data

MDPI

Data Descriptor

Genomic Epidemiology Dataset for the Important Nosocomial Pathogenic Bacterium *Acinetobacter baumannii*

Andrey Shelenkov *, Yulia Mikhaylova and Vasiliy Akimkin

Central Research Institute of Epidemiology, Novogireevskaya str., 3a, 111123 Moscow, Russia
* Correspondence: fallandar@gmail.com

Abstract: The infections caused by various bacterial pathogens both in clinical and community settings represent a significant threat to public healthcare worldwide. The growing resistance to antimicrobial drugs acquired by bacterial species causing healthcare-associated infections has already become a life-threatening danger noticed by the World Health Organization. Several groups or lineages of bacterial isolates, usually called 'the clones of high risk', often drive the spread of resistance within particular species. Thus, it is vitally important to reveal and track the spread of such clones and the mechanisms by which they acquire antibiotic resistance and enhance their survival skills. Currently, the analysis of whole-genome sequences for bacterial isolates of interest is increasingly used for these purposes, including epidemiological surveillance and the development of spread prevention measures. However, the availability and uniformity of the data derived from genomic sequences often represent a bottleneck for such investigations. With this dataset, we present the results of a genomic epidemiology analysis of 17,546 genomes of a dangerous bacterial pathogen, *Acinetobacter baumannii*. Important typing information, including multilocus sequence typing (MLST)-based sequence types (STs), intrinsic $bla_{OXA\text{-}51\text{-}like}$ gene variants, capsular (KL) and oligosaccharide (OCL) types, CRISPR-Cas systems, and cgMLST profiles are presented, as well as the assignment of particular isolates to nine known international clones of high risk. The presence of antimicrobial resistance genes within the genomes is also reported. These data will be useful for researchers in the field of *A. baumannii* genomic epidemiology, resistance analysis, and prevention measure development.

Dataset: The dataset is available at https://doi.org/10.5281/zenodo.10143377.

Keywords: *Acinetobacter baumannii*; healthcare-associated infections; genomic epidemiology; whole-genome sequencing; cgMLST; international clones of high risk; antibiotic resistance

Citation: Shelenkov, A.; Mikhaylova, Y.; Akimkin, V. Genomic Epidemiology Dataset for the Important Nosocomial Pathogenic Bacterium *Acinetobacter baumannii*. *Data* **2024**, *9*, 22. https://doi.org/10.3390/data9020022

Academic Editor: Juraj Gregáň

Received: 25 November 2023
Revised: 9 January 2024
Accepted: 12 January 2024
Published: 26 January 2024

1. Summary

Bacterial infections remain among the most serious problems in healthcare. The continuous spread of antimicrobial drug resistance in clinical pathogenic bacteria represents a serious threat to public health worldwide, leading to a limited set, if any, of available treatment options [1]. Making matters worse, antimicrobial resistance (AMR) has already outstepped hospitals and other healthcare institutions and become a significant matter in community settings [2], and AMR infections have become one of the top causes of death worldwide [3].

The spread of AMR within particular bacterial species can be driven by several lineages, usually called 'global clones' or 'international clones of high risk', which has been shown to be the case for some of the most successful and widespread pathogens, like *Klebsiella pneumoniae* [4] and *Acinetobacter baumannii* [5]. Thus, in order to perform epidemiological surveillance and develop effective prevention measures against multidrug-resistant

(MDR) bacteria, it is essential to track the spread of such global clones and to check whether particular isolates belong to these lineages. Simply stated, you should know your enemy before you fight it.

Isolate classification and assignment to a particular clone can be based on several characteristics revealed using molecular biology techniques, but currently, whole-genome sequencing is increasingly used for this and many other purposes due to the unprecedented amount of information it produces and its high cost-effectiveness [6,7].

Now, tens of thousands genomes of pathogenic bacteria are available in public databases, and this number will continue to increase. The representation in genomic databases is interrelated with the level of concern raised by a particular pathogen and the incidence of infections caused by it. *A. baumannii* is responsible for a significant share of nosocomial infections worldwide [8], and the World Health Organization (WHO) listed its isolates resistant to antibiotics from the carbapenem class as some of the most critical pathogens with the highest priority in new antibiotic development [9].

Currently, more than 17,000 draft *A. baumannii* genomes are available from NCBI (https://www.ncbi.nlm.nih.gov/datasets/taxonomy/470/, accessed on 23 October 2023), but the data provided there usually lack isolate typing information, which must be derived using various computational tools by users. Another commonly used database, PubMLST [10], contains epidemiologically related and typing information for more than 8700 isolates, but the number of available genomes is about 3300 (https://pubmlst.org/organisms/acinetobacter-baumannii, accessed on 23 October 2023). At the same time, as we described earlier, the assignment of a particular isolate to some international clone is not always straightforward, even when whole-genome data are available [11], and, to the best of our knowledge, no public database currently presents such assignment for large isolate datasets.

