



Draft genome sequence of *Paenibacillus algorifonticola* sp. nov., an antimicrobial-producing strain



Liyang Zhu^{a,1}, Qian Wu^{b,1}, Qing Xu^b, Xian Xu^b, Ling Jiang^{c,*}, He Huang^b

^a College of Sciences, Nanjing Tech University, Nanjing 210009, PR China

^b College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing 210009, PR China

^c College of Food Science and Light Industry, Nanjing Tech University, Nanjing 210009, PR China

ARTICLE INFO

Article history:

Received 1 June 2015

Received in revised form 11 June 2015

Accepted 16 June 2015

Available online 2 July 2015

Keywords:

Paenibacillus algorifonticola

Cold spring

Antimicrobial factors

Genome analysis

ABSTRACT

Paenibacillus algorifonticola sp. nov. is isolated from a cold spring sample from Xinjiang Uyghur Autonomous Region (China), a novel strain that can produce antimicrobial substance against human pathogenic bacteria and fungi, including *Staphylococcus aureus* and *Candida albicans*. Here we report a 7.60-Mb assembly of its genome sequence and other useful information, including the coding sequences (CDSs) responsible for the biosynthesis of antibacterial factors, anaerobic respiration and several immune-associated reactions. Also, prospective studies on *P. algorifonticola* sp. nov. in the cold spring might offer a potential source for the discovery of bioactive compounds with medical value. The data repository is deposited on the website <http://www.ncbi.nlm.nih.gov/nuccore/LAQO00000000> and the accession number is LAQO00000000.

© 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/tissue	<i>Paenibacillus algorifonticola</i> sp. nov.
Strain	XJ259 ^T
Sequencer or array type	Illumina Hiseq 2000
Data format	Raw and processed
Experimental factors	DNA extracted from a wild-type strain, no treatment
Experimental features	Draft genome sequencing of <i>Paenibacillus algorifonticola</i> , assembly and annotation
Consent	N/A
Sample source location	a cold spring sample from Xinjiang Uyghur Autonomous Region (China)

1. Direct link to deposited data

<http://www.ncbi.nlm.nih.gov/nuccore/LAQO00000000>

A large-scale metagenomic-based study revealed that the microbes with antibiotic resistance genes are abundant and diverse in nature [1]. However, the widespread misuse or abuse of antibiotics in clinical practice has led to a large number of antibiotic-resistant pathogenic bacteria such as *Staphylococcus aureus* and *Candida albicans*. Recently,

our group has isolated from a cold spring sample from Xinjiang Uyghur Autonomous Region (China) with a novel species of *Paenibacillus algorifonticola* sp. nov., named XJ259^T (= CGMCC 1.10223^T = JCM16598^T), which can produce water-soluble constituents with the significant inhibitory activity against both *S. aureus* and *C. albicans* [2]. Comparisons with 16S rRNA gene sequences as shown in Fig. 1 revealed that the novel strain (1519 nt) had the highest similarity to *Paenibacillus xinjiangensis* B538^T (96.6%) [3]. However, the phylogenetic distances from recognized species (Fig. 2) indicated that *P. algorifonticola* is not affiliated to any of these recognized species and the proportion of saturated straight-chain fatty acid C_{16:0} was relatively high [2]. We can therefore conclude that this strain represents a novel species of the genus *Paenibacillus*. In consequence, investigation of the genetic information and characteristics of *P. algorifonticola* is desired to further investigate its mechanism of metabolic regulation. Knowledge of the genome sequence and bioinformatics will be of great help in this regard.

Here we present the draft genome sequence of strain *P. algorifonticola* XJ259^T obtained using Illumina Hiseq 2000 system. The reads were assembled with SOAPdenovo [4,5], the version is 2.04, and the sequence was annotated using the RAST annotation server (Fig. 3) [6], and the KEGG metabolic pathway was also constructed (Fig. 4). A library containing 500-bp inserts was constructed. Sequencing was performed based on the paired-end strategy of 478-bp reads to produce 1161 Mb of filtered sequences, representing a 132-fold coverage of the genome. The sequence

* Corresponding author.

E-mail address: jiangling@njtech.edu.cn (L. Jiang).

¹ These authors contributed equally to this work.

of *P. alborfonticola* is 7,495,641 bases with a G + C content of 48.39%, which was assembled into 115 contigs and 101 scaffolds. It contains 7048 open reading frames (ORFs), 70 tRNA genes, and 4 rRNA genes (Table 1) identified by Glimmer 3.02 [7], Genemark [8], tRNAscan-SE [9], and RNAmmer [10].

