M, Golparian D, Shimuta K, et al. Is neisseria gonorrhoeae initiating a future era of untreatable gonorrhea?: Detailed characterization of the first strain with high-level resistance to ceftriaxone. *Antimicrob Agents Chemother*. 2011; **55**: 3538-3545.

87   Ezewudo MN, Joseph SJ, Castillo-Ramirez S, et al. Population structure of neisseria gonorrhoeae based on whole genome data and its relationship with antibiotic resistance. *PeerJ*. 2015; **3**: e806.

88   Demczuk W, Lynch T, Martin I, et al. Whole-genome phylogenomic heterogeneity of neisseria gonorrhoeae isolates with decreased cephalosporin susceptibility collected in canada between 1989 and 2013. *J Clin Microbiol*. 2015; **53**: 191-200.

89   Palmer HM, Young H, Winter A, Dave J. Emergence and spread of azithromycin-resistant neisseria gonorrhoeae in scotland. *J Antimicrob Chemother*. 2008; **62**: 490-494.

90   Chisholm SA, Neal TJ, Alawattegama AB, Birley HD, Howe RA, Ison CA. Emergence of high-level azithromycin resistance in neisseria gonorrhoeae in england and wales. *J Antimicrob Chemother*. 2009; **64**: 353-358.

91   Galarza PG, Alcala B, Salcedo C, et al. Emergence of high level azithromycin-resistant neisseria gonorrhoeae strain isolated in argentina. *Sex Transm Dis*. 2009; **36**: 787-788.

92   Allen VG, Seah C, Martin I, Melano RG. Azithromycin resistance is coevolving with reduced susceptibility to cephalosporins in neisseria gonorrhoeae in ontario, canada. *Antimicrob Agents Chemother*. 2014; **58**: 2528-2534.

93   Katz AR, Komeya AY, Soge OO, et al. Neisseria gonorrhoeae with high-level resistance to azithromycin: Case report of the first isolate identified in the united states. *Clin Infect Dis*. 2012; **54**: 841-843.

94   Chisholm SA, Dave J, Ison CA. High-level azithromycin resistance occurs in neisseria gonorrhoeae as a result of a single point mutation in the 23s rrna genes. *Antimicrob Agents Chemother*. 2010; **54**: 3812-3816.

95   Ng LK, Martin I, Liu G, Bryden L. Mutation in 23s rrna associated with macrolide resistance in neisseria gonorrhoeae. *Antimicrob Agents Chemother*. 2002; **46**: 3020-3025.

96   Starnino S, Stefanelli P, Neisseria gonorrhoeae Italian Study G. Azithromycin-resistant neisseria gonorrhoeae strains recently isolated in italy. *J Antimicrob Chemother*. 2009; **63**: 1200-1204.

97   Buono S, Wu A, Hess DC, et al. Using the neisseria gonorrhoeae multiantigen sequence-typing method to assess strain diversity and antibiotic resistance in san francisco, california. *Microb Drug Resist*. 2012; **18**: 510-517.

98   Galagan JE. Genomic insights into tuberculosis. *Nature reviews Genetics*. 2014; **15**: 307-320.

99   Köser CU, Feuerriegel S, Summers DK, Archer JA, Niemann S. Importance of the genetic diversity within the *mycobacterium tuberculosis* complex for the development of novel antibiotics and diagnostic tests of drug resistance. *Antimicrob Agents Chemother*. 2012; **56**: 6080-6087.

100   Martinez E, Holmes N, Jelfs P, Sintchenko V. Genome sequencing reveals novel deletions associated with secondary resistance to pyrazinamide in mdr *mycobacterium tuberculosis*. *J Antimicrob Chemother*. 2015; **70**: 2511-2514.

101   Koeck JL, Fabre M, Simon F, et al. Clinical characteristics of the smooth tubercle bacilli '*mycobacterium canettii*' infection suggest the existence of an environmental reservoir. *Clin Microbiol Infect*. 2011; **17**: 1013-1019.

102   Feuerriegel S, Köser CU, Baù D, et al. Impact of *fgd1* and *ddn* diversity in *mycobacterium tuberculosis* complex on *in vitro* susceptibility to pa-824. *Antimicrob Agents Chemother*. 2011; **55**: 5718-5722.

103   Feuerriegel S, Köser CU, Richter E, Niemann S. *Mycobacterium canettii* is intrinsically resistant to both pyrazinamide and pyrazinoic acid. *J Antimicrob Chemother*. 2013; **68**: 1439-1440.

