Bacterial genome sequencing and analysis: paving the way for a Swiss-wide molecular epidemiological surveillance platform

Egli, Adrian ; Blanc, Dominique S ; Greub, Gilbert ; Keller, Peter M ; Lazarevic, Vladimir ; Lebrand, Aitana ; Leib, Stephen ; Neher, Richard A ; Perreten, Vincent ; Ramette, Alban ; Schrenzel, Jaques ; Stephan, Roger ; Wagner, Karoline ; Wüthrich, Daniel ; Xenarios, Ioannis

Abstract: Facing multidrug resistant (MDR) bacterial pathogens is one of the most important challenges for our society. The spread of highly virulent and resistant pathogens can be described using molecular typing technologies; in particular, whole genome sequencing (WGS) data can be usedf or molecular typing purposes with high resolution. WGS data analysis can explain the spatiotemporal patterns of pathogen transmission. However, the transmission between compartments (human, animal, food, environment) is very complex. Interoperable and curated metadata are a key requirement for fully understanding this complexity. In addition, high quality sequence data are a key element between centres using WGS data for diagnostic and epidemiological applications. We aim to describe steps to improve WGS data analysis and to implement a molecular surveillance platform allowing integration of high resolution WGS typing data and epidemiological data.

DOI: https://doi.org/10.4414/smw.2019.14707

Established in 1871

Swiss Medical Weekly

Formerly: Schweizerische Medizinische Wochenschrift

An open access, online journal • www.smw.ch

Review article: Biomedical intelligence | Published 15 December 2018 | doi:10.4414/smw.2018.14693
Cite this as: Swiss Med Wkly. 2018;148:w14693

Improving the quality and workflow of bacterial genome sequencing and analysis: paving the way for a Switzerland-wide molecular epidemiological surveillance platform

Egli Adrian[ab], Blanc Dominique S.[c], Greub Gilbert[d], Keller Peter M.[e], Lazarevic Vladimir[f], Lebrand Aitana[g], Leib Stephen[h], Neher Richard A.[gi], Perreten Vincent[j], Ramette Alban[hk], Schrenzel Jacques[fl], Stephan Roger[m], Wagner Karoline[e], Wüthrich Daniel[ab], Xenarios Ioannis[g]

[a] Division of Clinical Microbiology, University Hospital Basel, Switzerland
[b] Applied Microbiology Research, Department of Biomedicine, University of Basel, Switzerland
[c] Hospital Preventive Medicine Service, Lausanne University Hospital, Switzerland
[d] Institute for Microbiology, Lausanne University Hospital, Switzerland
[e] Institute of Microbiology, University of Zurich, Switzerland
[f] Bacteriology Laboratory, Division of Laboratory Medicine, Department of Genetics Laboratory Medicine and Pathology, Geneva University Hospitals, Switzerland
[g] SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland
[h] Institute for Infectious Diseases, University of Bern, Switzerland
[i] Biozentrum, University of Basel, Switzerland
[j] Institute of Veterinary Bacteriology, Vetsuisse Faculty, University of Bern, Switzerland
[k] Swiss Centre for Antibiotic Resistance (anresis.ch), Institut für Infektionskrankheiten, University of Bern, Switzerland
[l] Genomic Research Laboratory, Division of Infectious Diseases, Geneva University Hospitals, Switzerland
[m] Institute for Food Safety, Vetsuisse Faculty, University of Zurich, Switzerland

Summary

Facing multidrug resistant (MDR) bacterial pathogens is one of the most important challenges for our society. The spread of highly virulent and resistant pathogens can be described using molecular typing technologies; in particular, whole genome sequencing (WGS) data can be used for molecular typing purposes with high resolution. WGS data analysis can explain the spatiotemporal patterns of pathogen transmission. However, the transmission between compartments (human, animal, food, environment) is very complex. Interoperable and curated metadata are a key requirement for fully understanding this complexity. In addition, high quality sequence data are a key element between centres using WGS data for diagnostic and epidemiological applications. We aim to describe steps to improve WGS data analysis and to implement a molecular surveillance platform allowing integration of high resolution WGS typing data and epidemiological data.

Keywords: *next generation sequencing, whole genome sequencing, molecular epidemiology, typing, quality assessment, quality control, ring trial, surveillance database, Swiss pathogen surveillance platform*

Author contributions
AE, GG, and JS provided the concept of the article. AE, GG, JS, DB, AL, VP, RN, AR, IX, PMK, KW, VL, SL, RS, and DW wrote the manuscript.

Correspondence:
Adrian Egli, MD PhD, Division of Clinical Microbiology, University Hospital Basel, Petersgraben 4, CH-4031 Basel, adrian.egli[at]usb.ch

Introduction

Facing multidrug resistant (MDR) bacterial pathogens is one of the most important challenges for our society. The rapid worldwide expansion of MDR pathogens may exceed 10 million fatal cases per year by 2050, exceeding today's cancer-related deaths [1]. In humans, MDR pathogens such as extended spectrum beta-lactamase (ESBL)-producing or carbapenemase-producing Enterobacteriaceae are increasingly detected in the community and in the hospital setting [2, 3]. Similarly, methicillin-resistant *Staphylococcus aureus* (MRSA) represents an important threat in the community and hospitals [4–6]. Specific virulence factors such as Panton-Valentine leucocidin [7], toxic shock syndrome toxin [8] or reduced susceptibility to vancomycin have been associated with more severe clinical courses and therapeutic failures [9, 10]. Overall, in humans MDR bacteria are associated with significant morbidity, mortality and healthcare costs [11]. Today, MRSA spread in hospitals could be reduced by a better understanding of transmission patterns [12] and the implementation of efficient surveillance strategies including screening of high-risk patients and immediate isolation of colonised or infected patients to interrupt the transmission chain [13]. However, these strategies have not been successful for all MDR pathogens, as indicated by the overall increase in resistance rates in Europe (http://www.ecdc.europe.eu [14, 15]) and specifically in Switzerland as reported by the national surveillance system for antibiotic resistance and consumption in Switzerland[16, 17].

The emergence of MDR bacteria in veterinary medicine [18] leads to a high economic burden and constitutes a reservoir of antibiotic resistance genes and of zoonoses that represent a serious public health concern [19, 20]. Transmission events between domestic or farm animals and humans have been described in the past by use of various bacterial typing methods, including pulsed field gel electrophoresis and, more recently, whole genome sequencing (WGS) [21–23]. These two compartments – humans and animals – are linked via environmental sources such as sewage waters, farming, livestock and domestic animals, and the food chain [24–26].

Recent technological advances in molecular epidemiology, including high-throughput sequencing and data analysis, now allow bacterial typing at higher resolution (whole genome), and a better understanding of transmission rates and routes within individual compartments [27]. In addition, WGS data show high portability, as genomic sequences can be shared and compared across centres, in contrast to other typing methods [28]. However, only a few studies have used WGS-based typing to link various compartments, such as human and veterinary specimens with environmental samples [29–33]. Such high-resolution transmission information may close the current knowledge gaps and help decipher transmission routes at higher spatiotemporal resolution.

With the introduction of a country-wide molecular surveillance database in Switzerland, we aim to (1) increase the overall WGS data quality by defining common best practices for WGS typing analyses based on ring trials results and (2) establish a Switzerland-wide molecular surveillance platform for selected pathogens present in different environments and circulating in both veterinary and human settings.

