

The Lifestyle of the Segmented Filamentous Bacterium: A Non-Culturable Gut-Associated Immunostimulating Microbe Inferred by Whole-Genome Sequencing

TOMOMI Kuwahara^{1,†}, YOSITOSHI Ogura^{2,3,†}, KENSHIRO Oshima⁴, KEN Kurokawa⁵, TADASUKE Ooka³,
HIDEKI Hirakawa⁶, TAKEHIKO Itoh⁵, HARUYUKI Nakayama-Imaohji⁷, MINORU Ichimura⁷, KIKUJI Itoh⁸,
CHIEKO Ishifune⁷, YOICHI Maekawa⁷, KOJI Yasutomo⁷, MASAHIRA Hattori⁴, and TETSUYA Hayashi^{2,3,*}

Department of Microbiology, Faculty of Medicine, Kagawa University, 1750-1 Miki-cho, Kita-gun, Kagawa 761-0793, Japan¹; Division of Bioenvironmental Science, Frontier Science Research Center, University of Miyazaki, 5200 Kiyotake, Miyazaki 889-1692, Japan²; Division of Microbiology, Department of Infectious Diseases, Faculty of Medicine, University of Miyazaki, 5200 Kiyotake, Miyazaki 889-1692, Japan³; Department of Computational Biology, Graduate School of Frontier Sciences, University of Tokyo, Chiba, Japan⁴; Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Yokohama, Kanagawa 226-0026, Japan⁵; Department of Plant Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan⁶; Department of Parasitology and Immunology, Institute of Health Biosciences, The University of Tokushima Graduate School, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan⁷ and Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-8657, Japan⁸

*To whom correspondence should be addressed. Tel. +81 985-85-0871. Fax. +81 985-85-6475.
E-mail: thayash@miyazaki-u.ac.jp

Edited by Katsumi Isono
(Received 26 May 2011; accepted 10 June 2011)

Abstract

Numerous microbes inhabit the mammalian intestinal track and strongly impact host physiology; however, our understanding of this ecosystem remains limited owing to the high complexity of the microbial community and the presence of numerous non-culturable microbes. Segmented filamentous bacteria (SFBs), which are clostridia-related Gram-positive bacteria, are among such non-culturable populations and are well known for their unique morphology and tight attachment to intestinal epithelial cells. Recent studies have revealed that SFBs play crucial roles in the post-natal maturation of gut immune function, especially the induction of Th17 lymphocytes. Here, we report the complete genome sequence of mouse SFBs. The genome, which comprises a single circular chromosome of 1 620 005 bp, lacks genes for the biosynthesis of almost all amino acids, vitamins/cofactors and nucleotides, but contains a full set of genes for sporulation/germination and, unexpectedly, for chemotaxis/flagella-based motility. These findings suggest a triphasic lifestyle of the SFB, which comprises two types of vegetative (swimming and epicellular parasitic) phases and a dormant (spore) phase. Furthermore, SFBs encode four types of flagellin, three of which are recognized by Toll-like receptor 5 and could elicit the innate immune response. Our results reveal the non-culturability, lifestyle and immunostimulation mechanisms of SFBs and provide a genetic basis for the future development of the SFB cultivation and gene-manipulation techniques.

Key words: segmented filamentous bacterium; non-culturable bacteria; genome; flagella

† These authors contributed equally to this study.

1. Introduction

The mammalian digestive tract harbours diverse microbes, with populations as high as 100 trillion organisms. Their metabolic activities, such as energy harvesting from otherwise indigestible dietary polysaccharides and the synthesis of essential vitamins, and colonization resistance against enteric pathogens are considered beneficial for the host.^{1,2} The gut microbiota play pivotal roles in the maturation and differentiation of the gut immune system. However, alteration of the gut microbial composition is often linked to various pathological conditions, such as inflammatory bowel diseases, colon cancer and obesity.^{3–7}

Our understanding of gut microbiota has been severely hampered by its extreme complexity; however, the recent advent of culture independent techniques that employ 16S rRNA gene sequencing has revealed that mammalian gut microbiota comprise several hundred or more species, and the majority of the constituents are yet to be cultured or characterized.^{8–11} Moreover, 20–80% of gut microbes are predicted to be non-culturable by current cultivation techniques. Therefore, genomic information of these non-culturable populations is essential to a more thorough understanding of the biology of gut microbiota; however, genome sequencing of non-culturable microbes is still a significant scientific challenge even with metagenome analysis or single cell sequencing using new sequencing technologies.^{8–12}

Segmented filamentous bacteria (SFBs), which are clostridia-related, spore-forming, Gram-positive bacteria, are well-known members of these non-culturable populations. SFBs inhabit the terminal ilea of a wide range of mammals.^{13,14} Their characteristic prevalence at weaning periods and unique morphology of being tightly attached to the epithelial cells around Peyer's patches strongly suggest their involvement in the maturation of the gut immune system.^{15,16} In fact, it was found that SFB colonization enhances the luminal IgA production.¹⁷ More recently, several groups have shown that SFBs play a crucial role in the post-natal maturation of gut immune function, especially in the induction of a specific subset of proinflammatory T helper lymphocytes, Th17 cells.^{18,19} Thus, more attention is being paid to SFB-derived molecules and the signal transduction pathways responsible for such immunomodulatory effects; however, no technique is available to cultivate and genetically manipulate the SFB *in vitro*.

Here, we report the genome sequence determination of mouse SFBs, the first complete genome of a non-culturable mammalian gut microbe, using SFB

cells grown in germ-free mice. An analysis of the genome sequence revealed the genomic features underlying the non-culturability of SFBs and their unique triphasic lifestyle, which would help us to establish an SFB cultivation protocol. In addition, our results are the first to provide insight into the mechanisms of the immunostimulating activities of SFBs and some of the genomic information relevant for human-associated SFBs.

2. Materials and methods

2.1. Preparation of the genomic DNA

SFB-gnotobiotic mice were prepared according to the method described by Umesaki *et al.*¹⁵ Briefly, the terminal ileum of an 8-week-old male BALB/c mouse (Clea Japan) was cut into 1-cm pieces. Mucosal epithelial cells were separated from the lamina propria (LP) as described previously.²⁰ The epithelial cells were transferred to a fresh tube containing 1 ml of phosphate-buffered saline (PBS) and homogenized using a Teflon grinder. The homogenate was treated with 3% chloroform to lyse the host epithelial cells and non-spore-forming bacteria. After removing the chloroform by bubbling with carbon dioxide, 5-week-old male germ-free BALB/c mice (Clea Japan) that had been kept in vinyl isolators were orally inoculated with the sample. At 2 weeks after inoculation, caecal contents were collected and bacterial DNA was extracted as described previously.²¹

2.2. Preparation of mouse LP lymphocytes

The small intestines were opened longitudinally, washed with PBS and incubated for 20 min at 37°C in a FACS buffer [PBS containing 10% fetal bovine serum (FBS), 10 mM EDTA, 20 mM HEPES, 1 mM sodium pyruvate, 10 µg/ml polymyxin B sulphate, 100 units/ml penicillin and 100 µg/ml streptomycin]. The intestines were then cut into small pieces and incubated with RPMI1640 containing 10% FBS, 400 units/ml collagenase D and 100 µg/ml DNase I (all from Roche Diagnostics) for 1 h at 37°C. The digested tissues were washed with the FACS buffer, resuspended in 10 ml of 40% Percoll (GE Healthcare) and overlaid on 5 ml of 80% Percoll to perform Percoll gradient separation. Interface cells [LP lymphocytes (LPLs)] were collected and suspended in PBS containing 2% FBS and 0.05% NaN₃. For IFN-γ and interleukin (IL)-17A detection, the LPLs were stimulated for 4 h with 50 ng/ml of phorbol 12-myristate 13-acetate, 1 µg/ml of ionomycin and 2 µM monensin (all from Sigma). Cells were first stained for surface CD4 and then for intracellular IFN-γ and IL-17A after fixation with a 4% paraformaldehyde phosphate buffer solution and washing with

a 0.1% saponin-containing buffer. Antibodies used in this part of the study included biotinylated anti-CD4 monoclonal antibody (mAb), FITC-labelled anti-IFN- γ mAb, PE-labelled anti-IL-17A mAb and streptavidin-PE-Cy7 (all from eBioscience). Flow cytometry analysis was performed using a FACSCanto II with the FACSDiva software (BD Biosciences).

