



OPINION ARTICLE

REVISED The challenges of designing a benchmark strategy for bioinformatics pipelines in the identification of antimicrobial resistance determinants using next generation sequencing technologies [version 2; peer review: 2 approved]

Alexandre Angers-Loustau ¹, Mauro Petrillo ¹, Johan Bengtsson-Palme ^{2,3}, Thomas Berendonk⁴, Burton Blais⁵, Kok-Gan Chan ^{6,7}, Teresa M. Coque⁸, Paul Hammer⁹, Stefanie Heß⁴, Dafni M. Kagkli¹, Carsten Krumbiegel⁹, Val F. Lanza⁸, Jean-Yves Madec¹⁰, Thierry Naas¹¹, Justin O'Grady¹², Valentina Paracchini¹, John W.A. Rossen ¹³, Etienne Ruppé¹⁴, Jessica Vamathevan ¹⁵, Vittorio Venturi¹⁶, Guy Van den Eede¹⁷

¹European Commission Joint Research Centre, Ispra, 21027, Italy

²Department of Infectious Diseases, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, SE-413 46, Sweden

³Centre for Antibiotic Resistance research (CARe), University of Gothenburg, SE-413 46, Gothenburg, Sweden

⁴Institute for Hydrobiology, Technische Universität Dresden, Dresden, 01307, Germany

⁵Canadian Food Inspection Agency, Ottawa Laboratory (Carling), Ottawa, ON, K1A 0Y9, Canada

⁶International Genome Centre, Jiangsu University, Zhenjiang, China

⁷Division of Genetics and Molecular Biology, Institute of Biological Sciences, University of Malaya, Kuala Lumpur, 50603, Malaysia

⁸Departamento de Microbiología, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, 28034, Spain

⁹BIOMES.world, c/o Technische Hochschule Wildau, Wildau, 15745, Germany

¹⁰Unité Antibiorésistance et Virulence Bactériennes, ANSES Site de Lyon, Lyon, F-69364, France

¹¹Service de Bactériologie-Hygiène, Hôpital de Bicêtre, Le Kremlin-Bicêtre, F-94275, France

¹²Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, UK

¹³Department of Medical Microbiology, University Medical Center Groningen, University of Groningen, Groningen, 9713 GZ, The Netherlands

¹⁴Laboratoire de Bactériologie, Hôpital Bichat, INSERM, IAME, UMR 1137, Université Paris Diderot, Paris, F-75018, France

¹⁵European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, CB10 1SD, UK

¹⁶International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, 34149, Italy

¹⁷European Commission Joint Research Centre, Geel, B-2440, Belgium

v2 First published: 13 Apr 2018, 7:459 (<https://doi.org/10.12688/f1000research.14509.1>)

Latest published: 07 Dec 2018, 7:459 (<https://doi.org/10.12688/f1000research.14509.2>)

Abstract

Next-Generation Sequencing (NGS) technologies are expected to play a crucial role in the surveillance of infectious diseases, with their unprecedented capabilities for the characterisation of genetic information underlying the virulence and antimicrobial resistance (AMR) properties of microorganisms. In the implementation of any novel technology for

Open Peer Review**Reviewer Status**  

Invited Reviewers

regulatory purposes, important considerations such as harmonisation, validation and quality assurance need to be addressed. NGS technologies pose unique challenges in these regards, in part due to their reliance on bioinformatics for the processing and proper interpretation of the data produced. Well-designed benchmark resources are thus needed to evaluate, validate and ensure continued quality control over the bioinformatics component of the process. This concept was explored as part of a workshop on "Next-generation sequencing technologies and antimicrobial resistance" held October 4-5 2017. Challenges involved in the development of such a benchmark resource, with a specific focus on identifying the molecular determinants of AMR, were identified. For each of the challenges, sets of unsolved questions that will need to be tackled for them to be properly addressed were compiled. These take into consideration the requirement for monitoring of AMR bacteria in humans, animals, food and the environment, which is aligned with the principles of a "One Health" approach.

Keywords

Antimicrobial resistance, bioinformatics, next-generation sequencing, benchmarking



This article is included in the **International Society for Computational Biology Community Journal gateway**.



This article is included in the **Antimicrobial Resistance collection**.

version 2

(revision)
07 Dec 2018


1

2

version 1

13 Apr 2018



1 **Enrico Lavezzo** , University of Padova, Padova, Italy

Giorgio Palù, University of Padova, Padova, Italy

2 **Jason C. Kwong** , University of Melbourne, Melbourne, Australia

Austin Health, Melbourne, Australia

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Alexandre Angers-Loustau (alexandre.angers@ec.europa.eu)

Author roles: **Angers-Loustau A:** Conceptualization, Writing – Original Draft Preparation; **Petrillo M:** Conceptualization, Writing – Original Draft Preparation; **Bengtsson-Palme J:** Investigation, Writing – Review & Editing; **Berendonk T:** Investigation, Writing – Review & Editing; **Blais B:** Investigation, Writing – Review & Editing; **Chan KG:** Investigation, Writing – Review & Editing; **Coque TM:** Investigation, Writing – Review & Editing; **Hammer P:** Investigation, Writing – Review & Editing; **Heß S:** Investigation, Writing – Review & Editing; **Kagkli DM:** Investigation, Writing – Review & Editing; **Krumbiegel C:** Investigation, Writing – Review & Editing; **Lanza VF:** Investigation, Writing – Review & Editing; **Madec JY:** Investigation, Writing – Review & Editing; **Naas T:** Investigation, Writing – Review & Editing; **O’Grady J:** Investigation, Writing – Review & Editing; **Paracchini V:** Investigation, Writing – Review & Editing; **Rossen JWA:** Investigation, Writing – Review & Editing; **Ruppé E:** Investigation, Writing – Review & Editing; **Vamathevan J:** Investigation, Writing – Review & Editing; **Venturi V:** Investigation, Writing – Review & Editing; **Van den Eede G:** Investigation, Writing – Review & Editing

Competing interests: JOG receives some research funding from Oxford Nanopore Technologies. ER is consultant for Pathoquest.

Grant information: The "Next-generation sequencing technologies and antimicrobial resistance - Working groups kick-off" meeting (4-5 October 2017) was funded by the European Commission's Joint Research Centre (JRC), Ispra, Italy.

Copyright: © 2018 Angers-Loustau A *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Angers-Loustau A, Petrillo M, Bengtsson-Palme J *et al.* **The challenges of designing a benchmark strategy for bioinformatics pipelines in the identification of antimicrobial resistance determinants using next generation sequencing technologies [version 2; peer review: 2 approved]** F1000Research 2018, 7:459 (<https://doi.org/10.12688/f1000research.14509.2>)

First published: 13 Apr 2018, 7:459 (<https://doi.org/10.12688/f1000research.14509.1>)

REVISED Amendments from Version 1

This version contains a few text additions following the suggestions and comments from Dr Kwong, Dr Lavezzo and Dr Palù in their referee reports.

These additions include:

1. A short overview of the experts' scientific backgrounds.
2. A short discussion of de-identification processes for human DNA sequences in section 2.1.
3. A reflection of the impacts of the purpose of the experiment (diagnostics, risk assessment...) on the benchmark datasets in section 2.2.
4. A clarification that the "truth" of a sample will depend on the objective of the analysis (i.e. detecting the genetic determinants of AMR, predicting AMR or predicting antimicrobial treatment outcomes).
5. A paragraph in the conclusions discussing the applicability of the discussions to microorganisms other than bacteria.

