AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Genome Sequence of *Enterococcus pernyi*, a Pathogenic Bacterium for the Chinese Oak Silkworm, *Antheraea pernyi*

Ying Sun,[a,b] Xisheng Li,[a,c] Guobao Wang,[b] Yong Wang,[b] Yiren Jiang,[b] Yanqun Liu,[b] Zhiguo Yu,[a] Li Qin[a,b]

College of Plant Protection, Shenyang Agriculture University, Shenyang, China[a]; College of Bioscience and Biotechnology, Shenyang Agriculture University, Shenyang, China[b]; Sericultural Research Institute of Liaoning Province, Fengcheng, China[c]

We report the draft genome assembly of *Enterococcus pernyi*. The genome sequence is 3.09 Mb in length with a G+C content of 38.35%. It covers 3,153 genes with an average length of 854 bp, and contains 65 tRNAs, 13 small RNAs, and 18 rRNAs. Moreover, it contains 9 genomic islands with an average length of 14,058 bp and 3 prophages with an average length of 37,430 bp.

Received 30 December 2015 **Accepted** 4 April 2016 **Published** 19 May 2016

Citation Sun Y, Li X, Wang G, Wang Y, Jiang Y, Liu Y, Yu Z, Qin L. 2016. Genome sequence of *Enterococcus pernyi*, a pathogenic bacterium for the Chinese oak silkworm, *Antheraea pernyi*. Genome Announc 4(3):01764-15. doi:10.1128/genomeA.01764-15.

Address correspondence to Li Qin, qinli1963@163.com, or Zhiguo Yu, yu_zhiguo2012@163.com.

The Chinese oak silkworm (*Antheraea pernyi* [Guérin-Méneville, 1855]), which belongs to *Lepidoptera*: *Saturniidae*, is the most well-known wild silkworm for insect food and silk production. Rearing of the Chinese oak silkworm has a history of about 400 years in China (1). Empty-gut disease is one of the most important diseases in *A. pernyi*, and this disease seriously affects the yield of the tussah cocoon and causes great economic losses. The pathogen of the disease was established as *Streptococcus pernyi* sp. nov., based on morphological, physiological, biochemical, and serological characteristics (2). However, *S. pernyi* sp. nov. has been reclassified and renamed as *Enterococcus pernyi*, based on phylogenic analysis of 16S rRNA and *tuf* gene sequences (3–5). Now, the pathogen of empty-gut disease in *A. pernyi* has been defined as *E. pernyi* in the taxonomy of the National Center for Biotechnology Information. The genome of *E. pernyi* was sequenced to gain a better understanding of the taxonomic status of this bacterium and provide more genomic information for further studies to prevent and cure the disease.

The complete genome sequence was determined by Illumina Solexa technology at Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). Sequence assembly was performed using SOAPdenovo version 2.04 (6). Coding sequences (CDSs) were predicted using GeneMarkS software (7) and further annotated into databases through BLASTp, including NCBInr, COG, GO, KEGG, Swiss-Prot, and TrEMBL. tRNAscan (8), RNAmmer (9), and Rfam (10) were used to predict tRNAs, rRNAs, and small RNAs, respectively. Gene islands and prophages were predicted using IslandPath-DIOMB (11) and PHAST software (12).

The genome size of *E. pernyi* is 3.09 Mb with a G+C content of 38.35%. A total of 626 Mb of clean data were generated, reaching a genome coverage depth of over 200-fold. Sequences were assembled into 23 contigs with a total length of 3,181,210 bp (largest, 603,828 bp, and smallest, 654 bp) and with an N_{50} contig size of 370,188 bp. Finally, there were a total of 9 scaffolds with a total length of 3,188,572 bp (largest, 3,086,269 bp, and smallest, 654 bp) and with an N_{50} scaffold size of 3,086,269 bp. The genome contains 3,153 CDSs with an average length of 854 bp, which represent 84.48% of the whole genome. The annotation results showed that only 224 CDSs (7.1%) were not annotated into any databases; there were 2,916, 1,537, 1,577, 1,487, 1,242, and 2,812 CDSs annotated into NCBInr, COG, GO, KEGG, Swiss-Prot, and TrEMBL, respectively. Meanwhile, 65 tRNAs, 18 rRNAs, and 13 small RNAs were identified. Furthermore, the genome contains 9 genomic islands with an average length of 14,058 bp, and contains 3 prophages with an average length of 37,430 bp. There was no clustered regularly interspaced short palindromic repeat identified in the genome.