104   Koser CU, Ellington MJ, Cartwright EJ, et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog*. 2012; **8**: e1002824.

105   Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. *Trends in genetics : TIG*. 2014; **30**: 401-407.

106   Pankhurst LJ, Del Ojo Elias C, Votintseva AA, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: A prospective study. *Lancet Respir Med*. 2016; **4**: 49-58.

107   European Centre for Disease Prevention and Control. Tuberculosis surveillance and monitoring in europe 2015. Http://ecdc.Europa.Eu/en/publications/publications/tuberculosis-surveillance-monitoring-europe-2015.Pdf (accessed 9.2.2016).

108   Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a *mycobacterium tuberculosis* outbreak: A longitudinal molecular epidemiological study. *PLoS medicine*. 2013; **10**: e1001387.

109   Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate *mycobacterium tuberculosis* outbreaks: A retrospective observational study. *Lancet Infect Dis*. 2013; **13**: 137-146.

110   Walker TM, Lalor MK, Broda A, et al. Assessment of *mycobacterium tuberculosis* transmission in oxfordshire, uk, 2007-12, with whole pathogen genome sequences: An observational study. *Lancet Respir Med*. 2014; **2**: 285-292.

111   Doughty EL, Sergeant MJ, Adetifa I, Antonio M, Pallen MJ. Culture-independent detection and characterisation of *mycobacterium tuberculosis* and *m. Africanum* in sputum samples using shotgun metagenomics on a benchtop sequencer. *PeerJ*. 2014; **2**: e585.

112   Brown AC, Bryant JM, Einer-Jensen K, et al. Rapid whole-genome sequencing of *mycobacterium tuberculosis* isolates directly from clinical samples. *J Clin Microbiol*. 2015; **53**: 2230-2237.

113   Casali N, Nikolayevskyy V, Balabanova Y, et al. Microevolution of extensively drug-resistant tuberculosis in russia. *Genome Res*. 2012; **22**: 735-745.

114   Clark TG, Mallard K, Coll F, et al. Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. *PLoS One*. 2013; **8**: e83012.

115   Farhat MR, Shapiro BJ, Kieser KJ, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *mycobacterium tuberculosis*. *Nature genetics*. 2013; **45**: 1183-1189.

116   Merker M, Kohl TA, Roetzer A, et al. Whole genome sequencing reveals complex evolution patterns of multidrug-resistant *mycobacterium tuberculosis* beijing strains in patients. *PLoS One*. 2013; **8**: e82551.

117   Zhang H, Li D, Zhao L, et al. Genome sequencing of 161 *mycobacterium tuberculosis* isolates from china identifies genes and intergenic regions associated with drug resistance. *Nature genetics*. 2013; **45**: 1255-1260.

118   Casali N, Nikolayevskyy V, Balabanova Y, et al. Evolution and transmission of drug-resistant tuberculosis in a russian population. *Nature genetics*. 2014; **46**: 279-286.

119   Eldholm V, Norheim G, von der Lippe B, et al. Evolution of extensively drug-resistant *mycobacterium tuberculosis* from a susceptible ancestor in a single patient. *Genome Biol*. 2014; **15**: 490.

120   Cohen KA, Abeel T, Manson McGuire A, et al. Evolution of extensively drug-resistant tuberculosis over four decades: Whole genome sequencing and dating analysis of *mycobacterium tuberculosis* isolates from kwazulu-natal. *PLoS medicine*. 2015; **12**: e1001880.

121   Eldholm V, Monteserin J, Rieux A, et al. Four decades of transmission of a multidrug-resistant *mycobacterium tuberculosis* outbreak strain. *Nat Commun*. 2015; **6**: 7119.

122   Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of *mycobacterium tuberculosis* drug susceptibility and resistance: A retrospective cohort study. *Lancet Infect Dis*. 2015; **15**: 1193-1202.

123   Steiner A, Stucki D, Coscolla M, Borrell S, Gagneux S. Kvarq: Targeted and direct variant calling from fastq reads of bacterial genomes. *BMC Genomics*. 2014; **15**: 881.

124   Coll F, McNerney R, Preston MD, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med*. 2015; **7**: 51.

125   Feuerriegel S, Schleusener V, Beckert P, et al. Phyresse: Web tool delineating mycobacterium tuberculosis antibiotic resistance and lineage from whole-genome sequencing data. *J Clin Microbiol*. 2015; **53**: 1908-1914.

