Draft genome sequence of *Paenibacillus* strain LK1, a phytohormone producing bacterium

Kun Liu^{1,3} · Le He¹ · Shiwei Li¹ · Fengshou Tian² · Zhongke Sun^{1,2} · Chengwei Li^{1,3}

Received: 14 November 2017 / Accepted: 8 December 2017 / Published online: 19 January 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

A high-quality draft genome sequence of a multi-antibiotic resistant strain, *Paenibacillus* strain LK1 isolated from the rhizosphere of *Auricularia auricular* is described. Strain LK1 produced several types of phytohormones, which included Indole-3-acetic acid (IAA) and Zeatin. Phylogenetic analysis of the *16S rRNA* gene of strain LK1 showed a close relationship to *P. taichungensis* and *P. pabuli* (99% gene sequence similarity). The genome size of strain LK1 was estimated to be 7.16 Mb that consisted of 82 contigs contained in 49 scaffolds. The G + C content of the genome was 45.63% and it encoded 6499 genes, 115 pseudo genes, 8 rRNAs, 33 tRNAs, and 4 ncRNAs, based on the Prokaryotic Genome Annotation Pipeline (PGAP). Further analysis of the genome of strain LK1 identified 16 genes encoding monooxygenase and six genes encoding cytochrome P450, which may be potentially involved in zeatin biosynthesis. The draft whole genome sequence and annotation of strain LK1 has been deposited at DDBJ/EMBL/GenBank under the accession number (PEII000000000).

Keywords *Paenibacillus* sp · NGS · WGS · Phytohormone · Monooxygenase

Paenibacillus spp. are ubiquitous environmental endospore-forming bacteria. Many species of this genus possess plant growth promoting properties or have been developed as biological agents against plant pathogens (Xie et al. 2016; Rybakova et al. 2016). Consequently, *Paenibacillus* spp. have been exploited widely as environmental-friendly alternatives to chemical fertilizers and pesticides (Soussi et al. 2016). Furthermore, *Paenibacillus* spp. have found wide applications in medicine, process manufacturing, cosmetics, and bioremediation as they produce antimicrobials, hydrolytic

enzymes, and exopolysaccharides (Liang and Wang 2015; Grady et al. 2016). In our previous work, we isolated strain LK1 that produced high levels of Indole-3-acetic acid (IAA) and zeatin (165 and 279 µg/mL, respectively) in cell free supernatants (unpublished data). *rpoB* sequence analysis indicated that strain LK1 had the highest sequence similarity (83%) to all members of the genus *Paenibacillus*. Further phylogenetic analysis of the *16S rRNA* gene of strain LK1 showed a very close relationship to *P. taichungensis* and *P. pabuli* (99% gene sequence similarity). As currently there is only one draft genome assembly for *P. taichungensis* and none for *P. pabuli*, we decided to sequence the genome of strain LK1 and report it here.

Paenibacillus strain LK1 was isolated from the rhizosphere of *Auricularia auricular* in China. It was cultivated overnight at 30 °C in nutrient broth with shaking (150 rpm/min). Genomic DNA was extracted using a Genomic DNA isolation kit (Tiangen, China) following the manufacturer's protocol. Genomic DNA or glycerol stocks of strain LK1 is available for academic purpose upon request from Dr. Zhongke Sun. Strain LK1 was identified by sequencing *ropB* and *16S rRNA* genes amplified with corresponding primers as described elsewhere (Sun et al. 2017). Insert libraries were prepared from genomic DNA of *Paenibacillus* strain LK1 using standard Illumina protocols and were sequenced

This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number PII000000000. The version described in this paper is version PII000000000.1.

