151225

Complete genome sequence and cell structure of *Limnochorda pilosa*, a Gram-negative spore former within the phylum *Firmicutes*

Miho Watanabe[1,2 *], Hisaya Kojima[1] and Manabu Fukui[1].

[1] The Institute of Low Temperature Science, Hokkaido University, Sappro, Japan

[2] Graduate School of Environmental Science, Hokkaido University, Sappro, Japan

* Corresponding author. Tel/fax number: +81 11 706 5460.

The Institute of Low Temperature Science, Hokkaido University, Nishi 8, Kita 19, Kita-ku Sapporo, Hokkaido 060-0819, Japan.

E-mail: m.watanabe@pop.lowtem.hokudai.ac.jp

Running head: Complete genome sequence and cell structure of *Limnochorda pilosa*

Subject category: Evolution, Phylogeny and Biodiversity

The GenBank/EMBL/DDBJ accession numbers for the complete genome sequence of HC45^T is AP014924.

Abstract

Limnochorda pilosa is a pleomorphic facultative anaerobe and the sole species in the class *Limnochordia*, which has tentatively been placed in the phylum *Firmicutes*. In the present study, the complete genome sequence of *L. pilosa* HC45^T was obtained and analyzed. The genome size was 3.82 Mbp and the G+C content was 69.73%. Phylogenetic analyses based on the 30S-50S ribosomal proteins and 23S rRNA gene consistently indicated that *L. pilosa* is phylogenetically isolated from the other members of the phylum *Firmicutes*. Ultrastructural observation revealed that *L. pilosa* possesses a Gram-negative-type cell wall and the capacity to form endospores. Accordingly, the *L. pilosa* genome has characteristics that are specific to Gram-negative bacteria and contains many genes that are involved in sporulation. On the other hand, several sporulation genes were absent in *L. pilosa* genome although they have been regarded as essential for endospore-forming system of the phylum *Firmicutes*. The *gyr*B gene of *L. pilosa* possesses an intein sequence. The genome has a high percentage of GTG start codons and lacks several conserved genes related to cell division.

Introduction

The phylum *Firmicutes* has been traditionally regarded as a taxon encompassing Gram-positive endospore-forming bacteria characterized by genomes with low G+C content (Ludwig et al., 1999; Schleifer et al., 2009). *Firmicutes* is also thought to be a possible ancestral lineage of *Bacteria* (Ciccarelli et al., 2013; Koch, 2003). *Firmicutes* consists of bacteria with an enormous genetic range and contains six established classes (*Bacilli*, *Clostridia*, *Negativicutes*, *Thermolithobacteria*, *Erysipelotrichia*, *Tissierellia* and *Limnochordia*) (Garrity et al., 2005; Ludwig et al., 2009; Marchandin et al., 2010; Sokolova et al., 2007; Alauzet et al., 2014, Watanabe et al, 2015). The class *Negativicutes* was defined as a class of Gram-negative spore-forming organisms in this phylum (Marchandin et al., 2010), but phylogenetic analyses indicated that the class could be incorporated into the class *Clostridia* (Galperin, 2013; Yutin et al., 2013). The class *Limnochordia* was recently proposed to accommodate the *Limnochorda pilosa* HC45^T, which was isolated from the sediment of a meromictic lake in Japan (Watanabe et al., 2015). Analysis of the 16S rRNA gene sequence revealed that *L. pilosa* is distantly related to the genera

Symbiobacterium, *Sulfobacillus* and *Thermaerobacter*, which are facultative anaerobic Gram-positive clostridial genera (Beppu et al.,2009; da Costa et al., 2009; Spanevello et al., 2009). These 3 genera are characterized by high-G+C content (exept for sulfobacilli [46-62 mol%]), and their taxonomic affiliations of still controversial. *L. pilosa* HC45^T has some characteristics in common with these genera; the capability of facultative anaerobic growth and high G+C content (71 mol%, as estimated by high-performance liquid chromatography analysis), but it is Gram-stain-negative (Watanabe et al., 2015).

In the present study, the complete genome sequence of *L. pilosa* HC45^T was obtained as a representative of a distinct phylogenetic lineage. Using the genome data, detailed phylogenetic analyses were performed. In addition, an ultrastructural analysis was conducted to elucidate the cell wall structure and other morphological characteristics of *L. pilosa*.