According to the genomic analysis of the strain *P. alborfonticola*, we identified 6 neutrophil extracellular traps (NETs)-associated proteins, including the calcium ion binding protein and cytoplasmic

calprotectin protein complex (e.g., Mrp8/14-complex or S100A8/A9) with potent antimicrobial properties [11]. We also identified 6 ORFs relate to the immune-associated reactions, including the nucleotide-binding oligomerization domain(NOD)-like receptor signaling pathway, antigen processing and immune-responsive domain IPR011614. The strain *P. alborfonticola* can also utilize other complex immune responses that lead to production of soluble effectors, including 14 antimicrobial peptides (AMPs), or to

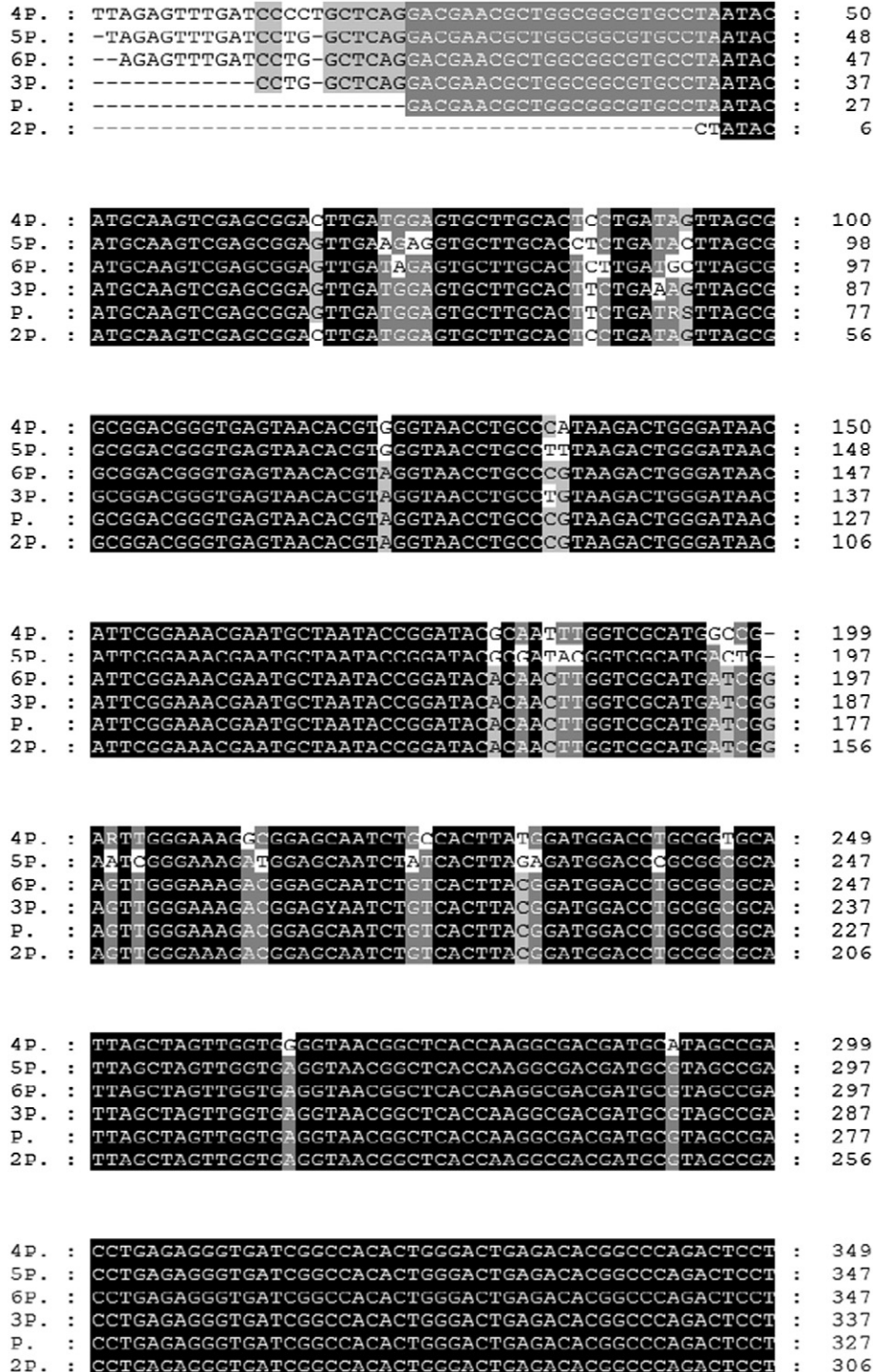


Fig. 1. Multiple sequences alignment of *Paenibacillus* (P: *Paenibacillus telluris*; 2P: *Paenibacillus* sp. LNUB461; 3P: *Paenibacillus telluris* strain JS01-08; 4P: *Paenibacillus xinjiangensis* strain B538; 5P: *Paenibacillus castaneae* strain Ch-32; 6P: *Paenibacillus alborfonticola*).

```

4P. : ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 399
5P. : ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 397
6P. : ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 397
3P. : ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 387
P. : ACGGCAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 377
2P. : ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGA : 356

4P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 449
5P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 447
6P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 447
3P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 437
P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 427
2P. : GCAACGCCCGGTGAGTGATGAAGGTTTTTCGGATCGTAAAGCTCTGTTGCC : 406

4P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAGGTGACGGTACCTGAG : 499
5P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAAGGTGACGGTACCTGAG : 497
6P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAGGTGACGGTACCTGAG : 497
3P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAGGTGACGGTACCTGAG : 487
P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAAGGTGACGGTACCTGAG : 477
2P. : AGGGAAGAAGCGCTAAGGAGAGTAAGTGCCTCCTAGGTGACGGTACCTGAG : 456

4P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 549
5P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 547
6P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 547
3P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 537
P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 527
2P. : AAGAAAGCCCCGGCTAACTACGTGCCAGCAGCCCGGGTAATACGTAGGGG : 506

4P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 599
5P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 597
6P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 597
3P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 587
P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 577
2P. : GCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCCTTG : 556

4P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 649
5P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 647
6P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 647
3P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 637
P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 627
2P. : TAAGTCTGTTGTTTAACTCGGAGCTCAACTTCGAGTCGCAATGGAACCT : 606

4P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 699
5P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 697
6P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 697
3P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 687
P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 677
2P. : GCAAAGCTTGAGTCCAGAAGAGGAAAGTGAATTCACGTTAGCGGGTGA : 656

```