2.3. Genome sequencing, annotation and comparative genomics

The genome sequence of SFBs was determined using a combined strategy of 454 pyrosequencing (GS FLX; Roche) and Sanger sequencing. A total of 234 915 GS FLX reads were assembled with the GS Assembler software into 242 large contigs (≥ 500 bp). These contigs were then reassembled using the Phred/Phrap/Consed software (<http://www.phrap.com>) with 25 257 Sanger sequencing reads from plasmid libraries (3 and 10 kb inserts). Gap closing and the re-sequencing of low-quality regions were performed by sequencing the PCR products and the appropriate plasmid clones. The genome sequence was automatically annotated with the Microbial Genome Annotation Pipeline²² and manually curated using the IMC-GE software (In Silico Biology, Inc.). The genome sequence of SFBs has been deposited in the DDBJ/EMBL/GenBank databases under accession number AP012202.

The SFB gene contents were compared with those of *Clostridium difficile* strain 630 (CD) and *C. novyi* strain NT (CN). The orthologous protein coding sequences (CDSs or CDS groups) for CD and CN were taken from the MGD (<http://mbgd.genome.ad.jp/>). On the basis of the results of a reciprocal BLASTP search between SFB and CD and between SFB and CN (threshold; $\geq 25\%$ sequence identity and $\geq 60\%$ aligned length coverage of a query sequence), we categorized the CDSs (or CDS groups) from the three genomes into groups of 'conserved in all', 'conserved in SFB and CD but not in CN', 'conserved in SFB and CN but not in CD', 'conserved in CD and CN but not in SFB' and those unique to each species.

2.4. Analysis of Toll-like receptor 5-linked NF- κ B signaling pathway activation

Four flagellin genes were PCR-amplified and cloned into pGEX-6P-1 (GE Healthcare) to generate glutathione-S-transferase (GST)-flagellin fusion genes. GST-fusion proteins were expressed in *Escherichia coli* BL21 via incubation with 1 mM IPTG at 20°C overnight and purified with Sepharose 4B GST resin (GE Healthcare). The Toll-like receptor 5 (TLR5)-stimulating activities of SFB flagellins were determined by a luciferase reporter assay system as described previously.²³ Briefly, the 293T cells were transfected with pNF- κ B-Luc (Clontech) and pRL-TK (Promega)

in addition to the mouse TLR5-encoding pGA vector or empty vector. pNF- κ B-Luc contains the firefly luciferase gene, expression of which is controlled by the upstream NF- κ B response elements, and pRL-TK contains the *Renilla* luciferase gene, which is controlled by the promoter of the thymidine kinase. At 20 h after transfection, the cells were stimulated with 1 μ g/ml of recombinant *Salmonella* flagellin (IMGEX Corp.), GST or GST-fusion SFB flagellins for 24 h. After stimulation, the cells were harvested and the luciferase activities in the cell lysates were measured by the Dual Luciferase Reporter Assay System (Promega) using a Lumat LB9507 (Berthold). Luciferase activity was expressed as a ratio of the NF- κ B-dependent firefly luciferase activity divided by the control *Renilla* luciferase activity (relative luciferase units).

2.5. The search for human SFB-derived sequences in human gut metagenome sequences

From the Illumina sequence reads produced in the MetaHIT project,¹¹ we removed reads containing one or more Ns. The remaining 7 345 361 234 reads were trimmed for low-quality sequences and used as queries in a BLAST search against the SFB genome sequence, 1334 RefSeq complete microbial genome sequences (<http://www.ncbi.nlm.nih.gov/RefSeq/>) and 16 276 contig sequences produced by the Human Microbiome Project (<http://www.hmpdacc.org/>). If a query sequence returned a top hit to the mouse SFB genome with a $\geq 95\%$ sequence identity for $\geq 95\%$ of the query sequence and the difference in the BLAST bit score between the top (to SFB) and second hits was larger than zero, we defined it as a human SFB-related sequence. In the 277 reads we identified, BLASTN bit score differences between the top and second hits ranged from 2.0 to 121.0 (hits for only SFB) and the median value was 11.9, indicating that the 277 sequences showed a significantly higher sequence identity to SFBs than to other known genomes.

2.6. Ethics

All animal experiments were performed at the University of Tokushima according to the university's guidelines for animal experiments. The experimental designs were approved by the university animal committee.

3. Results and discussion

3.1. Genome sequencing

To obtain genomic DNA for sequencing, we generated SFB-associated gnotobiotic mice by orally inoculating SFBs into germ-free BALB/c mice. The SFB cells used for inoculation (mostly in the endospore form)

were obtained from the ileal mucosa of an 8-week-old male BALB/c mouse. Monoassociation was confirmed by Gram-staining and sequencing of the 16S rRNA gene. For the latter, a chromatography comprising single peaks was detected. The SFB cells in the gnotobiotic mice had a segmented filamentous morphology and were tightly attached to the gut epithelia with a deep indentation of the epithelial cell surface where the SFB was attached (Fig. 1A). We further confirmed that the colonization of SFBs in the germ-free mouse intestine induced Th17 and Th1 subsets in the gut LP CD4⁺ lymphocyte preparation as reported previously^{18,19}; the IL-17-producing Th17 population increased from 0.6% (germ-free mouse) to 8.4% (SFB-gnotobiotic mouse) and the IFN- γ -producing Th1 population from 20.5 to 37.0% (Fig. 1B).

The genomic DNA of SFBs was extracted from the caecal contents of SFB-gnotobiotic mice and used for genome sequencing via a combination of 454 pyrosequencing and Sanger sequencing. Because it is currently impossible to generate SFB-gnotobiotic mice from a single SFB cell, the DNA sample contained DNA molecules derived from multiple SFB clones. In fact, we observed many single-nucleotide polymorphisms and

small indels. Therefore, when sequence polymorphisms were observed, we regarded the most frequently appearing sequence as representative of the SFB genome sequence.

3.2. General genome features

The genome of SFBs comprises a single circular chromosome of 1 620 005 bp with an average 28.8% GC content (Table 1). The chromosome encodes 1491 CDSs, 6 rRNA (*rrn*) operons, 37 tRNA genes and only 24 pseudogenes. Of the 1491 CDSs, 994 (66.7%) were functionally assigned. Like other low GC content Gram-positive bacteria, the SFB chromosome exhibits clear GC skew transitions at the replication origin and terminus and a strong coding bias with 82.4% of the CDSs being encoded on the leading strand (Fig. 2). Chromosome regions showing GC bias anomalies correspond to *rrn* or prophage regions.

