Health Biotechnology and Biopharma (2021), 5(1): 58-72

Original Research Article

Lessons from comparative genome analysis of *Acinetobacter baumannii* strains

Arezoo Piran [1], Sepideh Fereshteh [1] , Farzad Badmasti [1, 2, *]

[1]Department of Bacteriology, Pasteur Institute of Iran, Tehran, Iran

[2]Microbiology Research Center (MRC), Pasteur Institute of Iran, Tehran, Iran

***Corresponding author:** Farzad Badmasti, Department of Bacteriology, Pasteur Institute of Iran, Tehran, Iran. E-mail: f_badmasti@pasteur.ac.ir*

DOI: 10.22034/HBB.2020.12

Received: April 24, 2021; **Accepted:** May 22, 2021

ABSTRACT

The whole-genome sequencing method allows an analysis of genetic differences between bacterial isolates belonging to the same species. In this study, six strains, including *Acinetobacter baumannii* ACICU (ST2), *A. baumannii* ATCC 17978 (ST437), *A. baumannii* ATCC 19606 (ST57), *A. baumannii* BJAB0868 (ST2), *A. baumannii* MDR-TJ (ST2), and *A. baumannii* SDF (ST17) were investigated. Comparative genome analysis revealed different genomic events on genomes of identical STs. In contrast, the positions of the tRNAs on the genomes showed that the identical STs (ACICU, BJAB0868, and MDR-TJ as ST2) have the exact location map. It appeared that the genomic backbone of identical STs is conserved among them. Still, each strain in the same ST has undergone genomic events on its genomes which can be considered as genome plasticity process.

***Keywords*:** Genome analysis, comparative genome analysis, sequencing method, *Acinetobacter baumannii* strains

INTRODUCTION

Acinetobacter baumannii, as nosocomial Gram-negative coccobacilli, causes various infections, especially in hospitalized and ICU (Intensive Care Units) cases [1]. The molecular epidemiology of *A. baumannii* and the determination of lineage relationships among isolates are critical for establishing infection control measures in

hospitals [2]. Pulsed-Field Gel Electrophoresis (PFGE) and Multi-Locus Sequence Typing (MLST) are two classical typing methods. These methods have been developed to analyze the epidemiology and clonal relatedness of different microorganisms such as *A. baumannii* strains [3]. PFGE is a widely used method to distinguish bacterial strains from nosocomial outbreaks [4], and MLST is used to study population structures of bacterial pathogens [5]. However, these methods give us very little information about the genomic backbone of bacteria.

The recently available, rapid, and inexpensive Whole-Genome Sequencing (WGS) method allows a thorough analysis of genetic differences between bacterial isolates belonging to the same genus [6]. WGS has already been used to characterize bacterial isolates in many significant outbreaks. Likely, it will soon replace existing typing methods [7]. Therefore, this study aimed to examine the genome of *A. baumannii* in detail to highlight the minor differences in the same STs. It also looks at the evolutionary discrepancies that have occurred in the genome. This study attempts to demonstrate the WGS method as a typing technique for future use.

MATERIALS AND METHODS

Collection of the dataset

In this study, six strains of *A. baumannii*, including *A. baumannii* ACICU (ST2), *A. baumannii* ATCC 17978, *A. baumannii* ATCC 19606, *A. baumannii* BJAB0868 (ST2), *A. baumannii* MDR-TJ (ST2), and *A. baumannii* SDF were selected to comparative genome analysis. The MLST Sequence Type (ST) of each strain was characterized using the MLST 2.0 webserver. *A. baumannii* ATTC 17978 and ATCC 19606 are reference strains and have been considered genomic models in *A. baumannii* projects. The SDF strain is a non-pathogenic *A. baumannii* that is used as a control strain in this analysis.

Comparative genome evaluation and statistical analysis

BioCyc provides tools for navigation, visualization, analysis of underlying databases, and analysis of omics data, including genomes and metabolic pathways [8]. The BioCyc analysis program was used to compare the genome contents of the five selected strains and identify genome size, GC content, genes, pseudogenes, metabolic pathways, tRNA, rRNA, and plasmids.