Materials and methods

DNA preparation, genome sequencing and annotation

L. pilosa HC45^T was grown at 45°C for 10 days in R2A liquid medium supplemented with 2% NaCl under aerobic conditions. Genomic DNA was purified from collected cells using a Wizard® Genomic DNA Purification Kit (Promega; Madison, WI, USA). The extracted genomic DNA was sequenced at Takara Bio, Inc. (Otsu, Shiga, Japan). Library preparation (approximately 10 kb) was performed using the SMRTBell™ Template Prep Kit 1.0. Sequence reactions with four single-molecule real-time (SMRT)® cells were performed on the PacBio RS II sequencer (Pacific Biosciences; Menlo Park, CA, USA). SMRT Analysis portal version 2.2 was utilized for de novo assembly using the Hierarchical Genome Assembly Process (HGAP) (Chin et al., 2013) (PacBio DevNet; Pacific Biosciences), which is the execution program of SMRT Analysis packages, yielding a finished genome sequence.

The genome was automatically annotated using the Microbial Genome Annotation Pipeline (MiGAP) (Sugawara et al., 2009). In the pipeline, RNAmmer (Lagesen et al., 2007) and tRNAscan-SE (Lowe et al., 1997) were used to identify rRNA and tRNA genes, respectively. MetaGene Annotator (Noguchi et al., 2008) was used to predict open reading frames likely to encode proteins (coding sequences [CDSs]), and functional annotation was performed

74 based on reference databases, including Reference Sequence (RefSeq), TrEMBL, and Clusters of Orthologous
75 Groups (COG). Manual annotation was performed using IMC-GE software (In Silico Biology; Yokohama, Japan).
76 Putative CDSs were confirmed again by a sequence similarity search against the GenBank protein database using
77 the BLASTP tool. Putative CDSs possessing BLASTP matches with more than 70% coverage and 35% identity and
78 E-values less than $1 \times e^{-5}$ were considered potentially functional genes. When these standards were not satisfied,
79 the CDSs were annotated as hypothetical proteins. Transcription start sites of predicted proteins were corrected
80 based on multiple sequence alignments. Short putative CDSs (approximately <500 bp) overlapping with longer
81 CDSs or rRNAs were removed. Clustered regularly interspaced short palindromic repeat (CRISPR) loci were
82 distinguished using the CRISPR Recognition Tool (Bland et al., 2007). The complete genome sequence of *L. pilosa*
83 has been deposited into the DDBJ/EMBL/GenBank database and assigned the accession number AP014924.
84
85 Phylogenetic analyses
86 The ribosomal protein amino acid sequences and the 23S rRNA gene sequence were obtained from the *L. pilosa*
87 genome. Data of 79 bacterial genomes retrieved from the NCBI genome database were used as the source of
88 ribosomal proteins and 23S rRNA gene sequences. The reference sequences were included those of 5 classes
89 (*Clostridia*, *Bacilli*, *Negativicutes*, *Tissierellia* and *Erysipelotrichia*) of the phylum *Firmicutes*, as well as all orders
90 in the classes *Clostridia* (*Clostridiales*, *Thermoanaerobacterales*, *Natranaerobiales* and *Halanaerobiales*) and
91 *Bacilli* (*Bacillales*, *Lactobacillales*). Universal 32 ribosomal proteins in the three domains of life (S2, S3, S4, S5,
92 S7, S8, S9, S10, S11, S12, S13, S14, S15, S17, S19, L1, L2, L3, L4, L5, L6, L10, L11, L12, L13, L14, L15, L18,
93 L22, L23, L24, and L29) were selected for phylogenetic analysis (Yutin et al., 2012). Each ribosomal protein was
94 separately aligned, and the 32 alignments of the proteins were then concatenated to perform a multilocus sequence
95 analysis (MLSA). All sequences were aligned with reference sequences using the ClustalX version 2.1 program
96 (Larkin et al., 2007). Phylogenetic trees were constructed using the maximum-likelihood method, by using the
97 program *MEGA* version 5.1 (Tamura et al., 2011). Bootstrap analyses were performed with 1,000 replicates.
98

Ultrastructural analysis

Cell morphology and intracellular structures were observed by transmission electron microscopy (TEM). Cultivation and fixation of *L. pilosa* cells were performed as described previously (Watanabe et al., 2015). TEM was performed by post-fixing cell samples with osmium tetroxide, followed by dehydration via a graded ethanol series (50-100%). Samples were transferred to propylene oxide and embedded in epoxy resin. Ultrathin sections were cut with an ultramicrotome and double-stained with uranyl acetate and lead citrate. For negative-stained electron microscopy, cells were adsorbed onto a grid coated with a carbon support film and stained with uranyl acetate. Sample observation was performed using a JEM-1400 Plus transmission electron microscope (JEOL; Akishima, Tokyo, Japan).

Results

General features of the *L. pilosa* genome

Genome sequencing of *L. pilosa* HC45^{T} revealed a single circular chromosome of 3,817,036 bp with a mol% G+C content of 69.73. The genome contained 2 rRNA operons, 47 tRNAs, and 2 CRISPR loci. In the genome, 3,603 CDSs were predicted, 47% of which had ATG start codons. The second most frequent start codon was GTG (43%), and the TTG start codon was found in 10% of CDSs. The total length of CDSs was 3,387,686 bp, and the CDS density was 88%.