The SFB contains no IS elements but four prophages (Fig. 2). Among the four prophages, one (SFBMP01) is apparently intact, whereas the others (SFBMP02~SFBMP04) are highly degraded phage remnants and SFBMP02 is similar to a part of SFBMP01 (Supplementary Fig. S1). The presence of multiple prophages suggests that phages are the main driving force for horizontal gene transfer between SFBs and other microbes. Frequent exposure to invading foreign DNAs is also suggested by the presence of three loci of clustered, regularly interspaced and short palindromic repeats (CRISPRs) in the SFB genome (Supplementary Fig. S1). CRISPRs are involved in a recently discovered microbial adaptive immunity system that protects microbial cells from invading foreign genetic elements, such as bacteriophages and conjugative plasmids.^{24,25} CRISPR loci typically consist of several non-contiguous direct repeats separated by stretches of variable sequences called spacers, which correspond to segments of captured phage and plasmid sequences. The CRISPR sequences thus provide an adaptive, heritable record of past infections and express small RNAs that

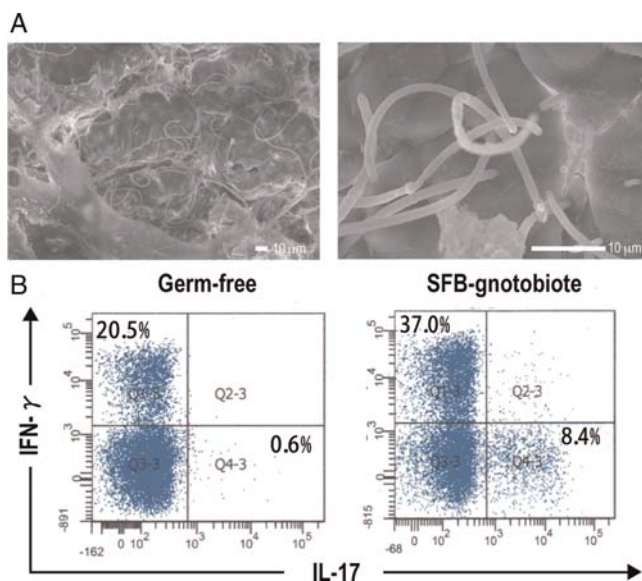


Figure 1. Morphology and immunostimulatory activity of SFBs grown in SFB-monoassociated gnotobiotic mice. (A) Electron microscopic examination of SFBs. BALB/c germ-free mice were colonized with the SFB, and the terminal ilea collected 14 days after inoculation were examined by scanning electron microscopy. SFBs show a characteristic long filamentous morphology and preferentially colonized around Peyer's patches. Epithelial cell surfaces to which SFB attach are deeply indented. (B) Induction of Th1 and Th17 cells by SFB colonization. Intestinal LPLs were isolated from BALB/c germ-free mice or SFB-monoassociated gnotobiotic mice and the percentages of IFN- γ ⁺ or IL-17⁺ cells within the CD4⁺ lymphocyte population in each LPL preparation were determined by FACS analysis.

Table 1. General features of the SFB chromosome

Size (bp)	1 620 005
G + C content (%)	28.3
CDSs	1491 ^a
Coding density	89.8
Average CDS length (bp)	960
rRNA operons	6
tRNA genes	37
Prophages	4

^aTwenty-four pseudogenes are not included.

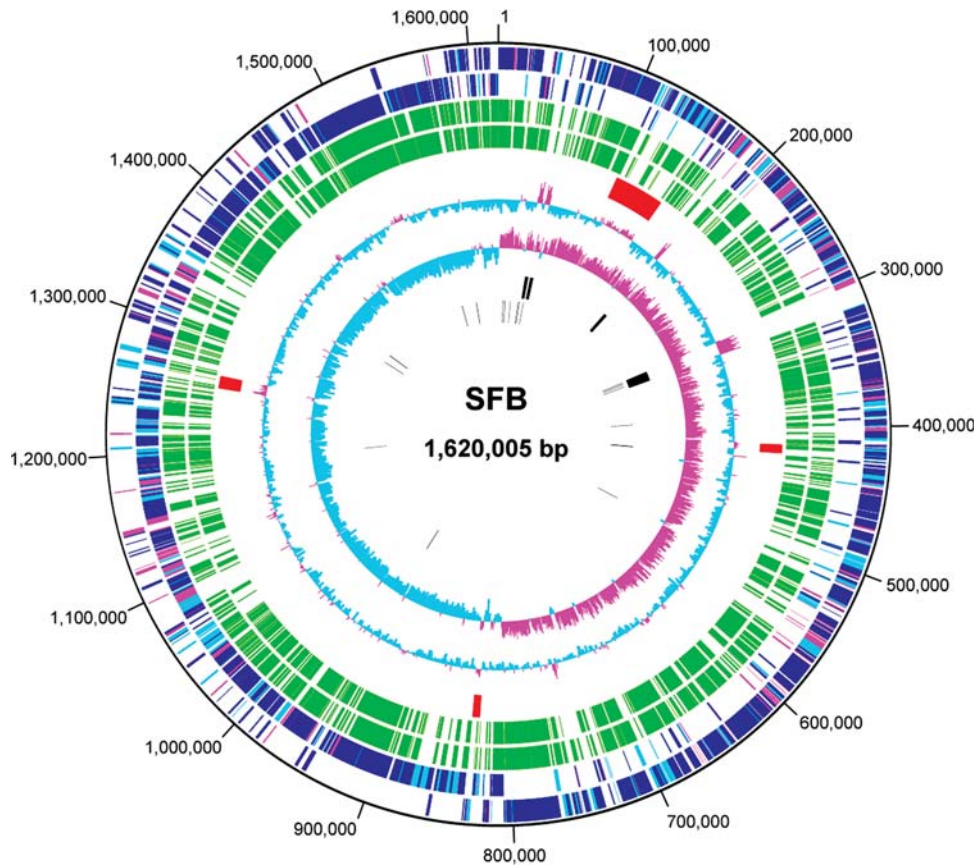


Figure 2. Circular map of the SFB chromosome. From the outside in: the first circle depicts the nucleotide sequence positions (in bp). The second and third depict CDSs that are transcribed clockwise and counterclockwise, respectively, and CDSs were coloured according to the results of the BLASTP homology search of each CDS (E -value $< 1e-5$) against the RefSeq database at the NCBI (blue, top hit to order Clostridiales; light blue, top hit to other species; red, no hit). The fourth and fifth depict CDSs conserved in *C. difficile* and *C. novyi*, respectively ($\geq 25\%$ sequence identity and $\geq 60\%$ aligned length coverage of a query sequence). The sixth depicts prophages, the seventh GC content, the eighth GC skew, the ninth rRNA operons and the tenth tRNA genes.

target invading nucleic acids. One of the CRISPR loci in the SFB genome (CRISPR3) contains 13 spacer-repeat units, which are preceded by a gene cluster encoding six CRISPR-associated (Cas) proteins. Interestingly, the other two loci (each contains three and four spacer-repeat units, respectively) are located on prophage SFBMP01. Neither is associated with *cas* genes, but they share the same 160-bp leader sequence and 32-bp repeat sequences with CRISPR3, suggesting that the precursor RNAs derived from these prophage-encoded CRISPRs may also be matured to short CRISPR RNAs by the CRISPR3-associated Cas proteins. Prophage-encoded CRISPRs were previously described in *C. difficile*.²⁶ Any spacer sequences found in the three CRISPR loci of SFBs showed no significant homology to the sequences in databases.

3.3. The phylogenetic position of SFBs

In our initial BLAST analysis against the RefSeq database, 66.2% of the SFB CDSs had best-hit homologues of the order Clostridiales but with relatively low levels

of sequence identity (Fig. 2 and Supplementary Table S1). These results are consistent with the phylogenetic position of SFBs previously inferred by 16S rRNA sequence analysis; SFBs belong to a distinct lineage within the *Clostridium* subphylum at the rank of genus, whose proposed genus name was *Candidatus Arthromitus* (Fig. 3A).²⁷ We further confirmed this phylogenetic position by constructing a genomic phylogenetic tree based on the sequences of 31 universal genes²⁸ from fully sequenced species belonging to the order Clostridiales (Fig. 3B).

