

# *Bifidobacterium animalis* subsp. *lactis* ATCC 27673 Is a Genomically Unique Strain within Its Conserved Subspecies

Joseph R. Loquasto,<sup>a</sup> Rodolphe Barrangou,<sup>a,b,c</sup> Edward G. Dudley,<sup>a</sup> Buffy Stahl,<sup>b</sup> Chun Chen,<sup>a</sup> Robert F. Roberts<sup>a</sup>

Department of Food Science, The Pennsylvania State University, University Park, Pennsylvania, USA<sup>a</sup>; DuPont Nutrition and Health, Madison, Wisconsin, USA<sup>b</sup>; Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina, USA<sup>c</sup>

Many strains of *Bifidobacterium animalis* subsp. *lactis* are considered health-promoting probiotic microorganisms and are commonly formulated into fermented dairy foods. Analyses of previously sequenced genomes of *B. animalis* subsp. *lactis* have revealed little genetic diversity, suggesting that it is a monomorphic subspecies. However, during a multilocus sequence typing survey of *Bifidobacterium*, it was revealed that *B. animalis* subsp. *lactis* ATCC 27673 gave a profile distinct from that of the other strains of the subspecies. As part of an ongoing study designed to understand the genetic diversity of this subspecies, the genome of this strain was sequenced and compared to other sequenced genomes of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis*. The complete genome of ATCC 27673 was 1,963,012 bp, contained 1,616 genes and 4 rRNA operons, and had a G+C content of 61.55%. Comparative analyses revealed that the genome of ATCC 27673 contained six distinct genomic islands encoding 83 open reading frames not found in other strains of the same subspecies. In four islands, either phage or mobile genetic elements were identified. In island 6, a novel clustered regularly interspaced short palindromic repeat (CRISPR) locus which contained 81 unique spacers was identified. This type I-E CRISPR-*cas* system differs from the type I-C systems previously identified in this subspecies, representing the first identification of a different system in *B. animalis* subsp. *lactis*. This study revealed that ATCC 27673 is a strain of *B. animalis* subsp. *lactis* with novel genetic content and suggests that the lack of genetic variability observed is likely due to the repeated sequencing of a limited number of widely distributed commercial strains.

Many *Bifidobacterium animalis* subsp. *lactis* strains are considered to be probiotic microorganisms and are commonly added to a variety of fermented and nonfermented foods (1, 2). Characteristics making strains of this subspecies desirable for use as probiotics include their perceived health benefits as well as technological advantages over organisms of the same genus. Health benefits attributed to this subspecies include modulation of the immune system (3), increased digestive comfort (4), and reduction of colonic transit time (5). Technological advantages claimed for this subspecies include tolerance to oxygen (6), resistance to acid (7, 8) and bile (9), as well as viability during extended refrigerated storage (10). *B. animalis* subsp. *lactis* DSM 10140 (originally identified as *B. lactis*) was first described in 1997 as a unique species of *Bifidobacterium* and was identified as an oxygen-tolerant isolate from a fermented milk sample (11).

In previous work, the genome of the type strain of the subspecies DSM 10140 and that of a commercial strain, BI-04, were completely sequenced and subjected to comparative analysis (12). This analysis revealed that the two genomes were highly conserved, as only 47 single nucleotide polymorphisms (SNPs) and 4 indels were identified between the two strains. Clustered regularly interspaced short palindromic repeats (CRISPRs) were found to be the major source of variation in *B. animalis* subsp. *lactis*, representing 43.67% of all differential nucleotides between DSM 10140 and BI-04 (12).

Sequencing of 8 additional genomes has confirmed the highly monomorphic nature of the *B. animalis* subsp. *lactis* genome (13–19). A possible reason for the lack of variability observed in the *B. animalis* subsp. *lactis* genomes to date is that only strains of commercial importance have been sequenced, reflecting sampling bias as opposed to natural genetic diversity. The custom of isolating strains from competitors' products for use in new starter systems has been a common practice in the industry. Even though 10 iso-

lates of *B. animalis* subsp. *lactis* have been completely sequenced, it is not clear that the true genomic diversity within the subspecies has been explored.

As part of an ongoing effort to understand the diversity within the *B. animalis* group, the genome of the type strain of the related subspecies *B. animalis* subsp. *animalis* (20) was sequenced. The genome of *B. animalis* subsp. *animalis* ATCC 25527 exhibited approximately 96.23% similarity with the genome of *B. animalis* subsp. *lactis* (20). Of note, *B. animalis* subsp. *animalis* had a CRISPR locus that differed in terms of spacer and repeat sequences and overall size from the CRISPR locus of *B. animalis* subsp. *lactis* and was classified as a type I-E CRISPR-*cas* system (21, 22).