Here, we present a dataset containing typing information, including assignment to international clones, for the whole set of *A. baumannii* isolates available from NCBI (accessed on 23 February 2023), which includes 17,546 genomes. The information available includes multilocus sequence typing (MLST)-based types (STs), intrinsic $bla_{OXA\text{-}51\text{-}like}$ gene variants, capsular (KL) and oligosaccharide (OCL) types, core genome MLST (cgMLST) profiles, and data regarding the presence of CRISPR-Cas systems in each of the isolates. Data regarding the presence of AMR genes providing resistance to various classes of antibiotics are also available.

We already used this dataset to study the representation of known international clones in Genbank [11]. Deriving all this information from the available genomes requires an advanced level of bioinformatics expertise and significant computational resources. As was noticed recently, proper global genomic epidemiology investigations of *A. baumannii* are very important for understanding the global dissemination of important clones [12]. We believe that this dataset will be useful for epidemiological studies of *A. baumannii*, including for the facilitation of selection of the proper reference isolate sets for any type of genome-based investigation.

These precomputed data will be especially helpful for the researchers starting their investigations in the emerging field of genomic epidemiology concerning this important and dangerous pathogen.

 Updates are scheduled to be provided at least twice a year.

2. Data Description

2.1. Data Structure

The dataset contains three tables provided in various formats: xlsx, txt (tab-delimited), and pdf (summary table only). When using the xlsx format in table processing software, users can create their own filters to select the subsets of interest, change or add sorting parameters, build graphs, etc. The text format (txt) is intended for computational processing

using bioinformatics tools, while the pdf format is provided for summary file containing main isolate typing results, enabling it to be presented in human-readable form.

The tables, which will be described below in more details, include:

- A summary table, which contains typing information for all isolates, such as MLST Pasteur ST, the OXA-51-like variant, KL- and OCL-based types, assignments of an isolate to known international clones of high risk (IC1-IC9), and the possible presence and type of CRISPR-Cas systems in the genome of an isolate.
- An AMR gene table containing information on the presence of genes known to confer, when properly expressed, AMR to various classes of antimicrobial drugs for particular isolates.
- A table containing cgMLST profiles for all isolates, which can be used for extended comparison and clustering purposes.

2.1.1. Typing Summary Table

The format and exemplary data for the typing summary table are given in Table 1. Generally speaking, this table contains the characteristics of the isolates, by which they can be typed, grouped, or selected for particular investigation goals. For example, studying isolates belonging to particular ICs or STs represents a frequent epidemiological task.

Table 1. Exemplary data and column information for typing summary table.

Assembly Code	IC	MLST Pasteur ST	OXA-51-Like Variant	KL Type	OCL Type	CRISPR-Cas
GCA_000184515.2	IC7	25	64	KL14	OCL5	I-F
GCA_000222265.2	NOIC	415	115	KL6	OCL1	NF [2]
GCA_000222225.2	NOIC	ND [1]	115	KL22	OCL3	NF
GCA_000297515.1	IC2	414	ND [1]	KL22	OCL1	NF

[1] 'ND' indicates ST (column 2) or OXA-51 variant (column 3) could not be determined; [2] 'NF' indicates the CRISPR-Cas system was not found in the genome. IC stands for 'international clone of high risk', and 'I-F' represents a particular type of CRISPR-Cas system.

The first column contains the assembly code assigned by Genbank, which uniquely identifies a particular *A. baumannii* genome assembly.

'IC' stands for 'international clone of high risk' and shows the assignment of a particular isolate to know international clones. The procedure for such an assignment is described in the Methods section and our previous publication [11]. If an isolate is not assigned to any IC, this column has the 'NOIC' designation.

The third column contains an ST defined as a combination of seven loci (*cpn60*, *fusA*, *gltA*, *pyrG*, *recA*, *rplB*, and *rpoB* genes) from a typing scheme known as 'Pasteur MLST' [13]. Each variant of a particular locus is numbered, and the combination of seven locus variants constitutes an ST, which has its own number. For example, the combination *cpn60*_3, *fusA*_29, *gltA*_30, *pyrG*_1, *recA*_9, *rplB*_1, and *rpoB*_4 was defined as ST17. Locus variants and the definitions of corresponding STs can be found in the PubMLST database (https://pubmlst.org/organisms/acinetobacter-baumannii, accessed on 23 October 2023). 'ND' in this column indicates that an ST was not determined due to either low sequencing quality or the presence of a novel MLST allele not uploaded to the databases yet.