Phylogeny

Phylogenetic trees based on ribosomal proteins and the 23S rRNA gene sequence are shown in Fig. 1 and 2. The taxonomic clusters of the class- or order-level were supported in ribosomal protein and the 23S rRNA gene phylogeny, but the general topology was not consistent with each result of phylogenetic analysis. The phylogenetic analyses showed that *L. pilosa* is coherently related to the 3 genera *Symbiobacterium*, *Sulfobacillus* and *Thermaerobacter.*

Genomic insights into the cell wall structure

The cell wall structure has been regarded as an important criterion for classification of bacteria at higher taxonomic levels. It has been suggested that *L. pilosa* HC45^T has Gram-negative type cell wall, on the basis of a Gram stain test and a trial to analyze amino acid component of the cell wall (Watanabe et al., 2015). Accordingly, the genome of *L. pilosa* HC45^T had characteristics that are specific to Gram-negative bacteria. Lipopolysaccharide (LPS) is a structural component of the outer membrane of Gram-negative bacteria, and Kdo_2-lipid A composes an essential part of LPS. Kdo_2-lipid A is synthesized in 9 enzymatic steps (Opiyo et al., 2010), and *L. pilosa* HC45^T was found to encode genes for 4 of the enzymes (*lpxA*: LIP_3063, *lpxC*: LIP_3065, *lpxD*: LIP_3069 and *lpxB*: LIP_3061) involved in these steps. The pathway that consists of these 4 enzymes was also identified in the genomes of the phyla *Cyanobacteria* and *Dictyoglomi* (Opiyo et al., 2010). The LPS-glycosyl transferase was also conserved in the *L. pilosa* HC45^T genome, although almost none of the genes for LPS transport machinery (*lptA*, *lptC*, *lptD*, and *lptE*) were encoded, with the exception of the *lptB* (LIP_3055) gene. S-layer-associated proteins were not identified. Genes for flagella biosynthesis were also identified in the *L. pilosa* HC45^T genome, and their composition suggests that *L. pilosa* HC45^T is Gram-negative bacterium. L-ring and P-ring proteins, components of flagella specifically observed in Gram-negative bacteria (Aizawa, 2014), were conserved in the genome of *L. pilosa* (LIP_3076–3077). The flagella motor (switch complex [C-ring]) of Gram-negative bacteria consists of three components (FliG, FliM and FliN), and *L. pilosa* HC45^T encoded genes for all of them (LIP_1652, LIP_2960 and LIP_2959, respectively).

Genes involved in sporulation and cell division

In a previous study, endospore-like structures were observed in *L. pilosa* HC45^T by phase-contrast microscopy (Watanabe et al., 2015). In another study, essential sporulation genes that are conserved in the *Bacilli* and *Clostridia* classes were identified (Galperin et al., 2012). The *L. pilosa* HC45^T genome harbored many of these genes, but some important genes were not detected (Table 1). *L. pilosa* HC45^T lacked half of the genes for stage II of spore formation (post-septation) and several genes for stages III-IV (post-engulfment) that are conserved in the

spore-forming bacteria of *Bacilli* and *Clostridia*. In addition to the genes listed Table 1, *L. pilosa* HC45^T lacked the *divIVA* gene. The *divIVA* gene is widely conserved among high-G+C and low-G+C spore-forming bacteria and is involved in sporulation, cell growth and cell division.

As in the case of *divIVA*, many sporulation genes are also relevant to cell division. Upon inspection of such genes, it was found that *L. pilosa* HC45^T lacked several *fts* genes, which are widely conserved genes involved in cell division. The *ftsA*, *B*, *L*, *N* and *Q* genes were not found in the *L. pilosa* HC45^T genome, although genes for FtsZ, FtsI (SpoVD), FtsW (SpoVE) and FtsK/SpoIIIE family protein were identified (LIP_1859, LIP_1869, LIP_1865, LIP_1851, respectively). The FtsA protein might be substituted by a protein coded by the *mreB* gene (LIP_2790), which included a conserved domain of an FtsA-like protein that is essential for the membrane attachment of FtsZ. The genes *ftsQ*, *ftsB* and *ftsL* encoding divisome proteins were absent in *L. pilosa* HC45^T, and their functions are known to be substituted by the *divIB*, *divIC*, and *yllD* genes in Gram-positive bacteria, respectively [12, 48]. However, these alternative genes were also absent from the *L. pilosa* HC45^T genome.