Improving whole genome sequencing quality on the way to an interoperable surveillance platform

In general, the similarity of pathogens can be described on a genomic, plasmid or protein level, and differences can be visualised in phylogenetic trees, allowing the identification of closely related isolates [28, 34]. Short-read (e.g., Illumina and IonTorrent) and long-read sequencing technologies (e.g. Oxford Nanopore Technologies and Pacific Biosciences) allow the whole pathogen genome to be reconstructed [35]. Thereby, WGS provides significantly higher resolution for determining transmission events compared with classical typing technologies such as pulse field gel electrophoresis or multi-locus sequencing typing (MLST) [36–39]. With WGS data, several typing approaches can be used for isolate comparison: (1) variant call of single nucleotide polymorphisms, (2) core genome multi-locus sequencing type (cgMLST), (3) whole genome multi-locus sequencing type (wgMLST), and (4) k-mer-based grouping of the closest genome matches [40–43]. The choice of a reference genome is critical when sequencing data are mapped to detect variants for whole genome reconstruction. In general, the choice of non-outbreak-related isolate might influence data interpretation. Differences in the selected reference genomes and sequencing technology used can lead to differences between centres in the quantification of isolate relatedness. Future needs are inter-laboratory agreements on minimal sequencing quality requirements and a set of common quality markers. With this we want to ensure that the critical quality aspects such as genome coverage, average read-depth and base-call quality are controlled within all submitted datasets.

The workflow to generate WGS-based phylogenetic data includes a wet- and a dry-lab component. At the wet-lab, specifically trained laboratory technicians register the sample and add available metadata in a laboratory information system or a biobanking information management system. Next, technicians prepare the sample, extract the DNA (or RNA) in sufficient quality and quantity, prepare the library and perform the sequencing. At the dry lab, bioinformaticians start the data processing with quality checks and analyse the WGS data using either a commercial or an in-house developed bioinformatics pipeline in order to detect and visualise differences in the sequence data. Quality checks (such as Phred Quality Scores, coverage rates, proportion of core genes) at each step are crucial for reliable analysis results [44, 45], and the frequency and length of contigs (e.g., N50) provide important baseline quality surrogate markers for overall WGS quality. Results will be interpreted together with clinical microbiologists, hospital epidemiologists, infectious disease physicians and public health experts.

For many pathogens, the analytical workflow is a major bottleneck, especially since best practices and pipelines are not available. The resulting delay can be associated with higher overall costs and have deleterious effects on healthcare systems [46–48], such as a delay in making connections among individual samples during an outbreak. Nowadays, WGS data are either *de novo* assembled or mapped against the genome of a particular reference strain, which is the basis for reconstructing a phylogeny. This process can be rather time-consuming and currently has a series of uncertainties, including differences between bioinformatics pipelines and pipeline specifications regarding quality controls, lack of epidemiological metadata, unavailability of non-outbreak-related isolates, etc. Because of these differences, the Swiss Institute of Bioinformatics has initiated a nationwide quality assessment ring trial focusing on bacterial phylogeny. Ten Swiss diagnostic and research laboratories, representing all Swiss university hospitals and several research groups, are currently participating in a WGS-based bacterial typing ring trial in three phases, assessing key aspects of sequencing, data analysis and clinical context. Results of this ring trial will be published soon.

Similarly, the European Centre for Disease Prevention and Control (ECDC) has recently published an expert opinion on WGS as a tool for public health surveillance of MDR pathogens. The ECDC concluded that the establishment of standards and systems enabling EU-wide use of WGS as the method of choice for typing microbial pathogens will replace other methods and that this will improve the accuracy and effectiveness of disease surveillance, outbreak investigation and evaluation of prevention policies [49]. Quality control of the sequencing procedure and standardisation of the analytical process will be a key element of the ECDC initiative. Similarly, in Switzerland, the recent ring trial initiative aims to improve the overall quality of sequencing and subsequent data analysis.

Molecular surveillance with spatiotemporal resolution integrating epidemiological metadata.

The increased knowledge on the biology of the pathogens and their transmission has already contributed to important insights regarding their pathogenicity and evolution. Still unresolved aspects can be highlighted:

1. Identification of potential sources of transmission. For example, which are the compartments serving as the sources of resistant isolates? What is the role of food?
2. Uncertainties about the potential routes of transmission. For example, what are the (molecularly confirmed) nodes within the network between humans and animals? What is the probability that resistance is transferred from animals or from the environment? What is the contribution of human-to-human spread in the community? How are the compartments livestock, community and hospital linked in a network of transmission? What is the role of domestic animals in the transmission chain?
3. Description of the consequences following pathogen transmission. For example, what is the prevalence of specific resistant subtypes such as. MRSA with decreased vancomycin susceptibilities? What are the most prevalent subtypes of, for example, ESBL- and carbapenemase-producing Enterobacteriaceae and why are they successfully persisting in the environment? Do particular virulent strains increase in frequencies over time?

Different settings including humans, animals and the environment play a varied and complex role in transmission events and dynamics. Phylogeny based on high-resolution WGS data allows a profound understanding of the relation between samples; however, for the transmission pathway other information needs to be included, such as clinical information or additional microbiological testing of vectors and surfaces. The transmission dynamics between compartments are very difficult to assess, in particular when additional, less frequently sampled compartments are introduced into the equation. Transmission rates and routes depend on factors such as antibiotic usage in hospitals and outpatient settings [50], social and behavioural aspects [51], immunity of the host [52, 53], the particular pathogenic species and its ability to survive and thrive in the environment [54, 55], which also involves genetic exchanges between species [56–58]. On a large spatial scale (region or nation), transmission events between the community, hospitals and animal compartments are therefore complex and poorly understood overall, notably because of limited data on the transmission events and tracking between the compartments. Thus, additional interoperable epidemiological and clinical metadata are required to provide a better understanding of transmission dynamics and to guide containment of such MDR pathogens. This includes information on the time and geographical location of detection and invasiveness (screening vs infection), host (animal or human), environment of detection (sewage, hospital or food), specific microbiological features of an isolate (phenotypic resistance or virulence), etc. (table 1). Such information should be integrated into genetic distance comparisons and phylogenetic analyses. Switzerland may serve as a good model for the evaluation of such tools because of its geographical location, small and medium size of hospitals and already implemented monitoring of resistance in both human and animals (anresis.ch).

Although WGS-based phylogeny and geographical data have been combined previously at specific geographical areas or around the globe (e.g., *Francisella tularensis* in Norway [59], *Mycobacterium tuberculosis* [60], *Salmonella enteritidis* subspecies [61], MRSA [62], and Zika [63, 64], influenza [65], Ebola [66, 67] and Lyssa viruses [68]), the geographical resolution and dynamic aspects were often not sufficient to have a meaningful and immediate public health impact. More recently, the *microreact tool* (microreact.org [69]) was used to describe the spatial distribution of MRSA isolates across Europe [70]. A total of 308 invasive *S. aureus* isolates were integrated into the analysis, and the distribution of particular frequent clonal complexes (CC5, CC22, and CC30) were discussed [70]. The *microreact* tool allowed identification of high-risk clones on the basis of population level properties such as clonal relatedness, abundance and spatial structuring. Epidemiological or mathematical models can be performed with this tool.

Table 1: Important minimum interoperable metadata for a surveillance database.