When identifying species within the genus *Bifidobacterium*, Delétoile et al. (23) subjected a number of species of bifidobacteria to a multilocus sequence typing (MLST) scheme based upon the housekeeping genes *clpC*, *fusA*, *gyrB*, *ileS*, *purF*, *rplB*, and *rpoB*. One strain, ATCC 27673, was separated from the other strains of *B. animalis* subsp. *lactis* on the basis of this MLST scheme. This was interesting, given the lack of diversity previously observed within the subspecies. Thus, the goal of this work was to sequence the genome of *B. animalis* subsp. *lactis* ATCC 27673 and compare

Received 31 May 2013 Accepted 27 August 2013

Published ahead of print 30 August 2013

Address correspondence to Robert F. Roberts, rfr3@psu.edu.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01777-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01777-13

TABLE 1 General genomic characteristics of sequenced genomes of *B. animalis* subspecies

Organism	GenBank accession no.	Length (bp)	Coding %	% G+C content	No. of genes	No. of rRNA operons	No. of tRNAs	Reference or source
<i>B. animalis</i> subsp. <i>lactis</i>								
ATCC 27673	CP003941	1,963,012	83.4	61.6	1,616	4	52	This work
DSM 10140	CP001606	1,938,483	90.3	60.5	1,629	4	51	12
Bl-04	CP001515	1,938,709	90.5	60.5	1,631	4	52	12
Bb-12	CP001853	1,942,198	89.9	60.5	1,642	4	52	14
V9	CP001892	1,944,050	86.0	60.5	1,636	4	52	15
AD011 <sup>a</sup>	CP001213	1,933,695	84.6	60.5	1,528	2	52	13
CNCM I-2494	CP002915	1,943,113	90.9	60.5	1,724	4	52	16
BLC1	CP003039	1,943,990	86.4	60.5	1,622	4	52	17
Bi-07	CP003498	1,938,822	86.5	60.5	1,661	4	52	18
B420	CP003497	1,938,595	86.1	60.5	1,625	4	52	18
Bl12	CP004053	1,938,605	86.1	60.5	1,607	4	52	19
<i>B. animalis</i> subsp. <i>animalis</i> ATCC 25527								
	CP002567	1,932,963	85.4	60.5	1,597	4	52	20

<sup>a</sup> This strain was excluded from bioinformatic analyses because of issues associated with assembly (13).

it with the genomes of other strains of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis*.

## MATERIALS AND METHODS

**Growth and identification.** *Bifidobacterium animalis* subsp. *lactis* ATCC 27673 was obtained from the American Type Culture Collection (ATCC; Manassas, VA) and was grown in de Man, Rogosa, and Sharpe medium (24) supplemented with 0.05% (wt/vol) cysteine hydrochloride (MRSc). All cultures were grown anaerobically (85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub>) at 37°C. Genomic DNA was isolated from 10 ml of an overnight culture using a Wizard genomic DNA purification kit (Promega, Madison, WI) according to the manufacturer's instructions. Subspecies-specific primers and conditions developed by Ventura and Zink (25) were used to verify that the culture was *B. animalis* subsp. *lactis*.

**Sequencing and assembly.** Genomic DNA was shotgun sequenced by 454 pyrosequencing on a GS-FLX sequencer at The Pennsylvania State University. A total of 551,519 reads and 195,302,731 bp of sequence data were generated. Sequencing was followed by *de novo* assembly in the Newbler software package (Roche, Branford, CT), and the contigs generated were ordered by alignment with the genome of *B. animalis* subsp. *lactis* DSM 10140 (12) using the pheromone trail-based genetic algorithm (26). To close gaps remaining in the sequence, PCR primers were designed on the end of each ordered contig, and PCR was conducted to amplify the sequence in the gaps. The resulting amplicons were sequenced at The Pennsylvania State University using 3' BigDye-labeled dideoxynucleotide triphosphates (version 3.1 dye terminators; Applied Biosystems, Foster City, CA) and an ABI 3730XL DNA analyzer with ABI sequence analysis software (version 5.1.1.).

A *de novo* KpnI optical map of *B. animalis* subsp. *lactis* ATCC 27673 DNA was constructed by OpGen (Gaithersburg, MD) and was compared to a KpnI *in silico* map of the assembled *B. animalis* subsp. *lactis* ATCC 27673 genome created using MapSolver software (OpGen). Following verification, the genome sequence of *B. animalis* subsp. *lactis* ATCC 27673 was submitted for automated annotation at NCBI using the Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP).

**Comparative genomic analysis.** Genomic alignments were conducted using the progressiveMauve tool (27) with default settings. Alignments for tree construction were conducted using the Sequence Search and Alignment by Hashing Algorithm (SSAHA) (28), and the tree was visualized using MEGA (version 5.1) (29). Differential gene content was defined as called genes having less than 60% amino acid identity over the length of the gene between *B. animalis* subsp. *lactis* ATCC 27673 and Bl-04 (12) and *B. animalis* subsp. *animalis* ATCC 25527 (20). Protein sequences

identified as unique genes present in a specific strain were evaluated for homologs using the BLASTP program in the nonredundant database (30). CRISPRs were identified using the Dotter (31) and CRISPRFinder (32) programs. Genomic islands were defined as four consecutive differential open reading frames (ORFs) with assigned functions or greater than four consecutive differential ORFs. Candidate regions for horizontal gene transfer (HGT) events were detected using the Alien Hunter software, available from the Sanger Institute, with default settings (33).

SNPs were identified from alignments conducted using progressiveMauve (27) and the "export SNPs" feature with default settings. Each genome was aligned pairwise with each other strain, and SNPs were identified. Synonymous, nonsynonymous, and intergenic SNPs were identified using a script generated in our laboratory at The Pennsylvania State University and publically available annotations from NCBI. All comparisons were conducted with complete and publically available genome sequences.