The fourth column shows the variant of a gene encoding intrinsic OXA-51-like β-lactamase, which is possessed by all *A. baumannii* isolates and is sometimes used for typing purposes [14]. 'NOT_FOUND' can appear in this column due to low sequencing quality or low similarity with known OXA-51-like variants.

KL and OCL type show the typing classes based on the corresponding sets of genes, respectively. Capsular polysaccharide is an essential factor determining bacterial virulence and susceptibility to phages, which makes it a useful epidemiological marker [15]. The capsular polysaccharide gene cluster consists of about 30 genes, while the OC locus includes only 5. Each distinct gene cluster found between the flanking genes is assigned a unique number identifying the locus type, and these data can be found in public databases [16].

The final column shows the presence of a CRISPR-Cas system in the isolate. Clustered regularly interspaced short palindromic repeat (CRISPR) arrays and CRISPR-associated genes (cas) constitute bacterial adaptive immune systems and function as variable genetic elements. Each CRISPR-Cas locus includes a strain-specific array of so-called spacer sequences, which can be used for strain subtyping [17]. CRISPR-Cas systems can be divided into six major types (I–VI) and several subtypes (A-I, K, U) based on a combination of evidence from phylogenetic, comparative genomic, and structural analysis [18,19]. 'UNKN' in this column denotes an incomplete system, in which some genes are absent, while 'NF' indicates that CRISPR-Cas system was not revealed in a particular isolate genome.

2.1.2. AMR Gene Table

Another part of the dataset includes information regarding the presence of various genes known to confer, when expressed properly, antimicrobial resistance in the investigated *A. baumannii* genomes. This information can be useful for the comparative analysis of AMR gene acquisition by particular groups of isolates, especially in cases where genetically and genomically similar isolates carry different AMR determinants. The first column shows the assembly code, which is the same as that in the typing summary table; the second column shows the number of AMR genes found in a particular isolate; and the other columns exhibit the presence of a particular gene in the first row and its sequence similarity level with the corresponding allele from the database. The absence of a gene is marked with a dot for better readability purposes.

An example is provided in Table 2. Only a few available genes are presented in this example.

Table 2. Exemplary data and column information for antimicrobial resistance gene table.

Assembly Code	NUM_FOUND [1]	aac(3)-Ia	aac(6′)-Iaf	...	blaOXA-72	blaOXA-82
GCA_000222245.2	10	89.51	100.00	...	.	100.00
GCA_000222265.2	11	97.38	100.00	...	100.00	

[1] 'NUM_FOUND' column shows the total number of AMR genes revealed in a particular isolate; three dots indicate that some genes were omitted in this exemplary table but are available online in the full version; single dots indicate that a particular gene was not revealed in a particular genome.

The presence of a particular gene by itself does not confirm resistance to the corresponding antimicrobial drug since this gene, for example, might not be expressed [20,21]. However, gene presence/absence information is very useful for the estimation of the AMR potential within bacterial population.

2.1.3. cgMLST Profiles

The third part of the dataset includes cgMLST profiles for all isolates. The cgMLST typing scheme is similar to MLST in that it enumerates the selected gene variants and uses their combination to form a profile, but the difference is that cgMLST relies on all conservative genes within a particular species. Such a scheme was proposed for *A. baumannii* in 2017 [22] and included 2390 loci. The allele variants for these loci are available in a regularly updated public database at cgmlst.org (https://www.cgmlst.org/ncs/schema/schema/3956907, accessed on 14 February 2023).

cgMLST profiles can be used in the cluster analysis of a set of isolates in order to estimate their genomic relatedness and, possibly, obtain some evolutionary or epidemiologically valuable insights. The threshold of three different cgMLST loci was proposed to check whether two *A. baumannii* isolates belonged to a single strain [23], but less stringent criteria can be used depending on the specific investigation.

In this dataset, cgMLST profiles are given in a table format. The first column contains the same assembly identifier as other dataset tables, and the other columns show the numbers representing the variants of genes given in a header row. Some special characters can appear besides numbers. Namely, 'N' indicates a novel allele variant not present in the

database; '0?' indicates a locus is missing in the assembly (probably due to low sequencing quality); '-' indicates an allele is partially covered; and '+' represents multiple possible alleles, in which case the most probable is shown.

2.2. General Statistics

Some general statistics based on the dataset are provided below. As was noted previously [11,24], a Genbank set of genomes cannot be considered representative of the whole *A. baumannii* population since it is strongly biased towards multidrug-resistant or other clinically relevant isolates. At the same time, the statistics on particular STs, ICs, AMR genes, etc., can provide useful information for reference set selection and comparison purposes.