3.4. Comparative genomics

Because no closely related species has yet been sequenced, we chose CD and CN for genomic comparison because both are gut-associated clostridia and they have the largest or smallest genome among the fully sequenced Clostridiales (4.3 and 2.5 Mb, respectively).^{26,29} Based on the orthologous gene group classification of CD and CN genes in the MGD database (<http://mbgd.genome.ad.jp/>), in

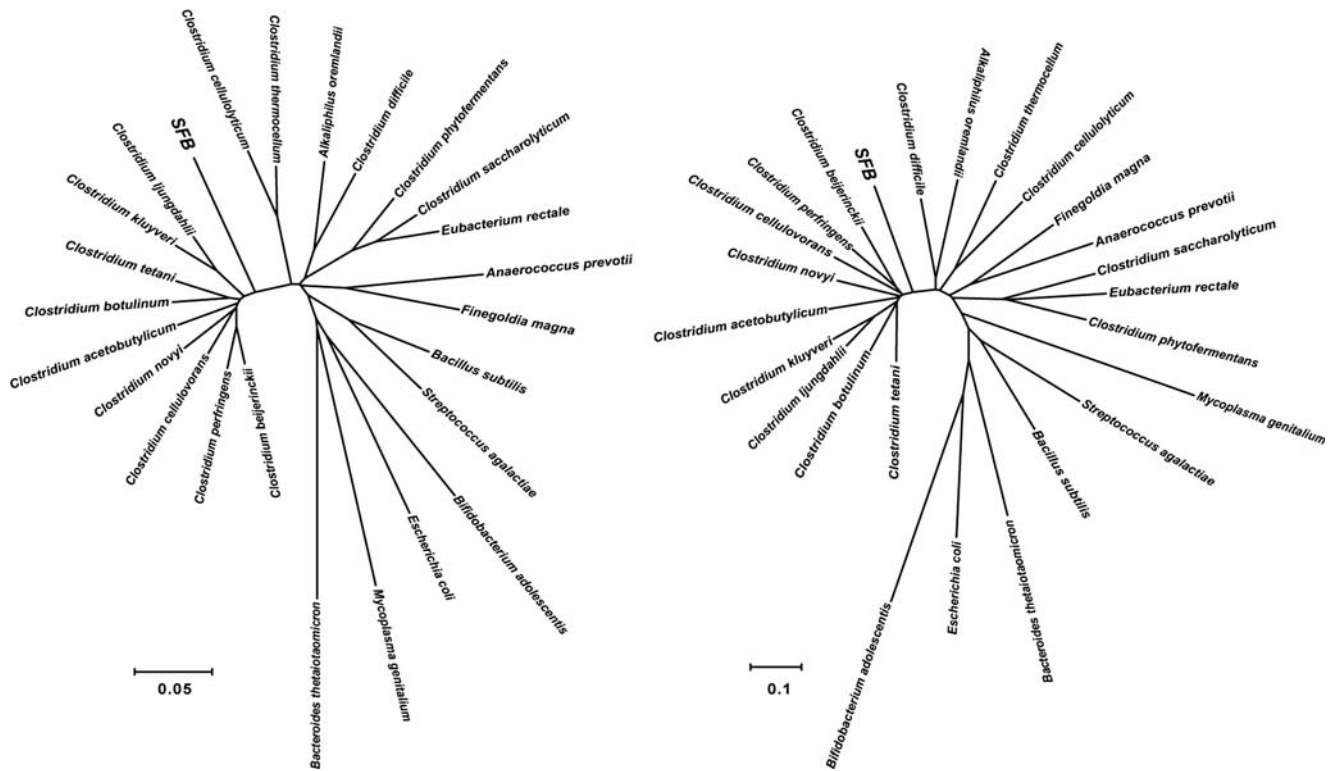


Figure 3. The phylogenetic position of SFBs among the fully sequenced *Clostridium* species. Phylogenetic trees constructed based on the 16S rDNA sequences (A) and the concatenated sequences of 31 universal genes²⁸ (B) are shown.

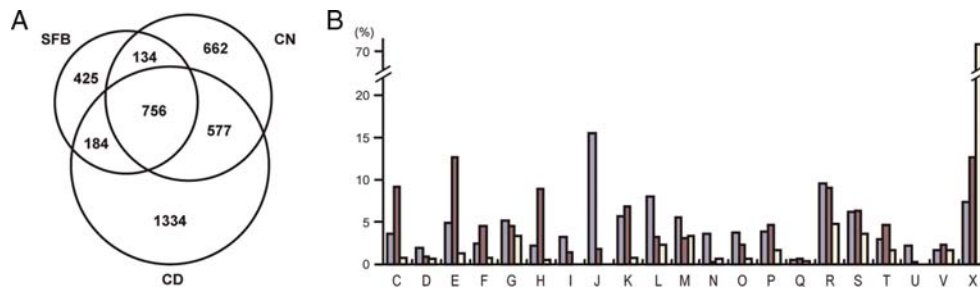


Figure 4. A comparison of SFB gene content to *C. difficile* and *C. novyi*. (A) A Venn diagram showing the conservation of orthologous CDSs (or CDS groups) between SFBs, CD and CN. The numbers of orthologous CDSs in each category are shown. (B) The functional distributions of orthologous CDSs belonging to three conservation categories; ‘conserved in the three bacteria’ (left bars), ‘conserved in CD and CN but not in SFB’ (middle) and ‘unique to SFB’ (right). The functional classification follows the COG categories (<http://www.ncbi.nlm.nih.gov/COG/>) (C: energy production and conversion; D: Cell cycle control, mitosis and meiosis; 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; K: transcription; L: replication, recombination and repair; M: cell wall/membrane biogenesis; N: cell motility; O: posttranslational modification, protein turnover, chaperones; P: inorganic ion transport and metabolism; Q: secondary metabolites biosynthesis, transport and catabolism; R: general function prediction only; S: function unknown; T: signal transduction mechanisms; U: intracellular trafficking and secretion; V: defense mechanisms) and ‘X’ indicates CDSs that are not assigned to any category. Each bar represents the percentage of CDSs belonging to each functional category in each of three conservation categories.

which 3741 CDSs of CD were grouped into 2851 CDSs (or CDS groups) and 2315 CDSs of CN were grouped into 2129 CDSs (or CDS groups), we compared the gene content of SFBs with these two gut-associated clostridia via an all-to-all BLAST analysis. The results revealed that 940 and 890 of the SFB

CDSs (or CDSs groups) are conserved in CD and CN, respectively; 756 are conserved in both strains, whereas 425 are unique to SFB (Fig. 4A). Most of the unique genes are of unknown function. More importantly, we identified 577 CDSs that are shared by CD and CN but not SFB. The distribution of the

577 CDSs into the functional categories of the clusters of orthologous groups (COGs) of proteins indicates that, in comparison to CD and CN, the SFB lacks remarkable numbers of genes involved in the following processes: 'energy production and conversion (C in Fig. 4B)', 'amino acid transport and metabolism (E)', 'nucleotide transport and metabolism (F)' and 'coenzyme metabolism (H)'. Significant portions of the CD/CN shared genes that are related to 'carbohydrate transport and metabolism (G)', 'transcription (K)' and 'inorganic ion transport and metabolism (P)' are also absent in SFBs. These results and the smaller genome size of SFBs suggest that during its strict niche specialization, genome reduction has taken place and many metabolic functions have been lost (see below).