**Nucleotide sequence accession number.** The genome sequence of *B. animalis* subsp. *lactis* ATCC 27673 was deposited in GenBank with the accession number CP003941.

## RESULTS

**General features.** The complete genome sequence of *B. animalis* subsp. *lactis* ATCC 27673 is 1,963,012 bp long, which is 24,529 bp longer than the genome of the type strain *B. animalis* subsp. *lactis* DSM 10140 and 24,303 bp longer than the genome of commercial strain Bl-04 (Table 1). Annotation using PGAAP at NCBI revealed that the ATCC 27673 genome contains 1,616 genes, which is 13 fewer than the number for the type strain and 15 fewer than the number for Bl-04. The genome also contained 4 rRNA operons (Table 1). The G+C content was found to be 61.55%, slightly higher than that of the other sequenced genomes of *B. animalis* subsp. *lactis*. The genome of ATCC 27673 exhibited a coding percentage of 83.4%, lower than that of Bl-04, which has a coding percentage of 90.5%. *B. animalis* subsp. *lactis* ATCC 27673 was verified to the subspecies level using subspecies-specific primers for *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* as a negative control (see Fig. S1 in the supplemental material) (25, 34). BLAST analysis showed 100% identity of the ATCC 27673 16S rRNA gene with the 16S rRNA genes of other *B. animalis* subsp. *lactis* strains. The SNPs in ATCC 27673 identified by the

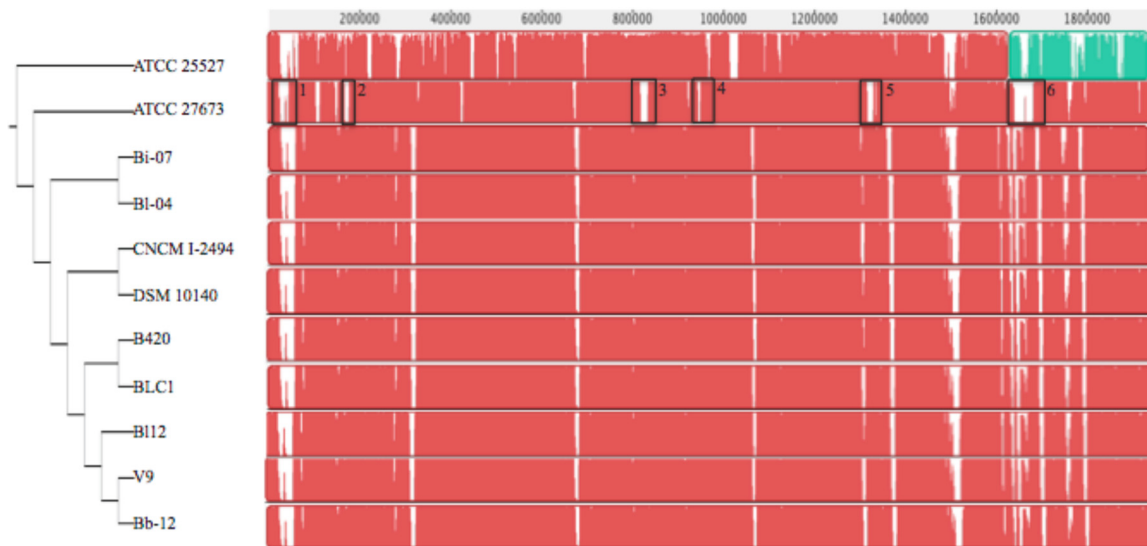


FIG 1 progressiveMauve alignments of sequenced strains of *B. animalis*. Pairwise alignments with the genome of *B. animalis* subsp. *lactis* ATCC 27673 were made using progressiveMauve. All strains are *B. animalis* subsp. *lactis*, except for ATCC 25527, which is *B. animalis* subsp. *animalis*. All sequences used in this analysis are complete and publicly available. Boxes 1 to 6, genomic islands 1 to 6, respectively.

MLST scheme developed by Delétoile et al. (23) were confirmed in the genome sequence.

**Assembly verification.** Whole-genome alignments of the fully sequenced genomes of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* verified the genetic homogeneity and synteny across strains (Fig. 1). Comparison of *in vivo* and *in silico* KpnI optical maps of ATCC 27673 revealed only minor differences and confirmed the genome assembly. Additionally, *in silico* KpnI optical maps of *B. animalis* subsp. *lactis* ATCC 27673 and Bl-04 and *B. animalis* subsp. *animalis* ATCC 25527 were also aligned (see Fig. S2 in the supplemental material). Considerable differences were observed throughout the length of the genome, indicating a high degree of difference in the DNA sequence between these three genomes.

**Comparative genomics and unique genes.** A maximum-parsimony tree generated from the SSAHA alignment provided insight into the relationship between strains of *B. animalis* (Fig. 1). As expected, *B. animalis* subsp. *animalis* ATCC 25527 is the most distantly related strain from the core group of commercial *B. animalis* subsp. *lactis* strains. *B. animalis* subsp. *lactis* ATCC 27673 is more closely related to *B. animalis* subsp. *animalis* than to strains of *B. animalis* subsp. *lactis*. Alignment of *B. animalis* subsp. *lactis* strains showed a very high degree of synteny and shared very similar genetic content with previously sequenced strains. Alignment of previously sequenced strains with the sequence of ATCC 27673 showed a high degree of synteny and very similar genetic content, despite the novel gene content. Whole-genome SNP tree analysis indicated that strains Bi-07 and Bl-04 are more closely related to ATCC 27673 than to the other strains of *B. animalis* subsp. *lactis*.