Regarding other genes related to cell division, *L. pilosa* was found to possess the *sepF* gene (LIP_2545), which is specifically found in Gram-positive bacteria. However, *L. pilosa* HC45^T was also found to possess the genes encoding the MinCDE system (*minCDE*: LIP_2784–2786), which is thought to be specific to Gram-negative bacteria, as most spore-forming Gram-positive bacteria encode the MinCDJ/DivIVA system instead of the MinCDE system (Barak, 2013). As mentioned above, *L. pilosa* HC45^T lacked the *divIVA* gene, and the *minJ* gene was also absent from the *L. pilosa* HC45^T genome.

Intein in DNA gyrase subunit B

As a candidate for a phylogenetic marker, the sequence of the *L. pilosa* HC45^T gene encoding DNA gyrase subunit B, *gyrB* (LIP_0006), was investigated. The *gyrB* gene was found to possess an intein sequence. The intein included the sequence of the LAGLIDADG3 homing-endonuclease, and the total length of the intein sequence was 506 amino acids. GyrB inteins have been identified in bacterial phyla including *Proteobacteria* (MacGregor et al., 2013; Soucy et al., 2014), *Chloroflexi* (Kiss et al., 2011), *Actinobacteria* (Soucy et al., 2014), *Firmicutes* (Soucy et

al., 2014) and *Cyanobacteria* (Nakamura et al., 2003) and in the archaeal phylum *Euryarchaeota* (Soucy et al., 2014).

Ultrastructure of *L. pilosa* HC45^T cells

As direct evidence for a Gram-negative-type cell wall in *L. pilosa* HC45^T, TEM analysis clearly showed a three-layer structure consisting of an outer membrane, a peptidoglycan layer and a plasma membrane (Fig. 3A). The cell surface of *L. pilosa* HC45^T was shaggy in appearance (Watanabe et al., 2015), which was also observed in the TEM images (Fig. 3). Different internal structures in round and filamentous cells were observed (Fig. 3B, 3C). Cell surfaces were often irregularly hollowed (Fig. 3D). TEM analysis revealed the presence of projecting cylinders (Fig. 4A-C). A complete septum was not observed at the division site of *L. pilosa* HC45^T cells, thus suggesting a constriction mode of cell division (Fig. 5). The presence of both forespores and endospores was also confirmed (Fig. 6A and 6B). TEM analysis of the negatively stained *L. pilosa* HC45^T revealed the presence of small spherical structures (ca. 200 nm in diameter) adhering to cells (Fig. 7A) that could not be detected using other methods.

Discussion

It had been already shown that *L. pilosa* is characterized by a high G+C content and is phylogenetically separated from low-G+C Gram-positive bacteria in the phylum *Firmicutes* (Watanabe et al., 2015). In this study, ultrastructural and genetic analyses confirmed that *L. pilosa* HC45^T is a Gram-negative bacterium representing a novel independent lineage. In the phylogenetic analyses based on ribosomal proteins and the 23S rRNA gene, phylogenetic uniqueness of *L. pilosa* within the phylum *Firmicutes* was consistently indicated (Fig. 1, Fig. 2). Because of the longer informative sequence, the 23S rRNA gene is considered a better phylogenetic tool than the 16S rRNA gene (Ludwig et al., 2001; Ludwig et al., 1999; Ludwig et al., 1998; Yarza et al., 2010). The sequences of ribosomal proteins have also been utilized as useful markers in phylogenetic analyses at levels of taxa ranging from domain to species (Ciccarelli et al., 2005; Jolley et al., 2012; Yutin et al., 2013).

As a representative of a novel lineage without close relatives, *L. pilosa* HC45^T possessed several unique features.

The *L. pilosa* HC45^T genome was characterized by a high percentage of GTG start codons (43%) in the coded proteins. A majority of known bacteria primarily use the ATG start codon, and the highest prevalence of the GTG start codon usage that has been reported so far is 30-35% in mycobacterial genomes, which have high G+C content (66-67%) (Magee et al., 2011; Newton-Foot et al., 2013). The high incidence of the GTG start codon in the *L. pilosa* HC45^T genome may be related to its high G+C content (70%). Another feature of the *L. pilosa* HC45^T genome was the lack of several conserved genes involved in cell division. The ultrastructural analysis suggested that the cell division system in *L. pilosa* HC45^T resembles the Gram-negative type without a complete septum, or the so-called constriction mode (Egan et al., 2013). *L. pilosa* HC45^T may have a unique, previously unidentified cell division system, which may be related to its pleomorphy.

An enormous number of small spherical structures were observed at the negatively stained cell surface of *L. pilosa* HC45^T (Fig. 7A). Their appearances were very similar to those of external membrane vesicles (MVs), produced by organisms in all 3 domains of life (Deatherage et al., 2012). The *L. pilosa* HC45^T MVs fell within the range of Gram-negative bacterial MVs (10-300 nm in diameter) and were larger than those of Gram-positive bacteria (50-150 nm) (Deatherage et al., 2012). In Gram-negative bacteria, MVs are often produced by the faster growth of the outer membrane than the peptidoglycan layer (Bernadac et al., 1998). Accordingly, MVs derived from overproduced outer membranes were clearly observed at the cell surface of *L. pilosa* HC45^T (Fig. 7B).

