High quality genome sequence of *Treponema phagedenis* KS1 isolated from bovine digital dermatitis

Hector M. Espiritu, Lovelia L. Mamuad, Su-jeong Jin, Seon-ho Kim, Sang-suk Lee and Yong-il Cho*

Department of Animal Science and Technology, Sunchon National University, Suncheon 57922, Korea

*Corresponding author
Yong-il Cho
Department of Animal Science and Technology, Sunchon National University, Suncheon 57922, Korea.
Tel: +82-61-750-3234
E-mail: ycho@scnu.ac.kr

ORCID
Hector M. Espiritu
https://orcid.org/0000-0001-9051-1995
Lovelia L. Mamuad
https://orcid.org/0000-0002-1866-0897
Su-jeong Jin
https://orcid.org/0000-0003-0224-2057
Seon-ho Kim
https://orcid.org/0000-0002-9350-1853
Sang-suk Lee
https://orcid.org/0000-0003-1540-7041
Yong-il Cho
https://orcid.org/0000-0001-7756-3416

Abstract

Treponema phagedenis KS1, a fastidious anaerobe, was isolated from a bovine digital dermatitis (BDD)-infected dairy cattle in Chungnam, Korea. Initial data indicated that *T. phagedenis* KS1 exhibited putative virulent phenotypic characteristics. This study reports the whole genome assembly and annotation of *T. phagedenis* KS1 (KCTC14157BP) to assist in the identification of putative pathogenicity related factors. The whole genome of *T. phagedenis* KS1 was sequenced using PacBio RSII and Illumina HiSeqXTen platforms. The assembled *T. phagedenis* KS1 genome comprises 16 contigs with a total size of 3,769,422 bp and an overall guanine-cytosine (GC) content of 40.03%. Annotation revealed 3,460 protein-coding genes, as well as 49 transfer RNA- and 6 ribosomal RNA-coding genes. The results of this study provide insight into the pathogenicity of *T. phagedenis* KS1.

Keywords: *Treponema phagedenis* KS1, Bovine digital dermatitis, High-quality genome assembly

Treponema phagedenis is a fastidious anaerobe recognized as a major bacterium in bovine digital dermatitis (BDD) [1], a polytreponemal foot disease that causes lameness in cattle due to painful lesions and has serious effects on both animal welfare and economic efficiency [2]. The persistence of *T. phagedenis* in the interface between affected and healthy tissues [3] and its frequent detection and abundance in BDD suggests that it is a key agent in BDD development [4,5]. The pathogenicity of *T. phagedenis* is unclear, suggesting the need for further investigation because only a limited number of studies are available. This study reports the whole genome of *T. phagedenis* KS1, which should assist in research aimed at identifying putative pathogenicity related factors in *T. phagedenis*.

T. phagedenis KS1, the first BDD-associated *Treponema* spp. strain identified in Korea, was isolated from cattle with BDD [6]. Biopsy material from a BDD lesion was cleaned, inoculated to oral *Treponema* enrichment broth (OTEB, AnaerobeSystems, Morgan Hill, CA, USA) containing 10% fetal bovine serum, rifampicin, and enrofloxacin (10 μg/mL), and then incubated anaerobically at 37℃ for 7 days. Purity was obtained by subculturing on fastidious anaerobe agar with 5% sheep blood and supplementations and growth conditions as above. The strain exhibited weak β-hemolysis, agar penetration due to its motile and chemotactic nature, and microscopic round body formation, which are rarely reported for other *T. phagedenis* strains and are considered putative virulence factors. The strain also exhibited increased resistance toward enrofloxacin and kanamycin. QIAamp DNA kit (Qiagen, Germantown, MD, USA) was used for DNA isolation performed by following the standard protocol.

Authors' contributions
Conceptualization: Lee Ss, Cho Yi.
Data curation: Espiritu HM, Mamuad LL, Jin Sj.
Formal analysis: Espiritu HM, Mamuad LL.
Investigation: Espiritu HM, Mamuad LL.
Writing - original draft: Espiritu HM, Mamuad LL.
Writing - review & editing: Kim Sh, Lee Ss, Cho Yi.

