Data in Brief

Genome sequencing and annotation of *Morganella* sp. SA36

Samy Selim [a,b,*], Sherif Hassan [a,c], Nashwa Hagagy [b]

[a] *Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Aljouf University, Sakaka, P.O. 2014, Saudi Arabia*
[b] *Microbiology and Botany Department, Faculty of Science, Suez Canal University, Ismailia, P.O.Box 41522, Egypt*
[c] *Department of Botany, Faculty of Science, University of Beni-Suef, Beni-Suef 62511, Egypt*

A B S T R A C T

We report draft genome sequence of *Morganella* sp. Strain SA36, isolated from water spring in Aljouf region, Saudi Arabia. The draft genome size is 2,564,439 bp with a G + C content of 51.1% and contains 6 rRNA sequence (single copies of 5S, 16S & 23S rRNA). The genome sequence can be accessed at DDBJ/EMBL/GenBank under the accession no. LDNQ00000000.

© 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	*Morganella* sp.
Strain(s)	SA36
Sequencer or array type	Sequencer; Roche 454
Data format	Processed
Experimental factors	Microbial strains
Experimental features	Draft genome sequence of *Morganella* sp. SA36 Assembly and annotation
Consent	N/A
Sample source location	Water spring in Aljouf, Saudi Arabia

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/bioproject/?term=LDNQ00000000.

2. Experimental design, materials and methods

Morganella sp. a motile gram-negative rod belonging to the family Enterobacteriaceae, has low pathogenicity, but compromised patients can develop diarrhoea, wound infections, urinary tract infections, bacteraemia, and sepsis due to *Morganella morganii* [1–2]. Genomic DNA was extracted from pure culture of bacterial strain and subsequently sequenced using Roche 454 GS (FLX Titanium) pyrosequencing.

All of the reads were assembled using GS De Novo Assembler version 2.8 (454 life science), which generated 1355 contigs with *N*50 2419 bp. The G + C content was calculated using the draft genome sequence. The G + C content for the draft genome is 51.1%. The genome contains 34 tRNA genes and 6 rRNA genes (5S-23S-16S) predicted by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP).

A total of 3229 protein coding sequences in 322 subsystems were functionally annotated by Rapid Annotation using the Subsystems Technology (RAST) [3] server (Fig. 1). Genome analysis revealed that the genome of *Morganella* sp. strain SA36 contains various gene clusters for biosynthesis of secondary metabolites and antimicrobial peptides. The draft genome was annotated using an automated annotation pipeline based on the Prodigal gene prediction algorithm [4], which predicted 2552 candidate protein-encoding gene models for strain SA36. The predicted CDSs were used to search the NCBI non-redundant database using BLAST tools [5]. Genome analysis revealed that the genome of *Morganella* sp. SA36 contains various gene clusters encoding putative virulence genes and heavy metal resistance proteins (cobalt, molybdenum, manganese and copper).

Functional comparison of genome sequences in the RAST server revealed the closest neighbours of *Providencia rustigianii* DSM 4541 (score 549) followed by *Providencia stuartii* ATCC 25827 (score 276), *Providencia alcalifaciens* DSM 30120 (score 251), *Providencia stuartii* MRSN 2154 (score 245) and *M. morganii* subsp. *morganii* KT (score 150). On the other hand, analysis of the complete 1355 contig sequence in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) under default settings (with matches only against cultured strains) *M. morganii* subsp. *morganii* KT. Overall the various in silico results confirmed that the present environmental isolate is a member of the genus

* Corresponding author at: Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Aljouf University, Sakaka, P.O. 2014, Saudi Arabia.
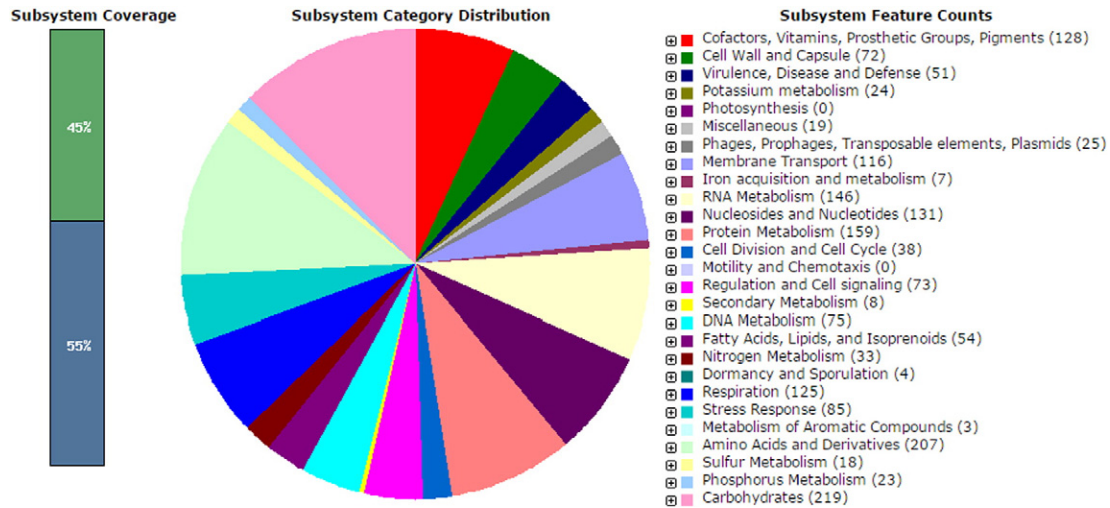*E-mail address:* sadomm2003@yahoo.com (S. Selim).

Fig. 1. Subsystem distribution of *Morganella* strain SA36 (based on RAST annotation server).

Morganella, though further characterization work is required to determine its species.

3. Nucleotide sequence accession number

The *Morganella* sp. SA36 whole genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no LDNQ00000000.

Conflict of interest

The authors declare that there is no conflict of interests on the work published in this paper.

Acknowledgements

References

[1] H.Y. Chang, S.M. Wang, N.C. Chiu, H.Y. Chung, H.K. Wang, Neonatal *Morganella morganii* sepsis: a case report and review of the literature. Pediatr. Int. 53 (2011) 121–123.

[2] V. Tucci, H.D. Isenberg, Hospital cluster epidemic with *Morganella morganii*. J. Clin. Microbiol. 14 (1981) 563–566.

[3] 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.

[4] 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 Bioinf. 11 (2010) 119, http://dx.doi.org/10.1186/1471-2105-11-119.

[5] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool. J. Mol. Biol. 215 (1990) 403–410.