Because of the high degree of similarity between the previously sequenced *B. animalis* subsp. *lactis* strains, comparisons were made using strain Bl-04 as a reference. There were 1,346 genes shared between ATCC 27673, Bl-04, and ATCC 25527, establishing a core genome for the *B. animalis* group (Fig. 2). The core genome included 83.29% of the genes in the genome of ATCC 27673. The genome of ATCC 27673 contained 96 unique genes

compared to the genomes of both Bl-04 and ATCC 25527. These 96 unique genes make up 5.94% of the genes present, and this number differs significantly from the numbers of unique genes previously reported within this subspecies. When comparing strains of *B. animalis* subsp. *lactis*, 1,430 shared genes were identified. Recently, it was determined that the core genome of *B. animalis* subsp. *lactis* was comprised of 1,518 genes (19). As expected, addition of a new strain with a unique gene content decreased the size estimate of the core genome. *B. animalis* subsp. *lactis* ATCC 27673 and Bl-04 were found to share 84 genes not present in *B. animalis* subsp. *animalis* ATCC 25527. Only 29 genes were found in both *B. animalis* subsp. *lactis* ATCC 27673 and *B. animalis* subsp. *animalis* ATCC 25527 but absent in *B. animalis* subsp. *lactis* Bl-04. *B. animalis* subsp. *animalis* ATCC 25527 contains 127 unique genes compared to the gene content of two strains of *B. animalis* subsp. *lactis*, thus showing a closer relation-

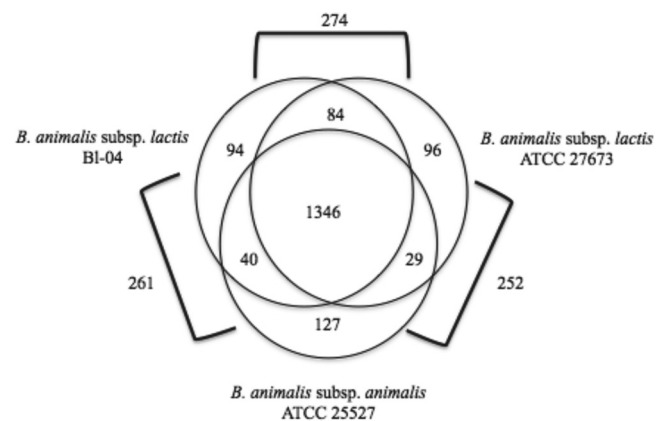
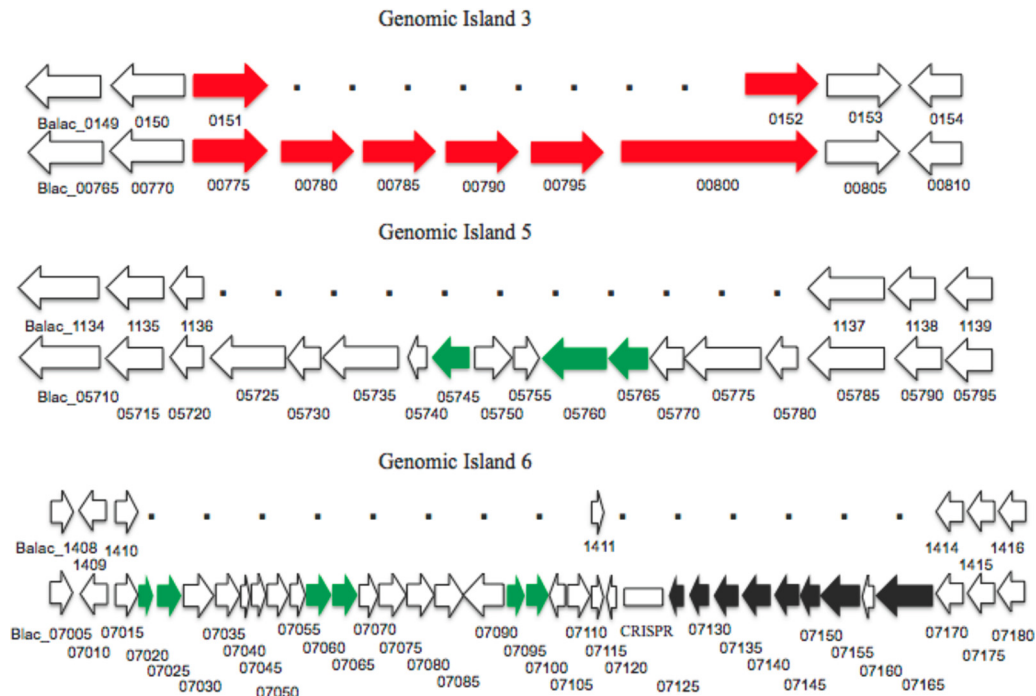


FIG 2 Shared and unique genes between three groups of *B. animalis*. Unique genes were identified as having less than 60% amino acid similarity between any pair of strains. On the basis of the similarity of the sequenced *B. animalis* subsp. *lactis* genomes, Bl-04 was chosen as a representative of the group.



