Data in Brief

Complete genome sequence of Acinetobacter baumannii XH386 (ST208), a multi-drug resistant bacteria isolated from pediatric hospital in China

Youhong Fang a, Jingjing Quan b, Xiaoting Hua b, Ye Feng c, Xi b, Jianfeng Wang b, Zhi Ruan d, Shiqiang Shang e, Yunsong Yu b,⁎

a The Children’s Hospital, College of Medicine, Zhejiang University, 57 Zhugan Lane, Yanan Road, Hangzhou 310003, People's Republic of China
b Department of Infectious Diseases, Sir Run Run Shaw Hospital, College of Medicine, Zhejiang University, Hangzhou, Zhejiang 310016, People's Republic of China
c Institute for Translational Medicine, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310016, People's Republic of China
d Clinical Laboratory, Sir Run Run Shaw Hospital, College of Medicine, Zhejiang University, Hangzhou, Zhejiang 310016, People's Republic of China
e Key Laboratory for Diagnosis and Therapy of Neonatal Diseases, Ministry of Education, Children’s Hospital, Zhejiang University, School of Medicine, Hangzhou 310003, People’s Republic of China

Abstract

Acinetobacter baumannii is an important bacterium that emerged as a significant nosocomial pathogen worldwide. The rise of A. baumannii was due to its multi-drug resistance (MDR), while it was difficult to treat multi-drug resistant A. baumannii with antibiotics, especially in pediatric patients for the therapeutic options with antibiotics were quite limited in pediatric patients. A. baumannii ST208 was identified as predominant sequence type of carbapenem resistant A. baumannii in the United States and China. As we knew, there was no complete genome sequence repotted for A. baumannii ST208, although several whole genome shotgun sequences had been reported. Here, we sequenced the 4087-kilobase (kb) chromosome and 112-kb plasmid of A. baumannii XH386 (ST208), which was isolated from a pediatric hospital in China. The genome of A. baumannii XH386 contained 3968 protein-coding genes and 94 RNA-only encoding genes. Genomic analysis and Minimum inhibitory concentration assay showed that A. baumannii XH386 was multi-drug resistant strain, which showed resistance to most of antibiotics, except for tigecycline. The data may be accessed via the GenBank accession number CP010779 and CP010780.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/nuccore/CP010779
http://www.ncbi.nlm.nih.gov/nuccore/CP010780

⁎ Corresponding authors.
E-mail addresses: shangsq33@sina.com (S. Shang), yvys119@163.com (Y. Yu).

2. Experimental design, materials and methods

2.1. Introduction

Acinetobacter baumannii is an important bacterium which emerged as a significant nosocomial pathogen worldwide [1]. It caused bloodstream infection, pneumonia, endocarditis and so on [2]. The rise of A. baumannii was due to its multi-drug resistance, while it was difficult to treat multi-drug resistant A. baumannii with antibiotics [3,4]. It caused by A. baumannii had a strong potential to develop antimicrobial resistance, which largely related to mobile genetic elements [5]. Carbapenem resistance in A. baumannii was mediated most by oxacillinases (OXAs) and less by metallo-β-lactamases (MBLs) [6]. Carbapenem resistance in A. baumannii was increasing worldwide, and was considered as a marker of emerging antibiotic resistance [7]. CRAB infection was also a growing problem in the pediatric population. The children were susceptible to infections while the therapeutic options with antibiotics were quite limited. However, the research focusing treatment options on CRAB infections in children was limited. The physicians were forced to use the data extrapolated from the adult literature [8].
For CRAB, sequence types (STs) belonging to the clonal complex 92 (CC92) and the pan-European clonal lineage II (EU2) were predominant in the United States. Of them, *A. baumannii* ST208 was one of the two most common STs of carbapenem-non-susceptible isolates [9]. Recently, ST 208 had been identified as predominant ST of Carbapenem Resistant *A. baumannii* (CRAB) in China [10,11]. These high prevalence of ST208 carrying *bla*OXA-23 indicated that ST 208 was an emerging lineage mediating the spread of carbapenem resistance via *bla*OXA-23 [10].

The mobility of the resistance genes was mainly mediated by insertions sequences and transposons. The complete genome would be very useful to study the horizontal transferred resistance genes. Most of *A. baumannii* strains that harbored complete genome were isolated from adult patients. *A. baumannii* strain XH386 reported in the paper was isolated from a pediatric patient. This would be helpful to understand whether there was difference between *A. baumannii* strains isolated from adult patient and pediatric patient. As we knew, there was no complete genome sequence of ST208, although several whole genome shotgun sequences had been reported [12]. Here, we present the complete genome sequence of *A. baumannii* XH386 (ST208), which was isolated from a pediatric hospital in China, together with a summary classification and a set of features.