126   Iwai H, Kato-Miyazawa M, Kirikae T, Miyoshi-Akiyama T. Castb (the comprehensive analysis server for the *mycobacterium tuberculosis* complex): A publicly accessible web server for epidemiological analyses, drug-resistance prediction and phylogenetic comparison of clinical isolates. *Tuberculosis (Edinb)*. 2015; **95**: 843-844.

127   Sekizuka T, Yamashita A, Murase Y, et al. Tgs-tb: Total genotyping solution for *mycobacterium tuberculosis* using short-read whole-genome sequencing. *PLoS One*. 2015; **10**: e0142951.

128   Canetti G, Froman S, Grosset J, et al. Mycobacteria: Laboratory methods for testing drug sensitivity and resistance. *Bull World Health Organ*. 1963; **29**: 565-578.

129   Werngren J, Sturegård E, Juréen P, Ängeby K, Hoffner S, Schön T. Reevaluation of the critical concentration for drug susceptibility testing of *mycobacterium tuberculosis* against pyrazinamide using wild-type mic distributions and *pnca* gene sequencing. *Antimicrob Agents Chemother*. 2012; **56**: 1253-1257.

130   Blakemore R, Story E, Helb D, et al. Evaluation of the analytical performance of the xpert mtb/rif assay. *J Clin Microbiol*. 2010; **48**: 2495-2501.

131   Folkvardsen DB, Thomsen VO, Rigouts L, et al. Rifampin heteroresistance in *mycobacterium tuberculosis* cultures as detected by phenotypic and genotypic drug susceptibility test methods. *J Clin Microbiol*. 2013; **51**: 4220-4222.

132   Folkvardsen DB, Svensson E, Thomsen VO, et al. Can molecular methods detect 1% isoniazid resistance in *mycobacterium tuberculosis*? *J Clin Microbiol*. 2013; **51**: 1596-1599.

133   Kendall EA, Fofana MO, Dowdy DW. Burden of transmitted multidrug resistance in epidemics of tuberculosis: A transmission modelling analysis. *Lancet Respir Med*. 2015; **3**: 963-972.

134   Eilertson B, Maruri F, Blackman A, Herrera M, Samuels DC, Sterling TR. High proportion of heteroresistance in *gyra* and *gyrb* in fluoroquinolone-resistant *mycobacterium tuberculosis* clinical isolates. *Antimicrob Agents Chemother*. 2014; **58**: 3270-3275.

135   Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. *Nat Rev Microbiol*. 2015; **13**: 787-794.

136   Clinical and Laboratory Standards Institute. Susceptibility testing of mycobacteria, nocardiae, and other aerobic actinomycetes, 2nd edition approved standard. Clsi document m24-a2. 2011.

137   World Health Organization. *Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis (2014)* *http://appswhoint/iris/bitstream/10665/130918/1/9789241548809_engpdf?ua=1&ua=1 (accessed 1382015)*.

138   Ängeby K, Juréen P, Kahlmeter G, Hoffner SE, Schön T. Challenging a dogma: Antimicrobial susceptibility testing breakpoints for *mycobacterium tuberculosis*. *Bull World Health Organ*. 2012; **90**: 693-698.

139   Pasipanodya JG, McIlleron H, Burger A, Wash PA, Smith P, Gumbo T. Serum drug concentrations predictive of pulmonary tuberculosis outcomes. *J Infect Dis*. 2013; **208**: 1464-1473.

140   Reeves AZ, Campbell PJ, Willby MJ, Posey JE. Disparities in capreomycin resistance levels associated with the rrs a1401g mutation in clinical isolates of *mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2015; **59**: 444-449.

141   Coeck N, de Jong BC, Diels M, et al. Correlation of different phenotypic drug susceptibility testing methods for four fluoroquinolones in *mycobacterium tuberculosis*. *J Antimicrob Chemother*. 2016; **71**: 1233-1240.

142   van der Burgt EP, Sturkenboom MG, Bolhuis MS, et al. End tb with precision treatment! *Eur Respir J*. 2016; **47**: 680-682.

143   EUCAST. Workshop on recommendations for pharmaceutical companies regarding data required for new antituberculous drugs. 11-12 november 2014, basel, switzerland. Http://www.Eucast.Org/fileadmin/src/media/pdfs/eucast_files/mycobacteria/mycobacterial_workshop_basel_11-12nov2014.Pdf (accessed 19.10.2015).