See referee reports

1. Introduction

Next-Generation Sequencing (NGS) technologies are increasingly regarded as an essential tool in modern regulatory frameworks. Monitoring schemes that rely on the characterisation of genetic information will gain considerably by utilising these technologies. Their importance for infectious diseases surveillance was highlighted by "The Review on Antimicrobial Resistance" in 2014, which stated that "advances in genetics, genomics and computer science will likely change the way that infections and new types of resistance are diagnosed, detected and reported worldwide, so that we can fight back faster when bacteria evolve to resist drugs"¹.

This interest can be observed in the rapid expansion in recent years of whole-genome sequencing capacities in national public health infectious diseases surveillance laboratories, as recently

reported in a European survey by the European Centre for Disease Prevention and Control (ECDC)². Antimicrobial resistance (AMR), i.e. the ability of a microorganism to resist the action of an antimicrobial agent, is of particular importance in this surveillance program. Its observed rise places heavy burdens on healthcare systems, leading to prolonged treatment times, higher mortality and high economic impacts (see 3). In March 2017, the Joint Research Centre organised a meeting in order to better understand the state-of-the-art of the application of NGS technologies in the fight against AMR⁴. Although it is clear that the uses of NGS vary according to the specific need (e.g. to guide clinical intervention or to evaluate the environmental and human health risks of AMR genetic determinants), these discussions highlighted overlaps in the needs and the challenges of implementing NGS for the monitoring of AMR in humans, animals, food and the environment. Some of these were also highlighted in previous workshops organized by the European Food Safety Authority (EFSA) and the ECDC^{5,6}.

A full regulatory implementation of NGS technologies to monitor AMR will need to address many standardisation challenges throughout the process, which broadly includes sample preparation and DNA extraction, library preparation for sequencing, the use of an NGS instrument for generating the sequences, the bioinformatics analysis, and interpretation and reporting of results (see Figure 1). Focusing on the bioinformatics step, an important shared challenge is the need to correctly and reliably identify the known genomic determinants of AMR from a set of NGS reads produced from sequencing a sample. The ECDC study reported the requirement for sufficient bioinformatics expertise as one of the important hurdles to a more general implementation of NGS for routine testing². This observation has also been made in recent case studies and reviews⁷⁻¹¹.

By contrast, within the scientific research community the recent literature reflects widespread enthusiasm for the application of NGS approaches to the determination of AMR

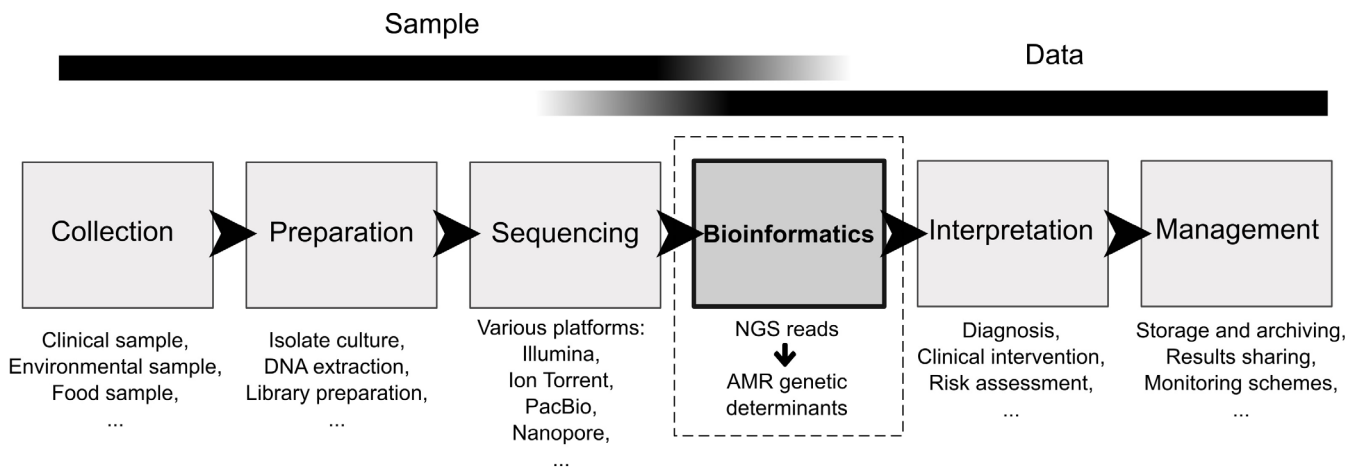


Figure 1. Overview of the different steps involved in the use of Next-Generation Sequencing technologies for the detection and monitoring of antimicrobial resistance. The benchmark strategy discussed in the current article focuses on the bioinformatics steps, the pipeline converting the output of the sequencing experiment into a list of identified antimicrobial resistance genetic determinants (dashed rectangle).

characteristics in bacteria. For the bioinformatics steps, many useful strategies have been published. These are, however, very varied in the approaches and resources they use. Some start with sequencing reads produced by the Illumina^{12,13}, Ion Torrent¹⁴, PacBio¹⁵ or Nanopore¹⁶ platforms, just to give a few examples. To predict the resistance profile, interesting results were reported with very different strategies, including k-mer analysis of the reads¹⁷, sequence comparisons of individual reads to databases^{12,16}, first assembling the reads into contigs using various software packages^{9,18} and building and comparing de Bruijn graphs of the sequenced sample reads and the reference database¹⁹. The reference set of genetic determinants of AMR used by the bioinformatics pipelines also varied, including databases such as ARGANNOT²⁰, CARD¹⁶, ResFinder⁹, Resqu¹², ARDB²¹, custom-generated from Genbank sequences^{18,22} or combinations of these¹⁴. Interestingly, the choice of the database was shown to greatly influence the interpretation of risk associated with AMR in public health^{23,24}. Even individual steps, such as mapping sequenced reads to a reference, can be done with different tools, each carrying their own compromises (see 25–28).

This complex - and dynamic - reality poses a challenge for the implementation of bioinformatics pipelines in regulatory settings, where the demonstration of reliability and reproducibility is crucial (see also 11,29). Harmonisation approaches must face the variability described above in terms of technologies, strategies, and software used, each with their demonstrated success, limitations and caveats. A further factor influencing the complexity of applying a given bioinformatics pipeline is that new versions of the individual tools that perform tasks such as quality-checking, trimming or assembling the reads, are constantly being released, which may have unanticipated impacts on pipeline performance. Ready-made and/or commercially available solutions that aim to facilitate the implementation of a NGS-based pipeline by lowering the technical skill required (see, for example, 30,31) face the attendant “black-box” issues when proposed for regulatory purposes.

In response to this complex state-of-the-art and the fast-moving environment in which these technologies are developing, efforts for the standardisation and development of best practices have avoided the prescription of restrictive guidelines, methods or technologies in favour of a more flexible approach emphasising quality metrics and fitness-for-purpose^{32,33}. For bioinformatics pipelines, the development of benchmark resources would play an important role in validating specific bioinformatics strategies and workflows, testing any update to the software underlying an established pipeline or allowing proficiency testing of individual laboratories^{33–35}. These resources would need to include a set of inputs for the bioinformatics pipelines (“*in silico* reference materials”) linked to a “correct” expected output, as well as consideration for the minimum performance requirements to be met by the pipelines. Different initiatives are ongoing to develop these benchmarking resources including, for example, the Critical Assessment of Metagenome Interpretation (CAMI) project for the evaluation methods for metagenome analysis³⁶.