L. pilosa HC45^T endospore formation was confirmed by TEM analysis. The phylogenetic expanse of endospore-forming organisms was once thought to be restricted to the classes *Bacilli* and *Clostridia*, but endospore-formers were reported in *Mycobacterium* spp. (Ghosh., 2009; Lamont et al., 2012; Singh et al., 2010) and *Streptomyces* sp. (Filippova et al., 2005) within the phylum *Actinobacteria* although disputed on the basis of genome comparisons (Traag et al, 2010; Galperin et al., 2012; Abecasis et al., 2013). Moreover, members of the phylum *Proteobacteria* were also reported to generate endospores although they are typical Gram-negatives. (Ajithkumar et al., 2003; Girija et al., 2010). As shown in Table 1, *L. pilosa* HC45^T was found to encode many genes involved in sporulation. The genes listed in Table 1 were identified by genome analysis of spore formers in the classes *Bacilli* and *Clostridia*, including the uncultivated *Candidatus* Arthromitus spp., which have a greatly

reduced gene set for regulating sporulation (Chase et al., 1976, Kuwahara et al., 2011). The list was used for the rapid screening of sporulation genes, and the absence of some genes in this list does not necessarily indicate a lack of sporulation capacity. For instance, *L. pilosa* HC45^T was found to lack the genes *spoIIE*, *spoIIGA*, *spoIIR*, *spoIIM* and *yqfD*, but the absence of these genes is a common property among actinobacterial spore-formers (Abecasis et al., 2013). However, the lack of the *divIVA* gene may be a unique property of *L. pilosa* HC45^T among spore-forming bacteria, as the gene is widely conserved in spore-forming *Firmicutes*, sporogenous *Actinobacteria* and exosporulating *Deltaproteobacteria* (Abecasis et al., 2013). *L. pilosa* HC45^T is phylogenetically far-distant from other known sporogenous bacteria, and other unidentified proteins may have roles in its endospore expression.

Conclusion

Polyphasic phylogenetic analyses showed that *L. pilosa* is assigned to an independent lineage within the phylum *Firmicutes*. These results were not in conflict with the 16S rRNA gene-based phylogeny results. Ultrastructural observation revealed that *L. pilosa* HC45^T has a Gram-negative-type cell wall and a faculty for endospore formation. The genomic information generally supported these results, but *L. pilosa* HC45^T is proposed to possess sporulation machinery that is somewhat different from that of other bacteria in the phylum *Firmicutes*.

Acknowledgement

This study was supported by a grant-in-aid for Research Fellow of Japan Society for the Promotion Science to Watanabe and JSPS KAKENHI Grant Number 22370005 to Fukui.

References

Abecasis, A.B., Serrano, M., Alves, R., Quintais, L., Pereira-Leal, J.B. & Henriques, A.O. (2013). A genomic signature and the identification of new sporulation genes. *J Bacteriol* **195**, 2101–2115.

Aizawa, S. (2014). The flagellar world: electron microscopic images of bacterial flagella and related surface structures, 1st ed. Academic Press, Elsevier Inc, Waltham, MA.

Ajithkumar, B., Ajithkumar, V.P., Iriye, R., Doi, Y. & Sakai, T. (2003). Spore-forming *Serratia marcescens* subsp. *sakuensis* subsp. nov., isolated from a domestic wastewater treatment tank. *Int J Syst Evol Microbiol* **53**, 253–258.

Alauzet, C., Marchandin, H., Courtin, P., Mory, F., Lemée, L., Pons J.-L., Chapot-Chartier, M. P., Lozniewski, A. & Jumas-Bilak, E. (2014). Multilocus analysis reveals diversity in the genus *Tissierella*: description of *Tissierella carlieri* sp. nov. in the new class *Tissierellia* classis nov. *Syst Appl Microbiol* **37**, 23–34.

Barak, I. (2013). Open questions about the function and evolution of bacterial Min systems. *Front Microbiol* **4**, 378.

Beppu, T. & Ueda, K. (2009). Family XVIII *Incertae Sedis* In: De Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 3, The *Firmicutes*, Springer, New York, pp.1188-1190.

Bernadac, A., Gavioli, M., Lazzaroni, J. C., Raina, S. & Lloubes, R. (1998). *Escherichia coli* tol-pal mutants form outer membrane vesicles. *J Bacteriol* **180**, 4872–4878.

Bland, C., Ramsey, T.L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N.C. & Hugenholtz, P. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC *Bioinformatics* **8**, 209.