Fig. 1 (continued).

activation of complement to inhibit activity of key metabolic enzymes (e.g., malate dehydrogenase and succinate dehydrogenase), which can directly damage the pathogen [12,13]. Additionally, *P. alborifonticola* have the ability to keep alive in the anaerobic condition since there were 8 ORFs related to the anaerobic dehydrogenases and anaerobic regulatory protein. Except for the biosynthesis of antimicrobial substances, the trehalose, vitamin B12 and B2 biosynthesis were also annotated in *P. alborifonticola*, further studies will be performed to confirm their functions, and a complete genome sequence will be included in the future to reveal the unique molecular characteristics of strain *P. alborifonticola*.

2. Nucleotide sequence accession numbers

This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession number LAQ00000000. The version described in this paper is the first version, with accession number LAQ001000000.

Conflict of interest

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

```

4P. : AATGCGTAGACATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 749
5P. : AATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 747
6P. : AATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 747
3P. : AATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 737
P. : AATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 727
2P. : AATGCGTAGACATGTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGGC : 706

4P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 799
5P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 797
6P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 797
3P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 787
P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 777
2P. : TGTAAGTACGCTGAGGCGCAAAGCGTGGGGAGCAAACAGGATTAGATA : 756

4P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCG : 849
5P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCA : 847
6P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCG : 847
3P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCG : 837
P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCG : 827
2P. : CCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTAGGGGTTTCG : 806

4P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 899
5P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 897
6P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 897
3P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 887
P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 877
2P. : ATACCCTTGGTGCCGAAGTTAACACATTAAGCATTCGCGCTGGGGAGTAC : 856

4P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 949
5P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 947
6P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 947
3P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 937
P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 927
2P. : GGTGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCAGT : 906

4P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 999
5P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 997
6P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 997
3P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 987
P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 977
2P. : GGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTG : 956

4P. : ACATCCCTCTGACCGTCCATAGAGATAGGACTTCCCTTCGGGACAGAGGAG : 1049
5P. : ACATCCCTCTGACCGTCCATAGAGATAGGACTTCCCTTCGGGACAGAGGAG : 1047
6P. : ACATCCCTCTGACCGGTACAGAGATGTACCTTCCCTTCGGGACAGAGGAG : 1047
3P. : ACATCCCTCTGATCCACTAGAGATAGTGGCGCCCTTCGGGACAGAGGAG : 1037
P. : ACATCCCTCTGATCCCTAGAGATAGCGGCGCCCTTCGGGACAGAGGAG : 1027
2P. : ACATCCCTCTGATCTCCATAGAGATAGCGGCGCCCTTCGGGACAGAGGAG : 1006
    
```

Fig. 1 (continued).