On the 5th of October 2017, the Joint Research Centre invited experts in the field of AMR monitoring in order to discuss the challenges involved in the development of such a benchmark strategy, for the specific purpose of evaluating the bioinformatics pipelines that transform a set of NGS reads to a characterised AMR profile. The experts were invited based on their recent publications on the use of NGS to detect the genetic determinants of AMR in diverse fields: human and veterinary health, the food chain and the environment. The conclusions of these discussions are summarised in Table 1, and discussed in this document.

2. The challenges

Although some of the challenges considered reflect the reality of NGS technologies in general, efforts were made to highlight the issues that are specific to the identification of AMR determinants. Broadly, the challenges can be grouped in different, often overlapping categories.

2.1. Nature of the benchmark datasets

How should a benchmark strategy handle the current and expanding universe of NGS platforms? What should be the quality profile (in terms of read length, error rate, etc.) of “in silico reference materials”? Should different sets of reference materials be produced for each platform? In that case, how to ensure no bias is introduced in the process?

As described in the Introduction, different NGS technology platforms exist for the generation of sequence data serving as inputs for the bioinformatics processes used in the analysis of AMR determinants. Moreover, the technology continues to evolve rapidly with the advent of what is now termed “third generation sequencing” methods that can read the nucleotide sequences at the level of single molecules³⁷. Focusing on validating the technology or the instrument itself is therefore not a useful approach to ensure the reliability of the bioinformatics steps, since it can reasonably be expected that sequencing technologies and protocols will undergo many changes over the coming years. Section 862.2265 of the FDA’s Code of Federal Regulations Title 21³⁸ regulates the general use of NGS instruments for clinical use; even when, in this context, devices are cleared as Class II exempt¹, laboratories using these instruments must still establish a bioinformatics pipeline for their intended use³⁹. Thus, an effective benchmark strategy will be independent of existing and upcoming NGS technologies, while avoiding any bias that would favour one technology to the detriment of others.

The proprietary nature of the different raw data outputs produced by the various technologies may not be a primary consideration for present purposes since standard file formats exist that can store raw reads and the associated metadata (ex. QC metrics) produced by the different sequencers. These include FASTQ⁴⁰ and BAM⁴¹, and they have been successfully used

¹See, for example, <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRL/rl.cfm?lid=427645>

Table 1. Summary of the challenges identified in the generation of benchmark datasets for the purpose of evaluating the bioinformatics pipelines that process a set of NGS reads into a characterised AMR profile. See text for details.

Section	Challenges	Questions to be addressed
2.1	Nature of the benchmark datasets - NGS platforms	How should a benchmark strategy handle the current and expanding universe of NGS platforms? What should be the quality profile (in terms of read length, error rate, etc.) of <i>in silico</i> reference materials? Should different sets of reference materials be produced for each platform? In that case, how to ensure no bias is introduced in the process?
	Nature of the benchmark datasets - datasets origin	Should <i>in silico</i> reference material be composed of the output of real experiments, or simulated read sets? If a combination is used, what is the optimal ratio? How is it possible to ensure that the simulated output has been simulated "correctly"? For real experiments datasets, how to avoid the presence of sensitive information?
	Nature of the benchmark datasets - quality metrics	Regarding the quality metrics in the benchmark datasets (e.g. error rate, read quality), should these values be fixed for all datasets, or fall within specific ranges? How wide can/should these ranges be?
2.2	Samples composition - resistance mechanisms	How should the benchmark manage the different mechanisms by which bacteria acquire resistance? What is the set of resistance genes/mechanisms that need to be included in the benchmark? How should this set be agreed upon?
	Samples composition - bacterial species	Should different sample types (isolated clones, environmental samples, ...) be included in the same benchmark? Is a correct representation of different bacterial species (host genomes) important?
2.3	Evaluation of pipeline performance - dataset characterisation	How can the "true" value of the samples, against which the pipelines will be evaluated, be guaranteed? What is needed to demonstrate that the original sample has been correctly characterised, in case real experiments are used?
	Evaluation of pipeline performance - performance thresholds	How should the target performance thresholds (e.g. specificity, sensitivity, accuracy, ...) for the benchmark suite be set? What is the impact of these targets on the required size of the sample set?
2.4	Generation, distribution and update of the benchmark - future proofing	How can the benchmark stay relevant when new resistance mechanisms are regularly characterised? How is the continued quality of the benchmark dataset ensured?
	Generation, distribution and update of the benchmark - ownership	Who should generate the benchmark resource? How can it be efficiently shared?

in laboratory proficiency testing^{34,35,42}. More recent platforms produce outputs using the HDF5 standard or variants of it; conversion into FASTQ would require an additional computational step, using one of the available tools. However, all platforms (as well as sequencer models and versions within each platform) have differences in the profile and amount of raw reads produced, with variations in their number, length, error rates, error types, etc.^{43,44}. Attempting to create a single set of *in silico* reference materials would either introduce a bias towards a specific platform and/or create a dataset which is not representative. Creating individual sets of reads would increase the work (with no end in sight as platforms appear or evolve) and require careful consideration to avoid, once again, bias.

All this highlights a clear challenge, which is how to address both the evolution of the platforms, differences amongst

instruments and run-to-run variabilities, in view of the need for benchmark datasets serving as the basis for the validation and harmonisation of NGS approaches in clinical and/or regulatory frameworks.

Should in silico reference material be composed of the output of real experiments, or simulated read sets? If a combination is used, what is the optimal ratio? How is it possible to ensure that the simulated output has been simulated "correctly"? For real experiment datasets, how to avoid the presence of sensitive information?

The core component of a benchmark resource is, by definition, a set of inputs representative of what the benchmarked bioinformatics pipeline is expected to receive in normal, real-life use. A logical source for this dataset, then, is the actual

output of laboratory sequencing experiments^{17,34}. However, using data generated by real experiments assumes a high level of quality that will need to somehow be assessed and demonstrated. These experiments will need to be properly characterised in terms of the “true” conclusions the benchmarked pipeline is expected to reach. In addition, although there can be actions taken to ensure that most of the host DNA is filtered from the dataset, real metagenomics experiments from a human source could lead to privacy problems, while samples from food should ensure the absence of information on patented genetically modified food potentially present in the sample^{8,45}. Careful filtering against a standard “exclusion database”, or other adequate strategies, may be necessary to solve this issue. For example, de-identification processes for human DNA sequences have been proposed for clinical datasets. However, the impact of applying tools that modify in any way the reads in a metagenomics dataset should be well understood, as there is a risk that the filtered dataset is no longer representative of a real experiment. Experimental data could also be generated using pure cultures of bacteria present as well-characterised strains in biorepositories (see, for example, 46)

These concerns could be addressed by *in silico*-generated datasets, where the exact quantity of reads and genes from each source in the composite dataset can be better controlled. Many tools have been developed for this purpose, simulating reads from the different available platforms (see, for example 47–51). Once again, it will be important to properly understand these tools, agree on their applicability for the purpose of generating the desired benchmark datasets, and correctly set their parameters so that the resulting simulations are a correct representation of the “real” samples.

Regarding the quality metrics in the benchmark datasets (e.g. error rate, read quality), should these values be fixed for all datasets, or fall within specific ranges? How wide can/should these ranges be?