144   Starks AM, Aviles E, Cirillo DM, et al. Collaborative effort for a centralized worldwide tuberculosis relational sequencing data platform. *Clin Infect Dis*. 2015; **61**: S141-S146.

145   Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert(r) mtb/rif assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev*. 2014; **1**: CD009593.

146   Sanchez-Padilla E, Merker M, Beckert P, et al. Detection of drug-resistant tuberculosis by xpert mtb/rif in swaziland. *N Engl J Med*. 2015; **372**: 1181-1182.

147   Rigouts L, Gumusboga M, de Rijk WB, et al. Rifampin resistance missed in automated liquid culture system for mycobacterium tuberculosis isolates with specific rpob mutations. *J Clin Microbiol*. 2013; **51**: 2641-2645.

148   Scorpio A, Zhang Y. Mutations in *pnca*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med*. 1996; **2**: 662-667.

149   Miotto P, Cabibbe AM, Feuerriegel S, et al. *Mycobacterium tuberculosis* pyrazinamide resistance determinants: A multicenter study. *MBio*. 2014; **5**: e01819-01814.

150   Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of *mycobacterium tuberculosis*. *Nature*. 1992; **358**: 591-593.

151   Bergval IL, Schuitema AR, Klatser PR, Anthony RM. Resistant mutants of *mycobacterium tuberculosis* selected *in vitro* do not reflect the *in vivo* mechanism of isoniazid resistance. *J Antimicrob Chemother*. 2009; **64**: 515-523.

152   Müller B, Borrell S, Rose G, Gagneux S. The heterogeneous evolution of multidrug-resistant *mycobacterium tuberculosis*. *Trends in genetics : TIG*. 2013; **29**: 160-169.

153   Böttger EC. The ins and outs of *mycobacterium tuberculosis* drug susceptibility testing. *Clin Microbiol Infect*. 2011; **17**: 1128-1134.

154   Rigouts L, Coeck N, Gumusboga M, et al. Specific *gyra* gene mutations predict poor treatment outcome in mdr-tb. *J Antimicrob Chemother*. 2016; **71**: 314-323.

155   Köser CU, Bryant JM, Becq J, et al. Whole-genome sequencing for rapid susceptibility testing of *m. Tuberculosis*. *N Engl J Med*. 2013; **369**: 290-292.

156   Peltier J, Courtin P, El Meouche I, et al. Genomic and expression analysis of the vang-like gene cluster of clostridium difficile. *Microbiology*. 2013; **159**: 1510-1520.

157   Leeds JA, Sachdeva M, Mullin S, Barnes SW, Ruzin A. In vitro selection, via serial passage, of clostridium difficile mutants with reduced susceptibility to fidaxomicin or vancomycin. *J Antimicrob Chemother*. 2014; **69**: 41-44.

158   Huang H, Weintraub A, Fang H, Nord CE. Antimicrobial resistance in clostridium difficile. *Int J Antimicrob Agents*. 2009; **34**: 516-522.

159   Pelaez T, Cercenado E, Alcala L, et al. Metronidazole resistance in clostridium difficile is heterogeneous. *J Clin Microbiol*. 2008; **46**: 3028-3032.

160   Indra A, Schmid D, Huhulescu S, et al. Characterization of clinical clostridium difficile isolates by pcr ribotyping and detection of toxin genes in austria, 2006-2007. *J Med Microbiol*. 2008; **57**: 702-708.

161   Lynch T, Chong P, Zhang J, et al. Characterization of a stable, metronidazole-resistant clostridium difficile clinical isolate. *PLoS One*. 2013; **8**: e53757.

162   Chong PM, Lynch T, McCorrister S, et al. Proteomic analysis of a nap1 clostridium difficile clinical isolate resistant to metronidazole. *PLoS One*. 2014; **9**: e82622.

163   Peleg AY, Seifert H, Paterson DL. Acinetobacter baumannii: Emergence of a successful pathogen. *Clin Microbiol Rev*. 2008; **21**: 538-582.

164   Roca I, Espinal P, Vila-Farres X, Vila J. The acinetobacter baumannii oxymoron: Commensal hospital dweller turned pan-drug-resistant menace. *Front Microbiol*. 2012; **3**: 148.

165   Zavascki AP, Carvalhaes CG, Picao RC, Gales AC. Multidrug-resistant pseudomonas aeruginosa and acinetobacter baumannii: Resistance mechanisms and implications for therapy. *Expert Rev Anti Infect Ther*. 2010; **8**: 71-93.