We will refer to any Genbank assembly record containing a complete or partial genome as an 'isolate' for simplicity's sake, although different assemblies can represent the same isolate, and some records may contain only a part of the isolate genome.

The distribution of ICs in Genbank is shown in Figure 1.

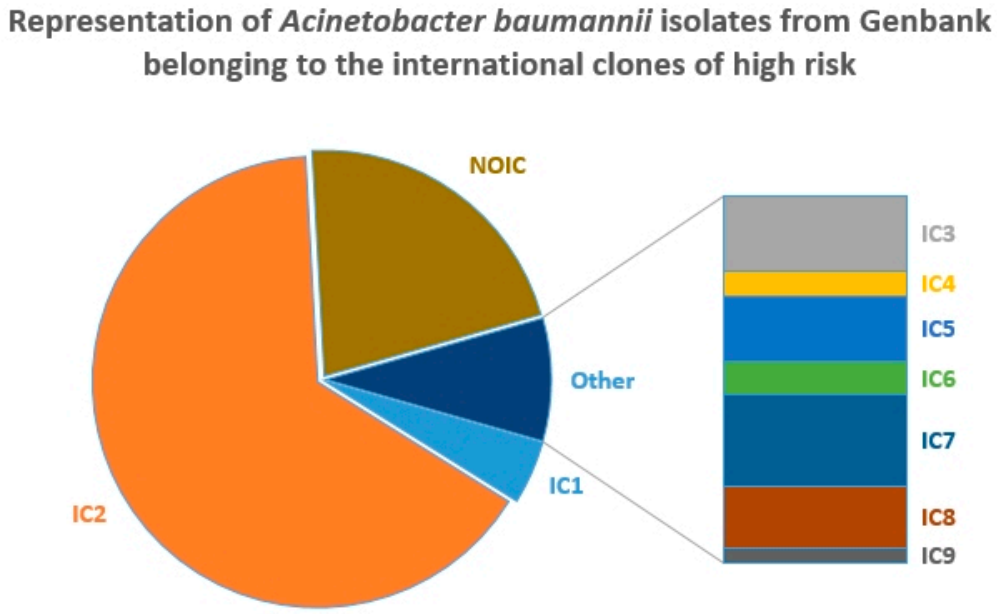

Figure 1. Representation of *A. baumannii* isolates belonging to the international clones of high risk in Genbank. 'NOIC' represents the isolates not belonging to IC1-IC9.

A summary of the top three dominating characteristics in each category is given in Table 3.

Table 3. Top three dominating representatives of all features investigated in *A. baumannii* dataset.

Feature	Top 3 Representatives
IC	IC2, IC1, IC7
ST^{Pas}	ST2, ST1, ST499
OXA-51-like variant	OXA-66, OXA-82, OXA-69
KL	KL2, KL3, KL22
OCL	OCL1, OCL3, OCL6
Acquired AMR genes	bla_{OXA-23}, *aph(6)-Id*, *aph(3″)-Ib*
CRISPR-Cas system	NF, I-F1, I-F2

In total, 78.5% of the isolates belonged to IC1-IC9, with IC2 accounting for 65% of all Genbank genomes and 83% of *A. baumannii* isolates belonging to ICs. The second largest IC-IC1- was revealed in about 3.6%/4.5% of all isolates and isolates belonging to ICs, respectively. These results are typical since IC2 is known to be the most successful

and widespread clone of *A. baumannii*, being responsible for the majority of outbreaks worldwide [25]. The dominating ST was, not surprisingly, ST2, which constituted most of IC2 and was revealed in 63.3% (11,108) of cases. ST1 (3.59%, IC1) and ST499 (3.4%, NOIC) were also in the top three. The number of distinct STs revealed was 482. However, 398 of these presented 10 or fewer isolates, with 236 STs featuring only a single isolate.

The dominating OXA-51-like variant was OXA-66 (about 60%) associated with IC2, followed by OXA-82 (also associated with IC2) and OXA-69 (IC1), each accounting for ten times fewer isolates than the former. In this case, the statistics are nearly complete since the genes encoding OXA-51-like beta-lactamases were revealed in 98% of the isolates due to low similarity, which was probably caused by bad sequencing quality or low coverage of particular genomic regions. The data on intrinsic beta-lactamases clearly correlate with the IC and ST distribution, which is as expected [11,14]. In total, 130 distinct OXA-51-like variants were revealed.