Available published studies of benchmarking NGS bioinformatics pipelines tend to focus on the performance of specific steps at various levels of input quality and/or complexity (SNP rate, GC content, error rate, quality of the reference sequences, contamination, etc.)^{26,52,53}. This is different from a fit-for-purpose evaluation of a complete pipeline under conditions where the quality of the input is guaranteed through the application of best practices and quality control of the laboratory component of the procedure. An important consideration is the extent to which the benchmark should challenge the pipeline robustness by including varying levels of, for example, error rates or reads quality. It is likely that a pipeline that works best under optimal conditions would be sensitive to variation of the sequencing run quality. The extent of desired variation should be agreed upon and captured in the *in silico* reference material included in the benchmark.

2.2. Samples composition

How should the benchmark manage the different mechanisms by which bacteria acquire resistance? What is the set of resistance

genes/mechanisms that need to be included in the benchmark? How should this set be agreed upon?

Several mechanisms for the development of resistance to antimicrobials have been characterised⁵⁴, including: 1) production of an enzyme that digests/metabolizes/modifies the antimicrobial; 2) production of efflux pumps that remove the drug from within the cell; 3) modification, through mutations or biochemical reactions, of the intracellular target of the antimicrobial so that their interaction is lost; 4) activation/upregulation/acquisition of alternate pathways that allow survival by bypassing the pathway disrupted by the antimicrobial; and 5) downregulation of the expression of the pores through which the drug enters the bacteria.

Mechanisms 1), 2) and 4), often involve the acquisition of novel genes by the bacteria from its environment (horizontal transfer) and may be detected, for example, by mapping reads to reference sequence databases that compile such genes. The genetic determinants of mechanisms 3) and 5), however, vary on a case-by-case basis, and may require the detection of Single Nucleotide Polymorphisms (SNPs), insertions/deletions (indels) or variations of copy numbers. These represent different types of bioinformatics determinations which a comprehensive pipeline must be able to resolve, and the benchmark needs to reflect this reality by ensuring that the various types of AMR determinants are correctly represented in the dataset.

Many recent evaluations of the use of NGS for the determination of AMR have emphasised the difficulty of establishing a curated knowledge base of drug resistance genetic determinants to be used as a reference database in NGS data analysis^{2,13,55}. The same problem is mirrored in the design of a benchmark that would ensure all determinants are correctly detected. It is also of foremost relevance to consider that certain genetic determinants such as efflux pumps (mechanism 2 above) are notorious for giving false positive results, as they perform a variety of export functions not necessarily related to antibiotic resistance (see, for example, 56). Eliminating these from the search parameters of bioinformatics pipelines was shown to improve positive predictive value⁵⁷. The results of testing a pipeline using a benchmark dataset involving all mechanisms must be interpreted with the aim of the pipeline in mind, and this should be taken into account when/if criteria are set (see also section 2.3).

Alternatively, choosing to focus a benchmark dataset on specific resistance mechanisms could simplify the task, but these choices would need to be agreed upon, justified and the limitations clearly stated. This reflection is to be linked to ongoing extensive discussions on the generation of appropriate databases of resistance genes and correct interpretation of resistome profiles (see 24,58). An *a priori* statement can be made that the benchmark dataset should focus on mechanisms of *acquired* bacterial resistance. Similarly, for lack of being exhaustive in terms of the AMR genetic determinants it includes, a set of *in silico* reference materials can be composed of the resistance mechanisms most relevant for public and environmental safety, for example, focusing on certain specific plasmids and AMR

genes which have been identified as being important in clinical infections.

The decisions through which specific resistance mechanisms are included in/excluded from the benchmark should be clear, transparent, agreed upon and justified in order to ensure that the benchmark is relevant to the types of risks considered. These will vary depending on the purpose of the experiment (e.g. environmental risk, clinical decision making), and will be important to evaluate whether the same resistance factors can be incorporated into a single benchmark dataset or if different resources will be needed. Transparency is important since these decisions will also guide the inclusion/exclusion of novel resistance mechanisms according to the changing epidemiology over time (see also section 2.4).

Should datasets representing different sample types (e.g. isolated clones, environmental samples) be included in the same benchmark? Is a correct representation of different bacterial species (host genomes) important?

The preceding section focused on the nature of the genetic determinants to be included in the *in silico* datasets. These sequences (i.e. AMR genes), however, represent a very small fraction of the overall totality of the sequence data generated from biological materials (e.g. bacterial genomes) in a given experiment. The nature of these majority “background” reads (bacterial host genomes, other contaminants in the sample etc.) in the components of a proper benchmark dataset thus needs to be carefully considered, as they can influence the accuracy of the pipelines.

The detection of drug resistance in clinical settings is often performed by sequencing pure cultured isolates^{18,59,60}. Pathogens of particular concern in the context of nosocomial infections will, accordingly, need to be properly represented in the *in silico* datasets. Lists of AMR pathogens presenting significant risks are maintained (see 61) and include the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* sp.) and *Escherichia coli*, among others.

Culture-dependent methods cannot be systematically applied to environmental samples for various reasons, including the fact that most environment bacteria are not recovered under standard culture conditions⁶². Culture-independent approaches (metagenomics) can then be used to analyse the human and environmental resistomes within complex bacterial populations^{13,25,63}. These approaches have also been proposed for clinical purposes, greatly reducing the time necessary for characterisation^{8,16}. For these samples, agreeing on a realistic genetic diversity within a benchmark⁶⁴ - a set of communities which can be considered “representative” - is a significant challenge as there is tremendous variability in the species composing the microbiomes of different communities^{13,65-67}.

2.3. Evaluation of pipeline performance

How can the “true” value of the samples, against which the pipelines will be evaluated, be guaranteed? What is needed to

demonstrate that the original sample has been correctly characterised, in case real experiments are used?

One of the objectives of validating a bioinformatics pipeline is to demonstrate that its accuracy is above an acceptable value, with low instances of false negative and false positive results⁶⁸. Antimicrobial susceptibility testing using traditional methods is, in itself, a complex procedure subject to differences in methodologies and interpretations⁶⁹; hence they have required (and will require) validation and standardisation⁷⁰⁻⁷². There have been reports where discrepancies between NGS-based predictions and susceptibility testing were caused by isolates with inhibition zones close to the susceptibility breakpoint. It was suggested that the results could have been concordant if the susceptibility testing had been performed under different culture conditions, for example, with a different culture medium⁷³. The extent to which these “borderline” cases should be included in the benchmark or not, and the final “correct” prediction that will be attached to them will need to be carefully considered. It should also be discussed what the most relevant endpoint in this context is, between, for example, the Minimum Inhibitory Concentration (MIC) prediction and resistance levels above wildtype/type strain.

The realities of veterinary medicine, with specific modalities of antimicrobial administration, mean that susceptibility MIC breakpoints may differ between humans and animals⁷⁴. Thus, the definition of science-based clinical MIC-breakpoints (CBPs) is relevant to interpret results and to harmonise the results of antimicrobial susceptibility testing of veterinary pathogens. Currently, this issue is being discussed in different working groups led by VETCAST. This may cause difficulties in assigning a universal “correct” label to some datasets that would apply to both humans and animals.

Reference samples of metagenomics experiments are even more complex in this regard, with each sample containing numerous instances of genetic AMR determinants^{12,14,75}. Metagenomics analyses can detect genes (genotype), which are not necessarily translated into resistance (phenotype); expression of the protein(s), which is not directly revealed by DNA sequencing, is important in this context. Assigning accurate profiles to components of a reference dataset will be challenging, as there is no existing pipeline recognised as the ‘gold standard’ to do so⁸. Spiked samples or simulated reads may be a necessary initial step in this context.