	Description details
Pathogen	
Species_type	Bacteria; viruses
Species_name	NCBI taxonomy name
Taxonomy_ID	NCBI taxonomy ID
Sample	
Isolation_source	Human; animal; environmental; food
Isolation_source_detailed	Site of isolation e.g., SNOMED CT for human
Isolation_country	Country where the sample was isolated
Isolation_date	Date of isolation
Laboratory_ID	Unique ID of the laboratory processing the sample
Patient (if isolation_source == "Human")	
Host_age	Age of the patient
Host_sex	Gender of the patient
Home_address_zip	ZIP code of the patient's home
Home_address_country	Country of the patient's home
Clinic	Name of hospital/clinic where patient is treated
Clinic_ward	Ward where patient is staying
Clinic_room	Room number where patient is staying
Study_ethical_consent	Describes if patient signed a study ethical consent
Sequencing	
Sequencing_lab_ID	Unique ID of the sequencing facility/laboratory
DNA_extraction_kit	Kit used for DNA extraction
Library_preparation_kit	Kit used for library preparation
Sequencing_platform	Technology used for sequencing, e.g., Illumina MiSeq
Bioinformatics	
Quality_filtering_tool	Name (version) of the tool used for quality filtering reads
Assembly_method	Mapping-based assembly; de-novo assembly
Assembly_tool	Name (version) of the tool used for genome assembly
Reference_genome	Reference genome used for mapping (FASTA format)

Recently, Hadfield and colleagues developed a phylodynamic analysis pipeline coupled to a web-browser based interactive visualisation for pathogens such as Zika, Ebola, influenza, dengue and measles viruses (nextstrain.org, see fig. 1) [71, 72]. *Nextstrain* was designed to facilitate near real-time analysis of pathogen sequence data. New data can be incorporated within hours and the results are readily disseminated via the web application. *Nextstrain* grew out of the real-time surveillance application *nextflu*, which provides phylodynamic analysis of influenza virus sequence and serological data in weekly updates (nextflu.org). *Nextstrain* infers likely transmission events, ancestral compartments, and divergence or introduction times. The phylogenetic analysis is coupled to a geographic display of sampling locations and transmission routes on a map. (See table 2 for a list of platforms and initiatives.)

Although these web-based platforms provide first important steps, data analysis remained somewhat static and did not allow prediction of the dynamic properties of an epidemic or outbreak, such as the basic reproduction number (R_0), predictions of expected total cases within the outbreak, modelling of spatiotemporal transmission events including both human and animal strain information, or identification of common genes associated with successful transmission in a highly automated way. In addition, both platforms are dependent on the upload of preprocessed curated and quality approved WGS and epidemic data. Ideally, a future surveillance platform would automatically preassess data quality (genomics and metadata), and then ask human experts for validation in all or selected (borderline) cases. These are important gaps in current tools, which have to be addressed to gain further insight into real-time transmission dynamics. The requested key features for a molecular surveillance database are summarised in supplementary table S1 in appendix 1.

Table 2: Selection of surveillance databases, initiatives, and analytical tools using whole genome sequence information from pathogens

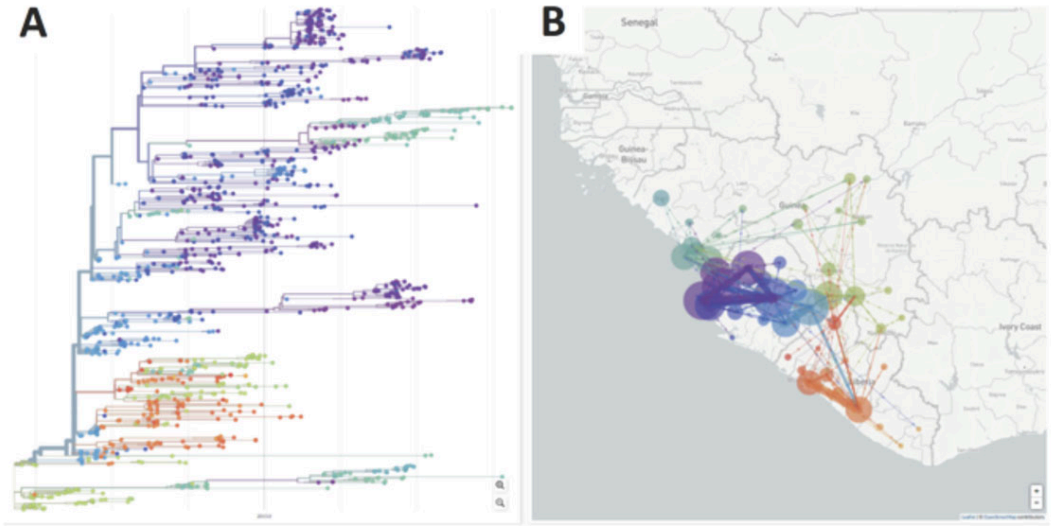
	Weblink
Database	
Global microbial identifier	http://www.globalmicrobialidentifier.org
Compare	http://www.compare-europe.eu/
IRIDA	http://www.irida.ca/
Innuendo	http://www.innuendoweb.org/
Center for Genomic Epidemiology	http://www.genomicepidemiology.org/
Enterobase	http://enterobase.warwick.ac.uk/
MLST databases	https://pubmlst.org/
Analytical tools	
Microreact	https://microreact.org
Nextstrain	https://nextstrain.org/

To date, there is no specific platform in Switzerland and Europe that allows the integration of WGS data from various diagnostic or reference laboratories, with a special focus on antibiotic resistance and virulence in combination with spatiotemporal information and visualisations. Considering the impact of MDR on our society and the rapidly evolving dynamics of MDR pathogens, there is a clear and urgent need for such a platform.

Steps to the future of dynamic visualisation using genomic and phylogenetic data.

A key feature of such an inter-institutional platform has to be the interoperability of sharable data. This can be achieved with a common syntax and nomenclature starting from data entry, species description and minimal datasets of documented microbiological information. Databases such as the NCBI taxonomy browser or the list of prokaryotic names with standing in nomenclature (LPSN) provide a definition for single bacterial species names. In addition, healthcare focused vocabularies and ontologies such as Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) or LOINC may offer detailed descriptions

Figure 1: NextStrain tool to analyse transmission of pathogens. (A) Phylogenetic tree of the Ebola virus outbreak in West Africa. Colours represent different geographic regions, and x-axis reflects a time scale during the outbreak. (B) Transmission dynamics modelled in real-time with new viral strains based on WGS data.



of the sample context, such as the anatomic site of isolation (e.g., abdominal structure SCTID: 113345001) and the source material (e.g., urine SCTID: 122575003). The usage of such controlled languages will be one of the most critical aspects in connecting different centres and databases to ensure efficient exchange [73]. As previously mentioned, we would envision using the recently defined biobanking minimal description set for microbes, which was established within the microbiology working group of the Swiss Biobanking Platform (see table 1).

Not only epidemiological relatedness can be extracted from WGS data. Virulence and antibiotic resistance could also be assessed. Various web-based tools and databases exist [74], providing data on virulence and resistance. Examples are SARG [75], VFDB [76–78], ResFinder [79, 80], PointFinder [81], or Card [82, 83]. Tools such as *PanGenome* allow rapidly searching and comparison of genetic information on a whole genome scale within single strains or larger clusters of isolates (www.pangenome.de [84]). Thereby the phylogeny of multiple resistance genes and virulence factors can be analysed. In combination with a spatiotemporal surveillance tool this would provide valuable information for rapid assessment of virulence and resistance, and prediction of the dynamics of potential outbreaks at an early stage. Next, it would allow the identification of specific epidemiological markers almost in real time. In addition, WGS data can be easily reanalysed once improved tools and databases are in place.