Multiple sequence alignments and phylogenic analysis

Mauve is software for generating multiple genome alignments in the presence of large-scale evolutionary events such as rearrangement and inversion [9]. Numerous genome alignments provide a basis for the study of comparative genomics and the analysis of genome-wide evolutionary dynamics. In this study, multiple sequence alignments of *A. baumannii* genomes were performed with Mauve to investigate differences and similarity arrangements of five selected genomes. Also, pan-genomic analysis has provided valuable insights into genome dynamics, species evolution, drug resistance, and many other features of the microbial world. In this project, the phylogenetic tree was drawn using BPGA software version 1.3 to investigate the differences and similarities among five bacterial genomes [10].

Comparison of ortholog genes in multiple genomes

Ortholuge Data Base (DB) was used to compare orthologous genes between genomes. Ortholuge DB contains ortholog-based predictions for fully sequenced bacterial and archaeal genomes. Ortholuge DB also has reciprocal best hit-based ortholog predictions, in-paralog predictions (recently duplicated genes), and suggested orthologues [11].

Detection of prophages and antimicrobial resistance genes in genomes

PHAST webserver was used to detect integrated prophages in DNA genomes of *A. baumannii* strains. PHAST (PHAge Search Tool) is a web server designed to rapidly and accurately identifying, annotating, and graphically displaying prophage sequences in bacterial genomes or plasmids [12]. The RGI can predict resistomes from protein or nucleotide data based on homology and SNP models. This web portal supports the analysis of genomes, genome assemblies, metagenomics contigs, or proteomes. The command-line tool additionally helps metagenomics reads analysis and k-mer prediction of pathogen-of-origin for AMR genes [13].

Characterization of genomic maps of selected strains

The CGView Server can be used to visualize features associated with any bacterial, plasmid, chloroplast, or mitochondrial genome. It can help identify conserved genome segments, horizontal gene transfer instances, and gene copy number differences [14].

RESULTS

Multi-locus sequence typing (MLST) and statistical results

An MLST analysis was carried out on the six strains to define the relatedness among them. The obtained result showed that the strains ACICU, BJAB0868, and MDR-TJ share the same Sequence Type (ST2), and the strains of ATTC 17978 and ATCC 19606 had ST437 and ST52, respectively. Besides, strain SDF belongs to ST17. The basic whole-genome sequencing statistics of selected genomes was shown in Table1.

Phylogenic genome analysis of A. baumannii strains

Phylogenic genome analysis showed that the three *A. baumannii* strains, including ACICU, MDR-TJ, and BJAB0868, are closer to each other, and they belong to the ST2 group. Moreover, *A. baumannii* ATCC 17978 and ATCC 19606 are categorized in the nearer region, and they belong to closely related STs (Figure 1). It seems that the SDF strain is far from the other strains investigated in this study.

Orthologous clusters in the A. baumannii genomes

We further analyzed three genomes, including *A. baumannii* ACICU and MDR-TJ, against *A. baumannii* SDF as the reference genome (non-pathogenic strain), ATCC 17978, and ACICU to identify possible strain-specific genes (Figure 2). We found that ACICU and MDR-TJ strains have a closed number of genes. These mentioned genes were not present in the reference genome (Genes are shown in Blue color). On the other hand, ATCC 19606 has more unique genes (831 genes). The genomes of ACICU (13 genes) and ATCC 17978 (16 genes) have more than one orthologous gene compared to the reference genome (Orange color). In contrast, the genome of MDR-TJ (63 genes) lost some orthologous and paralogous genes during evolution compared to the SDF genome (Red color).

Table 1. The basic whole-genome sequencing statistics of selected bacterial genomes

Database	**ACICU**	**ATCC 17978**	**ATCC 19606**	**BJAB0868**	**MDR-TJ**	**SDF**
Accession number	NC_010611	NZ_CP059041	NZ_CP045110	NC_021729	NC_017847	CU468230.2
Genome size	3.90 Mb	4.01 Mb	3.98 Mb	3.91 Mb	3.96 Mb	3.48 Mb
%GC Content	39.1	39	39.2	39	39.1	39.12
Pasteur's MLST	ST2	ST437	ST52	ST2	ST2	ST17
Genes	3,712	3,818	3,798	3,720	3,799	3,598
Pseudogenes	34	80	46	73	75	598
Protein	3,584	3,644	3,656	3,550	3,628	2,913
Genes of known or predicted molecular function	1,241	1,186	1,302	1,354	1,304	1,184
Enzymes that catalyze Small molecule reactions	575	681	623	688	631	633
tRNA Genes	72	72	74	75	75	64
rRNA Genes	18	18	18	18	18	15
Plasmids	2	2	1	3	2	3