Ultimately, the “true” values to be assigned to the samples in the dataset, and the challenges this will pose, will depend on what the validated pipelines will be required to achieve. For example, the benchmark has to be adaptable to whether the aim is detecting the genetic determinants of AMR, predicting AMR or (for human and veterinary health) predicting antimicrobial susceptibility and thus treatment outcomes.

How should the target performance thresholds (e.g. specificity, sensitivity, accuracy) for the benchmark suite be set? What is the impact of these targets on the required size of the sample set?

Validation of a process involves the determination of various performance parameters, such as specificity, sensitivity, accuracy, etc.³². When used specifically for the detection of antimicrobial resistance the benchmark resources need to include strict performance thresholds, and whether these should be set *a priori* along with the levels of these thresholds are subjects for consideration. One also needs to clarify how the process can cope with cases where more than one type of resistance needs to be identified in a single sample, in particular for metagenomics studies.

These performance parameters will be important, not only as information to be included in the benchmark, but also because they generally have a significant influence on the size of the *in silico* dataset needed (see, for example, 76,77). Understanding the target performance characteristics of a valid pipeline will be necessary to guide decisions as to how many samples will be needed in the *in silico* dataset, with respect to the presence or absence of AMR genetic determinants. Finally, not all parameters are equally important for all samples - for example, considerations of sensitivity are generally not relevant in the case of cultured isolates as the bacteria are present in high numbers, but may be crucial for metagenomics experiments where the proportion of the target(s) relative to the background is variable and unknown. Targeted metagenomics seem promising approaches for the accurate detection of minority genes in complex samples¹³, and challenging the sensitivity of bioinformatics pipelines with a benchmark dataset would be of added value in this context.

2.4. Generation, distribution and update of the benchmark

How can the benchmark stay relevant when new resistance mechanisms are regularly characterised? How is the continued quality of the benchmark dataset ensured?

An important fact concerning antimicrobial resistance - and one of the reasons it represents a global health emergency - is that novel mechanisms of resistance are constantly being reported and new genes and/or vectors of transmission regularly emerge^{58,78}. Assuming that a benchmark resource can be produced covering the existing complexity of AMR determinants (section 2.2), adapting this resource to new information is a challenge that will need to be addressed in order to ensure that its utility does not diminish with time. Criteria for inclusion of new *in silico* datasets, and the mechanisms by which these decisions should be taken, need to be discussed and agreed upon when developing the resource.

Newly identified genetic determinants can also impact the information linked to existing datasets in the benchmark resources. These datasets will need to be re-evaluated in view of new information to ensure that their AMR determinants are properly characterised. As an example, this issue was evidenced in 2015 with the identification of *mcr-1* as a plasmid-borne colistin resistance gene⁷⁹; re-analysis of existing NGS data from *E.coli* isolates from food, feed and hospitalised patients for the previous years in Denmark revealed previously characterised samples containing this gene^{80,81}.

Who should generate the benchmark resource? How can it be efficiently shared?

Current guidelines and recommendations place the responsibility of validating the bioinformatics pipelines (and ensuring reliability after update of any of its components) with the operator/quality manager of the test facility^{32,33,39}. In fact, thus far, many different sets of benchmark materials and resources have been produced for local use or within collaborative endeavours (see 34). Benchmark datasets have also been used to compare different methods or tools^{17,82,83}. The extent to which these datasets address the concerns described in this document is the subject of a case-by-case evaluation that may become crucial for a wide implementation of NGS technology for routine and regulatory use. An open and inclusive discussion on the different issues (described here or arising upon more detailed considerations) will be important for the development of a resource that can gain wide acceptance and use.

Conclusions

The aim of this document is to summarise a list of challenges that were identified at the meeting organised by the Joint Research Centre on the 4th and 5th of October 2017 for the creation of a benchmark resource. The specific objective of this benchmark would be to challenge the bioinformatics step of a workflow to identify antimicrobial resistance in samples, using NGS technologies. It is clear that this covers only a fraction of the work necessary to fully implement this technology in a regulatory context, which will also need to cover additional steps such as the sampling, library preparation, sequencing run, and interpretation of the AMR profiles (see Figure 1). However, this resource would facilitate the implementation of the NGS technology in routine laboratory analyses by:

- Ensuring confidence in the implementation of the bioinformatics component of the procedure, a step currently identified as limiting in the field^{2,8-10}.
- Allowing evaluation and comparisons of new/existing bioinformatics strategies, resources and tools.
- Contributing to the validation of specific pipelines and the proficiency testing of testing facilities.
- "Future-proofing" bioinformatics pipelines to updates and replacement of the tools and resources used in their different steps.

Some of the challenges in building such a resource are common to all NGS-based methods. Many reports on standardisation, quality management and good laboratory practice have focused on clinical testing and the detection of germline sequence variants linked to cancer or other diseases and could guide some of the decisions to be taken. In this context, reference materials were highlighted as necessary for test validation, QC procedures and proficiency testing⁶⁸. However, many of the challenges also reflect the reality of antimicrobial resistance monitoring and are specific to this framework. How much of the available resources can be directly applied or used to guide future

efforts in this field will need evaluation and, eventually, complementation.

While the present discussion focuses on the monitoring of bacteria, most of the challenges described herein and the means by which they will be approached should apply to AMR monitoring in other organisms, such as viruses, parasites, and fungi. The differences will be in the final solutions proposed for the composition of the benchmark dataset, due, of course, to the different biology in the mode of action of the antimicrobials and their associated resistance mechanisms.

As it was made apparent in the previous sections, many of the challenges are due to the large heterogeneity behind the reality of detecting AMR using NGS. Some of this heterogeneity will require the development of separate benchmark datasets (e.g. the different sequencing platforms) while some will obviously gain by being combined into a single resource (e.g. human and veterinary medicine). Other cases will require more discussions and evaluations of feasibility/added value in being considered together *vs* separately (e.g. samples composed of isolates *vs* metagenomics).

Whatever the final composition and number of the benchmark resource(s), the proper path will ensure a holistic view of the problem that also reflects current public health data. This decision-making process should include expertise in AMR characterisation in humans, animals, food and the environment,

in order to maximise its impact on the establishment of an AMR surveillance framework that is in line with the principles of a “One Health” approach.

Disclaimer

The contents of this article are the views of the authors and do not necessarily represent an official position of the European Commission.

Data availability

No data is associated with this article.

Grant information

The “Next-generation sequencing technologies and antimicrobial resistance - Working groups kick-off” meeting (4–5 October 2017) was funded by the European Commission’s Joint Research Centre (JRC), Ispra, Italy.

Acknowledgments

We would like to thank Marc Struelens (European Centre for Disease Prevention and Control, ECDC), Ernesto Liebana Criado and Valentina Rizzi (European Food Safety Authority, EFSA) for their participation to the workshop discussions. We are also grateful to Maddalena Querci and Alex Patak (Joint Research Centre) for their help during the workshop and reviewing the manuscript.