Chase, D.G. & Erlandsen, S.L. (1976). Evidence for a complex life cycle and endospore formation in the attached, filamentous, segmented bacterium from murine ileum. *J Bacteriol* **127**, 572–583.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W. & Korlach, J. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10**, 563–569.

Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C,J,, Snel, B. & Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**, 1283–1287.

da Costa, M.S., Rainey, F.A. & Albuquerque, L. (2009). Family XVII *Incertae Sedis* Genus I. *Sulfobacillus* In: De Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 3, The *Firmicutes*, Springer, New York, pp.1181-1184.

Daniel, R.A., Harry, E.J., Katis, V.L., Wake, R.G. & Errington, J. (1998). Characterization of the essential cell division gene *ftsL* (*yllD*) of *Bacillus subtilis* and its role in the assembly of the division apparatus. *Mol Microbiol* **29**,593–604.

Deatherage, B.L. & Cookson, B.T. (2012). Membrane vesicle release in bacteria, eukaryotes, and archaea: A conserved yet underappreciated aspect of microbial life. *Infect Immun* **80**, 1948−1957.

Egan, A.J.F. & Vollmer, W. (2013). The physiology of bacterial cell division. *Ann NY Acad Sci* **1277**, 8–28.

Filippova, S.N., Gorbatiuk, E.V., Poglazova, M.N., Soina, V.S., Kuznetsov, V.D. & El'-Registan, G.I. (2005). Endospore formation by *Streptomyces avermitilis* in submerged culture. *Microbiology* **74**, 169-178.

Galperin, M.Y. (2013). Genomic diversity of spore-forming Firmicutes. *Microbiol Spectrum* **1(2)** TBS-0015-2012.

Galperin, M.Y., Mekhedov, S.L., Puigbo, P., Smirnov, S., Wolf, Y.I. & other authors. (2012). Genomic determinants of sporulation in *Bacilli* and *Clostridia*: towards the minimal set of sporulation-specific genes. *Environ Microbiol* **14**, 2870–2890.

Garrity, G.M., Bell, J.A. & Lilburn, T. (2005). The revised road map of the manual In: Brenner, D.J., Krieg, N.R., Staley, J.T. (eds.), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 2, The *Proteobacteria*, Part A. Springer, New York, pp.159-220.

Ghosh, J., Larsson, P., Singh, B., Pettersson, B.M., Islam, N.M., Sarkar, S.N., Dasgupta, S. & Kirsebom, L.A. (2009). Sporulation in mycobacteria. *Proc Natl Acad Sci USA* **106**, 10781-10786.

Girija, K.R., Sasikala, C., Ramana, Ch.V., Sproer, C., Takaichi, S., Thiel, V. & Imhoff, J.F. (2010). *Rhodobacter johrii* sp. nov., an endospore-producing cryptic species isolated from semi-arid tropical soils. *Int J Syst Evol Microbiol* **60**, 2099–2107.

Jolley, K.A., Bliss, C.M., Bennett, J.S., Bratcher, H.B., Brehony, C.M., Colles, F.M., Wimalarathna, H.M., Harrison, O.B., Sheppard, S.K., & other authors (2012). Ribosomal multi-locus sequence typing: universal haracterization of bacteria from domain to strain. *Microbiology* **4**, 1005–1015.

Kiss, H., Nett, M., Domin, N., Martin, K., Maresca, J.A., Copeland, A., Lapidus, A., Lucas, S., Berry, K.W. & other authors (2011). Complete genome sequence of the filamentous gliding predatory bacterium *Herpetosiphon aurantiacus* type strain (114-95^{T}). *Stand Genomic Sci* **5**, 356–370.

Koch, A.L. (2003) Were Gram-positive rods the first bacteria? *Trends Microbiol* **11**, 166.

Kuwahara, T., Ogura, Y., Oshima, K., Kurokawa, K., Ooka, T., Hirakawa, H. & other authors. (2011) The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. *DNA Res* **18**, 291–303.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T. & Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**, 3100–3108.

Lamont, E.A., Bannantine, J.P., Armien, A., Ariyakumar, D.S. & Sreevatsan, S. (2012). Identification and characterization of a spore-like morphotype in chronically starved *Mycobacterium avium* subsp. *paratuberculosis* cultures. *PLoS One* **7**, e30648.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. & other authors. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.

Lowe, T.M. & Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955–64.

Ludwig, W. & Klenk, H.P. (2001). Taxonomic outline of the *Archaea* and *Bacteria*. In: Boone, D.R., Castenholz, R.W., Garrity, G.M.(eds), Bergey's Manual of Systematic Bacteriology, 2nd edn, vol. 1, The *Archaea* and the Deeply Branching and Phototrophic *Bacteria*, Springer, New York, pp. 49-65.

Ludwig, W. & Schleifer, K.H. (1999). Phylogeny of *Bacteria* beyond the 16S rRNA standard. *ASM News* **65**, 752-757.