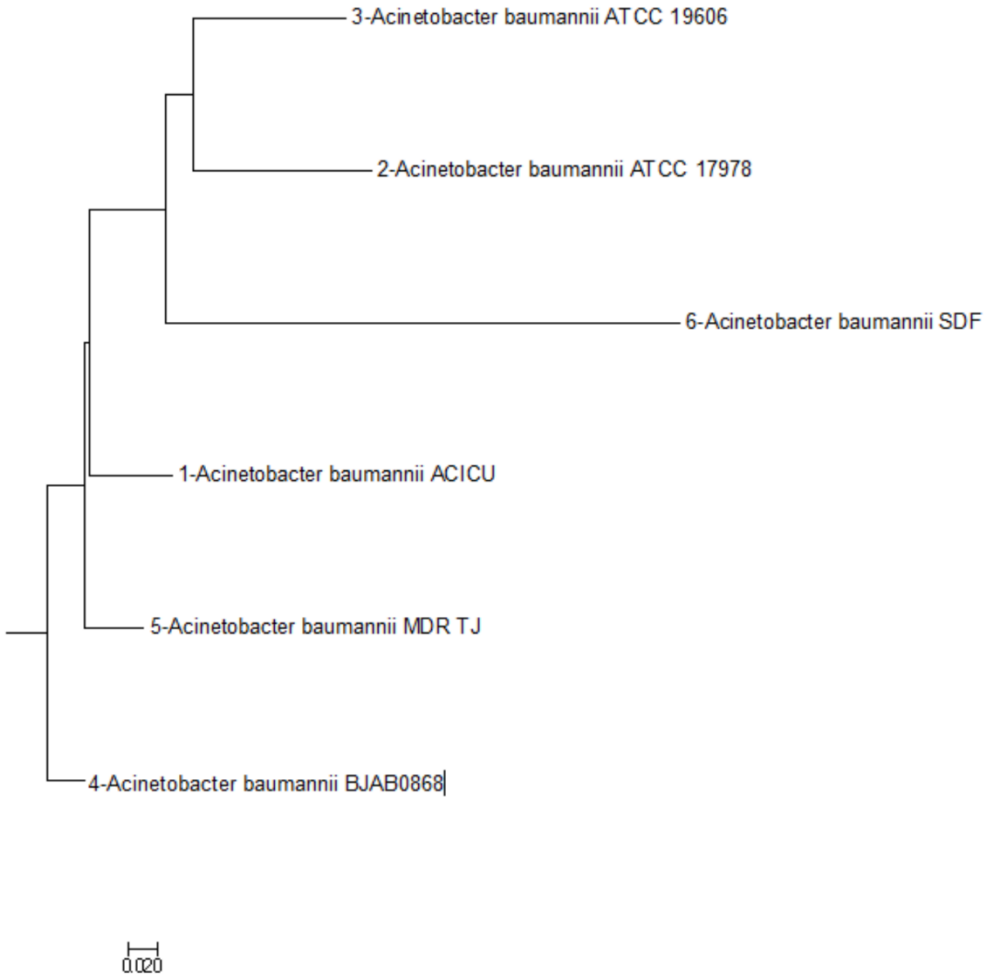

Figure 1. The neighbor-joining tree of five genomes of *A. baumannii*. ACICU, MDR-TJ, and BJAB0868 strains are close to each other and belong to ST2. *A. baumannii* ATCC 17978 and ATCC 19606 belong to ST437 and ST52, respectively.