166   Kempf M, Rolain JM. Emergence of resistance to carbapenems in acinetobacter baumannii in europe: Clinical impact and therapeutic options. *Int J Antimicrob Agents*. 2012; **39**: 105-114.

167   Diene SM, Rolain JM. Investigation of antibiotic resistance in the genomic era of multidrug-resistant gram-negative bacilli, especially enterobacteriaceae, pseudomonas and acinetobacter. *Expert Rev Anti Infect Ther*. 2013; **11**: 277-296.

168   Imperi F, Antunes LC, Blom J, et al. The genomics of acinetobacter baumannii: Insights into genome plasticity, antimicrobial resistance and pathogenicity. *IUBMB life*. 2011; **63**: 1068-1074.

169   Touchon M, Cury J, Yoon EJ, et al. The genomic diversification of the whole acinetobacter genus: Origins, mechanisms, and consequences. *Genome biology and evolution*. 2014; **6**: 2866-2882.

170   Antunes LC, Visca P, Towner KJ. Acinetobacter baumannii: Evolution of a global pathogen. *Pathogens and disease*. 2014; **71**: 292-301.

171   Klockgether J, Cramer N, Wiehlmann L, Davenport CF, Tummler B. Pseudomonas aeruginosa genomic structure and diversity. *Front Microbiol*. 2011; **2**: 150.

172   Mosquera-Rendon J, Rada-Bravo AM, Cardenas-Brito S, Corredor M, Restrepo-Pineda E, Benitez-Paez A. Pangenome-wide and molecular evolution analyses of the pseudomonas aeruginosa species. *BMC Genomics*. 2016; **17**: 45.

173   Viehman JA, Nguyen MH, Doi Y. Treatment options for carbapenem-resistant and extensively drug-resistant acinetobacter baumannii infections. *Drugs*. 2014; **74**: 1315-1333.

174   Karaiskos I, Giamarellou H. Multidrug-resistant and extensively drug-resistant gram-negative pathogens: Current and emerging therapeutic approaches. *Expert Opin Pharmacother*. 2014; **15**: 1351-1370.

175   Oliver A, Mulet X, Lopez-Causape C, Juan C. The increasing threat of pseudomonas aeruginosa high-risk clones. *Drug Resist Updat*. 2015; **21-22**: 41-59.

176   Zarrilli R, Pournaras S, Giannouli M, Tsakris A. Global evolution of multidrug-resistant acinetobacter baumannii clonal lineages. *Int J Antimicrob Agents*. 2013; **41**: 11-19.

177   Potron A, Poirel L, Nordmann P. Emerging broad-spectrum resistance in pseudomonas aeruginosa and acinetobacter baumannii: Mechanisms and epidemiology. *Int J Antimicrob Agents*. 2015; **45**: 568-585.

178   Bertini A, Poirel L, Mugnier PD, Villa L, Nordmann P, Carattoli A. Characterization and pcr-based replicon typing of resistance plasmids in acinetobacter baumannii. *Antimicrob Agents Chemother*. 2010; **54**: 4168-4177.

179   Lister PD, Wolter DJ, Hanson ND. Antibacterial-resistant pseudomonas aeruginosa: Clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clin Microbiol Rev*. 2009; **22**: 582-610.

180   Kos VN, Deraspe M, McLaughlin RE, et al. The resistome of pseudomonas aeruginosa in relationship to phenotypic susceptibility. *Antimicrob Agents Chemother*. 2015; **59**: 427-436.

181   Wright MS, Stockwell TB, Beck E, et al. Sispa-seq for rapid whole genome surveys of bacterial isolates. *Infect Genet Evol*. 2015; **32**: 191-198.

182   Liu F, Zhu Y, Yi Y, Lu N, Zhu B, Hu Y. Comparative genomic analysis of acinetobacter baumannii clinical isolates reveals extensive genomic variation and diverse antibiotic resistance determinants. *BMC Genomics*. 2014; **15**: 1163.

183   Turton JF, Wright L, Underwood A, et al. High-resolution analysis by whole-genome sequencing of an international lineage (sequence type 111) of pseudomonas aeruginosa associated with metallo-carbapenemases in the united kingdom. *J Clin Microbiol*. 2015; **53**: 2622-2631.

184   Rolain JM, Diene SM, Kempf M, Gimenez G, Robert C, Raoult D. Real-time sequencing to decipher the molecular mechanism of resistance of a clinical pan-drug-resistant acinetobacter baumannii isolate from marseille, france. *Antimicrob Agents Chemother*. 2013; **57**: 592-596.