Ludwig, W., Schleifer, K.H. & Whitman, W.B. (2009). Revised road map to the phylum *Firmicutes*. In: De Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 3, The *Firmicutes*, Springer, New York, pp.1-14.

Ludwig, W., Strunk, O., Klugbauer, S., Klugbauer, N., Weizenegger, M., Neumaier, J., Bachleitner, M. & Schleifer, K.H. (1998). Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* **19**, 554-568.

MacGregor, B.J., Biddle, J.F. & Teske, A. (2013). Mobile elements in a single-filament orange Guaymas Basin *Beggiatoa* (*Maribeggiatoa*) sp. draft genome: evidence for genetic exchange with cyanobacteria. *Appl Environ Microbiol* **79**, 3974–3985.

Magee, J.G. & Ward, A.C. (2011). Genus I. *Mycobacterium* In: Goodfellow, M., Kampfer, P., Busse, H.J., Trujillo, M., Suzuki, K-I., Ludwig, W., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology, 2nd Edn, Vol 5, The *Actinobacteria*, part A, Springer, New York, pp. 312-375.

Marchandin, H., Teyssier, C., Campos, J., Jean-Pierre, H., Roger, F., Gay, B., Carlier, J.-P. & Jumas-Bilak, E. (2010). *Negativicoccus succinicivorans* gen. nov., sp. nov., isolated from human clinical samples, emended description of the family *Veillonellaceae* and description of *Negativicutes* classis nov., *Selenomonadales* ord. nov. and *Acidaminococcaceae* fam. nov. in the bacterial phylum *Firmicutes*. *Int J Syst Evol Microbiol* **60**, 1271–1279.

Nakamura, Y., Kaneko, T., Sato, S. & other authors. (2003). Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids, *DNA Res* **10**, 137–145.

Newton-Foot, M. & Gey van Pittius, N.C. (2013). The complex architecture of mycobacterial promoters. *Tuberculosis* **93**, 60–74.

Noguchi, H., Taniguchi, T. & Itoh, T. (2008). MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res* **15**, 387–96.

Opiyo, S.O., Pardy, R.L., Moriyama, H. & Moriyama, E.N. (2010). Evolution of the Kdo_2-lipid A biosynthesis in bacteria. *BMC Evol Biol* **10**, 362.

Parrish, N.M., Dick, J.D. & Bishai, W.R. (1998). Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol* **6**, 107–112.

Schleifer, K.H. (2009). Phylum XIII. *Firmicutes* Gibbons and Murray 1978, 5 (Firmacutes [sic] Gibbons and Murray 1978, 5) In: De Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 3, The *Firmicutes*, Springer, New York, p.19.

Singh, B., Ghosh, J., Islam, N.M., Dasgupta, S. & Kirsebom, L.A. (2010). Growth, cell division and sporulation in mycobacteria. *Anton Leeuw Int J G* **98**, 165-177.

Sokolova, T., Hanel, J., Onyenwoke, R.U., Reysenbach, A.L., Banta, A., Geyer, R., Gonzalez, J.M., Whitman, W.B. & Wiegel, J. (2007). Novel chemolithotrophic, thermophilic, anaerobic bacteria *Thermolithobacter ferrireducens* gen. nov., sp. nov. and *chermolithobacter arboxydivorans* sp. nov. *Extremophiles* **11**, 145–157.

Soucy, S.M., Fullmer, M.S., Papke, R.T. & Gogarten, J.P. (2014). Inteins as indicators of gene flow in the halobacteria. *Front Microbiol* **5**, 299.

Spanevello, M.D. & Patel, B.K.C. (2009). Family XVII *Incertae Sedis* Genus II. *Thermaerobacter* In: De Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (eds), Bergey's Manual of Systematic Bacteriology. 2nd edn, vol. 3, The *Firmicutes*, Springer, New York, pp.1184-1187.

Sugawara, H., Ohyama, A., Mori, H. & Kurokawa, K. (2009). Microbial GenomeAnnotation Pipeline (MiGAP) for diverse users. Software Demonstrations S001-1-2L. In: 20th Int. Conf. Genome Inform. (GIW2009) Poster Software Demonstrations, Yokohama, Japan.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731-2739.

Traag, B.A., Driks, A., Stragier, P., Bitter, W., Broussard, G., Hatfull, G., Chu, F., Adams, K.N., Ramakrishnan, L. & Losick R. (2010). Do mycobacteria produce endospores? *Proc Natl Acad Sci USA* **107**, 878-881.

Veiga, H. & Pinho, M.G. (2012). Bacterial cell division: what it takes to divide a prokaryotic cell. *Canal BQ* **9**, 18–26.