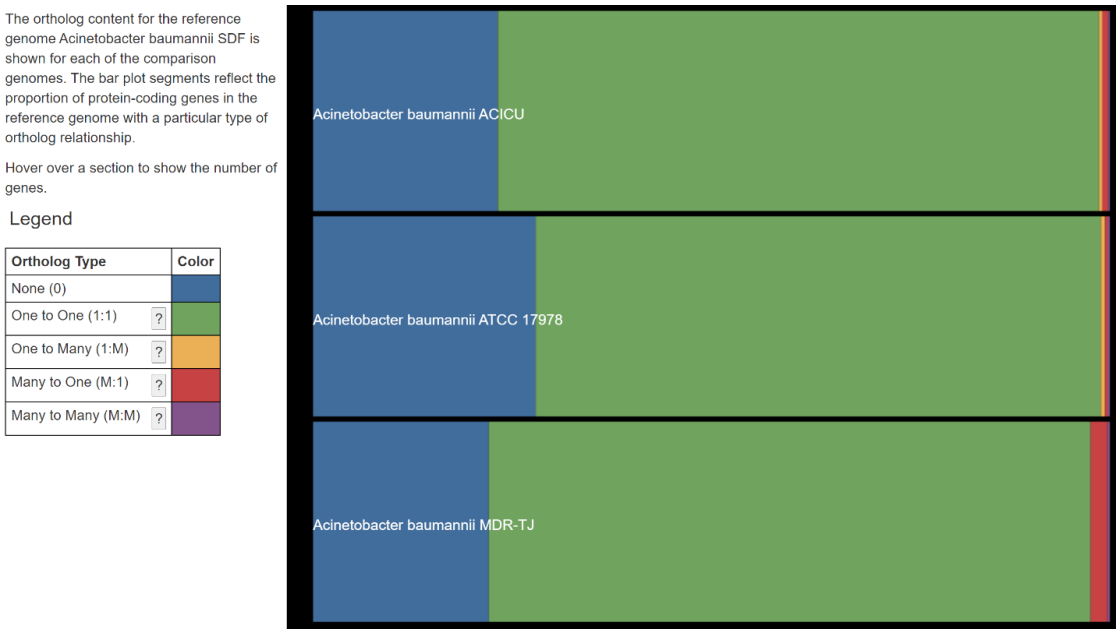

Figure 2. Orthologous analysis of three genomes, including *A. baumannii* ACICU and MDR-TJ against *A. baumannii* SDF. This graph showed unique genes compared to the reference strain (Blue color). Studied strains have more than one orthologous gene compared to the reference genome (Orange color). Three studied strains have lost some orthologous and paralogous genes than the SDF genome during evolution (Red color).

Comparative genome analysis

Antibiotic resistance genes

A comparative analysis of antibiotic resistance genes was performed on the six strains (Figure 3). The three isolates (ACICU, MDR-TJ, and BJAB0868) have the same STs (ST2) and possess the same antibiotic-resistant genes in the same relative positions. Antibiotic resistance mechanisms in these three strains include class D –lactamases, aminoglycoside modifying enzymes, and the Resistance-Nodulation-cell Division (RND) efflux pump. On the other hand, as a non-pathogenic strain, SDF has a poor repertoire of antimicrobial genes. It seems that the allele numbers of the *bla*OXA-51-like family and *bla*ADC are related to MLST sequence types.

Prophages on the genomes

The detection of prophages on genomes showed that MDR-TJ has four intact Acinetobacter phage Bphi-B1251 inserted into the genome. ACICU strain also has two intact Acinetobacter phage Bphi-B1251 and also one incomplete. However,

BJAJ0868 has one incomplete PHAGE _Acinet_Bphi_B1251 and two questionable prophages (Table 2).

Genome depiction

Genes (Gray vertical lines), location of tRNA (Orange arrows), rRNA (Green arrows), and GC skew (±) were characterized on the selected genomes. GC-Skew for genomes are conserved patterns within specific bacterial species. It appears that all strains (except ATCC 17978) have a conserved pattern in their GC-Skew. The positions of the tRNAs on the genomes show that the identical STs (ACICU, BJAB0868, and MDR-TJ as ST2) have the same location map. Moreover, MDR-TJ has a reversed complement arrangement. The closely related ST52 (ATCC 19606) and ST437 (ATCC 17978) are not identical but are similar (Figure 4).

Table 2. The integrated prophages among DNA genomes of five selected *A. baumannii* strains

Strain	#	Region length	Completeness	Score	#CDs	Region position	Possible phage	GC percentage
MDR-TJ	1	51.9Kb	intact	110	70	982451-1034354	Acinetobacter phage Bphi-B1251	39.37 %
	2	42.6Kb	intact	150	64	1574697-1617306	Acinetobacter phage Bphi-B1251	39.35 %
	3	37.1Kb	intact	110	55	2742745-2779895	Acinetobacter phage Bphi-B1251	39.99 %
	4	45.2Kb	intact	91	68	2787090-2832332	Acinetobacter phage Bphi-B1251	39.03 %
BJAB0868	1	28.9Kb	incomplete	20	29	997956-1026913	PHAGE_Acinet_Bphi_B1251	36.05 %
	2	25.5Kb	questionable	90	33	1245796-1271306	PHAGE_Pseudo_vB_PaeS_PMG1	40.44 %
	3	52.8Kb	questionable	90	71	3175076-3227902	PHAGE_Acinet_Bphi_B1251	39.46 %
ACICU	1	53.1Kb	intact	91	78	1114716-1167821	Acinetobacter phage Bphi-B1251	38.08 %
	2	15.2Kb	incomplete	20	28	2343200-2358451	Acinetobacter phage Bphi-B1251	37.13 %
	3	53.6Kb	intact	120	68	2878741-2932436	Acinetobacter phage Bphi-B1251	39.70 %