References

- O'Neill J: **Antimicrobial resistance: tackling a crisis for the health and wealth of nations.** *Rev Antimicrob Resist.* 2014; 1–16.
[Reference Source](#)
- Revez J, Espinosa L, Albiger B, *et al.*: **Survey on the use of Whole-Genome Sequencing for Infectious Diseases Surveillance: Rapid Expansion of European National Capacities, 2015–2016.** *Front Public Health.* 2017; 5: 347.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- G7 Health Ministers: **Berlin Declaration on Antimicrobial Resistance.** 2015.
[Reference Source](#)
- Angers A, Petrillo M, Patak A, *et al.*: **The role and implementation of next-generation sequencing technologies in the coordinated action plan against antimicrobial resistance.** Publications Office of the European Union; 2017.
[Publisher Full Text](#)
- European Food Safety Authority (EFSA): **EFSA's 20th Scientific Colloquium on Whole Genome Sequencing of food-borne pathogens for public health protection.** *EFSA Support Publ.* 2015; 12(2).
[Publisher Full Text](#)
- European Centre for Disease Prevention and Control: **Expert opinion on whole genome sequencing for public health surveillance.** Stockholm: ECDC; 2016.
[Reference Source](#)
- Deurenberg RH, Bathoorn E, Chlebowski MA, *et al.*: **Application of next generation sequencing in clinical microbiology and infection prevention.** *J Biotechnol.* 2017; 243: 16–24.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Forbes JD, Knox NC, Ronholm J, *et al.*: **Metagenomics: The Next Culture-Independent Game Changer.** *Front Microbiol.* 2017; 8: 1069.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lesho E, Clifford R, Onmus-Leone F, *et al.*: **The Challenges of Implementing Next Generation Sequencing Across a Large Healthcare System, and the Molecular Epidemiology and Antibiotic Susceptibilities of Carbapenemase-Producing Bacteria in the Healthcare System of the U.S.** Department of Defense. *PLoS One.* 2016; 11(5): e0155770.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Punina NV, Makridakis NM, Remnev MA, *et al.*: **Whole-genome sequencing targets drug-resistant bacterial infections.** *Hum Genomics.* 2015; 9(1): 19.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rossen JWA, Friedrich AW, Moran-Gilad J, *et al.*: **Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology.** *Clin Microbiol Infect.* 2018; 24(4): 355–360.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bengtsson-Palme J, Boulund F, Fick J, *et al.*: **Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India.** *Front Microbiol.* 2014; 5: 648.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lanza VF, Baquero F, Martínez JL, *et al.*: **In-depth resistome analysis by targeted metagenomics.** *Microbiome.* 2018; 6(1): 11.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Noyes NR, Yang X, Linke LM, *et al.*: **Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems.** *Sci Rep.* 2016; 6(1): 24645.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Yu CY, Ang GY, Chong TM, *et al.*: **Complete genome sequencing revealed novel genetic contexts of the *mcr-1* gene in *Escherichia coli* strains.** *J Antimicrob Chemother.* 2017; 72(4): 1253–5.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Schmidt K, Mwaigwisya S, Crossman LC, *et al.*: **Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing.** *J Antimicrob Chemother.* 2017; 72(1): 104–14.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Clausen PT, Zankari E, Aarestrup FM, *et al.*: **Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data.** *J Antimicrob Chemother.* 2016; 71(9): 2484–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
- 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**(10): 2234–44.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. 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.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 20. 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**(1): 212–20.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 21. Yang Y, Li B, Zou S, *et al.*: **Fate of antibiotic resistance genes in sewage treatment plant revealed by metagenomic approach.** *Water Res.* 2014; **62**: 97–106.
[PubMed Abstract](#) | [Publisher Full Text](#)
 22. Gordon NC, Price JR, Cole K, *et al.*: **Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing.** *J Clin Microbiol.* 2014; **52**(4): 1182–91.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 23. Bengtsson-Palme J, Larsson DG: **Antibiotic resistance genes in the environment: prioritizing risks.** *Nat Rev Microbiol.* 2015; **13**(2): 396.
[PubMed Abstract](#) | [Publisher Full Text](#)
 24. Martínez JL, Coque TM, Baquero F: **What is a resistance gene? Ranking risk in resistomes.** *Nat Rev Microbiol.* 2015; **13**(2): 116–23.
[PubMed Abstract](#) | [Publisher Full Text](#)
 25. Bengtsson-Palme J, Larsson DGJ, Kristiansson E: **Using metagenomics to investigate human and environmental resistomes.** *J Antimicrob Chemother.* 2017; **72**(10): 2690–703.
[PubMed Abstract](#) | [Publisher Full Text](#)
 26. Hatem A, Bozdağ D, Toland AE, *et al.*: **Benchmarking short sequence mapping tools.** *BMC Bioinformatics.* 2013; **14**(1): 184.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 27. Martínez JL, Coque TM, Lanza VF, *et al.*: **Genomic and metagenomic technologies to explore the antibiotic resistance mobilome.** *Ann N Y Acad Sci.* 2017; **1388**(1): 26–41.
[PubMed Abstract](#) | [Publisher Full Text](#)
 28. Zhou J, He Z, Yang Y, *et al.*: **High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats.** *mBio.* 2015; **6**(1): pii: e02288–14.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 29. Endrullat C: **Standardization in next-generation sequencing-Issues and approaches of establishing standards in a highly dynamic environment.** *PeerJ Preprints.* 2017; **5**: e2771v1.
[Publisher Full Text](#)
 30. Inouye M, Dashnow H, Raven LA, *et al.*: **SRST2: Rapid genomic surveillance for public health and hospital microbiology labs.** *Genome Med.* 2014; **6**(11): 90.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 31. Rowe W, Baker KS, Verner-Jeffreys D, *et al.*: **Search Engine for Antimicrobial Resistance: A Cloud Compatible Pipeline and Web Interface for Rapidly Detecting Antimicrobial Resistance Genes Directly from Sequence Data.** *PLoS One.* 2015; **10**(7): e0133492.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 32. Endrullat C, Glöckler J, Franke P, *et al.*: **Standardization and quality management in next-generation sequencing.** *Appl Transl Genom.* 2016; **10**: 2–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 33. Lambert D, Pightling A, Griffiths E, *et al.*: **Baseline Practices for the Application of Genomic Data Supporting Regulatory Food Safety.** *J AOAC Int.* 2017; **100**(3): 721–731.
[PubMed Abstract](#) | [Publisher Full Text](#)
 34. Davies KD, Farooqi MS, Gruidl M, *et al.*: **Multi-Institutional FASTQ File Exchange as a Means of Proficiency Testing for Next-Generation Sequencing Bioinformatics and Variant Interpretation.** *J Mol Diagn.* 2016; **18**(4): 572–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
 35. Duncavage EJ, Abel HJ, Pfeifer JD: **In Silico Proficiency Testing for Clinical Next-Generation Sequencing.** *J Mol Diagn.* 2017; **19**(1): 35–42.
[PubMed Abstract](#) | [Publisher Full Text](#)
 36. Sczyrba A, Hofmann P, Belmann P, *et al.*: **Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software.** *Nat Methods.* 2017; **14**(11): 1063–1071.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 37. Lee H, Gurtowski J, Yoo S, *et al.*: **Third-generation sequencing and the future of genomics.** *BioRxiv.* 2016; 048603.
[Publisher Full Text](#)
 38. Food and Drug Administration: **Code of Federal Regulations Title 21—Food and Drugs.** *Parts.* 1970; **10**: 121–200.
[Reference Source](#)
 39. Gargis AS, Kalman L, Bick DP, *et al.*: **Good laboratory practice for clinical next-generation sequencing informatics pipelines.** *Nat Biotechnol.* 2015; **33**(7): 689–93.
[PubMed Abstract](#) | [Publisher Full Text](#)
 40. Cock PJ, Fields CJ, Goto N, *et al.*: **The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants.** *Nucleic Acids Res.* 2010; **38**(6): 1767–71.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 41. Li H, Handsaker B, Wysoker A, *et al.*: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics.* 2009; **25**(16): 2078–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 42. Mellmann A, Andersen PS, Bletz S, *et al.*: **High Interlaboratory Reproducibility and Accuracy of Next-Generation-Sequencing-Based Bacterial Genotyping in a Ring Trial.** *J Clin Microbiol.* 2017; **55**(3): 908–13.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 43. Quail MA, Smith M, Coupland P, *et al.*: **A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.** *BMC Genomics.* 2012; **13**(1): 341.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 44. Suzuki S, Ono N, Furusawa C, *et al.*: **Comparison of sequence reads obtained from three next-generation sequencing platforms.** *PLoS One.* 2011; **6**(5): e19534.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 45. Gutmann A, Wagner J, Ali Y, *et al.*: **Privacy and progress in whole genome sequencing.** *Pres Comm Study Bioethical.* 2012; (2012).
[Reference Source](#)
 46. ATCC: **The Rise of Multidrug-Resistant Strains and Need for New Therapeutic Approaches (White Paper).** 2016.
[Reference Source](#)
 47. Huang W, Li L, Myers JR, *et al.*: **ART: a next-generation sequencing read simulator.** *Bioinformatics.* 2012; **28**(4): 593–4.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 48. Motro Y, Moran-Gilad J: **Microbial Metagenomics Mock Scenario-based Sample Simulation (M₃S₃).** *Clin Microbiol Infect.* 2018; **24**(3): 308.e1–308.e4.
[PubMed Abstract](#) | [Publisher Full Text](#)
 49. Ono Y, Asai K, Hamada M: **PBSIM: PacBio reads simulator—toward accurate genome assembly.** *Bioinformatics.* 2013; **29**(1): 119–21.
[PubMed Abstract](#) | [Publisher Full Text](#)
 50. Shcherbina A: **FASTQSim: platform-independent data characterization and in silico read generation for NGS datasets.** *BMC Res Notes.* 2014; **7**(1): 533.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 51. Yang C, Chu J, Warren RL, *et al.*: **NanoSim: nanopore sequence read simulator based on statistical characterization.** *Gigascience.* 2017; **6**(4): 1–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 52. Drezek E, Lavenier D: **Quality metrics for benchmarking sequences comparison tools.** In: *Brazilian Symposium on Bioinformatics.* Springer; 2014. 144–53.
[Publisher Full Text](#)
 53. Smolka M, Rescheneder P, Schatz MC, *et al.*: **Teaser: Individualized benchmarking and optimization of read mapping results for NGS data.** *Genome Biol.* 2015; **16**(1): 235.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 54. Lin J, Nishino K, Roberts MC, *et al.*: **Mechanisms of antibiotic resistance.** *Front Microbiol.* 2015; **6**: 34.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 55. Ellington MJ, Ekelund O, Aarestrup FM, *et al.*: **The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee.** *Clin Microbiol Infect.* 2017; **23**(1): 2–22.
[PubMed Abstract](#) | [Publisher Full Text](#)
 56. Deininger KN, Horikawa A, Kitko RD, *et al.*: **A requirement of ToIC and MDR efflux pumps for acid adaptation and GadAB induction in *Escherichia coli*.** *PLoS One.* 2011; **6**(4): e18960.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 57. Knowles M, Stinson S, Lambert D, *et al.*: **Genomic Tools for Customized Recovery and Detection of Foodborne Shiga Toxinigenic *Escherichia coli*.** *J Food Prot.* 2016; **79**(12): 2066–77.
[PubMed Abstract](#) | [Publisher Full Text](#)
 58. Bengtsson-Palme J, Kristiansson E, Larsson DGJ: **Environmental factors influencing the development and spread of antibiotic resistance.** *FEMS Microbiol Rev.* 2018; **42**(1): fux053.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 59. 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**(3): pii: e001124.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 60. Sherry NL, Porter JL, Seemann T, *et al.*: **Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory.** *J Clin Microbiol.* 2013; **51**(5): 1396–401.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 61. Garner MJ, Carson C, Lingohr EJ, *et al.*: **An assessment of antimicrobial resistant disease threats in Canada.** Conly J, editor. *PLoS One.* 2015; **10**(4): e0125155.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 62. Berendonk TU, Manaia CM, Merlin C, *et al.*: **Tackling antibiotic resistance: the environmental framework.** *Nat Rev Microbiol.* 2015; **13**(5): 310–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
 63. Crofts TS, Gasparrini AJ, Dantas G: **Next-generation approaches to understand and combat the antibiotic resistome.** *Nat Rev Microbiol.* 2017; **15**(7): 422–434.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 64. Lindgreen S, Adair KL, Gardner PP: **An evaluation of the accuracy and speed of metagenome analysis tools.** *Sci Rep.* 2016; **6**: 19233.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 65. Forster SC: **Illuminating microbial diversity.** *Nat Rev Microbiol.* Nature Publishing Group; 2017; **15**(10): 578.
[PubMed Abstract](#) | [Publisher Full Text](#)