**FIG 3** Genomic islands present in *B. animalis* subsp. *lactis* ATCC 27673. Genomic islands identified in *B. animalis* subsp. *lactis* ATCC 27673 (bottom row in each section) were compared to those identified in *B. animalis* subsp. *lactis* BI-04 (top row in each section). Red arrows, sugar transport genes; green arrows, mobile genetic elements; black arrows, CRISPR-associated genes; white arrows, hypothetical genes or genes not specified. Aligned genes from BI-04 and ATCC 27673 represent the gene homology flanking the genomic islands. Data for other genomic islands are presented in Fig. S3 in the supplemental material.

ship with BI-04 on the basis of the number of unique genes identified.

Compared to *B. animalis* subsp. *lactis* BI-04, the ATCC 27673 genome contained 6 genomic islands ranging in size from 6,747 to 40,966 bp. The gene content of these regions is provided in Table S1 in the supplemental material. Most of the genes found in the genomic islands have yet to be assigned functions. Genomic island 3 (Fig. 3A) contains four genes (Blac\_00780, Blac\_00785, Blac\_00790, and Blac\_00795) predicted to encode proteins related to sugar binding and transport. Three homologs of these genes were identified in the genome of *Bifidobacterium dentium* Bd1 by BLAST analysis of the nonredundant database (BDP\_125, BDP\_126, and BDP\_127), as Blac\_00780 and Blac\_00785 share a high degree of similarity to BDP\_125 (35).

Island 5 (Fig. 3B) contains 12 putative genes (Blac\_05725 to Blac\_05780), half with unassigned functions. In addition, island 5 contains the putative genes Blac\_05745 (partitioning protein ParA [36]), Blac\_05760, and Blac\_05765 (mobilization proteins MobA [37] and MobC [38]). Interestingly, the best BLAST matches of Blac\_05760 and Blac\_05765 were hits against *B. bifidum* PRL2010 MobA and MobC ( $E = 0.0$ ), respectively, which are also located on chromosomal DNA (39).

Island 6 (Fig. 3C), the largest island, is comprised of 29 ORFs. Eight of these are CRISPR-associated genes that are not found in other *B. animalis* subsp. *lactis* genomes but that have predicted functions similar to those of CRISPR-associated genes found in *B. animalis* subsp. *animalis* ATCC 25527. This locus contains 81 CRISPR spacers, which exhibit no homology with CRISPR spacers/repeats previously identified in other strains of *B. animalis* subsp. *lactis*. Another potentially interesting ORF is Blac\_07085,

which is annotated as an O-antigen polymerase (*wzy*) (40). BLASTP analysis showed that the best hit was to an O-antigen polymerase (*wzy*) identified in *Bacillus cereus* ATCC 10987 ( $E = 1e-14$ ) (41).

Four of the islands, islands 2, 4, 5, and 6, contain either insertion sequence (IS) elements, genes from bacteriophages putatively assigned functions, or other putative mobile elements. Islands 2 and 4 (see Fig. S3B and C in the supplemental material) contain genes predicted to encode phage integrase proteins (Blac\_00465 and Blac\_03510, respectively). Island 6 (Fig. 3C) contains four transposases (Blac\_07020, Blac\_07060, Blac\_07065, and Blac\_07095) as well as two transposase subunits (Blac\_07025 and Blac\_07100). Furthermore, out of the 125 unique genes identified between ATCC 27673 and BI-04, 83 (66.4%) were found to be in genomic islands.

In an effort to better understand gene differences between ATCC 27673 and other *B. animalis* subsp. *lactis* strains, BI-04 was chosen as a representative genome for unique gene content analysis. Six genomic islands were also identified in BI-04 (see Table S2 in the supplemental material). A total of 134 unique genes were identified, with 74 genes being observed in genomic islands. This number represents 55.2% of the total unique genes identified in BI-04. This is a lower percentage of genes present in genomic islands than the percentage identified in ATCC 27673. Island 1 is the largest island, with 25 ORFs, in the genome of BI-04. This region is in the same relative location in which genomic islands in *B. animalis* subsp. *lactis* ATCC 27673 and *B. animalis* subsp. *animalis* ATCC 25527 have been identified, indicating a region of variability across *B. animalis* genomes. Two islands contained mobile genetic elements: in island 3, an integrase (Balac\_1179) and a

TABLE 2 Total SNPs identified between strains of *B. animalis* subsp. *lactis*

Strain	Total no. of SNPs identified <sup>a</sup>								
	DSM 10140	Bl-04	Bb-12	B420	Bi-07	BLC1	CNCM I-2494	V9	Bl12
ATCC 27673	12,053	12,666	12,679	13,187	12,861	13,160	12,322	12,600	12,558
DSM 10140		47	387	43	58	55	150	56	59
Bl-04			358	12	11	24	123	25	28
Bb-12				340	352	365	407	374	358
B420					19	123	123	21	19
Bi-07						30	139	30	34
BLC1							153	37	32
CNCM I-2494								149	142
V9									31

<sup>a</sup> Total number of SNPs identified between pairs of genomes identified in the progressiveMauve program.

putative phage prohead protease (Balac\_1191) were identified, and island 6 contained a putative plasmid transfer protein (Balac\_1443) and a phage integrase (Balac\_1448). Island 4 contained the *cas* genes adjacent to the CRISPR spacer/repeat locus, representing a distinct CRISPR-*cas* system from *B. animalis* subsp. *lactis* ATCC 27673.

