



Data in Brief

Whole genome analysis of *Klebsiella pneumoniae* T2-1-1 from human oral cavity

Kok-Gan Chan *, Wai-Fong Yin, Xin-Yue Chan

Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:

Received 30 November 2015

Accepted 15 December 2015

Available online 17 December 2015

Keywords:

Human tongue surface

Oral cavity

Oral bacteria

Virulence

ABSTRACT

Klebsiella pneumoniae T2-1-1 was isolated from the human tongue debris and subjected to whole genome sequencing on HiSeq platform and annotated on RAST. The nucleotide sequence of this genome was deposited into DDBJ/EMBL/GenBank under the accession JACL000000000.

© 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/>).

Specifications	
Organism/cell line/tissue	<i>Klebsiella pneumoniae</i>
Strain	T2-1-1
Sequencer or array type	HiSeq (Illumina)
Data format	Analyzed
Experimental factors	Microbial strain
Experimental features	Whole genome analysis and gene annotation of <i>Klebsiella pneumoniae</i> T2-1-2
Consent	N/A
Sample source location	Debris on human tongue surface

and eluted with elution buffer (Qiagen) [3]. The purity of the extracted DNA was checked with Nanodrop (Thermo Scientific) and quantified using Qubit (Invitrogen) fluorometer. Subsequently the DNA sequencing library was prepared using TruSeq DNA PCR free library preparation kit and sequenced on HiSeq 2000 (Illumina) [3]. The raw reads generated by the sequencer was paired, trimmed and *de novo* assembled with CLC Genomic Workbench (v6) [4]. The assembled genome was subjected to gene prediction using Prodigal and annotated on RAST server [5,6].

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/assembly/GCF_000710855.1/

2. Materials and methods

Klebsiella pneumoniae is one of the pathogenic bacteria responsible for major cause of nosocomial infection [1]. In this study, *K. pneumoniae* strain T2-1-1 was isolated from human oral cavity as described in Yin et al. [2]. This bacterium was purified using KG medium and maintained on Luria-Bertani medium at 37 °C. Total genomic DNA was extracted using Qiagen DNA mini kit (Qiagen)

3. Data description

Whole genome sequencing has generated 63,236,675 raw reads. The filtered reads was assembled into 166 contigs and the estimated genome size is 5,398,530 bp. The average coverage of the assembled genome is 1211 fold. The N_{50} of the genome is 59,500 bp and the G + C content of 57%. Prodigal has predicted 5114 coding sequence within the genome. These genes were annotated and classified into 573 subsystems (Fig. 1). Most of the annotated genes were involved in carbohydrate metabolism (831), amino acids and derivatives (535), cofactors, vitamins, prosthetic groups and pigments formations (323), RNA metabolism (228), cell wall and capsule development (210) (Fig. 1) which is necessary for the survival of a bacteria. The RAST results also show that this bacteria genome carries 138 potential virulence, disease and defense related genes (Fig. 1) which allows the bacteria to cause lethal diseases in human. However, in depth studies in the future is needed in order to confirm the virulence gene and invasion strategies of *K. pneumoniae* strain T2-1-1.

* Corresponding author.

E-mail address: kokgan@um.edu.my (K.-G. Chan).

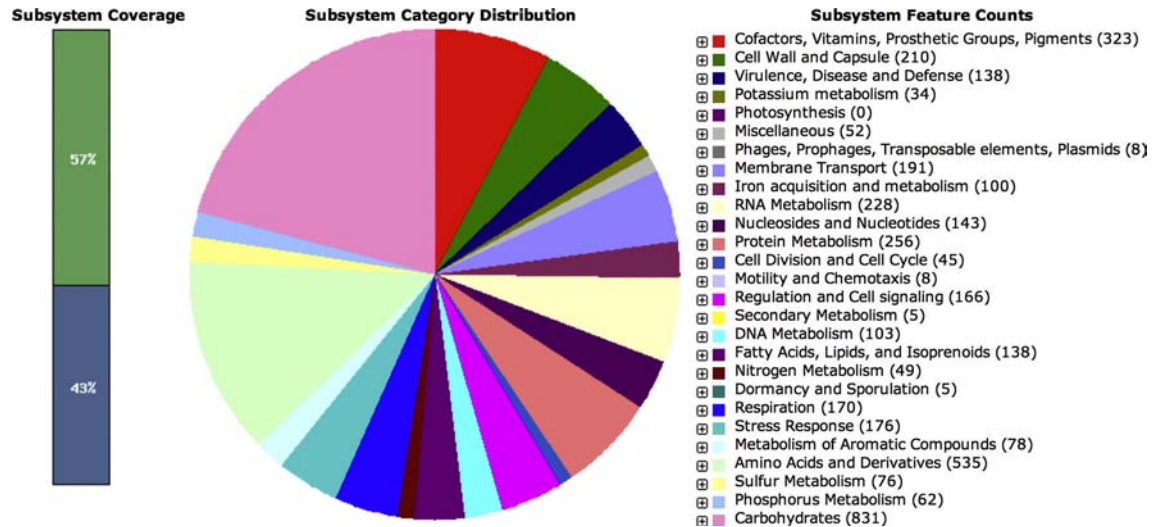


Fig. 1. Subsystems of *Klebsiella pneumoniae* strain T2-1-1 generated using RAST.

4. Nucleotide accession number

This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession no. JAQL00000000. The version described in this paper is the first version, JAQL01000000.

Conflict of interest

The authors clarified that this work and writing has no conflict of interest.

Acknowledgments

This research was financed by the University of Malaya High Impact Research (HIR) Grants (UM-MOHE HIR Grant UM.C/625/1/HIR/MOHE/CHAN/14/1, no. H-50001-A000027; UM-MOHE HIR Grant UM.C/625/1/HIR/MOHE/CHAN/01, no. A000001-50001) to Kok-Gan Chan which is gratefully acknowledged.

References

- [1] L. Tzouveleakis, A. Markogiannakis, M. Psychogiou, P. Tassios, G. Daikos, Carbapenemases in *Klebsiella pneumoniae* and other Enterobacteriaceae: an evolving crisis of global dimensions. *Clin. Microbiol. Rev.* 25 (4) (2012) 682–707.
- [2] W.F. Yin, K. Purmal, S. Chin, X.Y. Chan, C.L. Koh, C.K. Sam, K.G. Chan, *N*-acyl homoserine lactone production by *Klebsiella pneumoniae* isolated from human tongue surface. *Sensors* 12 (3) (2012) 3472–3483.
- [3] X.Y. Chan, K.H. Chua, S.D. Puthuchery, W.F. Yin, K.G. Chan, Draft genome sequence of an *Aeromonas* sp. strain 159 clinical isolate that shows quorum-sensing activity. *J. Bacteriol.* 194 (22) (2012) 6350.
- [4] H.M. Gan, T.H. Chew, A.O. Hudson, M.A. Savka, Genome sequence of *Methylobacterium* sp. strain GXF4, a xylem-associated bacterium isolated from *Vitis vinifera* L. grapevine. *J. Bacteriol.* 194 (18) (2012) 5157–5158.
- [5] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9 (1) (2008) 75.
- [6] D. Hyatt, G.L. Chen, P.F. LoCascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma.* 11 (1) (2010) 119.