185   Lee JY, Na IY, Park YK, Ko KS. Genomic variations between colistin-susceptible and -resistant pseudomonas aeruginosa clinical isolates and their effects on colistin resistance. *J Antimicrob Chemother*. 2014; **69**: 1248-1256.

186   Turton JF, Ward ME, Woodford N, et al. The role of isaba1 in expression of oxa carbapenemase genes in acinetobacter baumannii. *FEMS microbiology letters*. 2006; **258**: 72-77.

187   Lopes BS, Amyes SG. Role of isaba1 and isaba125 in governing the expression of blaadc in clinically relevant acinetobacter baumannii strains resistant to cephalosporins. *J Med Microbiol*. 2012; **61**: 1103-1108.

188   Breidenstein EB, de la Fuente-Nunez C, Hancock RE. Pseudomonas aeruginosa: All roads lead to resistance. *Trends Microbiol*. 2011; **19**: 419-426.

189   Murray JL, Kwon T, Marcotte EM, Whiteley M. Intrinsic antimicrobial resistance determinants in the superbug pseudomonas aeruginosa. *MBio*. 2015; **6**: e01603-01615.

190   Partridge SR. Analysis of antibiotic resistance regions in gram-negative bacteria. *FEMS Microbiol Rev*. 2011; **35**: 820-855.

191   Merker M, Blin C, Mona S, et al. Evolutionary history and global spread of the mycobacterium tuberculosis beijing lineage. *Nature genetics*. 2015; **47**: 242-249.

192   Nicolas-Chanoine MH, Blanco J, Leflon-Guibout V, et al. Intercontinental emergence of escherichia coli clone o25:H4-st131 producing ctx-m-15. *J Antimicrob Chemother*. 2008; **61**: 273-281.

193   Villa L, Feudi C, Fortini D, Garcia-Fernandez A, Carattoli A. Genomics of kpc-producing klebsiella pneumoniae sequence type 512 clone highlights the role of ramr and ribosomal s10 protein mutations in conferring tigecycline resistance. *Antimicrob Agents Chemother*. 2014; **58**: 1707-1712.

194   Carattoli A, Zankari E, Garcia-Fernandez A, et al. In silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother*. 2014; **58**: 3895-3903.

195   Lanza VF, de Toro M, Garcillan-Barcia MP, et al. Plasmid flux in escherichia coli st131 sublineages, analyzed by plasmid constellation network (placnet), a new method for plasmid reconstruction from whole genome sequences. *PLoS Genet*. 2014; **10**: e1004766.

196   Ashton PM, Nair S, Dallman T, et al. Minion nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat Biotechnol*. 2015; **33**: 296-+.

197   Conlan S, Thomas PJ, Deming C, et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing enterobacteriaceae. *Science Translational Medicine*. 2014; **6**.

198   Tsafnat G, Coiera E, Partridge SR, Schaeffer J, Iredell JR. Context-driven discovery of gene cassettes in mobile integrons using a computational grammar. *BMC bioinformatics*. 2009; **10**: 281.

199   Tsafnat G, Schaeffer J, Clayphan A, Iredell JR, Partridge SR, Coiera E. Computational inference of grammars for larger-than-gene structures from annotated gene sequences. *Bioinformatics*. 2011; **27**: 791-796.

200   Arena F, Rolfe PA, Doran G, et al. Rapid resistome fingerprinting and clonal lineage profiling of carbapenem-resistant klebsiella pneumoniae isolates by targeted next-generation sequencing. *J Clin Microbiol*. 2014; **52**: 987-990.

201   Veenemans J, Overdevest IT, Snelders E, et al. Next-generation sequencing for typing and detection of resistance genes: Performance of a new commercial method during an outbreak of extended-spectrum-beta-lactamase-producing escherichia coli. *J Clin Microbiol*. 2014; **52**: 2454-2460.

202   Ginn AN, Wiklendt AM, Zong Z, et al. Prediction of major antibiotic resistance in escherichia coli and klebsiella pneumoniae in singapore, USA and china using a limited set of gene targets. *Int J Antimicrob Agents*. 2014; **43**: 563-565.

203   Ginn AN, Zong ZY, Wiklendt AM, et al. Limited diversity in the gene pool allows prediction of third-generation cephalosporin and aminoglycoside resistance in escherichia coli and klebsiella pneumoniae. *Int J Antimicrob Ag*. 2013; **42**: 19-26.