✉ Zhongke Sun
sunzh@daad-alumni.de

✉ Chengwei Li
lichw_pgmb@sina.com

¹ College of Life Sciences and Agronomy, Zhoukou Normal University, Room 407, Life Science Building, East Wenchang Road, Chuanhui District, Zhoukou 466001, China

² Institute of Food and Drug Inspection, Zhoukou Normal University, Zhoukou 466001, China

³ Key Laboratory of Plant Genetics and Molecular Breeding, Zhoukou Normal University, Zhoukou 466001, China

Table 1 Statistic index for the genome assembly of *Paenibacillus* strain LK1

Index	No. or length (in base pair)
No. of all scaffolds	49
Bases in all scaffolds	7,164,189
No. of large scaffolds (> 1000 bp)	39
Bases in large scaffolds	7,158,792
Largest length	824,104
Scaffold N50	398,963
Scaffold N90	101,651
G + C content (%)	45.63
N rate (%)	0.00
No. of all contigs	82
No. of large contigs (> 1000 bp)	60
Bases in large contigs	7,153,668
Largest length	791,490
Contig N50	367,511
Contig N90	53,205

on an Illumina Hiseq Xten system (Illumina, USA). The quality of the raw sequences was assessed using PRINSEQ lite version 0.20.4, and adapters as well as fragments smaller than 25 bp were discarded. After removing low-quality

reads, the whole genome shotgun sequences were assembled using SOAPdenovo v2.04 with default parameters (Luo et al. 2012). GapCloser v1.12 was used to fill local inner gaps and correct miss-incorporated bases.

The G + C content of the assembled genome of strain LK1 was 45.63% and the size was estimated to be 7.16 Mbp (Table 1). The assembled genome (coverage of 280 ×) was composed of 82 contigs contained in 49 scaffolds. The sizes of scaffold N50 and contig N50 were 398,963 and 367,511 bp, respectively. Comparison of the genome of strain LK1 with the only publicly released genome of *P. tai-chungensis* (accession number GCA_002264305.1), showed that the length of the later was 7.07 Mb and was composed of 218 scaffolds, whereas the draft genome of strain LK1 is slightly larger and is of better quality. Further Average Nucleotide Identity blast (ANIb) analysis of both genomes by JSpeciesWS revealed an identity index of 90.12% and which is below the cutoff value for differentiating species (Richter et al. 2016) suggesting that *Paenibacillus* strain LK1 is a potential novel species (Colston et al. 2014).

Genes in the genome of strain LK1 were predicted using Glimmer v3.02 (Delcher et al. 2007) and annotated using blastp (BLAST 2.2.28 +) against four databases, namely, non-redundant (NR) database, Genes database from GenBank, String (<https://string-db.org/>), and Gene Ontology (<http://www.geneontology.org/>). rRNAs and

Fig. 1 Functional classification of proteins with gene ontology in *Paenibacillus* strain LK1 based on the orthologous grouping of prokaryotes

tRNAs were identified using Barrnap v0.4.2 and tRNAscan-SE v1.3.1 software, respectively. All predicted proteins were analyzed using Clusters of Orthologous Groups (COG, <http://string-db.org/>, v9.05), Kyoto Encyclopedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>), and Gene Ontology (<http://www.geneontology.org/>), through blastp, blastx/blastp, and blast2go, respectively. Automatic annotation was performed using Prokaryotic Genome Annotation Pipeline (PGAP, ver. 4.2).

Glimmer prediction yielded 6995 genes with an average gene length of 877 bp. This number is slightly higher than the 6499 genes and 6339 coding sequences (CDS) predicted by PGAP. The gene density was determined to 0.976, and the G + C content in gene regions was 46.7%. There were 8 rRNAs, 33 tRNAs, and 4 ncRNAs in the genome. Blastp of predicted genes showed that of the 3641 sequences with Gene Ontology (Fig. 1), 2803 fell into the Gene Ontology category of metabolic processes, the most predominant subsystem (76.98%, GO:0008152). A further 2246 genes were identified to be associated with the molecular function of catalytic activity. The fact that so many genes are taking part in the process of metabolism and catalysis suggests that strain LK1 may be a candidate for producing complex and diverse metabolites. Additionally, 1645 genes were predicted to be associated with molecular function of binding, which are exceptionally more than other closely related bacteria, such as *P. tai-chungensis* VTT E-133285 (Delcher et al. 2007).