The top three KL types were KL2 (15.5% of the isolates), KL3 (11.6%), and KL22 (5.4%). KL2 was recently reported to be the most common type in *A. baumannii* and was associated with increased AMR [26], which conforms with the previously noted bias of Genbank *A. baumannii* isolates towards more resistant ones. In total, 229 types were present.

The diversity of OCL types was lower, with OCL1 accounting for 70% of the isolates. OCL3 and OCL6 were found in about 11% and 5%, respectively. In total, 22 types were revealed. These results correspond to a previous study, in which OCL1 was the most common type among *A. baumannii* isolates belonging to ST1, ST2, ST3, and ST78, which together constituted about 71% of the isolates available in Genbank [27].

The median number of AMR genes possessed by the isolates was 10, with the number ranging from 1 to 25. The number was never equal to zero since intrinsic $bla_{OXA\text{-}51\text{-}like}$ and bla_{ADC} genes were also included. In several cases, intrinsic genes were not revealed due to low similarity or insufficient genome region coverage. We revealed 1236 (7.0%) isolates possessing only the intrinsic genes denoted above.

The most abundant genes, excluding intrinsic ones, were $bla_{OXA\text{-}23}$, *aph(6)-Id*, and *aph(3″)-Ib*, which were revealed in about 64% of the isolates each, although not always they were presented simultaneously. The first gene, encoding OXA-23 carbapenemase, was shown to be associated with IC2 [28], and our study confirmed that it was found in about 95% of IC2 isolates. The latter two genes, usually being of plasmid origin [29], encode aminoglycoside phosphotransferases, conferring AMR to streptomycin.

CRISPR-Cas systems were revealed in about 8% (1342) of the isolates. The dominating type in this case was I-F, with the I-F1 subtype accounting for 67% and I-F2–for about 20% of the CRISPR-Cas-positive isolates, which agrees with a previous RefSeq analysis, in which it was found that most CRISPR-Cas systems in *A. baumannii* belonged to the types I-F1 and I-F2 [30]. The systems were predominantly found in the IC1 (39%), NOIC (29%), and IC7 (17%) groups and were not revealed in IC2, which also conforms with previous investigations [30]. The relatively low fraction of isolates in Genbank containing apparently functional CRISPR-Cas systems could also be attributed to the overrepresentation of IC2, which usually does not possess such a system.

3. Methods

We retrieved 17,546 genomic sequences of *A. baumannii* from Genbank (https://www.ncbi.nlm.nih.gov/genbank/, accessed on 14 January 2023), for which the assembly level was defined as 'Complete Genome', 'Chromosome', or 'Scaffold'.

Multilocus sequence typing (MLST) was performed using the PubMLST database (https://pubmlst.org/bigsdb?db=pubmlst_abaumannii_seqdef, accessed on 14 February 2023) using the Pasteur [13] typing scheme. When we tried to obtain STs based on the Oxford scheme [31], reliable assignment was obtained for only 22.8% (4053 of 17,546) of the isolates due to the known issue of *gdhB* gene paralogy and technical artifacts [32]. It was also shown that the Pasteur scheme is more appropriate for population biology and epidemiological studies of *A. baumannii* than the Oxford one since it can be used for more

precise isolate classification in clonal groups [33]. For these reasons, the data for Oxford STs are not presented in our dataset.

The AMR genes and intrinsic *bla*$_{OXA\text{-}51\text{-}like}$ gene variants were detected using Resfinder 4.3.0 software (https://cge.cbs.dtu.dk/services/ResFinder/, accessed on 20 February 2023, using default parameters).

The detection of capsule synthesis loci (KL) and lipooligosaccharide outer core loci (OCL) was conducted using Kaptive version 2.0.3 [34] with the default parameters (the last update of the databases was on 13 February 2023).

The presence of CRISPR-Cas systems in the genomes analyzed was investigated using CRISPRCasFinder [35] version 4.2.20 with the following parameters: '-fast -rcfowce -ccvRep -vicinity 1200 –cas -useProkka'. Recent classifications based on the multiparametric analysis reported type I loci with *cas3* as a signature gene and type I-F with fused *cas3* and *cas2* genes [36,37]. The types I-F1 and I-F2 both contain the *cas1* and *cas2-3* genes, but the rest of the loci are different and include four genes, *csy1* (*cas8f1*), *csy2* (*cas5f1*), *csy3* (*cas7f1*), and *csy4* (*cas6f*), in I-F1 and three genes, *cas5fv* (*cas5f2*), *cas6f*, and *cas7fv* (*cas7f2*), in I-F2 [30].

The cgMLST profiles were built with MentaList [38] (https://github.com/WGS-TB/MentaLiST, version 0.2.4, default parameters, accessed on 25 May 2023) using the scheme including 2390 loci obtained from cgmlst.org (https://www.cgmlst.org/ncs/schema/schema/3956907/, last update on 20 February 2023 [22]).