### 3. Organism information

#### 3.1. Classification and features

*A. baumannii* XH386 is a non-fermentative, strictly aerobic, non-motile, non-pigmented, catalase-positive and oxidase-negative Gram-negative coccobacilli (Fig. 1). The strain grew on simple microbiological media optimally at –37 °C, forming smooth colonies of ~2 mm diameter. To evaluate the phylogenomic relationships between *A. baumannii* XH386 and other strain in this genus, Phylogenetic tree was generated with MEGA 6.0 using Neighbor-Joining method with 500 bootstraps and standard settings. 16S rRNA gene sequences of *Acinetobacter* spp. were derived from NCBI GenBank. The phylogenetic neighborhood of *A. baumannii* XH386 in a 16S rDNA gene sequence based tree was showed in Fig. 2.

To evaluate the phylogenomic relationships between *A. baumannii* XH386 and other strains in this species *A. baumannii*, comparisons between all the strains were calculated as percentages of similarity using Gegennes (version 2.2.1). Then, the percentage of similarity was used

**Table 1**

Summary of genome: one chromosome and one plasmid.

| Label            | Size (Mb) | Topology | INSDC identifier | RefSeq ID  |
|------------------|-----------|----------|------------------|------------|
| Chromosome 1     | 4.08      | Circular | PRJNA273343      | CP010778.1 |
| Plasmid 1        | 0.11      | Circular | PRJNA273343      | CP010780.1 |
to generate a phylogenomic tree with SplitsTree (version 4.13.1). The phylogenomic relationship in *A. baumannii* was shown in Fig 4A.

4. Genome sequencing information

4.1. Genome project history

The genome was selected based on the isolation site of the strain XH386. *A. baumannii* XH386 was a multi-drug resistant bacteria isolated from a female patient, 10Y3M, with acute bronchopneumonia in a pediatriic hospital in Hangzhu, China on May 29, 2014. The genome sequence was completed on 25 Jan., 2015. Annotation was performed by the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP).

4.2. Growth conditions and genomic DNA preparation

*A. baumannii* XH386 was cultured to mid logarithmic phase in 50 ml of LB medium at 37 °C. DNA for sequencing was extracted via a QIAamp DNA minikit (Qiagen Valencia, CA) followed the protocol of the manufacturer. The quality of DNA was determined by gel electrophoresis and NanoDrop 2000 spectrophotometer (Nano-drop Technologies, Wilmington, DE).

4.3. Genome sequencing and assembly

The genome of *A. baumannii* XH386 was sequenced at Meiji Biotechnology Company (Shanghai, China) using a hybrid of the Illumina and Pacific Biosciences (PacBio) technologies. An Illumina standard shotgun library was constructed, and then was sequenced using the Illumina HiSeq 2000 platform. 3,798,266 reads totaling 953 Mb were generated.

| Code | Value | % of total a | Description |
|------|-------|--------------|-------------|
| J    | 235   | 5.79         | Translation |
| A    | 1     | 0.02         | RNA processing and modification |
| K    | 269   | 6.62         | Transcription |
| L    | 131   | 3.23         | Replication, recombination and repair |
| B    | 0     | 0.00         | Chromatin structure and dynamics |
| D    | 39    | 0.96         | Cell cycle control, mitosis and meiosis |
| Y    | 0     | 0.00         | Nuclear structure |
| V    | 66    | 1.62         | Defense mechanisms |
| T    | 117   | 2.88         | Signal transduction mechanisms |
| M    | 186   | 4.58         | Cell wall/membrane biogenesis |
| N    | 55    | 1.35         | Cell motility |
| Z    | 0     | 0.00         | Cytoskeleton |
| W    | 3     | 0.07         | Extracellular structures |
| U    | 55    | 1.35         | Intracellular trafficking and secretion |
| O    | 121   | 2.98         | Posttranslational modification, protein turnover, chaperones |
| C    | 201   | 4.95         | Energy production and conversion |
| G    | 153   | 3.77         | Carbohydrate transport and metabolism |
| E    | 263   | 6.47         | Amino acid transport and metabolism |
| F    | 82    | 2.02         | Nucleotide transport and metabolism |
| H    | 143   | 3.52         | Coenzyme transport and metabolism |
| I    | 221   | 5.44         | Lipid transport and metabolism |
| P    | 183   | 4.51         | Inorganic ion transport and metabolism |
| Q    | 67    | 1.65         | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 238   | 5.86         | General function prediction only |
| S    | 210   | 5.17         | Function unknown |
| -    | 1023  | 25.18        | Not in COGs |

Table 3 Number of genes associated with the 25 general COG functional categories.

Table 2 Nucleotide content and gene count levels of the genome.

| Attribute | Genome (total) | Value | % of total a |
|-----------|----------------|-------|--------------|
| Genome size (bp) | 4,087,343 | 100 |
| DNA coding (bp) | 3,627,022 | 88.7 |
| DNA G + C (bp) | 1,596,791 | 39.1 |
| DNA scaffolds | 1 | 100 |
| Total genes | 4062 | 100 |
| Protein coding genes | 3908 | 97.7 |
| RNA genes | 94 | 2.3 |
| Pseudo genes | 26 | 0.6 |
| Genes in internal clusters | Not determined | Not determined |
| Genes with function prediction | 3887 | 95.7 |
| Genes assigned Pfam domains | 3209 | 74.9 |
| Genes with signal peptides | 322 | 80.4 |
| Genes with transmembrane helices | 864 | 2.3 |
| CRISPR repeats | 2 |

a The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome. Also includes 26 pseudogenes and 6 frameshifted genes.