Acknowledgments

This work was supported by the National Science Foundation for Young Scholars of China (21225626, 21406111), the National High Technology Research and Development Program of China (2012AA022101, 2012AA021700), the National Natural Science Foundation of China for Young Scholars (21406111), the National Science Foundation of Jiangsu Province (BK20131406, BK20130917), and the Natural Science Foundation for Colleges and Universities in Jiangsu Province (14KJB530003).

References

- [1] N. Joseph, C. Se'bastien, S. Pascal, et al., Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr. Biol.* 24 (2014) 1096–1100.
- [2] Q.Y. Tang, N. Yang, J. Wang, et al., *Paenibacillus algorifonticola* sp. nov., isolated from a cold spring. *Int. J. Syst. Evol. Microbiol.* 61 (2011) 2167–2172.
- [3] J.M. Lim, C.O. Jeon, D.J. Park, et al., *Paenibacillus xinjiangensis* sp. nov., isolated from Xinjiang province in China. *Int. J. Syst. Evol. Microbiol.* 56 (2006) 2579–2582.
- [4] Li, et al., De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20 (2010) 265–272.

```

4P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1099
5P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1097
6P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1097
3P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1087
P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1077
2P. : ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGCCTGAGATGTTGGGTTAA : 1056

4P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCACTTTGGG-TG : 1148
5P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCACTTCGGG-TG : 1146
6P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCAGGTAGAGCTG : 1147
3P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCAGGTAAAGCTG : 1137
P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCAGGTAAAGCTG : 1127
2P. : GTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCAGGTAAAGCTG : 1106

4P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1198
5P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1196
6P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1197
3P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1187
P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1177
2P. : GGCACCTTAGGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC : 1156

4P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1248
5P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1246
6P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1247
3P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1237
P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1227
2P. : GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTAACAATGG : 1206

4P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1298
5P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1296
6P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1297
3P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1287
P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1277
2P. : CCGATACAACGGGAAAGCGAAACCGCGAGGTGGAGCCAATCCTATCAAAGT : 1256

4P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1348
5P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1346
6P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1347
3P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1337
P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1327
2P. : CGGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAAT : 1306

4P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1398
5P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1396
6P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1397
3P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1387
P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1377
2P. : TGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTT : 1356

```

Fig. 1 (continued).

- [5] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18 (2008) 821–829.
- [6] R.K. Aziz, D. Bartels, A.A. Best, et al., The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9 (75) (2008), <http://dx.doi.org/10.1186/1471-2164-9-75>.
- [7] A.L. Delcher, K.A. Bratke, E.C. Powers, et al., Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23 (2007) 673–679.
- [8] J. Besemer, A. Lomsadze, M. Borodovsky, GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29 (2001) 2607–2618.
- [9] K. Lagesen, P. Hallin, E.A. Rødland, et al., RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35 (2007) 3100–3108.
- [10] T.M. Lowe, S.R. Eddy, TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25 (1997) 955–964.
- [11] C.F. Urban, D. Ermert, M. Schmid, et al., Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against *Candida albicans*. *PLoS Pathog.* 5 (10) (2009) e1000639.
- [12] S. Marc, F.E. Joachim, Interplay between *Candida albicans* and the antimicrobial peptide armory. *Eukaryot. Cell* 13 (8) (2014) 950–957.
- [13] G. Mariusz, N. Joanna, F. Grzegorz, et al., Antimicrobial photodynamic therapy with fulleropyrrolidine: photoinactivation mechanism of *Staphylococcus aureus*, in vitro and in vivo studies. *Appl. Microbiol. Biotechnol.* 99 (9) (2015) 4031–4043.

```

4P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1448
5P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1446
6P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1447
3P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1437
P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1427
2P. : GTACACACCGCCCGTCACACCACGAGAGTTTACAACACCCGAAGCCGGTG : 1406

4P. : GGGTAACC-GCAAGG-AGCCAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1496
5P. : GGGTAACC-GCAAGG-GAGC-TAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1496
6P. : GGGTAACC-GCAAGG-AGCCAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1495
3P. : GGGTAACC-GCAAGG-AGCCAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1485
P. : GGGTAACC-GCAAGG-AGCCAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1475
2P. : GGGTAACC-GCAAGG-AGCCAGCCGTCGAAGGTGGGGTAGATGATTGGGG : 1454

4P. : TGAAGTCGTAACAAGGTAGCCAGGG----- : 1521
5P. : TGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCC : 1546
6P. : TGAAGTCGTAACAAGGTAGCCGTA----- : 1519
3P. : TGAAGTCGTAACAAGGTAGCCG----- : 1507
P. : TGAAGTCGTAACAAGGTAGCCGTATCGGAAGG----- : 1507
2P. : TGAAGTCG----- : 1462
    
```