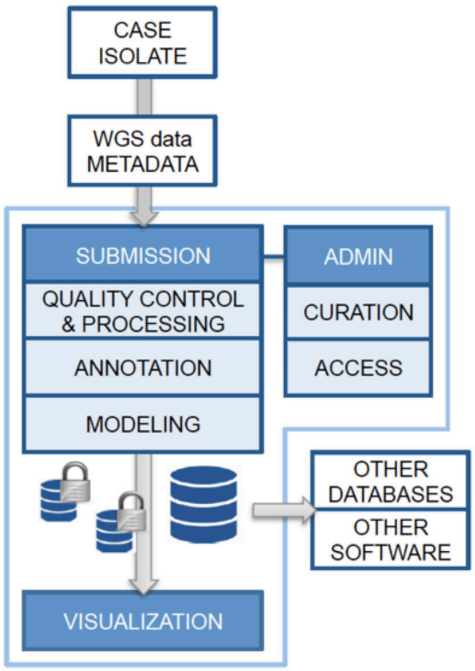
We anticipate that a shared interoperable surveillance platform for molecular epidemiology between human and veterinary medicine coupled to state-of-the-art automated phylogenetic analyses will enable more complete and detailed surveillance of MDR pathogens and generate actionable results for public health policy.

Our goal is therefore to establish, as a first step, a Switzerland-wide surveillance database combining the *NextStrain* and *PanGenome* tools. This will be started with *S. aureus* as a proof-of-concept pathogen, because a large strain collection and WGS data are already available. We then aim to gradually expand to include different pathogens such as ESBL-producing and carbapenemase-producing Enterobacteriaceae, stepwise inclusion of geographical regions beyond Switzerland and incorporating specific public health aspects such as food safety.

Such a platform will be of immense benefit to public health and research institutions in Switzerland. It will allow us to: (1) identify risks of MDR bacterial pathogens for hospitals and public health in Switzerland, (2) describe and explore these risks by predicting dynamics of spread, and (3) produce outcome control measurements of interventions, for example by monitoring the reduction of particular clones. In contrast to other existing resources, our project envisions real-time surveillance functionality, including the automated modelling of transmission events, epidemiological warning functions and linkage with high-resolution geographical data. Although a Switzerland-wide initiative, this database could clearly expand outside the Swiss border and might provide a unique tool enabling other participating countries to identify risks associated with MDR pathogens in other European (and non-European) countries.

Figure 2 shows the envisioned workflow for the surveillance platform. After isolating a strain, collecting epidemiological data and sequencing it, the user would upload raw data and metadata into a central or local/private database using the submission portal. As much as possible, this will make use of controlled vocabularies and ontologies to ensure high quality microbiological, clinical and epidemiological metadata. After submission, the data are quality controlled and then (semi)-automatically processed, with genome assembly, as well as MLST, single nucleotide polymorphisms (SNPs) and resistance calling. The resulting data are used for phylogenetic comparison and predictive modelling. All data will be stored in a structured interoperable database with controlled access, which would be queried by *NextStrain*/*PanGenome*-MRSA for the users to visualise and interact with the data in a user-friendly fashion. If open access, the structured database may also be used/queried by other apps or pushed to external databases and knowledge bases. Special attention will be paid to the choice of the operational platform technology, to the backwards compatibility of the analysis tools and to technological developments such as upcoming novel sequencing technologies and file formats. In addition to the submission portal, an administration portal will be developed, to manage users and rights, as well as a curation portal to edit and control vocabularies and to define curated reference strains and genes/variants (e.g., for resistance).

Figure 2: Envisioned workflow for the surveillance platform. After isolating a strain, collecting epidemiological data and sequencing it, the user uploads raw data and metadata into a central or local/private database using the submission portal. After submission, these data are quality controlled and then (semi-)automatically processed, to perform e.g. genome assembly, and MLST, SNP and resistance calling. The resulting data are used for phylogenetics and predictive modelling of e.g. transmission dynamics. All these data are stored in a structured database with controlled access, and visualised in a user-friendly interface. For open-data, the structured database may also be used/queried by other applications or pushed to external databases and knowledge bases. In addition to the submission portal, an administration portal will be developed, notably to manage users and rights, as well as a curating portal to edit controlled vocabularies and define curated reference strains and variants (e.g., for resistance).



Ethical considerations regarding epidemiological surveillance

The ethical framework and, especially, crises regarding the standard of care in disease outbreaks and pandemics are increasingly discussed in the literature [85]. The recent Ebola virus outbreak in West Africa [86, 87] and the global spread of Zika virus [88] have raised a series of ethical questions about outbreak preparedness and control.

From an individual's point of view, ethical and patient privacy aspects need to be considered in the collection and analysis of clinical and epidemiological patient data and molecular typing of pathogens. Switzerland participates in the Nagoya protocol and the declaration of Taipei of the World Medical Association (www.bafu.admin.ch und www.wma.net). In addition, the collection and use of data for research is regulated via the Human Research Act (www.swissethics.ch). These regulations define how patient-related data can be accessed and analysed. Detailed geo-referencing data has to be protected to reduce the chance of identifying a single individual hospital. However, public health specialists should be able to access data in the case of epidemic danger for a larger population.

From a societal point of view, the transmission and spread of resistant and virulent pathogens have to be controlled at an early stage [89]. Various stakeholder interests create a broad range of needs ranging from quality control in local hospitals to nationwide public health aspects. Although the purpose of an epidemiological surveillance database has to be focused on data sharing, the requirements of individual data protection and public health focusing on the potential impact of MDR and virulent pathogens have to be carefully analysed and discussed [90].

From the scientist's perspective, interoperable databases collecting pathogen sequencing data allow prospective and retrospective data analysis using state-of-the-art bioinformatic methods. By design, molecular surveillance platforms are multicentric and offer a broad geographic coverage. Publically funded research institutions should aim for open access publications and should be willing to share raw experimental data. Since in the past single bacterial or fungal isolates have been claimed as the "property" of the collecting institution, in some centres collaborative work with such isolates has been regulated by material transfer agreements. In order to protect the intellectual property of the participating institutions and scientists, data exchange and publication policies for work derived from a common interoperable molecular surveillance platform has to be discussed, in a contractual framework where necessary. However, a global cultural change regarding data ownership as more society focused is ongoing. In cases of a public health interest, such as an outbreak, transparent data sharing would be a crucial element to stop further spreading of a virulent pathogen.

Resulting requirements on data safety and data granularity should be reflected in the architecture of a molecular surveillance platform. For example, (1) the surveillance platform should be accessible only via secured logins; (2) only encoded data should be entered; (3) the patients informed or general consent could be added into such a database; (4) data access to sensitive information could be restricted to specific user profiles, such as clinical microbiology, public health, research, hospital epidemiology; (5) data access for research should only be provided on ethical approval of a study; (6) a governance board could be introduced to reflect different stakeholder perspectives; and (7) the geographic resolution of data should be reduced, for example from place of living to a larger geographic area such as a city district, when patient anonymity can be compromised.

Scientific significance of shared molecular surveillance data.

The developed platform will allow exchanging of relevant molecular epidemiological data, thereby enabling a One Health surveillance system for the transmission of data on MDR pathogens of different species and environments. Although mainly focusing on MDR pathogens, susceptible strains will also be included to give some background controls and to potentially investigate other aspects, such as specific gene transmission. The detailed information on relevant strains would be interesting for various stakeholders:

1. Human and veterinary microbiologists could better understand transmission routes and develop new prevention strategies at the most critical cross-points.
2. Public health authorities would have a nearly real-time surveillance tool for a detailed assessment and tracking of single strains but also large scale outbreaks.
3. Hospital epidemiologists, infection control specialists and clinicians could connect outbreaks to potential sources inside and outside of hospitals.
4. Regulatory agencies could link potential contaminated products to improve patient safety.
5. Basic researchers could access the database and use the available WGS data to develop more complex transmission models, which go beyond an automated surveillance.