3.5. Metabolic capabilities

Consistent with the prediction that SFBs are strict anaerobes, no CDS for tricarboxylic acid cycle- or respiratory chain-related proteins were identified, whereas the SFB contains two catalases and one peroxidase. They may confer some level of oxygen tolerance, which is required for the SFB to pursue an epicellular life on intestinal epithelia where a microaerobic environment is formed. However, SFBs contain a complete enzyme set for the glycolytic pathway from glucose to pyruvate (Supplementary Fig. S2A). SFBs can further metabolize pyruvate using the pyruvate:ferredoxin oxidoreductase–hydrogenase system to produce acetate or ethanol but not butyrate and butanol. Pyruvate can also be metabolized to lactate. These are the primary energy production pathways for the SFB, like many other clostridia.³⁰ Furthermore, SFBs possess a D-ribose ATP-binding cassette (ABC) transporter, ribokinase and transketolase and can, therefore, uptake extracellular ribose and metabolize it to glyceraldehyde-3-phosphate, an intermediate of the glycolytic pathway. Although the pentose phosphate pathway (PPP) is conserved, components for the *de novo* synthesis of purine and pyrimidine nucleotides from phosphoribosyl pyrophosphate (PRPP) are missing. Thus, it is not likely that the PPP of SFBs is involved in nucleotide biosynthesis.

A more striking metabolic feature of SFBs is the almost complete lack of *de novo* biosynthesis pathways for amino acids, vitamins/cofactors and nucleotides (Table 2). As for amino acid biosynthesis, we identified only 20 genes related to this function. Genes for lysine biosynthesis are relatively well conserved, but they probably just function to synthesize diamino-pimeric acid as a component of peptidoglycan. Similarly, SFBs possess only 14 genes for vitamin and cofactor biosynthesis, and almost all the genes required for the biosynthesis of thiamin, riboflavin,

pyridoxine, nicotinamide, pantothenate, coenzyme A, biotin, folate and vitamin B₁₂ are absent. As mentioned above, SFBs are unable to synthesize purines and pyrimidines from PRPP. In addition, key enzymes for nucleotide interconversion, such as thymidylate synthase and guanine monophosphate reductase, are also absent in the SFB.

Severe deficiency in the biosynthesis capability of amino acids, vitamins/cofactors and nucleotides suggests that SFBs uptake these compounds from extracellular environments. Consistent with this, at least 103 genes are devoted to transport functions (Supplementary Table S2). They include ABC transporter systems for oligopeptides, amino acids, sugars, phosphate, phosphonate, iron compounds and cobalt. SFBM_1365, SFBM_1366 and SFBM_1367 may constitute an energy-coupling factor transporter to uptake vitamins and cofactors,³¹ although the substrate-binding components therein have yet to be identified. In addition, SFBs possess nine phosphotransferase systems to uptake various simple sugars and several other types of transport systems for nucleotides, amino acids and other compounds.

Although the Tat protein secretion system is not present in the SFB, the Sec system is conserved. It may be noteworthy that, among the eight glycosyl hydrolases of SFBs (Supplementary Table S3), only SFBM_0066 was predicted as an extracellular enzyme by the PSORT program (<http://psort.hgc.jp/>). This suggests that SFBs have a very limited ability to digest food- or host-derived polysaccharides by themselves and, instead, primarily uptake simple sugars present in ready-for-use states in food or provided by the host or other gut bacteria. In contrast, 41 protease/peptidases are encoded by SFB. Of these, 4 are extracellular enzymes and 20 are membrane-associated (Supplementary Table S3). In SFBs, it is very likely that these enzymes, together with transporters for oligopeptides and amino acids, contribute to the acquisition of amino acids derived from external polypeptides. Similarly, two extracellular nucleases (SFBM_0242 and SFBM_0633) probably contribute to nucleotide acquisition from external sources.

In summary, SFB growth heavily relies on the uptake of essential compounds from extracellular environments, some of which may be supplied by host cells that SFB tightly attaches to. These metabolic features form the basis of the non-culturability of SFBs *in vitro* and would be valuable information for developing culture media and protocols for the SFB.

3.6. Sporulation and germination

During the epicellular parasitic phase, SFB probably obtain various nutritional components via intimate interactions with host cells; however, once they are

Table 2. Gene repertoires for amino acid and vitamin/cofactor biosynthesis in *C. difficile* (CD), *C. novyi* (CN) and SFB

Organisms	CD	CN	SFB
Amino acids			
Glycine, serine and threonine	<i>serA</i> <i>serB</i> <i>glyA</i> <i>thrBC</i> <i>hom</i>	<i>serA</i> <i>serB</i> <i>glyA</i> <i>thrBC</i> <i>hom</i>	<i>serA</i> <i>glyA</i>
Cysteine and methionine	<i>cysEK</i> <i>metAY</i> <i>metC</i> <i>metH</i> <i>metK</i>	<i>cysEK</i> <i>metAY</i> <i>metH</i> <i>metK</i>	 <i>metC</i> <i>metK</i>
Valine, leucine and isoleucine	<i>ilvA</i> <i>ilvBCD</i> <i>leuABCD</i>	<i>ilvA</i> <i>ilvBCD</i> <i>ilvE</i>	 <i>ilvE</i>
Lysine	<i>dapABCDEFG</i> <i>lysA</i> <i>glnA</i>	<i>dapABCDEFG</i> <i>lysA</i> <i>glnA</i>	<i>dapABDFG</i> <i>lysA</i> <i>glnA</i>
Alanine, aspartate and glutamate	<i>asnA</i> <i>asnB</i> <i>gltD</i>	<i>asnA</i> <i>gltD</i>	 <i>asnB</i>
Arginine and proline	<i>aspC</i> <i>aspD</i> <i>proC</i> <i>argBCDEFJ</i> <i>argGH</i>	 <i>aspD</i> <i>proABC</i> <i>argBCDFJ</i> <i>argGH</i>	 <i>aspC</i>
Histidine	<i>hisABCDFGHIZ</i> <i>hisK</i>	<i>hisABCDFGIZ</i> <i>hisK</i>	
Phenylalanine, tyrosine and tryptophan	<i>aroABCEGK</i> <i>aroD</i> <i>trpB</i> <i>tyrA</i> <i>pheA</i>	<i>aroABCEGK</i> <i>aroD</i> <i>trpB</i> <i>tyrA</i> <i>pheA</i>	<i>aroABCEGK</i>
Vitamins and cofactors			
Thiamin	<i>thiCFGH</i> <i>thiD</i> <i>thiI</i> <i>thiN</i> <i>isc</i>	<i>thiCFGH</i> <i>thiD</i> <i>thiI</i> <i>thiN</i> <i>isc</i>	<i>thiH</i> <i>thiN</i> <i>isc</i>
Riboflavin	<i>ribABDH</i> <i>ribC</i>	<i>ribABDH</i> <i>ribC</i>	 <i>ribC</i>
Vitamin B ₆	 <i>pdxK</i>	 <i>pdxA</i> <i>pdxS</i>	 <i>pdxS</i>
Nicotinate and nicotinamide	<i>nadABC</i> <i>nadD</i> <i>nadE</i>	<i>nadABC</i> <i>nadD</i>	 <i>nadD</i>
Pantothenate	<i>panBC</i> <i>panE</i>	<i>panBCD</i> <i>panE</i>	
Coenzyme A	<i>coaX</i> <i>coaBC</i> <i>coaD</i> <i>coaE</i>	<i>coaX</i> <i>coaBC</i> <i>coaD</i> <i>coaE</i>	 <i>coaD</i> <i>coaE</i>
Biotin	<i>bioA</i> <i>bioB</i> <i>birA</i>	<i>bioAD</i> <i>bioB</i> <i>birA</i>	 <i>birA</i>

Continued

Table 2. Continued

Organisms	CD	CN	SFB
Folate	<i>pabABC</i>	<i>pabABC</i> <i>folA</i>	<i>pabC</i> <i>folA</i>
	<i>folBEKP</i> <i>folC</i>	<i>folBEKP</i> <i>folC</i>	
Vitamin B ₁₂	<i>hemBCD</i> <i>hemN</i> <i>cbiABCDEFGHIJKLPT</i> <i>cobCDSTU</i>	<i>hemABCDL</i> <i>hemN</i> <i>cbiABCDEFGHIJKLPT</i> <i>cobDSTU</i> <i>cobO</i>	<i>cbiA</i> <i>cobC</i> <i>cobO</i>
	<i>eutT</i>		

detached from the host cells or SFB-attached host cells slough off in the process of epithelial cell turnover, such nutrition supply from the host cells would be shut-down, which probably induces the sporulation process in the SFB. We identified 66 genes related to sporulation or germination in the SFB, which are dispersed throughout the chromosome (Supplementary Table S4). They include key proteins for the sporulation signalling cascade, such as sporulation sigma factors, and other stage-specific sporulation proteins (Supplementary Fig. S2B); however, in common with other sequenced clostridia, SFBs lack several proteins that have been shown in *B. subtilis* to be involved in the phosphorelays that trigger sporulation, such as Spo0F, Spo0B and the Kin proteins.^{32,33} In *B. subtilis*, all the five sensor histidine kinases (KinA to KinE) are orphan kinases. Among the 10 sensor kinases identified in the SFB (Supplementary Table S5), SFBM_0504 is the sole orphan kinase and has a PAS domain-like domain, like the *B. subtilis* KinA protein³² and, thus, could be a candidate sensor kinase for sporulation initiation.