- 1023 25.18 Not in COGs
from the standard shotgun library. A PacBio SMRTbell™ was constructed and sequenced on the PacBio RS platform. 150,292 raw PacBio reads yielded 76,398 adapter trimmed and quality filtered subreads totalling 355 Mb. De novo assembly of the read sequences was performed using continuous long reads following the Hierarchical Genome Assembly Process (HGAP) workflow (PacBio DevNet; Pacific Biosciences) as available in SMRT Analysis v2.3.0, and then Bresq v0.25b with Illumina short reads. The final assembly is based on 953 Mb of Illumina standard PE and 355 Mb of PacBio post filtered data, which provides an average 232 × Illumina coverage and 54.76 × Pacbio coverage of the genome, respectively (Table 1).

4.4. Genome annotation

Annotation of A. baumannii XH386 was finished using the NCBI PGAAP annotation pipeline and manually checked. The pipeline uses Genemark to predict open reading frames (ORF) and searches against Proteins Clusters. Protein coding genes were searched against the NCBI RefSeq database using BLASTp. COG functional categories assignment of the ORFs were archived by BLAST against the COG database. InterPro searches were also done to identify conserved domains in each ORF.

5. Genome properties

The genome of A. baumannii XH386 is 4,199,500 nucleotides 39.1% GC content and contain one 4,087,343 bp circular chromosome and one 112,157 bp circular plasmid (Fig. 3). Among of the 4062 genes, predicted 3968 were protein-coding genes, and 94 RNAs; 26 pseudogenes were also identified. The genome summary and distribution of genes into COG functional categories are listed in Tables 2 and 3.
The abundance of the resistance genes among A. baumannii strains XH386 and other strains in this species were detected by ResFinder (https://cge.cbs.dtu.dk/services/ResFinder/). The phylogenetic tree, MLST and resistance genes of A. baumannii strains was combined showed in Fig. 4A. The distribution of antibiotic resistance genes in A. baumannii XH386 was also shown in Table 4. Fig. 4B showed the distribution of resistance genes in the plasmids harbored by the A. baumannii strains. The difference of the distribution of antibiotic resistance genes between chromosome and plasmid demonstrate that the antibiotic resistance genes more often appeared in chromosome. A. baumannii XH386 was showed resistance to all antibiotics tested except tigecycline, namely tobramycin, gentamicin, levofloxacin, ciprofloxacin, cefoperazone-sulbactam, amoxicillin-clavulanic acid, piperacillin-tazobactam, ampicillin, ceftriaxone, cepime, cefoxitin, imipenem, Astreanom, Cefazolin, Nitrofurantoin, Sulfamethoxazole-trimethoprim (Table 5).

6. Insights from the genome sequence

The detection of blaOXA-23 explained the resistance to carbapenem. The existence of aac(6′)-Ib-cr, aacA4, aadA1, aph(3′)-Ic and armA showed good correlation to the resistance of tobramycin and gentamicin. A. baumannii XH386 demonstrated more resistance genes than sensitive strains, but not the other resistance ST strains, that indicated the emergence of ST208 had affected by other factors, e.g. show high fitness in clinical environment, more virulence.

7. Conclusions

A. baumannii ST208 was identified predominant ST of Carbapenem Resistant A. baumannii in the United States and China. Although several whole genome shotgun sequences of A. baumannii ST208 had been reported, there was not complete genome sequence of ST208 so far. In current study, a complete genome of A. baumannii ST208 was reported. And the genomic analysis showed that multiple antibiotic resistance genes were detected in the genome, including resistance to aminoglycoside, beta-lactam, fluoroquinolone, macrolide, sulphonamide and tetracycline. The genome sequence of A. baumannii XH386 would provide deeper insight into the molecular resistance mechanisms and it might facilitate the development of clinical research to control the antibiotic resistance in A. baumannii.
[11] M. Deng, M.H. Zhu, J.J. Li, S. Bi, Z.K. Sheng, F.S. Hu, J.J. Zhang, W. Chen, X.W. Xue, J.F. Sheng, L.J. Li, Molecular epidemiology and mechanisms of tigecycline resistance in clinical isolates of *Acinetobacter baumannii* from a Chinese University Hospital. Antimicrob. Agents Chemother. 58 (1) (2014) 297–303.

[12] H. Wen, K. Wang, Y. Liu, M. Tay, F.M. Lauro, H. Huang, H. Wu, H. Liang, Y. Ding, M. Givskov, Y. Chen, L. Yang, Population dynamics of an *Acinetobacter baumannii* clonal complex during colonization of patients. J. Clin. Microbiol. 52 (9) (2014) 3200–3208.