The establishment of a Switzerland-wide, scalable and flexible platform for molecular epidemiology would allow monitoring of different MDR pathogens in near real-time. Importantly, combining animal, food, human and environmental isolates with geographic and phylogenic data will allow the exchange rates between different hosts to be modelled. As a first step, our interdisciplinary group will develop and test a prototype and implement a production version for a prospective evaluation period. Special attention will be devoted to ensuring that genomes can also be analysed in a global context by comparison with existing databases such as BIGSdb, potentially contributing to global monitoring of MDR pathogens. In a next step, additional MDR pathogens would be added to this database after a final improvement phase and consultation with additional diagnostic divisions at the university hospitals and in the private sector.

Such a platform covers a series of important aspects: (1) knowledge gain regarding transmission events between humans and animals, which will allow the identification of critical nodes in the network and thus help the development of improved and better targeted screening strategies and reduce the overall burden of MDR pathogens; (2) generation of the first molecular epidemiological surveillance database for Switzerland, linking the key players in the

field; (3) public health institutions will have a high-quality, near real-time surveillance tool for highly pathogenic and multidrug-resistant bacteria; (4) general public awareness about transmission events between compartments will be raised; and (5) due to the high interest of all involved institutions in such a platform, there is a very high potential to develop it into a sustainable tool for future epidemiological surveillance.

A common shared database infrastructure, which allows modelling of spatiotemporal distribution of a pathogen is a key missing element across European countries including Switzerland [49] and our initiative aims at filling this gap.

Thanks to this article, we hope to encourage a number of federal and public health institutions and researchers to join our initiative in order to make a unique, large and useful platform for MDR bacteria tracking and to study the evolution and the spread of these major public health threats.

Financial disclosure

GG, DB, JS, IX, RN, VP, and AE have received a research grant of the Swiss National Science Foundation in the NRP72 program (SNF Grant Nr. 407240_177504 / 1) entitled “Development of a Swiss surveillance database for molecular epidemiology of multi-drug resistant pathogens”.

Potential competing interests

No conflict of interest has been declared.

References:

1 O’Neill J, et al. Tackling drug-resistant infections globally: Final report and recommendations. London: Wellcome Trust, HM government; 2016.

2 Grundmann H, Glasner C, Albiger B, Aanensen DM, Tomlinson CT, Andrasević AT, et al.; European Survey of Carbapenemase-Producing Enterobacteriaceae (EuSCAPE) Working Group. Occurrence of carbapenemase-producing Klebsiella pneumoniae and Escherichia coli in the European survey of carbapenemase-producing Enterobacteriaceae (EuSCAPE): a prospective, multinational study. Lancet Infect Dis. 2017;17(2):153–63. doi: http://dx.doi.org/10.1016/S1473-3099(16)30257-2. PubMed.

3 Albiger B, Glasner C, Struelens MJ, Grundmann H, Monnet DL; European Survey of Carbapenemase-Producing Enterobacteriaceae (EuSCAPE) working group. Carbapenemase-producing Enterobacteriaceae in Europe: assessment by national experts from 38 countries, May 2015. Euro Surveill. 2015;20(45):30062. doi: http://dx.doi.org/10.2807/1560-7917.ES.2015.20.45.30062. PubMed.

4 Aramburu C, Harbarth S, Liassine N, Girard M, Gervaix A, Scherenzel J, et al. Community-acquired methicillin-resistant Staphylococcus aureus in Switzerland: first surveillance report. Euro Surveill. 2006;11(1):42–3. doi: http://dx.doi.org/10.2807/esm.11.01.00594-en. PubMed.

5 Boyce JM. Methicillin-resistant Staphylococcus aureus. Detection, epidemiology, and control measures. Infect Dis Clin North Am. 1989;3(4):901–13. PubMed.

6 Moran GJ, Krishnadasan A, Gorwitz RJ, Fosheim GE, McDougal LK, Carey RB, et al.; EMERGEncy ID Net Study Group. Methicillin-resistant S. aureus infections among patients in the emergency department. N Engl J Med. 2006;355(7):666–74. doi: http://dx.doi.org/10.1056/NEJMoa055356. PubMed.

7 Gijón M, Bellusci M, Petraitiene B, Noguera-Julian A, Zilinskaite V, Sanchez Moreno P, et al. Factors associated with severity in invasive community-acquired Staphylococcus aureus infections in children: a prospective European multicentre study. Clin Microbiol Infect. 2016;22(7):643.e1–6. doi: http://dx.doi.org/10.1016/j.cmi.2016.04.004. PubMed.

8 Maeda M, Shoji H, Shirakura T, Takuma T, Ugajin K, Fukuchi K, et al. Analysis of Staphylococcal Toxins and Clinical Outcomes of Methicillin-Resistant Staphylococcus aureus Bacteremia. Biol Pharm Bull. 2016;39(7):1195–200. doi: http://dx.doi.org/10.1248/bpb.b16-00255. PubMed.

9 Rossi F, Diaz L, Wollam A, Panesso D, Zhou Y, Rincon S, et al. Transferable vancomycin resistance in a community-associated MRSA lineage. N Engl J Med. 2014;370(16):1524–31. doi: http://dx.doi.org/10.1056/NEJMoa1303359. PubMed.

10 Britt NS, Patel N, Shireman TI, El Atrouni WI, Horvat RT, Steed ME. Relationship between vancomycin tolerance and clinical outcomes in Staphylococcus aureus bacteraemia. J Antimicrob Chemother. 2017;72(2):535–42. doi: http://dx.doi.org/10.1093/jac/dkw453. PubMed.

11 World Health Organization. Global action plan on antimicrobial resistance. Geneva: WHO Library Cataloguing-in-Publication Data; 2015.

12 Baier C, Ipaktchi R, Ebadi E, Limbourg A, Mett TR, Vogt PM, et al. A multimodal infection control concept in a burn intensive care unit - lessons learnt from a meticillin-resistant Staphylococcus aureus outbreak. J Hosp Infect. 2018;98(2):127–33. doi: http://dx.doi.org/10.1016/j.jhin.2017.07.023. PubMed.

13 Olearo F, Albrich WC, Vernaz N, Harbarth S, Kronenberg A; Swiss Centre For Antibiotic Resistance Anresis. Staphylococcus aureus and methicillin resistance in Switzerland: regional differences and trends from 2004 to 2014. Swiss Med Wkly. 2016;146:. doi: http://dx.doi.org/10.4414/smw.2016.14339. PubMed.

14 Bielicki JA, Cromwell DA, Johnson A, Planche T, Sharland M; ARPEC project. Surveillance of Gram-negative bacteria: impact of variation in current European laboratory reporting practice on apparent multidrug resistance prevalence in paediatric bloodstream isolates. Eur J Clin Microbiol Infect Dis. 2017;36(5):839–46. doi: http://dx.doi.org/10.1007/s10096-016-2869-4. PubMed.

15 Freeman R, Ironmonger D, Puleston R, Hopkins KL, Welfare W, Hope R, et al. Enhanced surveillance of carbapenemase-producing Gram-negative bacteria to support national and international prevention and control efforts. Clin Microbiol Infect. 2016;22(10):896–7. doi: http://dx.doi.org/10.1016/j.cmi.2016.07.020. PubMed.

16 Egli A, Schmid H, Kuenzli E, Widmer AF, Battegay M, Plagge H, et al. Association of daptomycin use with resistance development in Enterococcus faecium bacteraemia-a 7-year individual and population-based analysis. Clin Microbiol Infect. 2017;23(2):118.e1–7. doi: http://dx.doi.org/10.1016/j.cmi.2016.10.003. PubMed.

17 Ruppé E, Olearo F, Pires D, Baud D, Renzi G, Cherkaoui A, et al. Clonal or not clonal? Investigating hospital outbreaks of KPC-producing Klebsiella pneumoniae with whole-genome sequencing. Clin Microbiol Infect. 2017;23(7):470–5. doi: http://dx.doi.org/10.1016/j.cmi.2017.01.015. PubMed.