66. Gibson MK, Forsberg KJ, Dantas G: **Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology.** *ISME J.* 2015; **9**(1): 207–16.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
67. Schlaberg R, Chiu CY, Miller S, *et al.*: **Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection.** *Arch Pathol Lab Med.* 2017; **141**(6): 776–86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
68. Gargis AS, Kalman L, Berry MW, *et al.*: **Assuring the quality of next-generation sequencing in clinical laboratory practice.** *Nat Biotechnol.* 2012; **30**(11): 1033–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
69. Var SK, Hadi R, Khardori NM: **Evaluation of regional antibiograms to monitor antimicrobial resistance in hampton roads, Virginia.** *Ann Clin Microbiol Antimicrob.* 2015; **14**(1): 22.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
70. Brown D, Canton R, Dubreuil L, *et al.*: **Widespread implementation of EUCAST breakpoints for antibacterial susceptibility testing in Europe.** *Euro Surveill.* 2015; **20**(2): pii: 21008.
[PubMed Abstract](#) | [Publisher Full Text](#)
71. Kahlmeter G: **Defining antibiotic resistance-towards international harmonization.** *Ups J Med Sci.* 2014; **119**(2): 78–86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
72. Leclercq R, Cantón R, Brown DF, *et al.*: **EUCAST expert rules in antimicrobial susceptibility testing.** *Clin Microbiol Infect.* 2013; **19**(2): 141–60.
[PubMed Abstract](#) | [Publisher Full Text](#)
73. Aanensen DM, Feil EJ, Holden MT, *et al.*: **Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive *Staphylococcus aureus* in Europe.** *MBio.* 2016; **7**(3): pii: e00444–16.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
74. Toutain PL, Bousquet-Mélou A, Damborg P, *et al.*: **En Route towards European Clinical Breakpoints for Veterinary Antimicrobial Susceptibility Testing: A Position Paper Explaining the VetCAST Approach.** *Front Microbiol.* 2017; **8**: 2344.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
75. Caucci S, Karkman A, Cacace D, *et al.*: **Seasonality of antibiotic prescriptions for outpatients and resistance genes in sewers and wastewater treatment plant outflow.** *FEMS Microbiol Ecol.* 2016; **92**(5): fiw060.
[PubMed Abstract](#) | [Publisher Full Text](#)
76. Bujang MA, Adnan TH: **Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis.** *J Clin Diagn Res.* 2016; **10**(10): YE01–YE06.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. Rao Y, Lee Y, Hsu JC: **Determination of sample size for validation study in pharmacogenomics.** preparation. 2009.
[Reference Source](#)
78. Blair JM, Webber MA, Baylay AJ, *et al.*: **Molecular mechanisms of antibiotic resistance.** *Nat Rev Microbiol.* 2015; **13**(1): 42–51.
[PubMed Abstract](#) | [Publisher Full Text](#)
79. 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**(2): 161–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
80. Hasman H, Hammerum AM, Hansen F, *et al.*: **Detection of *mcr-1* encoding plasmid-mediated colistin-resistant *Escherichia coli* isolates from human bloodstream infection and imported chicken meat, Denmark 2015.** *Euro Surveill.* 2015; **20**(49).
[PubMed Abstract](#) | [Publisher Full Text](#)
81. Kluytmans-van den Bergh MF, Huizinga P, Bonten MJ, *et al.*: **Presence of *mcr-1*-positive *Enterobacteriaceae* in retail chicken meat but not in humans in the Netherlands since 2009.** *Euro Surveill.* 2016; **21**(9): 30149.
[PubMed Abstract](#) | [Publisher Full Text](#)
82. Phelan J, O'Sullivan DM, Machado D, *et al.*: **The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs.** *Genome Med.* 2016; **8**(1): 132.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
83. Xavier BB, Das AJ, Cochrane G, *et al.*: **Consolidating and Exploring Antibiotic Resistance Gene Data Resources.** Kraft CS, editor. *J Clin Microbiol.* 2016; **54**(4): 851–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 29 June 2018