Additional data processing and output formatting were performed using the computational pipeline developed earlier by us [39,40].

Assignment of the isolates to ICs was based on Pasteur MLST ST and the *bla*$_{OXA\text{-}51\text{-}like}$ gene variant and supported by the cgMLST profile as needed, as described earlier by us [11]. In order to make the classification reliable, we performed cluster analysis of STs using currently available data and compared it with the information derived from the literature analysis of experimentally verified IC assignments.

4. User Notes

The regularly updated description of possible dataset applications for genomic epidemiology investigations, including detailed scripting commands for UNIX-based systems, will be provided on the dataset webpage in a how_to_use.txt file.

Author Contributions: Conceptualization, A.S. and Y.M.; data curation, A.S.; formal analysis, A.S. and Y.M.; funding acquisition, V.A.; investigation, A.S., Y.M. and V.A.; methodology, A.S. and Y.M.; project administration, V.A.; resources, A.S.; software, A.S.; supervision, V.A.; validation, A.S. and Y.M.; visualization, A.S.; writing—original draft, A.S.; writing—review and editing, A.S. and Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of a grant in the form of a subsidy for the creation and development of the 'World-class Genomic Research Center for Ensuring Biological Safety and Technological Independence under the Federal Scientific and Technical Program for the Development of Genetic Technologies', agreement No. 075-15-2019-1666.

Data Availability Statement: The dataset is available at https://doi.org/10.5281/zenodo.10143377.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Salam, M.A.; Al-Amin, M.Y.; Salam, M.T.; Pawar, J.S.; Akhter, N.; Rabaan, A.A.; Alqumber, M.A.A. Antimicrobial Resistance: A Growing Serious Threat for Global Public Health. *Healthcare* **2023**, *11*, 1946. [CrossRef]
2. Denissen, J.; Reyneke, B.; Waso-Reyneke, M.; Havenga, B.; Barnard, T.; Khan, S.; Khan, W. Prevalence of ESKAPE pathogens in the environment: Antibiotic resistance status, community-acquired infection and risk to human health. *Int. J. Hyg. Environ. Health* **2022**, *244*, 114006. [CrossRef]