18 Kraemer JG, Pires J, Kueffer M, Semaani E, Endimiani A, Hilty M, et al. Prevalence of extended-spectrum β-lactamase-producing Enterobacteriaceae and Methicillin-Resistant Staphylococcus aureus in pig farms in Switzerland. Sci Total Environ. 2017;603-604:401–5. doi: http://dx.doi.org/10.1016/j.scitotenv.2017.06.110. PubMed.

19 Seiffert SN, Hilty M, Perreten V, Endimiani A. Extended-spectrum cephalosporin-resistant Gram-negative organisms in livestock: an emerging problem for human health? Drug Resist Updat. 2013;16(1-2):22–45. doi: http://dx.doi.org/10.1016/j.drup.2012.12.001. PubMed.

20 Hindermann D, Gopinath G, Chase H, Negrete F, Althaus D, Zurfluh K, et al. *Salmonella enterica* serovar Infantis from Food and Human Infections, Switzerland, 2010-2015: Poultry-Related Multidrug Resistant Clones and an Emerging ESBL Producing Clonal Lineage. Front Microbiol. 2017;8:1322. doi: http://dx.doi.org/10.3389/fmicb.2017.01322. PubMed.

21 Lahti E, Löfdahl M, Ågren J, Hansson I, Olsson Engvall E. Confirmation of a Campylobacteriosis Outbreak Associated with Chicken Liver Pâté Using PFGE and WGS. Zoonoses Public Health. 2017;64(1):14–20. doi: http://dx.doi.org/10.1111/zph.12272. PubMed.

22 Davis GS, Waits K, Nordstrom L, Weaver B, Aziz M, Gauld L, et al. Intermingled Klebsiella pneumoniae Populations Between Retail Meats and Human Urinary Tract Infections. Clin Infect Dis. 2015;61(6):892–9. doi: http://dx.doi.org/10.1093/cid/civ428. PubMed.

23 Paterson GK, Harrison EM, Craven EF, Petersen A, Larsen AR, Ellington MJ, et al. Incidence and characterisation of methicillin-resistant Staphylococcus aureus (MRSA) from nasal colonisation in participants attending a cattle veterinary conference in the UK. PLoS One. 2013;8(7):. doi: http://dx.doi.org/10.1371/journal.pone.0068463. PubMed.

24 Doulgeraki AI, Di Ciccio P, Ianieri A, Nychas GE. Methicillin-resistant food-related Staphylococcus aureus: a review of current knowledge and biofilm formation for future studies and applications. Res Microbiol. 2017;168(1):1–15. doi: http://dx.doi.org/10.1016/j.resmic.2016.08.001. PubMed.

25 Laaksonen S, Oksanen A, Julmi J, Zweifel C, Fredriksson-Ahomaa M, Stephan R. Presence of foodborne pathogens, extended-spectrum β-lactamase -producing Enterobacteriaceae, and methicillin-resistant Staphylococcus aureus in slaughtered reindeer in northern Finland and Norway.

Acta Vet Scand. 2017;59(1):2. doi: http://dx.doi.org/10.1186/s13028-016-0272-x. PubMed.

26 Osman K, Badr J, Al-Maary KS, Moussa IM, Hessain AM, Girah ZM, et al. Prevalence of the Antibiotic Resistance Genes in Coagulase-Positive-and Negative-*Staphylococcus* in Chicken Meat Retailed to Consumers. Front Microbiol. 2016;7:1846. doi: http://dx.doi.org/10.3389/fmicb.2016.01846. PubMed.

27 Revez J, Espinosa L, Albiger B, Leitmeyer KC, Struelens MJ; ECDC National Microbiology Focal Points and Experts Group. 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. doi: http://dx.doi.org/10.3389/fpubh.2017.00347. PubMed.

28 Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl J, Laurent F, et al.; ESCMID Study Group of Epidemiological Markers (ESGEM). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Euro Surveill. 2013;18(4):20380. doi: http://dx.doi.org/10.2807/ese.18.04.20380-en. PubMed.

29 Crowe SJ, Bottichio L, Shade LN, Whitney BM, Corral N, Melius B, et al. Shiga Toxin-Producing E. coli Infections Associated with Flour. N Engl J Med. 2017;377(21):2036–43. doi: http://dx.doi.org/10.1056/NEJMoa1615910. PubMed.

30 Fischer J, Hille K, Ruddat I, Mellmann A, Köck R, Kreienbrock L. Simultaneous occurrence of MRSA and ESBL-producing Enterobacteriaceae on pig farms and in nasal and stool samples from farmers. Vet Microbiol. 2017;200:107–13. doi: http://dx.doi.org/10.1016/j.vetmic.2016.05.021. PubMed.

31 Ford L, Carter GP, Wang Q, Seemann T, Sintchenko V, Glass K, et al. Incorporating Whole-Genome Sequencing into Public Health Surveillance: Lessons from Prospective Sequencing of Salmonella Typhimurium in Australia. Foodborne Pathog Dis. 2018;15(3):161–7. doi: http://dx.doi.org/10.1089/fpd.2017.2352. PubMed.

32 Timms VJ, Rockett R, Bachmann NL, Martinez E, Wang Q, Chen SC, et al. Genome Sequencing Links Persistent Outbreak of Legionellosis in Sydney (New South Wales, Australia) to an Emerging Clone of Legionella pneumophila Sequence Type 211. Appl Environ Microbiol. 2018;84(5):. doi: http://dx.doi.org/10.1128/AEM.02020-17. PubMed.

33 Waldram A, Dolan G, Ashton PM, Jenkins C, Dallman TJ. Epidemiological analysis of Salmonella clusters identified by whole genome sequencing, England and Wales 2014. Food Microbiol. 2018;71:39–45. doi: http://dx.doi.org/10.1016/j.fm.2017.02.012. PubMed.

34 Oberle M, Wohlwend N, Jonas D, Maurer FP, Jost G, Tschudin-Sutter S, et al. The Technical and Biological Reproducibility of Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) Based Typing: Employment of Bioinformatics in a Multicenter Study. PLoS One. 2016;11(10):. doi: http://dx.doi.org/10.1371/journal.pone.0164260. PubMed.

35 Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clin Microbiol Infect. 2018;24(4):335–41. doi: http://dx.doi.org/10.1016/j.cmi.2017.10.013. PubMed.

36 van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al.; European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group on Epidemiological Markers (ESGEM). Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 2007;13(Suppl 3):1–46. doi: http://dx.doi.org/10.1111/j.1469-0691.2007.01786.x. PubMed.

37 Köser CU, Ellington MJ, Cartwright EJ, Gillespie SH, Brown NM, Farrington M, et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog. 2012;8(8):. doi: http://dx.doi.org/10.1371/journal.ppat.1002824. PubMed.

38 Reuter S, Ellington MJ, Cartwright EJ, Köser CU, Török ME, Gouliouris T, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. JAMA Intern Med. 2013;173(15):1397–404. doi: http://dx.doi.org/10.1001/jamainternmed.2013.7734. PubMed.

39 Struelens MJ, Brisse S. From molecular to genomic epidemiology: transforming surveillance and control of infectious diseases. Euro Surveill. 2013;18(4):20386. doi: http://dx.doi.org/10.2807/ese.18.04.20386-en. PubMed.

40 Chattaway MA, Schaefer U, Tewolde R, Dallman TJ, Jenkins C. Identification of Escherichia coli and Shigella Species from Whole-Genome Sequences. J Clin Microbiol. 2017;55(2):616–23. doi: http://dx.doi.org/10.1128/JCM.01790-16. PubMed.