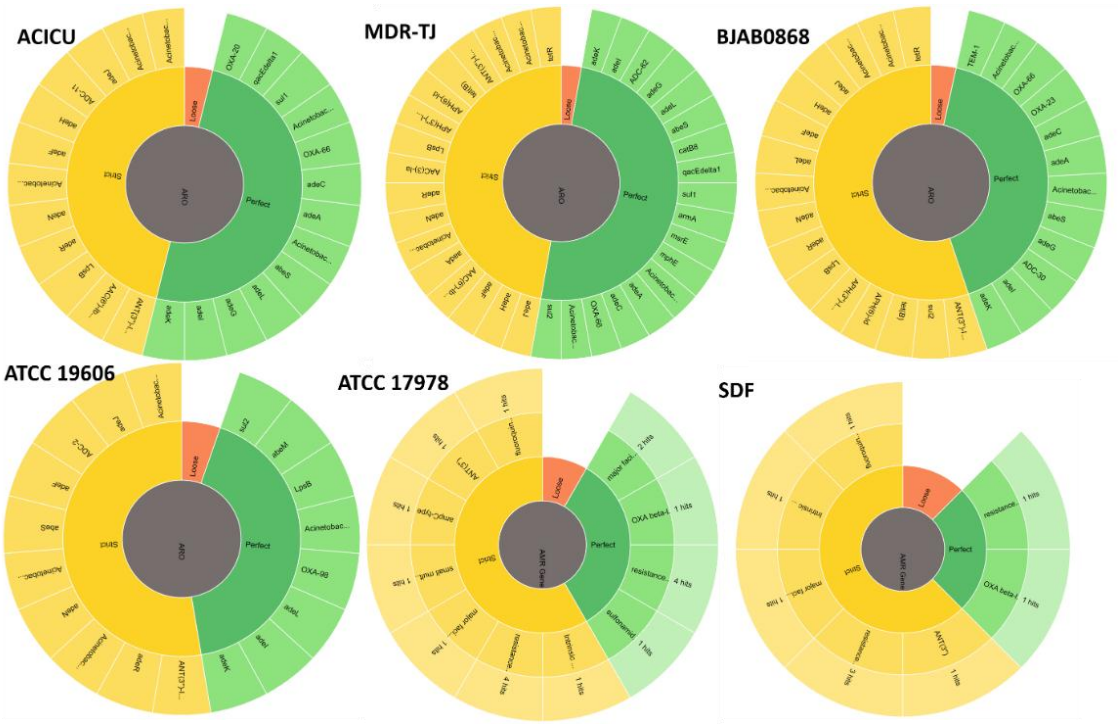

Figure 3. The resistome profile of six selected genomes. Data analysis of antimicrobial enzymes on genomes showed there is a high heterogeneity of antimicrobial resistance genes. Class D Beta-lactamase, aminoglycoside modifying enzymes, and resistance-nodulation-cell division efflux pump are widespread among the three ST2 strain collection (ACICU, MDR-TJ, and BJAB0868). On the other hand, as a non-pathogenic strain, SDF has a poor repertoire of antimicrobial resistance genes. It seems that allele numbers of the *bla*OXA-51-like family and *bla*ADC are related to MLST sequence types.

Figure 4. The genomic map of selected strains of *A. baumannii*. Obtained data on the genomic maps showed that tRNA, rRNA, and GC skewns (+ and -). Analysis of tRNA on the genomic maps showed that this feature could be considered a marker to type strains. It seems that these markers will soon be helpful probes to determine the rearrangement of genomes and DNA shuffling during the evolution of *A. baumannii*.

DISCUSSION

A. baumannii is a common opportunistic pathogen responsible for various nosocomial infections [15]. This bacterium becomes resistant to dry and harsh conditions; thus, it can persist in the hospital environment and contaminate hospital equipment [16,17]. This phenomenon leads to outbreaks that are difficult to confine and could affect hospitalized patients, especially those with compromised immune systems [18,19].