The DNA sample was submitted to Macrogen (Seoul, Korea) for sequencing with DNA condition and quantity pre-assessed. Libraries were constructed using the 20 kb SMRTbell and TruSeqNano DNA Kit for the PacBio RSII and Illumina HiSeqXTen platforms, respectively, and the results were quality checked using Qubit and Picogreen. The filtered RSII and HiSeqXTen raw reads produced 133,060 single-pass reads and 6,909,284 high-quality short-reads, respectively. De-novo assembly was done by mapping the PacBio RSII single-pass reads to seed reads using six assemblers and were compared based on total number of bases, N_{50}, number and minimum and maximum lengths of contigs. Assembly by Hierarchical Genome Assembly Process (HGAP3) was selected based on the mentioned criteria. Subsequently, the filtered Illumina high-quality short-reads were assembled to the RSII scaffolds for sequence compensation and error correction using Pilon v1.21 [7], thereby constructing the contigs more accurately. By mapping the Illumina subreads against the PacBio assembled contigs, the consensus sequence with depth of coverage data was generated. Post-assembly validations were conducted, including K-mer analysis, reads mapping, BLAST

Table 1. Genome properties and functional annotation of *Treponema phagedenis* KS1

Property	Value
Genome size (bp)	3,769,422
Contigs (n)	16
GC content (%)	40.03%
Protein-coding genes (CDS)	3,460
rRNA	6
tRNA	49
CDS with a functional category (EggNOG)	3,342
Translation, ribosomal structure and biogenesis	152
Transcription	111
Replication, recombination and repair	294
Cell cycle control, cell division, chromosome partitioning	17
Defense mechanisms	76
Signal transduction mechanisms	64
Cell wall/membrane/envelope biogenesis	121
Cell motility	45
Intracellular trafficking, secretion, and vesicular transport	22
Posttranslational modification, protein turnover, chaperones	78
Energy production and conversion	106
Carbohydrate transport and metabolism	197
Amino acid transport and metabolism	149
Nucleotide transport and metabolism	58
Coenzyme transport and metabolism	49
Lipid transport and metabolism	40
Inorganic ion transport and metabolism	121
Secondary metabolites biosynthesis, transport and catabolism	10
General function prediction only	581
Function unknown	1,073
CDS with single function	3,320
CDS with multiple functions	22
No hit	118

GC, guanine-cytosine; CDS, coding sequence ; tRNA, transfer RNA; rRNA, ribosomal RNA.
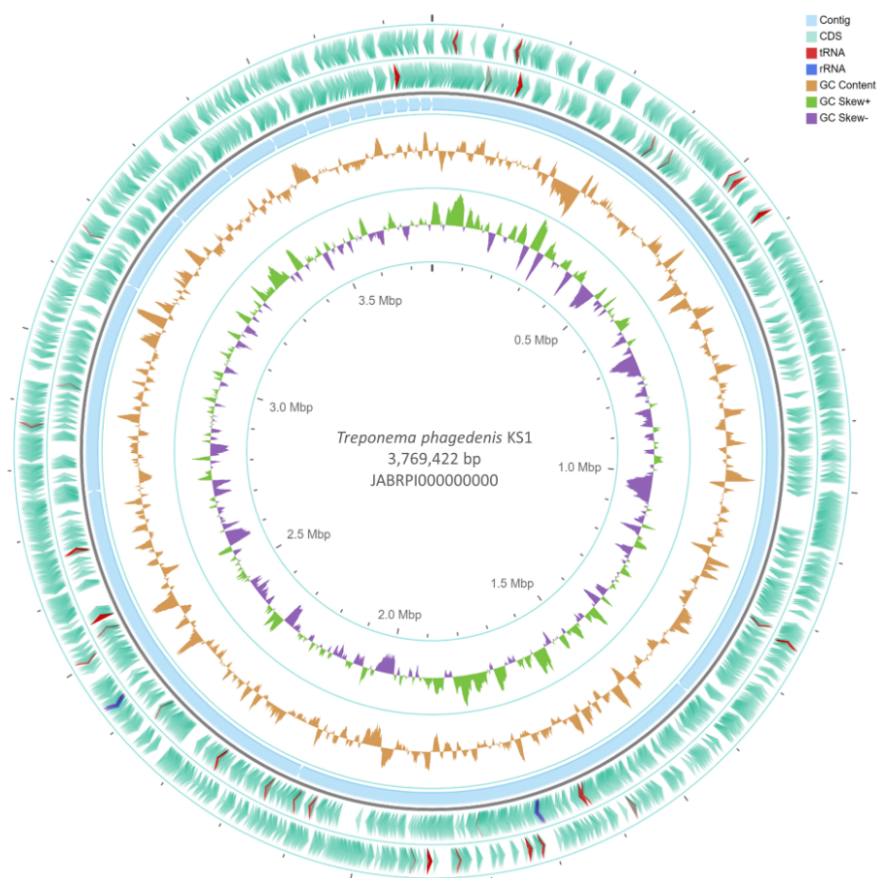
Fig. 1. Circular genome representation of *Treponema phagedenis* KS1. Figure was generated using CGView application [8]. From the outermost ring inward: CDSs, tRNA, and rRNA on forward and reverse strand, contigs map, GC ratio, and GC skew. CDS, coding sequence; CG, circular genome; GC, guanine-cytosine; tRNA, transfer RNA; rRNA, ribosomal RNA.

analysis, and BUSCO analysis. Homology with database sequences was analyzed to infer the structural annotations derived by using Prokka v1.12b [9]. Functional annotation based on ortholog identification was carried out by using EggNOG v5.0 [10].