41 Meinel DM, Kuehl R, Zbinden R, Boskova V, Garzoni C, Fadini D, et al. Outbreak investigation for toxigenic Corynebacterium diphtheriae wound infections in refugees from Northeast Africa and Syria in Switzerland and Germany by whole genome sequencing. Clin Microbiol Infect. 2016;22(12):1003.e1–8. doi: http://dx.doi.org/10.1016/j.cmi.2016.08.010. PubMed.

42 Piso RJ, Käch R, Pop R, Zillig D, Schibli U, Bassetti S, et al. A Cross-Sectional Study of Colonization Rates with Methicillin-Resistant Staphylococcus aureus (MRSA) and Extended-Spectrum Beta-Lactamase (ESBL) and Carbapenemase-Producing Enterobacteriaceae in Four Swiss Refugee Centres. PLoS One. 2017;12(1):. doi: http://dx.doi.org/10.1371/journal.pone.0170251. PubMed.

43 Sommerstein R, Führer U, Lo Priore E, Casanova C, Meinel DM, Seth-Smith HM, et al.; On Behalf Of Anresis; On Behalf Of Swissnoso. *Burkholderia stabilis* outbreak associated with contaminated commercially-available washing gloves, Switzerland, May 2015 to August 2016. Euro Surveill. 2017;22(49). doi: http://dx.doi.org/10.2807/1560-7917.ES.2017.22.49.17-00213. PubMed.

44 Gargis AS, Kalman L, Bick DP, da Silva C, Dimmock DP, Funke BH, et al. Good laboratory practice for clinical next-generation sequencing informatics pipelines. Nat Biotechnol. 2015;33(7):689–93. doi: http://dx.doi.org/10.1038/nbt.3237. PubMed.

45 Gargis AS, Kalman L, Lubin IM. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. J Clin Microbiol. 2016;54(12):2857–65. doi: http://dx.doi.org/10.1128/JCM.00949-16. PubMed.

46 McNerney R, Clark TG, Campino S, Rodrigues C, Dolinger D, Smith L, et al. Removing the bottleneck in whole genome sequencing of Mycobacterium tuberculosis for rapid drug resistance analysis: a call to action. Int J Infect Dis. 2017;56:130–5. doi: http://dx.doi.org/10.1016/j.ijid.2016.11.422. PubMed.

47 Sjödin A, Broman T, Melefors Ö, Andersson G, Rasmusson B, Knutsson R, et al. The need for high-quality whole-genome sequence databases in microbial forensics. Biosecur Bioterror. 2013;11(S1, Suppl 1):S78–86. doi: http://dx.doi.org/10.1089/bsp.2013.0007. PubMed.

48 Mellmann A, Bletz S, Böking T, Kipp F, Becker K, Schultes A, et al. Real-Time Genome Sequencing of Resistant Bacteria Provides Precision Infection Control in an Institutional Setting. J Clin Microbiol. 2016;54(12):2874–81. doi: http://dx.doi.org/10.1128/JCM.00790-16. PubMed.

49 Struelens M, et al. (ed ECDC) 1-14 (ECDC, Stockholm, 2016).

50 Agodi A, Auxilia F, Barchitta M, Brusaferro S, D'Errico MM, Montagna MT, et al.; SPIN-UTI network of the GISIOWorking Group of the Italian Society of Hygiene, Preventive Medicine and Public Health (SItI). Antibiotic consumption and resistance: results of the SPIN-UTI project of the GISIO-SItI. Epidemiol Prev. 2015;39(4, Suppl 1):94–8. PubMed.

51 Conceição T, Diamantino F, Coelho C, de Lencastre H, Aires-de-Sousa M. Contamination of public buses with MRSA in Lisbon, Portugal: a possible transmission route of major MRSA clones within the community. PLoS One. 2013;8(11):. doi: http://dx.doi.org/10.1371/journal.pone.0077812. PubMed.

52 Mortensen R, Christensen D, Hansen LB, Christensen JP, Andersen P, Dietrich J. Local Th17/IgA immunity correlate with protection against intranasal infection with Streptococcus pyogenes. PLoS One. 2017;12(4):. doi: http://dx.doi.org/10.1371/journal.pone.0175707. PubMed.

53 Turner RD, Chiu C, Churchyard GJ, Esmail H, Lewinsohn DM, Gandhi NR, et al. Tuberculosis Infectiousness and Host Susceptibility. J Infect Dis. 2017;216(suppl_6):S636–43. doi: http://dx.doi.org/10.1093/infdis/jix361. PubMed.

54 Islam MA, Islam M, Hasan R, Hossain MI, Nabi A, Rahman M, et al. Environmental Spread of New Delhi Metallo-β-Lactamase-1-Producing Multidrug-Resistant Bacteria in Dhaka, Bangladesh. Appl Environ Microbiol. 2017;83(15):. doi: http://dx.doi.org/10.1128/AEM.00793-17. PubMed.

55 Surette MD, Wright GD. Lessons from the Environmental Antibiotic Resistome. Annu Rev Microbiol. 2017;71(1):309–29. doi: http://dx.doi.org/10.1146/annurev-micro-090816-093420. PubMed.

56 Furness LE, Campbell A, Zhang L, Gaze WH, McDonald RA. Wild small mammals as sentinels for the environmental transmission of antimicrobial resistance. Environ Res. 2017;154:28–34. doi: http://dx.doi.org/10.1016/j.envres.2016.12.014. PubMed.

57 Parker D, Sniatynski MK, Mandrusiak D, Rubin JE. Extended-spectrum β-lactamase producing Escherichia coli isolated from wild birds in Saskatoon, Canada. Lett Appl Microbiol. 2016;63(1):11–5. doi: http://dx.doi.org/10.1111/lam.12589. PubMed.

58 Marques C, Belas A, Franco A, Aboim C, Gama LT, Pomba C. Increase in antimicrobial resistance and emergence of major international high-risk clonal lineages in dogs and cats with urinary tract infection: 16 year retrospective study. J Antimicrob Chemother. 2018;73(2):377–84. doi: http://dx.doi.org/10.1093/jac/dkx401. PubMed.

59 Afset JE, Larssen KW, Bergh K, Larkeryd A, Sjodin A, Johansson A, et al. Phylogeographical pattern of Francisella tularensis in a nationwide outbreak of tularaemia in Norway, 2011. Euro Surveill. 2015;20(19):9–14. doi: http://dx.doi.org/10.2807/1560-7917.ES2015.20.19.21125. PubMed.

60 Stucki D, Brites D, Jeljeli L, Coscolla M, Liu Q, Trauner A, et al. Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. Nat Genet. 2016;48(12):1535–43. doi: http://dx.doi.org/10.1038/ng.3704. PubMed.

61 Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events. Nat Genet. 2015;47(6):632–9. doi: http://dx.doi.org/10.1038/ng.3281. PubMed.

62 Glaser P, Martins-Simões P, Villain A, Barbier M, Tristan A, Bouchier C, et al. Demography and Intercontinental Spread of the USA300 Community-Acquired Methicillin-Resistant Staphylococcus aureus Lineage. MBio. 2016;7(1):. doi: http://dx.doi.org/10.1128/mBio.02183-15. PubMed.

63 Faria NR, Quick J, Claro IM, Thézé J, de Jesus JG, Giovanetti M, et al. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. Nature. 2017;546(7658):406–10. doi: http://dx.doi.org/10.1038/nature22401. PubMed.

64 Grubaugh ND, Ladner JT, Kraemer MUG, Dudas G, Tan AL, Gangavarapu K, et al. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. Nature. 2017;546(7658):401–5. doi: http://dx.doi.org/10.1038/nature22400. PubMed.