**SNPs.** In an effort to further assess the diversity of *B. animalis* subsp. *lactis*, SNPs were identified in a pairwise approach (Table 2). As expected, *B. animalis* subsp. *lactis* ATCC 27673 showed the greatest number of SNPs, clearly differentiating it from strains of the subspecies. In comparison, *B. animalis* subsp. *animalis* ATCC 25527 was reported to contain 73,021 SNPs compared to the sequence of DSM 10140 (20), whereas ATCC 27673 contained 12,053 SNPs compared to the sequence of DSM 10140. SNP analysis among non-ATCC 27673 strains of *B. animalis* subsp. *lactis* revealed Bb-12 to be the most diverse strain. Bb-12 was found to contain between 340 and 407 SNPs compared to the sequences of all other previously sequenced *B. animalis* subsp. *lactis* strains. Further analysis of the SNPs identified between the sequences of Bb-12 and the other sequenced strains revealed that the majority of SNPs reside in regions of relatively high density, suggesting regions of hypervariability or regions of imperfect sequence. The high degree of similarity between strains of *B. animalis* subsp. *lactis* is highlighted by analysis of Bi-07 and Bl-04, which revealed

only 11 SNPs, and B420 and Bl-04, which revealed only 12 SNPs. The recently sequenced strain Bl12 showed the most similarity with B420, with only 19 SNPs being exhibited between the two strains.

SNPs were also evaluated to determine whether they were synonymous or nonsynonymous using the publicly available annotations. The majority of SNPs identified were within open reading frames and were nonsynonymous. This is in accordance with the findings of a previous comparison between DSM 10140 and Bl-04 (12), where greater numbers of SNPs were found to be nonsynonymous.

**CRISPR-*cas* system.** A novel CRISPR locus was identified in the genome of *B. animalis* subsp. *lactis* ATCC 27673. This locus contained 81 spacers/82 repeats, 8 CRISPR-associated (*cas*) genes belonging to the type I-E CRISPR-*cas* system, and the signature type I *cas3* gene (22) with repeat sequence 5'-GTGTTCCCCGCAAGCGGGGATGATCCC-3'. Some degeneracy existed in the repeat sequence, with the alternative repeat sequence 5'-GTGTTCCCCGCAAGCGGGGATGATCCT-3' existing as 26 repeats toward the terminal end of the locus and an additional 8 distinct repeats existing once as the last eight repeats of the locus (Fig. 4). Although *B. animalis* subsp. *animalis* ATCC 25527 also contains a type I-E CRISPR-*cas* system, the presence of a type I-E system in ATCC 27673 stands in stark contrast to the system present in all

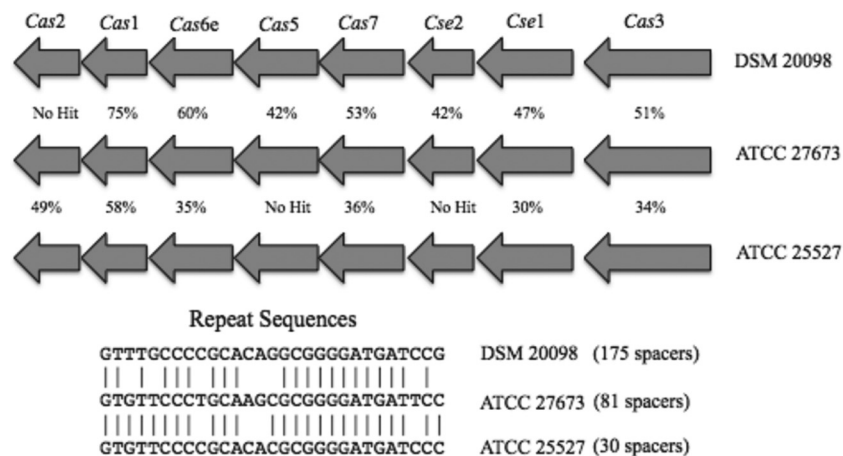


FIG 4 (Top) CRISPR-*cas* system of *B. animalis* subsp. *lactis* ATCC 27673 and comparison to similar systems. Gene similarity represents the percent amino acid similarity between genes of the same assigned function. (Bottom) In addition, repeats from the CRISPR array were aligned with matches to ATCC 27673, highlighted by connecting lines. ATCC 27673, *B. animalis* subsp. *lactis* ATCC 27673; ATCC 25527, *B. animalis* subsp. *animalis* ATCC 25527; DSM 20098, *B. animalis* subsp. *animalis* DSM 20098.