Therefore, strict adherence to infection control practices and understanding the lineage relationships between isolates is critical to interrupt *A. baumannii* transmission.

Recently, molecular typing of bacteria has made incredible progress. Many microbiology laboratories and researchers are now better trained and equipped to perform these new techniques worldwide. Briefly, molecular typing methods can be defined as strategies used to differentiate bacteria based on the composition of biological molecules, such as nucleic acids [20].

Third-generation sequencing technology, also called Next-Generation Sequencing (NGS), is defined by direct sequencing of single-sequence templates. This method allows for much longer sequence reads, although the sequence quality is lower than second-generation short-read methods. NGS has revolutionized genomic research by lowering the cost per mega-base and dramatically increasing throughput. Despite all its advantages, it has not yet been widely adopted in routine molecular typing laboratories [21]. In this study, phylogenetic analysis of *A. baumannii* isolates showed that the MLST sequence type matched the phylogenetic data of the genome. Moreover, the neighbor-joining tree of five genomes of *A. baumannii* agreed with the MLST results.

Whole-genome sequencing analysis of the SDF strain showed that this strain has the smallest genome size and the fewest genes with known function and protein properties. Fournier *et al.* compared the genome of AYE and SDF *Acinetobacter* using the whole-shotgun genome sequencing method. Their comparison results showed that the genome of the virulent AYE strain contained an 86 kb region called the resistance island, which included a cluster of 45 resistance genes. The homologous site in the SDF strain had a 20 kb genomic island without these resistance markers [22]. This ability to switch its genomic structure roughly explains the speed with which *Acinetobacter* captures resistance markers under antibacterial pressure.

Sequencing of *A. baumannii* genomes has uncovered an extensive collection of antimicrobial resistance genes related to transposable elements and insertion sequences found in Genomic Islands (GIs) called AbaR. Various systems transposases, recombinases, and integrases can resize and redesign a few AbaR islands. Resistance genes are also found within

plasmids that can be exchanged intra- and interspecies and even by prophages [23]. In the present study, a comparative analysis of antibiotic resistance genes showed that the three isolates (ACICU, MDR-TJ, and BJAB0868) have similar STs (ST2) and similar antibiotic resistance genes in similar relative positions. Nevertheless, the non-pathogenic strain, SDF, has a low repertoire of antimicrobial genes.

Consistent with our results, a comparative genomic analysis on MDR and drug-sensitive strains of *A. baumannii* showed a high degree of variation in SNPs, AbaR, and prophage well as Type 1 secretion system [24]. Indeed, the plasticity of the *A. baumannii* genome is critical for the development of antibiotic-resistant phenotype among clinical isolates.

As we know, *A. baumannii* strain; ATCC 19606 is sensitive to most antibiotics due to the absence of MDR genes. However, the complete genome sequence and genome-wide metabolic modeling of *A. baumannii* strain ATCC 19606 revealed that this strain contains the genes associated with resistance to various antibiotics and several multidrug resistance efflux pumps. Indeed, two insertion sequences, including ISCR2 and ISAba11, are embedded in a 36.1-Kb genomic island and mediated DNA recombination [22].

CONCLUSION

This study aimed to discover the differences in bacterial genomes, especially with identical sequence types. The author of this study suggested that instead of using the old typing method, researchers can use the NGS method to detect the tRNA position on the genome. Three clinical and two reference isolates of *A. baumannii* have the same number of tRNA genes. This result showed that the localization associated with tRNA genes could be used as a new typing method. Indeed, the alignment of these regions can be helpful for bacterial sorting.

ACKNOWLEDGMENTS

The authors thank the personnel and staff in the Pasteur Institute of Iran for their support.

REFERENCES

[1]. Dijkshoorn L, Nemec A, Seifert H. An increasing threat in hospitals: multidrug-resistant *Acinetobacter baumannii. Nat. Rev. Microbiol.* 2007; 5(12): 939-51.

[2]. Piran A, Shahcheraghi F, Solgi H, Rohani M, Badmasti F. A reliable combination method to identification and typing of epidemic and endemic clones among clinical isolates of *Acinetobacter baumannii*. *Infect. Genet. Evol.* 2017; 54: 501-507.