As mentioned above, *Paenibacillus* strain LK1 produces phytohormones, such as IAA and zeatin, because both of which have been detected in the culture supernatant. Zeatin is a derivative of adenine and belongs to the family of cytokinins involved in a wide range of plant growth and developmental processes (Keshishian and Rashotte 2015). It was reported that the fungus *Claviceps purpurea* formed trans-zeatin through a pathway mainly catalyzed by a CK-specific cytochrome P450 monooxygenase (Hinsch et al. 2015). Therefore, we determined the numbers of genes in the genome of strain LK1 which were annotated as either cytochrome P450 or monooxygenase and identified 16, 15, 12 open reading frames (ORFs) in the genome of *Paenibacillus* strain LK1 which were annotated as putative monooxygenases using blast against the database NR, Swissprot, and String, respectively. Furthermore, another 6 genes encoding cytochrome P450 were identified using blast against the NR database. These genes may provide an important genetic insight into the biosynthesis of zeatin in *Paenibacillus* strain LK1.

Acknowledgements This work was partially supported by a grant from Zhoukou Normal University (no. ZKNU2015107) and a grant from the Education Department of Henan Province (no. 17A180017).

Compliance with ethical standards

Conflict of interest We declare no conflict of interest with respect to this paper.

References

- Colston SM, Fullmer MS, Beka L et al (2014) Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *MBio* 5(6):e02136. <https://doi.org/10.1128/mBio.02136-14>
- Delcher AL, Bratke KA, Powers EC et al (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23(6):673–679. <https://doi.org/10.1093/bioinformatics/btm009>
- Grady EN, MacDonald J, Liu L et al (2016) Current knowledge and perspectives of *Paenibacillus*: a review. *Microb Cell Fact* 15:203. <https://doi.org/10.1186/s12934-016-0603-7>
- Hinsch J, Vrabka J, Oeser B et al (2015) De novo biosynthesis of cytokinins in the biotrophic fungus *Claviceps purpurea*. *Environ Microbiol* 17(8):2935–2951. <https://doi.org/10.1111/1462-2920.12838>
- Keshishian EA, Rashotte AM (2015) Plant cytokinin signalling. *Essays Biochem* 58:13–27. <https://doi.org/10.1042/bse0580013>
- Liang TW, Wang SL (2015) Recent advances in exopolysaccharides from *Paenibacillus* spp.: production, isolation, structure, and bioactivities. *Mar Drugs* 13:1847–1863. <https://doi.org/10.3390/md13041847>
- Luo R, Liu B, Xie Y et al (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18. <https://doi.org/10.1186/2047-217X-1-18>
- Richter M, Rosselló-Móra R, Glöckner FO et al (2016) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32(6):929–931. <https://doi.org/10.1093/bioinformatics/btv681>
- Rybakova D, Cernava T, Köberl M et al (2016) Endophytes-assisted biocontrol: novel insights in ecology and the mode of action of *Paenibacillus*. *Plant Soil* 405:125–140. <https://doi.org/10.1007/s11104-015-2526-1>
- Soussi A, Ferjani R, Marasco R et al (2016) Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential. *Plant Soil* 405(1–2):357–370
- Sun Z, Liu K, Zhang J et al (2017) IAA producing *Bacillus altitudinis* alleviates iron stress in *Triticum aestivum* L. seedling by both bioleaching of iron and up-regulation of genes encoding ferritins. *Plant Soil* 419(1–2):1–11. <https://doi.org/10.1007/s11104-017-3218-9>
- Xie J, Shi H, Du Z et al (2016) Comparative genomic and functional analysis reveal conservation of plant growth promoting traits in *Paenibacillus polymyxa* and its closely related species. *Sci Rep* 6:21329. <https://doi.org/10.1038/srep21329>