other *B. animalis* subsp. *lactis* strains, which contain a type I-C CRISPR-*cas* system. This is the first report of the presence of a unique (non-type I-C) CRISPR-*cas* system in *B. animalis* subsp. *lactis*. To date, three versions of the CRISPR loci have been identified in *B. animalis* subsp. *lactis*. The previously identified CRISPR loci in *B. animalis* subsp. *lactis* were type I-C and contained either 23 repeats, 20 repeats, or 19 repeats (12, 19, 42). The CRISPR locus in ATCC 27673 contained 82 repeats and 81 spacers, but none of these matched any of the repeat/spacers present in any other strain of *B. animalis* subsp. *lactis*. This locus was interrupted by 280 bp of seemingly random sequence, which is possibly a degenerative repeat/spacer sequence. BLAST analysis of the ATCC 27673 *cas3* DNA sequence resulted in 263 nucleotides of perfect homology in an intergenic region present in the other *B. animalis* subsp. *lactis* genomes. Also of note, this region of perfect homology was immediately adjacent to a transposase (Balac\_1413) at the same relative position in the genome of BI-04 (and other *B. animalis* subsp. *lactis*) strains, possibly suggesting a deletion event.

Although there are no known phages that possess the ability to infect *Bifidobacterium*, the importance of CRISPR is highlighted by the general lack of other complete phage resistance mechanisms in bifidobacteria (19, 43). The number of spacers suggests that either at some time in the past or currently, CRISPR played an important role as a defense mechanism. Due to the lack of information about phages able to interact with bifidobacteria, spacer sequences were BLAST analyzed for sequence homology. Most spacer sequences exhibited no significant similarity to known sequences in the nonredundant (nr/nt), whole-genome shotgun contigs (wgs), or genomic survey sequences (gss) databases. One spacer (S55) with the sequence 5'-GCCACGCGTGAAATCGA TGCGTGTGGCCGTG-3' was a perfect hit against a genome sequence in *B. animalis* subsp. *animalis* ATCC 25527. The spacer sequence was found in Banan\_07350, which is annotated as a hypothetical protein. Interestingly, this spacer was a perfect match against the genome sequence of a closely related organism and not that of a bacteriophage or a plasmid, perhaps indicating that in addition to involvement in immunity against invasive genetic elements, CRISPR-*cas* systems may also play a role in the horizontal transfer of DNA, as previously documented (44, 45).

## DISCUSSION

The overall architecture of the genome of ATCC 27673 was similar to that of the genomes of other sequenced strains of *B. animalis* subsp. *lactis*, but significant sequence diversity was also seen across the genome. The genomes of previously sequenced strains of *B. animalis* subsp. *lactis* exhibited tremendous sequence similarity, and the subspecies has been termed “monomorphic” or “monophyletic” (19). The monomorphic nature observed in the previously sequenced genomes may exist for biological or nonbiological reasons. One possible nonbiological reason for the high degree of genome similarity observed is that most strains chosen for genome sequencing have been of industrial significance as probiotics. Because of this, many industrial strains would have been selected from environmental samples on the basis of similar criteria. This may have artificially reduced the diversity detected within the subspecies because of selection bias. It also a possibility that new strains simply represent existing strains that have been reisolated and renamed. ATCC 27673 has not been associated with use as a probiotic food additive, which lends credibility to this suggestion.

Some authors have argued that isolates of *B. animalis* subsp. *lactis* recovered from human subjects represent unique strains (15, 19). While this is potentially true, the possibility that human isolates simply represent isolates of widely distributed commercial strains consumed by these individuals that have been reisolated remains likely. Isolation of new strains is also complicated by the fact that the natural habitat of *B. animalis* subsp. *lactis* has yet to be established.

Biological reasons for genetic monomorphism include the possibility that *B. animalis* subsp. *lactis* is well adapted to its specific environment. Additionally, there may have been a drastic alteration in this organism's ecological niche that allowed the survival of only a specific lineage. Without knowing the organism's natural reservoir, this hypothesis cannot be tested. Perhaps the simplest explanation for the monomorphism observed is a recent evolutionary bottleneck that eliminated diversity, as previously suggested (46).

Monomorphic lineages have previously been defined to “have such low levels of sequence diversity that only few polymorphisms, or even none at all, are found upon sequencing a few genes” (46). As shown with previously sequenced strains, *B. animalis* subsp. *lactis* certainly fits this criterion. Several bacterial lineages have previously been identified as monomorphic pathogens: *Mycobacterium leprae*, *Burkholderia mallei*, *Bordetella pertussis*, *Yersinia pestis*, *Salmonella enterica* serovar Typhi, *Bacillus anthracis*, and *Mycobacterium tuberculosis* (47). SNPs within the core genomes of each lineage were identified and ranged from 7 to 226 SNPs per 100 kb (47–54). In comparison, across nine full genomes of *B. animalis* subsp. *lactis* (not including the ATCC 27673 genome), 29 SNPs per 100 kb were identified. The addition of *B. animalis* subsp. *lactis* strain ATCC 27673 to this analysis resulted in 685 SNPs per 100 kb, highlighting the highly monomorphic nature of the previously sequenced strains. Even with the addition of ATCC 27673, the group would still be considered monomorphic.

Unique genes have been identified between ATCC 27673, BI-04, and ATCC 25527, and the majority of the unique content was found to be contained in genomic islands. Four of the genomic islands identified in ATCC 27673 contained either transposases, phage elements, or other mobile genetic elements, suggesting that much of the diversity within this subspecies was acquired via horizontal gene transfer. All six genomic islands identified by unique gene analysis were selected as strong candidates for HGT by Alien Hunter software. This suggests that much of the diversity observed in ATCC 27673 has been acquired via horizontal evolution.