Nucleotide sequence accession number. The whole-genome sequences of *E. pernyi* have been deposited at DDBJ/EMBL/GenBank under the accession number LPVT00000000.

ACKNOWLEDGMENTS

This work was supported by the National Modern Agriculture Industry Technology System Construction Project (silkworm and mulberry) (CARS-22), the Cultivation Plan for Youth Agricultural Science and Technology Innovative Talents of Liaoning Province (2014040), the Scientific Research Project for the Education Department of Liaoning Province (2014476), the Magnitude Science and Technology Projects of Liaoning Province, and the Liaoning Province Climbing Scholars Support Plan (1102).

FUNDING INFORMATION

This work, including the efforts of Li Qin, was funded by National Modern Agriculture Industry Technology System Construction Project (Silkworm and Mulberry) (CARS-22). This work, including the efforts of Zhiguo Yu, was funded by Liaoning Province Climbing Scholars Support Plan (1102). This work, including the efforts of Yiren Jiang, was funded by Cultivation Plan for Youth Agricultural Science and Technology Innovative Talents of Liaoning Province (2014040). This work, including the efforts of Yiren Jiang, was funded by Scientific Research Project for Education Department of Liaoning Province (2014476). This work, including the efforts of Li Qin, was funded by Magnitude Science and Technology Projects of Liaoning Province.

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. **Liu Y, Li Y, Li X, Qin L**. 2010. The origin and dispersal of the domesticated Chinese oak silkworm, *Antheraea pernyi*, in China: a reconstruction based on ancient texts. J Insect Sci **10:**180. http://dx.doi.org/10.1673/031.010.14140.
2. **Wang DS, Zhan LK, Yu XB, Zhang CF**. 1980. Studies on empty-gut disease of tussah II. Identification of the causative agent of empty-gut disease of tussah, *Streptococcus pernyi* sp. nov. Acta Microbiol Sin **20:**225–229. http://dx.doi.org/10.13343/j.cnki.wsxb.1980.03.001.
3. **Wang LL, Qin F, Song C, Zhou ZY**. 2010. Reclassification of the pathogen for empty-gut disease of Chinese oak silkworm, *Antheraea pernyi*. J Food Agric Environ **8:**156–158.
4. **Shang CF, Qin L, Zhao ZJ, Song C, Li SY, Jiang DF, Fan Q**. 2011. Identification of the pathogen for *Antheraea pernyi* empty-gut disease by using 16S rRNA gene sequence. Sci Sericulture **37:**931–936. http://dx.doi.org/10.13441/j.cnki.cykx.2011.05.026.
5. **Li XR, Xing J, Li BY, Wang P, Liu JX**. 2012. Use of tuf as a target for sequence-based identification of Gram-positive cocci of the genus *Enterococcus*, *Streptococcus*, coagulase-negative *Staphylococcus*, and *Lactococcus*. Ann Clin Microbiol Antimicrob **11:**31. http://dx.doi.org/10.1186/1476-0711-11-31.
6. **Li RQ, Zhu HM, Ruan J, Qian WB, Fang XD, Shi ZB, Li YR, Li ST, Shan G, Kristiansen K, Li SG, Yang HM, Wang J, Wang J**. 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res **20:**265–272. http://dx.doi.org/10.1101/gr.097261.109.
7. **Besemer J, Lomsadze A, Borodovsky M**. 2001. 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:**2607–2618. http://dx.doi.org/10.1093/nar/29.12.2607.
8. **Lowe TM, Eddy SR**. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25:**955–964. http://dx.doi.org/10.1093/nar/25.5.0955.
9. **Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW**. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.
10. **Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR, Bateman A**. 2009. Rfam: updates to the RNA families database. Nucleic Acids Res **37:**D136–D140. http://dx.doi.org/10.1093/nar/gkn766.
11. **Hsiao W, Wan I, Jones SJ, Brinkman FS**. 2003. IslandPath: aiding detection of genomic islands in prokaryotes. Bioinformatics **19:**418–420. http://dx.doi.org/10.1093/bioinformatics/btg004.
12. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS**. 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39:**W347–W352. http://dx.doi.org/10.1093/nar/gkr485.