Key elements in the germination process involve germinant receptors that sense appropriate germinants, such as specific amino acids, to trigger germination. The GerA family of germinant receptors are encoded by tricistronic *ger* operons,³⁴ and bacilli and clostridia often contain multiple *ger* operons to sense different germinants. For example, CN and *C. botulinum* (CB) possess two and three operons, respectively (Fig. 5), although the *ger* operon was

not identified in the CD genome.²⁶ In SFBs, we identified only one *ger* operon (Fig. 5), indicating that SFBs can sense limited types of germinants. A comparison of the analogous *ger* operon loci between SFB, CN and CB have revealed that, although many genes downstream of the *ger* operon are missing in the SFB, the overall genetic organization of this locus is conserved.

As shown in Fig. 5, the *ispE* and *cphAB* genes are located just upstream of the *ger* operon in the SFB. The presence of these genes may be noteworthy in terms of the metabolism of SFBs. The *ispE* gene encodes 4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase, a component of the methyl-erythritol phosphate (MEP) pathway to synthesize terpenoids that are essential for membrane and peptidoglycan biogenesis. Because other components for the pathway were also identified (Supplementary Fig. S2C), the MEP pathway is completely conserved in the SFB. The *cphAB* genes encode cyanophycin synthase and cyanophycinase (Fig. 5). Cyanophycin is a prokaryote-specific copolymer of aspartate and arginine, which has been shown in many cyanobacteria to serve as a storage compound for nitrogen, carbon and energy.³⁵ The presence of the *cphAB* and isoaspartyl dipeptidase (SFBM_1358) genes suggests that SFBs can synthesize cyanophycin and may utilize it as an amino acid pool in the germination process. In a previous analysis of 570 microbial genomes, it was predicted that only 24 strains,

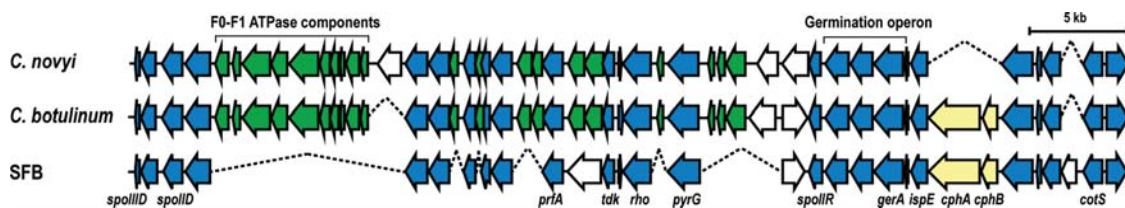


Figure 5. Comparison of the SFB *ger* operon locus with *C. novyi* (CN) and *C. botulinum* (CB). SFB contains a single *ger* operon encoding a germinant receptor, whereas CN and CB possess two and three *ger* operons, respectively. Among these, only the loci analogous to that of SFBs are shown. The genes conserved in the three species are depicted by blue arrows, those conserved in the CN and CB but not in SFB by green and those unique to each species by white. Regions missing in each genome are indicated by a dotted line.

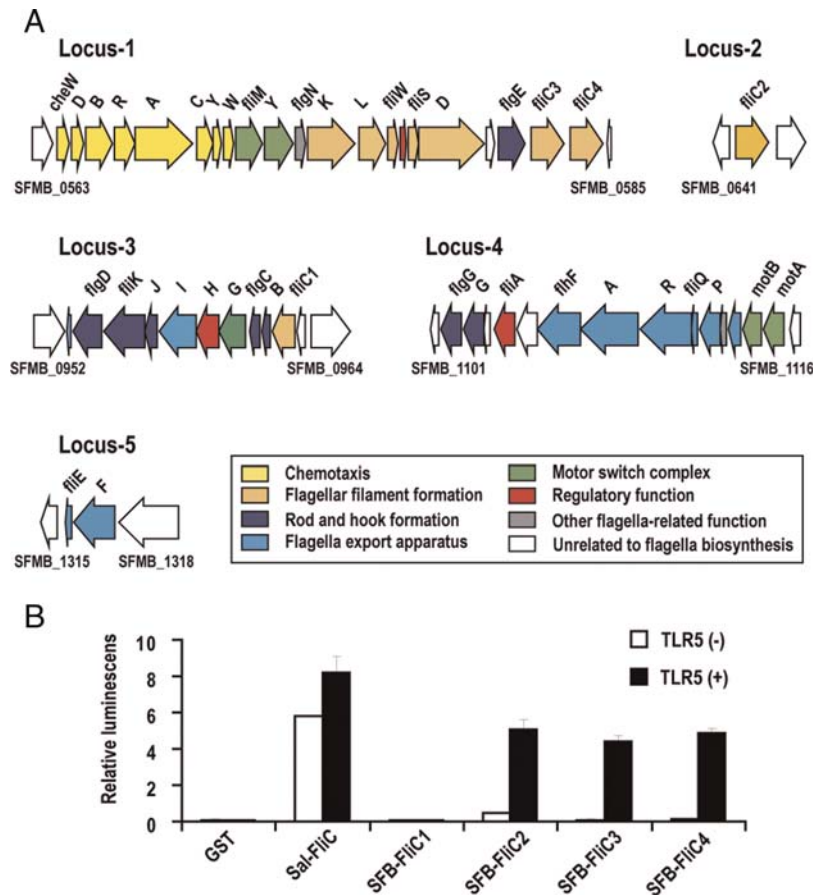


Figure 6. Chemotaxis/flagella biosynthesis gene loci and the activation of the TLR5-linked NF- κ B signaling pathway by SFB flagellins. (A) Genetic structures of five SFB genomic loci containing genes for chemotaxis and flagella-mediated motility are shown. Two loci containing genes for methyl-accepting chemotaxis proteins are not included in this figure. See Supplementary Table S4 for detail on each gene. (B) Luciferase activities in the 293T cells that express (black bar) or do not express mouse TLR5 (white bar) after stimulation with the recombinant *Salmonella* flagellin, recombinant GST fusion proteins of SFB flagellins (FliC1, FliC2, FliC3 or FliC4) or GST are shown. Results are presented as fold-increase relative to the control value.

which include several clostridia (*C. beijerinckii*, *C. botulinum*, *C. perfringens* and *C. thermocellum*), can synthesize cyanopycin.³⁶

3.7. Chemotaxis and flagella-mediated motility

SFBs possess a full set of genes for the chemotaxis and flagella-mediated motility system. These genes are separately located in five genomic loci (Locus-1 to Locus-5; Fig. 6A and Supplementary Table S4). In addition, two methyl-accepting chemotaxis proteins (receptors for attractants) are encoded in other loci (SFBM_0442 and SFBM_1032). Flagellated SFB cells were not observed in our electron microscopic examination of SFB-monoassociated gnotobiotic mice (from 3 days to 2 weeks after oral inoculation) as well as in previous examinations.^{13,37,38} Thus, it is possible that in SFBs, the system has been inactivated or is used for some other purposes, such as a type III secretion system (flagella type); however, the presence of a complete set for chemotaxis genes and four flagellin genes and the absence of recognizable pseudogenes in the

gene set (except for one flagellin gene; see below) strongly suggest that SFBs possess an active chemotaxis and flagella-mediated motility system. This information provides us a previously unknown aspect of SFB biology and life cycle. It is unknown at which life cycle stage SFBs produce flagella, but the most plausible scenario may be that after germination, SFBs swim in the intestine, by sensing some chemoattractant(s), to the distal parts of the small intestine where they colonize. In this regard, it is worth mentioning that SFBs contain only two methyl-accepting chemotaxis proteins (SFBM_0442 is probably the intracellular type). This may suggest that SFBs can sense very limited types of attractants.