3. Murray, C.J.L.; Ikuta, K.S.; Sharara, F.; Swetschinski, L.; Robles Aguilar, G.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; et al. Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *Lancet* **2022**, *399*, 629–655. [CrossRef]
4. Shaidullina, E.R.; Schwabe, M.; Rohde, T.; Shapovalova, V.V.; Dyachkova, M.S.; Matsvay, A.D.; Savochkina, Y.A.; Shelenkov, A.A.; Mikhaylova, Y.V.; Sydow, K.; et al. Genomic analysis of the international high-risk clonal lineage Klebsiella pneumoniae sequence type 395. *Genome Med.* **2023**, *15*, 9. [CrossRef]
5. Zhang, X.; Li, F.; Awan, F.; Jiang, H.; Zeng, Z.; Lv, W. Molecular Epidemiology and Clone Transmission of Carbapenem-Resistant *Acinetobacter baumannii* in ICU Rooms. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 633817. [CrossRef] [PubMed]
6. Rafei, R.; Osman, M.; Dabboussi, F.; Hamze, M. Update on the epidemiological typing methods for *Acinetobacter baumannii*. *Future Microbiol.* **2019**, *14*, 1065–1080. [CrossRef] [PubMed]
7. Atxaerandio-Landa, A.; Arrieta-Gisasola, A.; Laorden, L.; Bikandi, J.; Garaizar, J.; Martinez-Malaxetxebarria, I.; Martinez-Ballesteros, I. A Practical Bioinformatics Workflow for Routine Analysis of Bacterial WGS Data. *Microorganisms* **2022**, *10*, 2364. [CrossRef] [PubMed]
8. Castanheira, M.; Mendes, R.E.; Gales, A.C. Global Epidemiology and Mechanisms of Resistance of *Acinetobacter baumannii*-calcoaceticus Complex. *Clin. Infect. Dis.* **2023**, *76*, S166–S178. [CrossRef]
9. Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D.L.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y.; et al. Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **2018**, *18*, 318–327. [CrossRef]
10. Jolley, K.A.; Bray, J.E.; Maiden, M.C.J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* **2018**, *3*, 124. [CrossRef]
11. Shelenkov, A.; Akimkin, V.; Mikhaylova, Y. International Clones of High Risk of *Acinetobacter baumannii*-Definitions, History, Properties and Perspectives. *Microorganisms* **2023**, *11*, 2115. [CrossRef] [PubMed]
12. Castillo-Ramirez, S. Genomic epidemiology of *Acinetobacter baumannii* goes global. *mBio* **2023**, *14*, e0252023. [CrossRef]
13. Diancourt, L.; Passet, V.; Nemec, A.; Dijkshoorn, L.; Brisse, S. The population structure of *Acinetobacter baumannii*: Expanding multiresistant clones from an ancestral susceptible genetic pool. *PLoS ONE* **2010**, *5*, e10034. [CrossRef]
14. Pournaras, S.; Gogou, V.; Giannouli, M.; Dimitroulia, E.; Dafopoulou, K.; Tsakris, A.; Zarrilli, R. Single-locus-sequence-based typing of blaOXA-51-like genes for rapid assignment of *Acinetobacter baumannii* clinical isolates to international clonal lineages. *J. Clin. Microbiol.* **2014**, *52*, 1653–1657. [CrossRef]
15. Roshini, J.; Patro, L.P.P.; Sundaresan, S.; Rathinavelan, T. Structural diversity among *Acinetobacter baumannii* K-antigens and its implication in the in silico serotyping. *Front. Microbiol.* **2023**, *14*, 1191542. [CrossRef]
16. Cahill, S.M.; Hall, R.M.; Kenyon, J.J. An update to the database for *Acinetobacter baumannii* capsular polysaccharide locus typing extends the extensive and diverse repertoire of genes found at and outside the K locus. *Microb. Genom.* **2022**, *8*, e000878. [CrossRef] [PubMed]
17. Tyumentseva, M.; Mikhaylova, Y.; Prelovskaya, A.; Tyumentsev, A.; Petrova, L.; Fomina, V.; Zamyatin, M.; Shelenkov, A.; Akimkin, V. Genomic and Phenotypic Analysis of Multidrug-Resistant *Acinetobacter baumannii* Clinical Isolates Carrying Different Types of CRISPR/Cas Systems. *Pathogens* **2021**, *10*, 205. [CrossRef]
18. Makarova, K.S.; Koonin, E.V. Annotation and Classification of CRISPR-Cas Systems. *Methods Mol. Biol.* **2015**, *1311*, 47–75. [CrossRef]
19. Hryhorowicz, M.; Lipinski, D.; Zeyland, J. Evolution of CRISPR/Cas Systems for Precise Genome Editing. *Int. J. Mol. Sci.* **2023**, *24*, 14233. [CrossRef]
20. Stasiak, M.; Mackiw, E.; Kowalska, J.; Kucharek, K.; Postupolski, J. Silent Genes: Antimicrobial Resistance and Antibiotic Production. *Pol. J. Microbiol.* **2021**, *70*, 421–429. [CrossRef] [PubMed]
21. Shelenkov, A.; Mikhaylova, Y.; Yanushevich, Y.; Samoilov, A.; Petrova, L.; Fomina, V.; Gusarov, V.; Zamyatin, M.; Shagin, D.; Akimkin, V. Molecular Typing, Characterization of Antimicrobial Resistance, Virulence Profiling and Analysis of Whole-Genome Sequence of Clinical Klebsiella pneumoniae Isolates. *Antibiotics* **2020**, *9*, 261. [CrossRef]
22. Higgins, P.G.; Prior, K.; Harmsen, D.; Seifert, H. Development and evaluation of a core genome multilocus typing scheme for whole-genome sequence-based typing of *Acinetobacter baumannii*. *PLoS ONE* **2017**, *12*, e0179228. [CrossRef]
23. Schurch, A.C.; Arredondo-Alonso, S.; Willems, R.J.L.; Goering, R.V. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. *Clin. Microbiol. Infect.* **2018**, *24*, 350–354. [CrossRef]
24. Hamidian, M.; Nigro, S.J. Emergence, molecular mechanisms and global spread of carbapenem-resistant *Acinetobacter baumannii*. *Microb. Genom.* **2019**, *5*, e000306. [CrossRef] [PubMed]
25. Cain, A.K.; Hamidian, M. Portrait of a killer: Uncovering resistance mechanisms and global spread of *Acinetobacter baumannii*. *PLoS Pathog.* **2023**, *19*, e1011520. [CrossRef]
26. Yang, J.L.; Yang, C.J.; Chuang, Y.C.; Sheng, W.H.; Chen, Y.C.; Chang, S.C. Association of capsular polysaccharide locus 2 with prognosis of *Acinetobacter baumannii* bacteraemia. *Emerg. Microbes Infect.* **2022**, *11*, 83–90. [CrossRef]
27. Sorbello, B.M.; Cahill, S.M.; Kenyon, J.J. Identification of further variation at the lipooligosaccharide outer core locus in *Acinetobacter baumannii* genomes and extension of the OCL reference sequence database for Kaptive. *Microb. Genom.* **2023**, *9*, e001042. [CrossRef] [PubMed]