Fig. 1 (continued).

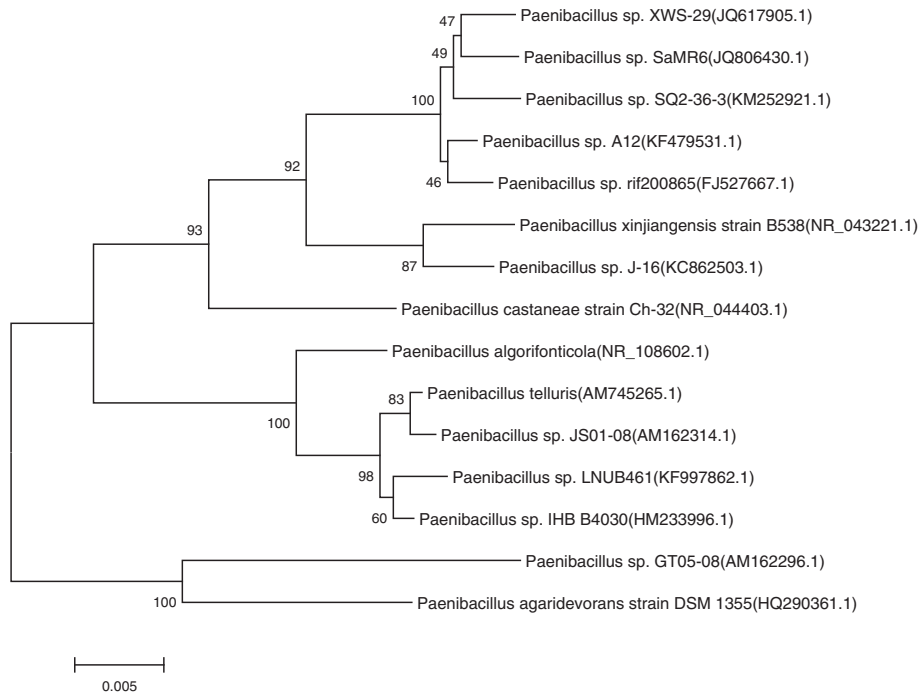


Fig. 2. Phylogenetic tree analysis of *Paenibacillus algorifonticola* in this study and other strains belonging to the genus *Paenibacillus* (MEGA 3.1).

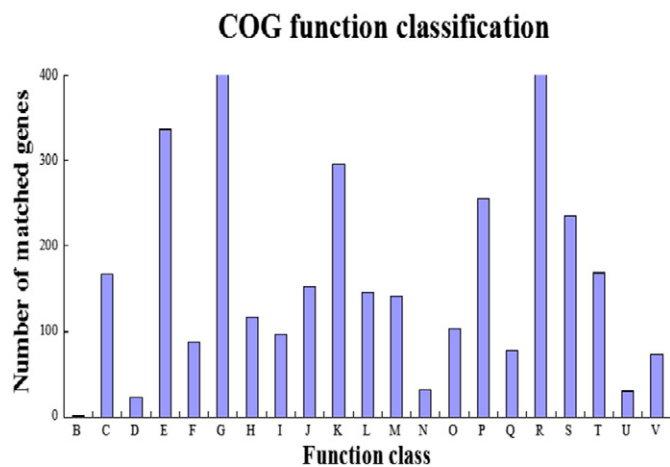


Fig. 3. COG function classification of *Paenibacillus algorifonticola* (B: Chromatin structure and dynamics; C: Energy production and conversion; D: Cell cycle control, cell division, chromosome partitioning; E: Amino acid transport and metabolism; F: Nucleotide transport and metabolism; G: Carbohydrate transport and metabolism; H: Coenzyme transport and metabolism; I: Lipid transport and metabolism; J: Translation, ribosomal structure and biogenesis; K: Transcription; L: Replication, recombination and repair; M: Cell wall/membrane/envelope biogenesis; N: Cell motility; O: Posttranslational modification, protein turnover, chaperones; P: Inorganic ion transport and metabolism; Q: Secondary metabolite biosynthesis, transport and catabolism; R: General function prediction only; S: Function unknown; T: Signal transduction mechanisms; U: Intracellular trafficking, secretion, and vesicular transport; V: Defense mechanisms).