Watanabe, M., Kojima, H., & Fukui, M. (2015). *Limnochorda pilosa* gen. nov., sp. nov., a moderately thermophilic, facultative anaerobic pleomorphic bacterium and proposal of *Limnochordaceae* fam. nov.,

Limnochordales ord. nov. and *Limnochordia* classis nov. in the phylum *Firmicutes*. *Int J Syst Evol Microbiol* **65**, 2378-2394.

Yarza, P., Ludwig, W., Euzeby, J., Amann, R., Schleifer, K.H., Glöckner, F.O. & Rossello-Mora, R. (2010). Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses. *Syst Appl Microbiol* **33**, 291–299.

Yutin, N. & Galperin, M.Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* **15**, 2631–2641.

Yutin, N., Puigbo, P., Koonin, E.V., & Wolf, Y.I. (2012). Phylogenomics of prokaryotic ribosomal proteins. *PLoS ONE* **7** e36972.

437 **Figure legends**

438 Fig. 1 Phylogenetic tree based on amino acid sequences of 30S-50S ribosomal proteins. The tree was generated by
439 the Maximum-Likelihood method. Bootstrap values (percentages of 1000 replications) only 50% or more are
440 shown at nodes.

441

442 Fig. 2 Maximum-Likelihood tree showing the phylogenetic position of *L. pilosa* based on 23S rRNA gene sequence.
443 Bootstrap values (percentages of 1000 replications) only 50% or more are shown at nodes.

444

445 Fig. 3 Transmission electron microscope images of A: Gram-negative type cell wall; B: round vacuolated cell; low
446 electron-dense granules and D: cell surface hollows of *L. pilosa* HC45^{T}. Arrows indicate outer membrane (OM),
447 peptidoglycan layer (PG) and plasma membrane (PM). Solid triangle indicates a low electron-dense granule. Open
448 triangles indicate asymmetric invaginations.

449

450 Fig. 4 Transmission electron microscope image of A: bipolarly; B: sub-polarly and C: laterally projecting cylinders
451 of *L. pilosa* HC45^{T}. Cells were negatively-stained in image B and C.

452

453 Fig. 5 Transmission electron microscope image of division site of *L. pilosa* HC45^{T} cell. An arrow indicate the
454 division site with constriction of envelope.

455

456 Fig. 6 Transmission electron microscope image of A: forespore-like structures and B: endospore-like structure of *L.*
457 *pilosa* HC45^{T}.

458

459 Fig. 7 Transmission electron microscope images of negatively-stained cells of *L. pilosa* HC45^{T}. Image A shows
460 MVs attached to cell surface, and image B shows MVs generated from redundant outer-membrane.

461 Table 1 List of conserved or absent essential sporulation genes in *L. pilosa* genome. For details see the text.

Sporulation stage	Identified sporulation genes in *L. pilosa* genome	Absent sporulation genes in *L. pilosa* genome
Stage 0	*spo0A* (LIP_2584), *sigH* (LIP_3293), *obgE* (LIP_2777), *spo0J* (LIP_3647)	
Stage II	*spoIIAA* (LIP_1706), *spoIIAB* (LIP_1707), *sigE* (LIP_1858), *spoIID* (LIP_2799), *spoIIP* (LIP_1847)	*sigF*, *spoIIE*, *spoIIGA*, *spoIIM*, *spoIIB*
Stage III-VI	*cwlD* (LIP_1224), *dapA* (LIP_2564), *dapB* (LIP_2569), *spmA* (LIP_0753), *spmB* (LIP_0754), *spoIIIAA* (LIP_2607), *spoIIIAB* (LIP_2606), *spoIIIAC* (LIP_2605), *spoIIIAD* (LIP_2604), *spoIIIAE* (LIP_2603), *spoIIID* (LIP_3082), *spoIIIE* (LIP_1851), *spoIIIAG* (LIP_2601), *spoIIIJ* (LIP_3652), *jag* (LIP_3651), *spoIVA* (LIP_1775), *spoIVB* (LIP_2585), *sigG* (LIP_1856), *sigK* (LIP_2614), *spoVAC* (LIP_1799), *spoVAD* (LIP_1798), *spoVAEB* (LIP_1797), *spoVC* (LIP_3596), *spoVD* (LIP_1869), *spoVT* (LIP_3594), *stoA* (LIP_0685), *yabP* (LIP_3590), *yabQ* (LIP_3589), *ylbJ* (LIP_1610), *ylmC* (LIP_1854), *yqfC* (LIP_2648), *ytvl* (LIP_2524), *yyaC* (LIP_3643)	*dacB*, *spoIIIAF*, *spoIIIAH*, *spoVB*, *spoVG*, *yqfD*
Spore coat	*spoIVA* (LIP_1775), *yncD* (LIP_1264)	
Germination	*gpr* (LIP_2672), *lgt* (LIP_1407)	

462