<https://doi.org/10.5256/f1000research.15794.r35036>

© 2018 Kwong J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Jason C. Kwong 

¹ University of Melbourne, Melbourne, Vic, Australia

² Austin Health, Melbourne, Vic, Australia

The opinion article by Alexandre Angers-Loustau and colleagues addresses an increasingly encountered question by clinical and public health microbiology laboratories - what are the considerations in developing a benchmark for validating methods to identify antimicrobial resistance determinants from high-throughput sequencing data? This is an important step in using sequencing data for this purpose, and this article reports back on the discussions by a panel of experts in the field of antimicrobial resistance surveillance.

In general, the article is well-suited to publication in *F1000Research*, offering insight into an issue of global importance, and I thank the authors for making the discussions from the October 2017 meeting organised by the Joint Research Centre public. It is well written and explores each of the discussion questions with supporting literature. My suggestions for improvement are minor and addressing them is left up to the authors' discretion.

My overall comment is that while many of the relevant considerations are raised, as a clinician and researcher, I am left with the burden of all these concerns, but no clear direction in how to address the challenges (other than a few comments in the conclusion). I appreciate the purpose of the paper was to summarise the challenges identified at the meeting, but I think this (opinion) article could be enhanced and have a greater impact if more expert opinions and ideas on how to proceed were included.

Minor comments:

Introduction (page 4)

It would be nice to know a little more about the panel of experts - for example, what mix of background or fields did the experts come from e.g. clinical or public health microbiology, bioinformaticians, epidemiologists, veterinary microbiology etc.?

Section 2.1: real experiments vs simulated read sets (page 5)

This section highlights some important concerns (e.g. filtering host DNA or patented genetically modified food DNA), but these are more relevant to identification of antimicrobial resistance from metagenomic

sequencing data - this is alluded to in the text, but not explicitly stated. Given the other technical challenges with this, I wonder if the discussion should focus more on the "real experimental" data from cultured bacterial organisms – the more likely initial validation before validation proceeds to metagenomic datasets.

Section 2.2: how should the set of resistance genes/mechanisms be agreed upon (page 6)
The considerations here are well explored, but I wonder if there could be some mention of variation in benchmark datasets depending on the purpose of resistance gene/mechanism detection. For example, detection of AMR genes from environmental or livestock samples may have a different focus, and thus different benchmark targets to detection of AMR genes in organisms from human samples. Even among human samples, there may be differences in the mechanisms prioritised for detection between testing for clinical treatment and testing for epidemiological surveillance (*mcr-1* is a good example).

Section 2.3: how can the "true" value of the samples be guaranteed (page 7)
Great section and discussion. Along the lines of the discussion, I am interested to hear the authors' thoughts on whether it is sufficient to validate NGS for predicting AMR against standard susceptibility testing (e.g. MIC), or whether clinical studies evaluating treatment outcomes based on NGS detection are required (i.e. treating based on genotype vs phenotype). My only comment for this section is that perhaps some of the commentary is more directed at validation of NGS for "identifying antimicrobial resistance" rather than "identifying antimicrobial resistance determinants" as stated in the title - two subtly different targets. Perhaps some of the discussion could be clarified here.

"HTF5 format" (page 4) should actually be "HDF5 format".

MIC should be spelt out at the first usage - minimum inhibitory concentration.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical infectious diseases, microbial pathogen genomics, antimicrobial agents and resistance, clinical metagenomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 11 May 2018

<https://doi.org/10.5256/f1000research.15794.r33450>

© 2018 Palù G et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Enrico Lavezzo 

University of Padova, Padova, Italy

Giorgio Palù

Department of Molecular Medicine, University of Padova, Padova, Italy

In this opinion article, Alexandre Angers-Loustau and colleagues address a hot topic in Diagnostic Microbiology regarding the implementation of next generation sequencing platforms for the identification of antimicrobial resistances. In particular, they focus on a specific step of the long pipeline which goes from sample collection to data interpretation and management, namely the bioinformatics analysis of sequencing data.

It's been more than 10 years since the introduction of the first NGS instrument, but an adequate standardization and harmonization level of the many different platforms is not yet accomplished.

The article is very well written; it is concise but covers all the important issues that need to be considered when willing to implement such technologies for the purpose of AMR detection.

I only have some minor points for the authors:

1. In section 2.1, page 5, regarding the potential presence of sensitive information (e.g. reads of human origin) when including real experiment datasets in the benchmark. I don't see the problem here. How the availability of anonymous DNA sequence information could compromise the patient's privacy? Of course, patient's identity must be separated from sample's ID, but the de-identification process is a standard procedure and others methods have also been developed for this purpose (see for example this paper from Qamar et al. (2014)¹). Please discuss about this.
2. Since this is an opinion article, I would like the authors to 'expose' themselves a little bit more on some points. The challenges are presented by explaining the state-of-the-art, but it is not easy to me to understand authors' thinking about the possible solutions. Obviously they can't have answers for all the challenges, but they could express some personal positions (for example, what about "Who should generate the benchmark resource"? To me, there should be a centralized action aimed, as far as possible, at the definition of a global standard (similar for example to the WHO HPV LabNet Proficiency Studies for Evaluating HPV DNA Typing Methods). What do the authors think about this?
3. Finally, no reference is made to microorganisms other than bacteria, but AMR may be related also to viruses and eukaryotic microorganisms. Are there any difference or peculiarity regarding the bioinformatics analysis of sequencing data coming from them?

References

1. Qamar N, Yang Y, Nadas A, Liu Z, et al.: Anonymously Analyzing Clinical Datasets. *arXiv preprint arXiv:1501.05916*. 2014.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Partly

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Microbiology, bioinformatics, NGS data analysis

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research