Genomics Data 12 (2017) 69–70

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Draft genome of *Bordetella pseudohinzii* BH370 isolated from trachea and lung tissues of a laboratory mouse

Shih Keng Loong [a], Kim-Kee Tan [a,b], Syuhaida Sulaiman [a,b], Pooi Fong Wong [c,d], Sazaly AbuBakar [a,b,*]

[a] Tropical Infectious Diseases Research & Education Centre (TIDREC), Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Department of Medical Microbiology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Animal Experimental Unit, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[d] Department of Pharmacology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 2 March 2017
Accepted 19 March 2017
Available online 21 March 2017

Keywords:
Animal pathogen
Bordetella hinzii
ICR mouse
Malaysia
Tropical infectious disease

ABSTRACT

In this study, we present the draft genome sequence of *B. pseudohinzii* BH370 recovered from the trachea and lung tissues of an ICR mouse in Malaysia. The genome consists of 4,474,040 bp with a GC content of 66.4%. Annotation using RAST algorithm displayed 5119 protein encoding and 52 RNA genes. The CRISPR-*cas* genomic sequences previously reported in *B. pseudohinzii* were identified. The nucleotide sequences of BH370 was deposited into the European Nucleotide Archive under the genome assembly accession number FPJN01000000.

Specifications	
Organism	*Bordetella pseudohinzii*
Strain	BH370
Sequencer or array type	Ion Torrent
Data format	Analyzed
Experimental factors	Microbial strain
Experimental features	Whole genome analysis of *B. pseudohinzii* BH370
Consent	N/A
Sample source location	Trachea and lung tissues of ICR mouse from the Animal Experimental Unit, Faculty of Medicine, University of Malaya, Malaysia

1. Direct link to deposited data

http://www.ebi.ac.uk/ena/data/view/FPJN01000000

2. Experimental design, materials and methods

Isolate BH370 obtained from the trachea and lung tissues of an apparently healthy ICR mouse was cultured in Mueller-Hinton broth overnight under aerobic condition at 37 °C. Bacterial genomic DNA was extracted using the Nucleospin Tissue Kit (Macherey-Nagel, Germany) according to the manufacturer's protocol. Whole genome sequencing was performed as previously described [1], with minor modifications. Briefly, the genome library preparation was carried out using the Ion Xpress™ Plus Fragment Library Kit (Thermo Fisher Scientific, USA). Genome libraries of 200-base read fragments were prepared using *E*-Gel® SizeSelect™ Agarose Gel, 2% (Thermo Fisher Scientific, USA). The sequencing template was prepared using Ion OneTouch™ 200 Template Kit V2 DL (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. Amplified Ion Sphere Particles were enriched using IonPGM Enrichment beads (Thermo Fisher Scientific, USA). Genome sequencing was undertaken using the IonTorrent PGM sequencer (Life Technologies, USA). The raw sequence reads were assembled *de novo* using SPAdes V3.1.0 [2] as implemented in Torrent Suite V5.0.0. The assembled contigs were functionally annotated with Rapid Annotation using Subsystem Technology (RAST) [3].

3. Genomic analysis

The non-classical *Bordetella* species, *B. hinzii*, has been suggested to cause respiratory infection among laboratory mice [4], and can hence interfere with studies using these animals [5]. Our recent study has suggested that a closely related species could be

* Corresponding author at: Tropical Infectious Diseases Research & Education Centre (TIDREC), Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia.
E-mail address: sazaly@um.edu.my (S. AbuBakar).

http://dx.doi.org/10.1016/j.gdata.2017.03.004

Table 1
General genome features of *Bordetella pseudohinzii* BH370.

Attribute	Chromosome
Genome size (bp)	4,474,040
GC content (%)	66.4
Contigs	390
ORFs	5119
Number of RNAs	52

responsible for some of these infections [5], including laboratory mice infections previously attributed to *B. hinzii*. The discovery of a new species of the *Bordetella* genogroup, *B. pseudohinzii*, isolated from the bronchoalveolar lavage fluid of a C57BL/6 mouse in the USA [6] led us to investigate the species identity of a *Bordetella* isolate recovered from an ICR mouse. Unlike *B. hinzii* and other members of the *Bordetella* genus, *B. pseudohinzii* is unique in that its genome contains the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas (CRISPR-associated) system [7]. Here, we describe the full genome sequencing of *B. pseudohinzii* BH370 recovered from the trachea and lung tissues of an ICR mouse.

The draft genome of *B. pseudohinzii* BH370 was 4,474,040 bp in length, comprising 390 contigs with N50 of 20,945 bp (Table 1). The GC content of the genome sequence was approximately 66.4%. A total of 5119 protein-coding genes and 52 RNAs were predicted using RAST. Out of the 5119 protein-coding genes, 54% of the translated proteins were assigned to subsystem categories according to function (Fig. 1). The CRISPR-*cas* locus containing the *cas9*, *cas1* and *cas2* genes was present in the genome sequence of isolate BH370, confirming its species identity as *B. pseudohinzii*. Isolate BH370 showed antimicrobial resistance, mainly to β-lactams [2]. This could be influenced by a number of potential virulence genes noted in the genome, including aminoglycoside modifying enzymes (3 genes), β-lactamases (2 genes), fluoroquinolone resistant genes (4 genes), multidrug resistant efflux pumps (8 genes), and the multidrug resistance protein, MarC. *In silico* analyses also found bacteriocin (7 genes), invasion genes (9 genes), and heavy metal resistant genes (22 genes) to be present in the genome.