[3]. Rafei R, Kempf M, Eveillard M, Dabboussi F, Hamze M, Joly-Guillou M-L. Current molecular methods in epidemiological typing of *Acinetobacter baumannii*. *Future Microbiol*. 2014; 9(10): 1179-94.

[4]. Trikaliotis C, Gogou V, Pournaras S, Tsorlini H, Katsifa H, Ramantani E. Identification of multi-resistant strains of *Acinetobacter baumannii* from an ICU in Greece using sequence-based typing and PFGE. *Ann. Microbiol*. 2011; 61(4): 871-77.

[5]. 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(4): 10034.

[6]. Zarrilli R, Pournaras S, Giannouli M, Tsakris A. Global evolution of multidrug-resistant *Acinetobacter baumannii* clonal lineages. *Int. J. Antimicrob. Agents*. 2013; 41(1): 11-19.

[7]. Sabat A, Budimir A, Nashev D, Sá-Leão R, Van Dijl J, Laurent F, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill.* 2013; 18(4): 20380.

[8]. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, et al. The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* 2019; 20(4): 1085-93.

[9]. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignments of conserved genomic sequence with rearrangements. *Genome Res*. 2004; 14(7): 1394-1403.

[10]. Chaudhari NM, Gupta VK, Dutta C. BPGA-an ultra-fast pan-genome analysis pipeline. *Sci. Re*. 2016; 6(1): 1-10.

[11]. Whiteside MD, Winsor GL, Laird MR, Brinkman FS. OrthologueDB: a bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res.* 2013; 41(1): 366-76.

[12]. Zhou M. Meeting strangers in a globalized city: Chinese attitudes toward black africans in Guangzhou china. Wah

Ching centre of research on education in China, Hong Kong. 2011.

[13]. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* 2013; 57(7): 3348-57.

[14]. Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res.* 2008; 36: 181-84.

[15]. Fereshteh S, Abdoli S, Shahcheraghi F, Ajdary S, Nazari M, Badmasti F. New putative vaccine candidates against *Acinetobacter baumannii* using the reverse vaccinology method. *Microb. Pathog*. 2020; 143: 104114.

[16]. Nazari M, Azizi O, Solgi H, Fereshteh S, Shokouhi S, Badmasti F. Emergence of carbapenem resistant *Acinetobacter baumannii* clonal complexes CC2 and CC10 among fecal carriages in an educational hospital. *Int J Environ Health Res*. 2021: 1-11.

[17]. Espinal P, Marti S, Vila J. Effect of biofilm formation on the survival of *Acinetobacter baumannii* on dry surfaces. *J Hosp Infect.* 2012; 80(1): 56-60.

[18]. Gottesman T, Fedorowsky R, Yerushalmi R, Lellouche J, Nutman A. An out break of carbapenem-resistant *Acinetobacter baumannii* in a COVID-19 dedicated hospital. *IPIP*. 2021; 3(1): 100113.

[19]. Hamidian M, Ambrose SJ, Blackwell GA, Nigro SJ, Hall RM. An outbreak of multiply antibiotic-resistant ST49: ST128: KL11: OCL8 *Acinetobacter baumannii* isolates at a Sydney hospital. *J Antimicro Chemotherapy*, 2021.

[20]. Farber J. An introduction to the hows and whys of molecular typing. *J Food Prot.* 1996; 59(10): 1091-1101.

[21]. Fürst D, Tsamadou C, Neuchel C, Schrezenmeier H, Mytilineos J, Weinstock C. Next-generation sequencing technologies in blood group typing. *Transfus Med Hemother*. 2020; 47(1): 4-13.

[22]. Fournier P-E, Vallenet D, Barbe V, Audic S, Ogata H, Poirel L. Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. *PLoS Gen,* 2006; 2(1): 7.

[23]. Leal NC, Campos TL, Rezende AM, Docena C, Mendes-Marques CL, de Sá Cavalcanti FL. Comparative genomics of *Acinetobacter baumannii* clinical strains from Brazil reveals polyclonal dissemination and selective exchange of

mobile genetic elements associated with resistance genes. *Front. Microbiol.* 2020; 11: 1176.

[24]. Liu F, Zhu Y, Yi Y, Lu N, Zhu B, Hu Y. Comparative genomic analysis of *Acinetobacter baumannii* clinical isolates reveals extensive genomic variation and diverse antibiotic resistance determinants. *BMC genomics.* 2014; 15(1): 1-14.