28. Khuntayaporn, P.; Kanathum, P.; Houngsaitong, J.; Montakantikul, P.; Thirapanmethee, K.; Chomnawang, M.T. Predominance of international clone 2 multidrug-resistant *Acinetobacter baumannii* clinical isolates in Thailand: A nationwide study. *Ann. Clin. Microbiol. Antimicrob.* **2021**, *20*, 19. [CrossRef]
29. Kyriakidis, I.; Vasileiou, E.; Pana, Z.D.; Tragiannidis, A. *Acinetobacter baumannii* Antibiotic Resistance Mechanisms. *Pathogens* **2021**, *10*, 373. [CrossRef]
30. Yadav, G.; Singh, R. In silico analysis reveals the co-existence of CRISPR-Cas type I-F1 and type I-F2 systems and its association with restricted phage invasion in *Acinetobacter baumannii*. *Front. Microbiol.* **2022**, *13*, 909886. [CrossRef] [PubMed]
31. Bartual, S.G.; Seifert, H.; Hippler, C.; Luzon, M.A.; Wisplinghoff, H.; Rodriguez-Valera, F. Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii*. *J. Clin. Microbiol.* **2005**, *43*, 4382–4390. [CrossRef]
32. Hamidian, M.; Nigro, S.J.; Hall, R.M. Problems with the Oxford Multilocus Sequence Typing Scheme for *Acinetobacter baumannii*: Do Sequence Type 92 (ST92) and ST109 Exist? *J. Clin. Microbiol.* **2017**, *55*, 2287–2289. [CrossRef]
33. Gaiarsa, S.; Batisti Biffignandi, G.; Esposito, E.P.; Castelli, M.; Jolley, K.A.; Brisse, S.; Sassera, D.; Zarrilli, R. Comparative Analysis of the Two *Acinetobacter baumannii* Multilocus Sequence Typing (MLST) Schemes. *Front. Microbiol.* **2019**, *10*, 930. [CrossRef] [PubMed]
34. Wyres, K.L.; Cahill, S.M.; Holt, K.E.; Hall, R.M.; Kenyon, J.J. Identification of *Acinetobacter baumannii* loci for capsular polysaccharide (KL) and lipooligosaccharide outer core (OCL) synthesis in genome assemblies using curated reference databases compatible with Kaptive. *Microb. Genom.* **2020**, *6*, e000339. [CrossRef]
35. Couvin, D.; Bernheim, A.; Toffano-Nioche, C.; Touchon, M.; Michalik, J.; Neron, B.; Rocha, E.P.C.; Vergnaud, G.; Gautheret, D.; Pourcel, C. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* **2018**, *46*, W246–W251. [CrossRef]
36. Makarova, K.S.; Wolf, Y.I.; Iranzo, J.; Shmakov, S.A.; Alkhnbashi, O.S.; Brouns, S.J.J.; Charpentier, E.; Cheng, D.; Haft, D.H.; Horvath, P.; et al. Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **2020**, *18*, 67–83. [CrossRef] [PubMed]
37. Shmakov, S.A.; Barth, Z.K.; Makarova, K.S.; Wolf, Y.I.; Brover, V.; Peters, J.E.; Koonin, E.V. Widespread CRISPR-derived RNA regulatory elements in CRISPR-Cas systems. *Nucleic Acids Res.* **2023**, *51*, 8150–8168. [CrossRef] [PubMed]
38. Feijao, P.; Yao, H.T.; Fornika, D.; Gardy, J.; Hsiao, W.; Chauve, C.; Chindelevitch, L. MentaLiST—A fast MLST caller for large MLST schemes. *Microb. Genom.* **2018**, *4*, e000146. [CrossRef]
39. Shelenkov, A.; Petrova, L.; Fomina, V.; Zamyatin, M.; Mikhaylova, Y.; Akimkin, V. Multidrug-Resistant Proteus mirabilis Strain with Cointegrate Plasmid. *Microorganisms* **2020**, *8*, 1775. [CrossRef]
40. Shelenkov, A.; Petrova, L.; Zamyatin, M.; Mikhaylova, Y.; Akimkin, V. Diversity of International High-Risk Clones of *Acinetobacter baumannii* Revealed in a Russian Multidisciplinary Medical Center during 2017–2019. *Antibiotics* **2021**, *10*, 1009. [CrossRef] [PubMed]