65 Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. Proc Natl Acad Sci USA. 2016;113(12):E1701–9. doi: http://dx.doi.org/10.1073/pnas.1525578113. PubMed.

66 Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, et al. Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. Nature. 2015;524(7563):97–101. doi: http://dx.doi.org/10.1038/nature14594. PubMed.

67 Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. Nature. 2017;544(7650):309–15. doi: http://dx.doi.org/10.1038/nature22040. PubMed.

68 Hayman DT, Fooks AR, Marston DA, Garcia-R JC. The Global Phylogeography of Lyssaviruses - Challenging the 'Out of Africa' Hypothesis. PLoS Negl Trop Dis. 2016;10(12):. doi: http://dx.doi.org/10.1371/journal.pntd.0005266. PubMed.

69 Argimón S, Abudahab K, Goater RJ, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genom. 2016;2(11):. doi: http://dx.doi.org/10.1099/mgen.0.000093. PubMed.

70 Aanensen DM, Feil EJ, Holden MT, Dordel J, Yeats CA, Fedosejev A, et al.; European SRL Working Group. Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive Staphylococcus aureus in Europe. MBio. 2016;7(3):. doi: http://dx.doi.org/10.1128/mBio.00444-16. PubMed.

71 Hadfield J, et al. Nextstrain: real-time tracking of pathogen evolution. bioRxiv. 2017. doi: http://dx.doi.org/10.1101/224048.

72 Neher RA, Bedford T. nextflu: real-time tracking of seasonal influenza virus evolution in humans. Bioinformatics. 2015;31(21):3546–8. doi: http://dx.doi.org/10.1093/bioinformatics/btv381. PubMed.

73 Griffiths E, Dooley D, Graham M, Van Domselaar G, Brinkman FSL, Hsiao WWL. Context Is Everything: Harmonization of Critical Food Microbiology Descriptors and Metadata for Improved Food Safety and Surveillance. Front Microbiol. 2017;8:1068. doi: http://dx.doi.org/10.3389/fmicb.2017.01068. PubMed.

74 Brhelova E, Antonova M, Pardy F, Kocmanova I, Mayer J, Racil Z, et al. Investigation of next-generation sequencing data of Klebsiella pneumoniae using web-based tools. J Med Microbiol. 2017;66(11):1673–83. doi: http://dx.doi.org/10.1099/jmm.0.000624. PubMed.

75 Yin X, Jiang XT, Chai B, Li L, Yang Y, Cole JR, et al. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics. 2018;34(13):2263–70. doi: http://dx.doi.org/10.1093/bioinformatics/bty053. PubMed.

76 Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. Nucleic Acids Res. 2016;44(D1):D694–7. doi: http://dx.doi.org/10.1093/nar/gkv1239. PubMed.

77 Gámez G, Castro A, Gómez-Mejia A, Gallego M, Bedoya A, Camargo M, et al. The variome of pneumococcal virulence factors and regulators. BMC Genomics. 2018;19(1):10. doi: http://dx.doi.org/10.1186/s12864-017-4376-0. PubMed.

78 Rasmussen LH, Højholt K, Dargis R, Christensen JJ, Skovgaard O, Justesen US, et al. In silico assessment of virulence factors in strains of Streptococcus oralis and Streptococcus mitis isolated from patients with Infective Endocarditis. J Med Microbiol. 2017. doi: http://dx.doi.org/10.1099/jmm.0.000573. PubMed.

79 Kleinheinz KA, Joensen KG, Larsen MV. Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. Bacteriophage. 2014;4(2):. doi: http://dx.doi.org/10.4161/bact.27943. PubMed.

80 Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640–4. doi: http://dx.doi.org/10.1093/jac/dks261. PubMed.

81 Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J Antimicrob Chemother. 2017;72(10):2764–8. doi: http://dx.doi.org/10.1093/jac/dkx217. PubMed.

82 Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45(D1):D566–73. doi: http://dx.doi.org/10.1093/nar/gkw1004. PubMed.

83 McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother. 2013;57(7):3348–57. doi: http://dx.doi.org/10.1128/AAC.00419-13. PubMed.

84 Ding W, Baumdicker F, Neher RA. panX: pan-genome analysis and exploration. Nucleic Acids Res. 2018;46(1):. doi: http://dx.doi.org/10.1093/nar/gkx977. PubMed.

85 Leider JP, DeBruin D, Reynolds N, Koch A, Seaberg J. Ethical Guidance for Disaster Response, Specifically Around Crisis Standards of Care: A Systematic Review. Am J Public Health. 2017;107(9):e1–9. doi: http://dx.doi.org/10.2105/AJPH.2017.303882. PubMed.

86 Smith MJ, Silva DS. Ethics for pandemics beyond influenza: Ebola, drug-resistant tuberculosis, and anticipating future ethical challenges in pandemic preparedness and response. Monash Bioeth Rev. 2015;33(2-3):130–47. doi: http://dx.doi.org/10.1007/s40592-015-0038-7. PubMed.

87 Alirol E, Kuesel AC, Guraiib MM, de la Fuente-Núñez V, Saxena A, Gomes MF. Ethics review of studies during public health emergencies - the experience of the WHO ethics review committee during the Ebola virus disease epidemic. BMC Med Ethics. 2017;18(1):43. doi: http://dx.doi.org/10.1186/s12910-017-0201-1. PubMed.

88 Jamrozik E, Selgelid MJ. Ethics, health policy, and Zika: From emergency to global epidemic? J Med Ethics. 2018;44(5):343–8. doi: http://dx.doi.org/10.1136/medethics-2017-104389. PubMed.

89 Derpmann S. Ethical reasoning in pandemic preparedness plans: Southeast Asia and the Western Pacific. Bioethics. 2011;25(8):445–50. doi: http://dx.doi.org/10.1111/j.1467-8519.2011.01922.x. PubMed.

90 Rump B, Cornelis C, Woonink F, VAN Steenbergen J, Verweij M, Hulscher M. Using typing techniques in a specific outbreak: the ethical reflection of public health professionals. Epidemiol Infect. 2017;145(7):1431–6. doi: http://dx.doi.org/10.1017/S0950268817000127. PubMed.

Appendix 1

Minimum functional features for a future surveillance tool

Table S1: Minimum functional features for a future surveillance tool

Submission portal
Submission of genomic data
Upload raw reads
Upload genome assemblies
Submission of metadata
Data entry using an online form
Data entry via uploading a structured text file
Use of controlled vocabularies, and ontologies where they exist
Possibility to enter missing, nonmandatory, metadata at a later stage
Quality control, processing, annotation and modelling of data
Quality control raw reads and assemblies to flag poor quality data
Typing (spa, cgMLST, SNPs)
Modelling of transmission dynamics
Annotation of resistance mutations and genes
Annotation of virulence factors
Visualisation portal
Colour-code strains according to different criteria
Phylogenetic tree
cgMLST-based tree
SNP-based tree
Branches lengths represent either time or divergence
Confidence intervals on branch lengths
Zoom in/out of the tree
Map of cases
Display cases and transmission links
Zoom in/out of the map
Resistance, virulence
Colour-code according to pre-defined resistant genotypes
Display presence/absence of resistance genes
Display presence/absence of virulence factors
Filtering
Filter-out displayed strains based on a combination of criteria
Search function
Interpretations
Annotate related strains (clusters)
Administration portal
Data curation
Maintain curated list of reference genomes
Integrate validated genotype-phenotype resistance associations
Define and update controlled vocabularies
Access
Several levels of access rights
Audit log
Secure login
Versioning