3.8. Multiple flagellins and immunostimulation activity

A unique feature of the SFB flagella system is the presence of four types of flagellins (FliC1 to FliC4). Of these, FliC3 and FliC4 are encoded in tandem in Locus-1 (Fig. 6A) and show a high sequence similarity

(82.3% identity). All contain conserved N- and C-terminal flagellin domains, but FliC1 lacks a variable region between the N- and the C-terminal domains.

Flagellins are recognized by TLR5 as pathogen-associated molecular patterns (PAMPs) and induce cytokine production to activate the host innate immune system.³⁹ TLR5 is highly expressed in LP in the small intestine, wherein dendritic cells are the dominant antigen-presenting cells and only CD11c^{hi}CD11b^{hi} LPDCs express TLR5.^{40–42} These flagellin-exposed LPDCs produce IL-6 and IL-12, which results in the generation of IgA⁺ plasma cells and polarization towards the Th17 and Th1 functions.⁴² Therefore, the identified flagellins could be the key molecules responsible for the unique immunostimulation activity of SFBs. To test this presumption, we prepared recombinant proteins for each type of the SFB flagellins and examined whether they could activate the TLR5-linked NF- κ B signalling pathway. As shown in Fig. 6B, although FliC1 was unable to activate the pathway, the other three activated it like that achieved by *Salmonella* flagellin. These results indicate that the immunostimulation activity of SFBs is attributable to the FliC2, FliC3 and FliC4 flagellins, although other molecules, such as cyanophycin and sporulation-related byproducts, may also be involved.

The tight association of SFBs to Peyer's patches, which could make LPDCs more easily accessible to SFB-derived PAMPs, may also directly or indirectly contribute to innate immunity activation. In this regard, the presence of a fibronectin-binding protein (SFBM_0986) and a phosphatidylinositol-specific phospholipase C (PI-PLC; SFBM_0755) is noteworthy. The elevation of intracellular free calcium levels induced by PI-PLC from various pathogens plays key roles in various processes of microbe–host cell interaction, such as bacterial internalization and host cell actin rearrangement.^{43,44} Therefore, SFB PI-PLC may also be involved in the formation of dense aggregates consisting of actin-like microfibrils beneath the host cell membrane where the SFB attached.¹³ Other unidentified factors are also probably required for the intimate binding of SFBs to host epithelial cells, and many of such factors may be encoded by SFB-specific genes, most of which are of unknown function (Fig. 4).

3.9. Human SFB-related sequences in the human gut metagenome sequence data

SFBs have also been found in humans, but SFBs are known to have strict host specificities and the SFB from a different host species may represent a different species.^{13,14} The availability of the mouse SFB genome sequence provided us a first opportunity to

search for human SFB-related sequences in the large data set of the human gut metagenome sequences.¹¹ By searching 7 345 361 234 high-quality Illumina sequences in the data set, we identified 277 reads that are highly homologous to the mouse SFB sequence (sequence identity; $\geq 95\%$, alignment length; $\geq 95\%$ of the read, excluded reads related to rRNA and tRNA sequences) (Supplementary Table S6). Only a small number of reads corresponding to 0.31% of the mouse SFB genome were identified. This may be because non-mapped regions exhibit significant sequence divergence between mouse and human SFB (our search was conservative) or, more likely, because the currently available metagenome sequence data set does not include samples from children at weaning periods. In mice, SFBs become prevalent at this age. Thus, the metagenome sequence data from such children would be required to obtain more genomic sequence information on human SFB.

4. Conclusion

The whole-genome sequence of mouse SFBs revealed that this unique gut microbe has a very limited range of metabolic capabilities similar to those of obligate intracellular parasites. In particular, SFBs lack nearly all *de novo* biosynthesis pathways for amino acids, vitamins and cofactors and nucleotides. These features well explain the non-culturability of SFBs and further suggest that they acquire these compounds, probably from the host intestinal epithelia to which SFBs are tightly attached, via well-conserved transport systems. In addition to a full set of genes for sporulation and germination, SFBs possess a chemotaxis and flagella-based motility system. This finding suggests a triphasic lifestyle of SFBs that comprises swimming, epicellular and dormant (spore) phases. Moreover, SFBs encode four different types of flagellins, and three of these are capable of activating the TLR5-linked NF- κ B signalling pathway, which could promote the luminal IgA production and the induction of particular T helper cell subclasses (Th1 and Th17). Thus, SFB genome sequence determination discloses the genetic basis for their non-culturability, which could help us to develop SFB-specific cultivation techniques. In addition, it reveals a previously unknown lifestyle including a flagellated swimming phase, and it provides a first glimpse into understanding the molecular basis for SFB-induced immunostimulation. Furthermore, through a systematic search of the human gut metagenome sequence database, the first pieces of the genomic information on human-associated SFB were obtained.

Acknowledgements: We thank Drs K. J. Ishii (Osaka University) and F. Takeshita (Yokohama City University) for providing the TLR5 cDNA and the NF- κ B reporter plasmid, respectively. The 16 276-contig sequences used in this study were generated in part with US federal funds from the NIH Human Microbiome Project, the Common Fund, National Institutes of Health and the Department of Health and Human Services.

Supplemental data: Supplemental data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was supported by grants-in-aid for Scientific Researches (C) and (B) (to T.K. and T.H., respectively), for Young Scientists (B) (to H.N-I.) and for Scientific Research on Priority Areas 'Applied Genomics' (to T.H.) and 'Comprehensive Genomics' (to M.H.) from the Ministry of Education, Science and Technology of Japan and by that from the Institute for Bioinformatics Research and Development, the Japan Science and Technology Agency (BIRD-JST) (to K.K.).