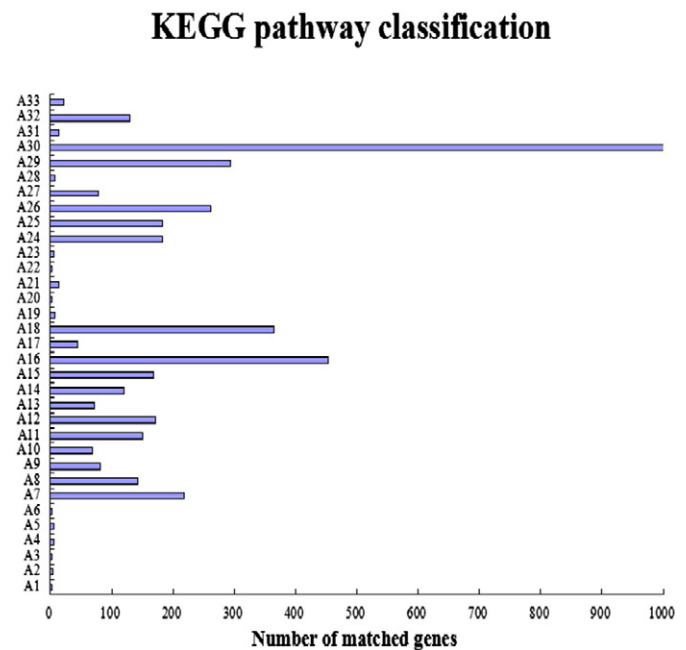


Fig. 4. KEGG metabolic pathway classification of *Paenibacillus algorifonticola* (A1: Organismal systems–Nervous system; A2: Organismal systems–Immune system; A3: Organismal systems–Excretory system; A4: Organismal systems–Environmental adaptation; A5: Organismal systems–Endocrine system; A6: Organismal systems–Digestive system; A7: Metabolism–Xenobiotics biodegradation and metabolism; A8: Metabolism–Nucleotide metabolism; A9: Metabolism–Metabolism of terpenoids and polyketides; A10: Metabolism–Metabolism of other amino acids; A11: Metabolism–Metabolism of cofactors and vitamins; A12: Metabolism–Lipid metabolisms; A13: Metabolism–Glycan biosynthesis and metabolism; A14: Metabolism–Enzyme metabolism; A15: Metabolism–Energy metabolism; A16: Metabolism–Amino acid metabolism; A17: Metabolism–Biosynthesis of other secondary metabolism; A18: Metabolism–Carbohydrate metabolism; A19: Human diseases–Neurodegenerative diseases; A20: Human diseases–Metabolic diseases; A21: Human diseases–Infectious diseases; A22: Human diseases–Immune system disease; A23: Human diseases–Cancers; A24: Genetic information processing–Translation; A25: Genetic information processing–Transcription; A26: Genetic information processing–Replication and repair; A27: Genetic information processing–Folding, sorting and degradation; A28: Environmental information processing–Signaling molecules and interaction; A29: Environmental information processing–Signal transduction; A30: Environmental information processing–Membrane transport; A31: Cellular processes–Transport and catabolism; A32: Cellular processes–Cell motility; A33: Cellular processes–Cell growth and death).

Table 1

General features of *Paenibacillus algorifonticola* sp. nov. draft genome.

Attributes	Value
Total sequence length (bp)	7,495,641
Gene number	7048
Gene length/Genome (%)	85.86
Gene average length (bp)	913
GC content (%)	48.39
N50 contig length (bp)	200,917
N90 contig length (bp)	57,078
Number of contig	115
Number of scaffolds	101
tRNAs	70
rRNAs (5S, 16S, 23S)	4
sRNA	5
Protein-coding genes	5364
Genes assigned to COGs	3376
Number of TFR	143