The general properties and functional annotation of *T. phagedenis* KS1 draft genome were summarized in Table 1 and the circular genome representation was illustrated in Fig. 1. Several spirochetes are classified as pathogenic because of their capability to invade a wide range of tissues in its host. The distinct shape and motility of spirochetes modulated by its periplasmic flagellum allows the bacterium to move through tissues making it to be highly invasive in mammals [11]. Out of 3,364 coding sequence (CDS) categorized, 1.3% are under cell motility, and 1.9% were categorized under signal transduction mechanisms. These included several methyl-accepting chemotaxis proteins (*Mcp*'s), chemotaxis proteins (*Che*'s), flagellar motor switch proteins (*fli*'s), and flagellar protein (*Fla*'s), and motility protein (*motB*) for cell motility and signal transduction mechanisms, respectively. Also, 2.26% of the CDS were categorized under defense mechanism which included multidrug resistance protein (*norM*), multiple antibiotic resistance (*MarC*)-related protein, methicillin resistance (*femX*), Beta-lactamases (*flp*) and desiccation/radiation resistance gene. Other putative pathogenicity-related factors predicted include one CDS encoding hemolysin III family protein, as well as three CDS encoding hypothetical proteins for Mu-like prophage

(Flumu protein gp29) and 58 hypothetical proteins for transposases. A total of eight hypothetical proteins encoding for surface antigens with unknown functions, and three genes (*tpf1*, *smc1*, and *tpd*) encoding for surface antigenic proteins were also predicted. This report provided useful insights on the putative pathogenicity-related factors of *T. phagedenis* KS1, as well as its antigenic components which could be used for diagnosis and prevention, and treatment strategies of BDD.

Nucleotide accession number

This Whole Genome Shotgun project for *Treponema phagedenis* KS1 (KCTC14157BP) has been deposited at DDBJ/ENA/GenBank under the accession JABRPI000000000. The version described in this paper is version JABRPI010000000.

REFERENCES

1. Kuhnert P, Brodard I, Alsaaod M, Steiner A, Stoffel MH, Jores J. Treponema phagedenis (ex Noguchi 1912) Brumpt 1922 sp. nov., nom. rev., isolated from bovine digital dermatitis. Int J Syst Evol Microbiol. 2020;70:2115-23. https://doi.org/10.1099/ijsem.0.004027

2. Evans NJ, Murray RD, Carter SD. Bovine digital dermatitis: current concepts from laboratory to farm. Vet J. 2016;211:3-13. https://doi.org/10.1016/j.tvjl.2015.10.028

3. Moter A, Leist G, Rudolph R, Schrank K, Choi BK, Wagner M, et al. Fluorescence in situ hybridization shows spatial distribution of as yet uncultured treponemes in biopsies from digital dermatitis lesions. Microbiology. 1998;144:2459-67. https://doi.org/10.1099/00221287-144-9-2459

4. Mamuad LL, Seo BJ, Al Faruk S, Espiritu HM, Jeong S, Kim W, et al. Treponema spp., the dominant pathogen in the lesion of bovine digital dermatitis and its characterization in dairy cattle. Vet Microbiol. 2020;245:108696. https://doi.org/10.1016/j.vetmic.2020.108696

5. Moreira TF, Facury Filho EJ, Carvalho AU, Strube ML, Nielsen MW, Klitgaard K, et al. Pathology and bacteria related to digital dermatitis in dairy cattle in all year round grazing system in Brazil. PLOS ONE. 2018;13:e0193870. https://doi.org/10.1371/journal.pone.0193870

6. Espiritu HM, Mamuad LL, Jin S, Kim S, Kwon S, Lee S, et al. Genotypic and Phenotypic Characterization of Treponema phagedenis from Bovine Digital Dermatitis. Microorganisms. 2020;8:1520.

7. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLOS ONE. 2014;9:e112963. https://doi.org/10.1371/journal.pone.0112963

8. Grant JR, Stothard P. The CGview server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 2008;36 Suppl 2:W181-4. https://doi.org/10.1093/nar/gkn179

9. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068-9. https://doi.org/10.1093/bioinformatics/btu153

10. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019;47:D309-14. https://doi.org/10.1093/nar/gky1085

11. Wolgemuth CW. Flagellar motility of the pathogenic spirochetes. Semin Cell Dev Biol. 2015;46:104-12. https://doi.org/10.1016/j.semcdb.2015.10.015