4. Nucleotide accession number

The genome sequences generated in this study are available from the European Nucleotide Archive under the genome assembly accession number FPJN01000000.

Conflict of interest

The authors declare that we have no conflict of interest.

Acknowledgements

The sequencing was made possible using sequencing infrastructures funded by the Ministry of Higher Education - Long Term Research Grant Scheme (LRGS/TD/2011/UM/Penyakit Berjangkit) and the High Impact Research – Ministry of Higher Education Grants (H-20001-00-E000011 and E000013-20001).

References

[1] N. Hayashimoto, M. Yasuda, K. Goto, A. Takakura, T. Itoh, Study of a *Bordetella hinzii* isolate from a laboratory mouse. Comp. Med. 58 (2008) 440–446.
[2] S.K. Loong, N.H. Mahfodz, H.A. Wali, S.A. Talib, S.A. Nasrah, P.F. Wong, S. AbuBakar, Molecular and antimicrobial analyses of non-classical *Bordetella* isolated from a laboratory mouse. J. Vet. Med. Sci. 78 (2016) 715–717.
[3] T. Spilker, R. Darrah, J.J. LiPuma, Complete genome sequences of *Bordetella flabilis*, *Bordetella bronchialis*, and "*Bordetella pseudohinzii*". Genome Announc. 4 (2016) (e01132-16).
[4] Y.V. Ivanov, N. Shariat, K.B. Register, B. Linz, I. Rivera, K. Hu, E.G. Dudley, E.T. Harvill, A newly discovered *Bordetella* species carries a transcriptionally active CRISPR-Cas with a small Cas9 endonuclease. BMC Genomics 16 (2015) 863.
[5] K.K. Tan, Y.C. Tan, L.Y. Chang, K.W. Lee, S.S. Nor'e, W.Y. Yee, M.N. Mat Isa, F.L. Jafar, C.C. Hoh, S. AbuBakar, Full genome SNP-based phylogenetic analysis reveals the origin and global spread of *Brucella melitensis*. BMC Genomics 16 (2015) 93.
[6] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19 (2012) 455–477.
[7] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9 (2008) 75.

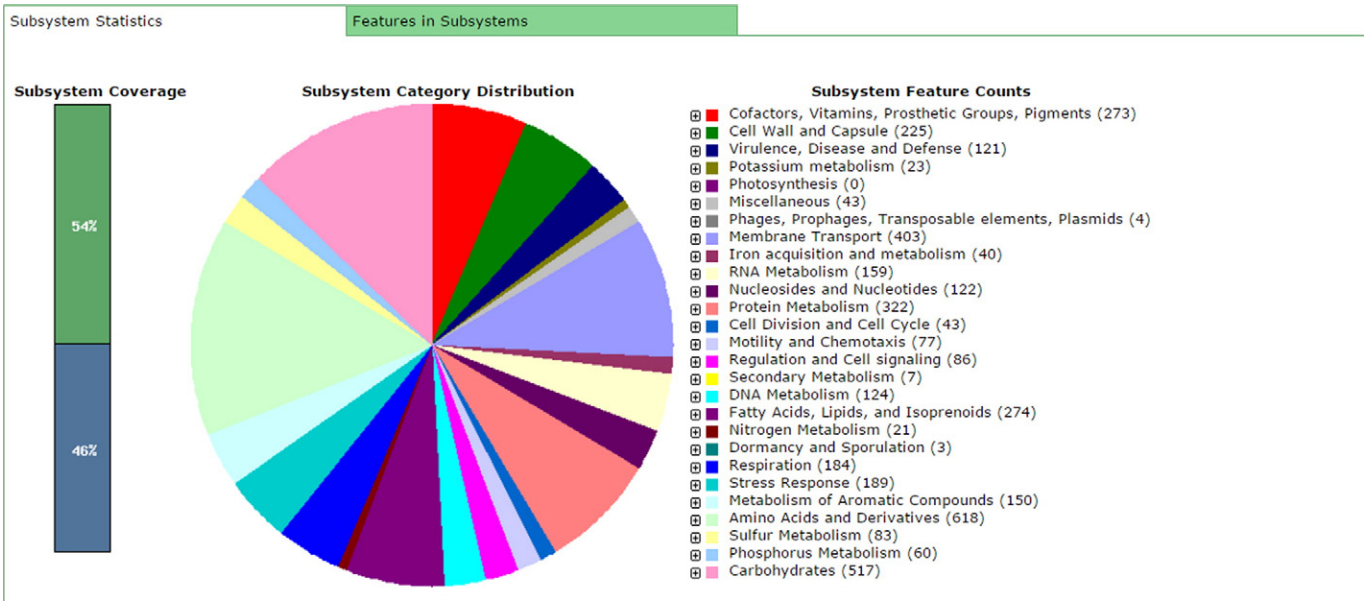

Fig. 1. Subsystem category distribution of *Bordetella pseudohinzii* BH370.