References

- Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. and Gordon, J.I. 2005, Host-bacterial mutualism in the human intestine, *Science*, **307**, 1915–20.
- Hooper, L.V., Midtvedt, T. and Gordon, J.I. 2002, How host-microbial interactions shape the nutrient environment of the mammalian intestine, *Ann. Rev. Nutr.*, **22**, 283–307.
- Elson, C.O., Cong, Y., McCracken, V.J., Dimmitt, R.A., Lorenz, R.G. and Weaver, C.T. 2005, Experimental models of inflammatory bowel disease reveal innate, adaptive, and regulatory mechanisms of host dialogue with the microbiota, *Immunol. Rev.*, **206**, 260–76.
- Sartor, R.B. 2008, Microbial influences in inflammatory bowel diseases, *Gastroenterology*, **134**, 577–94.
- Chu, F.F., Esworthy, R.S., Chu, P.G., et al. 2004, Bacteria-induced intestinal cancer in mice with disrupted *gpx1* and *gpx2* genes, *Cancer Res.*, **64**, 962–8.
- Ley, R.E., Turnbaugh, P.J., Klein, S. and Gordon, J.I. 2006, Microbial ecology: human gut microbes associated with obesity, *Nature*, **444**, 1022–3.
- Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R. and Gordon, J.I. 2006, An obesity-associated gut microbiome with increased capacity for energy harvest, *Nature*, **444**, 1027–31.
- Gill, S.R., Pop, M., Deboy, R.T., et al. 2006, Metagenomic analysis of the human distal gut microbiome, *Science*, **312**, 1355–9.
- Kurokawa, K., Itoh, T., Kuwahara, T., et al. 2007, Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes, *DNA Res.*, **14**, 169–81.
- Castello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I. and Knight, R. 2009, Bacterial community variation in human body habitats across space and time, *Science*, **326**, 1694–7.
- Qin, J., Li, R., Raes, J., et al. 2010, A human microbial gene catalogue established by metagenomic sequencing, *Nature*, **464**, 59–67.
- Marcy, Y., Ouverney, C., Bik, E.M., et al. 2007, Dissecting biological 'dark matter' with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth, *Proc. Natl Acad. Sci. USA*, **104**, 11889–94.
- Klaasen, H.L.B.M., Koopman, J.P., Poelma, F.G.J. and Beynen, A.C. 1992, Intestinal, segmented, filamentous bacteria, *FEMS Microbiol. Rev.*, **88**, 165–80.
- Klaasen, H.L.B.M., Koopman, J.P., Van den Brink, M.E., Bakker, M.H., Poelma, F.G. and Beynen, A.C. 1993, Intestinal, segmented, filamentous bacteria in a wide range of vertebrate species, *Lab. Anim.*, **27**, 141–50.
- Umesaki, Y., Okada, Y., Matsmoto, S., Imaoka, A. and Setoyama, H. 1995, Segmented filamentous bacteria are indigenous intestinal bacteria that activate intraepithelial lymphocytes and induce MHC class II molecules and fucosyl asialo GM1 glycolipids on the small intestinal epithelial cells in the ex-germ-free mouse, *Microbiol. Immunol.*, **39**, 555–62.
- Klaasen, H.L.B.M., Van der Heijden, P.J., Stok, W., et al. 1993, Apathogenic, intestinal, segmented, filamentous bacteria stimulate the mucosal immune system of mice, *Infect. Immun.*, **61**, 303–6.
- Suzuki, K., Meek, B., Doi, Y., et al. 2004, Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut, *Proc. Natl Acad. Sci. USA*, **101**, 1981–6.
- Ivanov, I.I., Atarashi, K., Manel, N., et al. 2009, Induction of intestinal Th17 cells by segmented filamentous bacteria, *Cell*, **139**, 485–98.
- Gaboriau-Routhiau, V., Rakotobe, S., Lécuyer, E., et al. 2009, The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell response, *Immunity*, **31**, 677–89.
- Hase, K., Ohshima, S., Kawano, K., et al. 2005, Distinct gene expression profiles characterize cellular phenotypes of follicle-associated epithelium and M cells, *DNA Res.*, **12**, 127–37.
- Morita, H., Kuwahara, T., Ohshima, K., et al. 2007, An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine, *Microb. Environ.*, **22**, 214–22.
- Sugawara, H., Ohyama, A., Mori, H. and Kurokawa, K. 2009, Microbial Genome Annotation Pipeline (MiGAP) for diverse users. The 20th International Conference on Genome Informatics (GIW2009) Poster and Software Demonstrations (Yokohama), S001-1-2.
- Hayashi, F., Smith, K.D., Ozinsky, A., et al. 2001, The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5, *Nature*, **410**, 1099–103.

24. Horvath, P. and Barrangou, R. 2010, CRISPR/Cas, the immune system of bacteria and archaea, *Science*, **327**, 167–70.
25. Marraffini, L.A. and Sontheimer, E.J. 2010, CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea, *Nat. Rev. Genet.*, **11**, 181–90.
26. Sebaihia, M., Wren, B.W., Mullany, P., et al. 2006, The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome, *Nat. Genet.*, **38**, 779–86.
27. Snel, J., Heinen, P.P., Blok, H.J., et al. 1995, Comparison of 16S rRNA sequences of segmented filamentous bacteria isolated from mice, rats, and chickens and proposal of 'Candidatus arthromitus', *Int. J. Syst. Bacteriol.*, **133**, 445–50.
28. Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B. and Bork, P. 2006, Toward automatic reconstruction of a highly resolved tree of life, *Science*, **311**, 1283–7.
29. Bettegowda, C., Huang, X., Lin, J., et al. 2006, The genome and transcriptomes of the anti-tumor agent *Clostridium novyi*-NT, *Nat. Biotechnol.*, **24**, 1573–80.
30. Desvaux, M. 2005, *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia, *FEMS Microbiol. Rev.*, **29**, 741–64.
31. Rodionov, D.A., Hebbeln, P., Eudes, A., et al. 2009, A novel class of modular transporters for vitamins in prokaryotes, *J. Bacteriol.*, **191**, 42–51.
32. Piggot, P.J. and Hilbert, D.W. 2004, Sporulation of *Bacillus subtilis*, *Curr. Opin. Microbiol.*, **7**, 579–86.
33. Paredes, C.J., Alsaker, K.V. and Papoutsakis, T.A. 2005, Comparative genomic view of clostridial sporulation and physiology, *Nat. Rev. Microbiol.*, **3**, 969–78.
34. Moir, A., Corfe, B.M. and Behravan, J. 2002, Spore germination, *Cell Mol. Life Sci.*, **59**, 403–9.
35. Allen, M.M. 1984, Cyanobacterial cell inclusions, *Annu. Rev. Microbiol.*, **38**, 1–25.
36. Fuser, G. and Steinbuchel, A. 2007, Analysis of genome sequences for genes of cyanophycin metabolism: identifying putative cyanophycin metabolizing prokaryotes, *Macromol. Biosci.*, **7**, 278–96.
37. Meyerholz, D.K., Stabel, T.J. and Cheville, N.F. 2002, Segmented filamentous bacteria interact with intraepithelial mononuclear cells, *Infect. Immun.*, **70**, 3277–80.
38. Yamauchi, K. and Snel, J. 2000, Transmission electron microscopic demonstration of phagocytosis and intracellular processing of segmented filamentous bacteria by intestinal epithelial cells of the chick ileum, *Infect. Immun.*, **68**, 6496–504.
39. Smith, K.D., Andersen-Nissen, E., Hayashi, F., et al. 2003, Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility, *Nat. Immunol.*, **4**, 1247–53.
40. Uematsu, S., Jang, M.H., Chevrier, N., et al. 2006, Detection of pathogenic intestinal bacteria by Toll-like receptor 5 on intestinal CD11c⁺ lamina propria cells, *Nat. Immunol.*, **7**, 868–74.
41. Uematsu, S., Fujimoto, K., Jang, M.H., et al. 2008, Regulation of humoral and cellular gut immunity by lamina propria dendritic cells expressing Toll-like receptor 5, *Nat. Immunol.*, **9**, 769–76.
42. Uematsu, S. and Akira, S. 2009, Immune responses of TLR5⁺ lamina propria dendritic cells in enterobacterial infection, *J. Gastroenterol.*, **44**, 803–11.
43. Wadsworth, S.J. and Goldfine, H. 1999, *Listeria monocytogenes* phospholipase C-dependent calcium signaling modulates bacterial entry into J774 macrophage-like cells, *Infect. Immun.*, **67**, 1770–8.
44. Ginocchio, C., Pace, J. and Galan, J.E. 1992, Identification and molecular characterization of *Salmonella typhimurium* gene involved in triggering the internalization of salmonellae into cultured epithelial cells, *Proc. Natl Acad. Sci. USA*, **89**, 5976–80.