One region identified as a strong candidate for HGT is genome island 6, which contains the CRISPR-*cas* system. Because they share the same CRISPR type, the *cas* proteins of *B. animalis* subsp. *lactis* ATCC 27673 were compared to those of *B. animalis* subsp. *animalis* ATCC 25527 and *B. angulatum* DSM 20098. Despite belonging to the same CRISPR type, similarity between proteins was found to be low, reflecting divergent evolution. CRISPR repeat sequences were also compared between these three strains (ATCC 27673, ATCC 25527, and DSM 20098) as well as to the CRISPR repeat sequences from *Escherichia coli* CRISPR1/2 and CRISPR3/4 (55, 56). Alignment of ATCC 27673 CRISPR repeats revealed 4 SNPs compared with the sequence of ATCC 25527 and 8 SNPs compared with the sequence of DSM 20098, whereas alignment

with either set of *E. coli* CRISPR repeat sequences revealed very little homology.

Recently, strain Bl12 was selected for sequencing in an attempt to assess diversity in the *B. animalis* subsp. *lactis* group (19). This strain, reportedly isolated from a person who did not consume probiotics, was chosen for sequencing on the basis of its increased susceptibility to tetracycline (MIC, 16 µg/ml, half of the MIC observed for Bb-12). Despite selection on the basis of phenotypic characteristics, the addition of this new genome sequence did not add to the genetic variability of the subspecies.

Despite the lack of genetic variability, strains of *B. animalis* subsp. *lactis* have been shown to differ in resistance to oxidative stress (57). This suggests that small differences in phenotypic characteristics may not be sufficient when selecting strains to assess diversity within this subspecies. In addition, sequenced strains isolated from various corners of the globe (e.g., Asia, North America, Europe) have not previously yielded genetic variability. The strain ATCC 27673 was isolated from sewage, most likely in Europe (58), where strains have been previously isolated, but this strain has yielded a great amount of additional genetic diversity. None of the previously sequenced strains were isolated from sewage, an environment that may have allowed increased genetic variation.

**Conclusions.** We report the complete sequence and analysis of the genome of *B. animalis* subsp. *lactis* ATCC 27673, a genetically distinct strain within this genetically monomorphic subspecies. Six genomic islands were identified in comparison to the genome of Bl-04 (representative of all other sequenced strains of *B. animalis* subsp. *lactis*). In most of the genomic islands, transposases, phage elements, or other mobile genetic elements were identified, suggesting that HGT is a possible driver of diversity in this subspecies. A novel CRISPR locus which differs from other *B. animalis* subsp. *lactis* loci in terms of CRISPR-*cas* system type, as well as spacer number and content, was identified. Analysis of ATCC 27673 provides evidence that the true diversity within this subspecies has yet to be explored and potentially unattributed phenotypes in this subspecies which may be beneficial for industrial or human health purposes exist. We propose that the most probable explanation for the lack of diversity within *B. animalis* subsp. *lactis* is the intense focus on commercially relevant strains and the likely re-isolation of these strains and their assignment as new strains.

## ACKNOWLEDGMENTS

We thank The Pennsylvania State University Genomics Core Facility for assistance with Sanger sequencing. Also, we thank Stephan Schuster for assistance with 454 sequencing and assembly and Jihye Park for bioinformatics assistance.

## REFERENCES

- Gueimonde M, Delgado S, Mayo B, Ruas-Madiedo P, Margolles A, de los Reyes-Gavilán CG. 2004. Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. *Food Res. Int.* 37:839–850.
- Masco L, Ventura M, Zink R, Huys G, Swings J. 2004. Polyphasic taxonomic analysis of *Bifidobacterium animalis* and *Bifidobacterium lactis* reveals relatedness at the subspecies level: reclassification of *Bifidobacterium animalis* as *Bifidobacterium animalis* subsp. *animalis* subsp. nov. and *Bifidobacterium lactis* as *Bifidobacterium animalis* subsp. *lactis* subsp. nov. *Int. J. Syst. Evol. Microbiol.* 54:1137–1143.
- Rizzardini G, Eskesen D, Calder PC, Capetti A, Jespersen L, Clerici M. 2012. Evaluation of the immune benefits of two probiotic strains *Bifidobacterium animalis* ssp. *lactis*, BB-12(R) and *Lactobacillus paracasei* ssp. *paracasei*, *L. casei* 431(R) in an influenza vaccination model: a randomized, double-blind, placebo-controlled study. *Br. J. Nutr.* 107:876–884.
- Guyonnet D, Woodcock A, Stefani B, Trevisan C, Hall C. 2009. Fermented milk containing *Bifidobacterium lactis* DN-173 010 improved self-reported digestive comfort amongst a general population of adults. A randomized, open-label, controlled, pilot study. *J. Dig. Dis.* 10:61–70.
- Marteau P, Cuillerier E, Meanace S, Gerhardt MF, Myara A, Bouvier M, Bouley C, Tondou F, Bommelaer G, Grimaud JC. 2002. *Bifidobacterium animalis* strain DN-173 010 shortens the colonic transit time in healthy women: a double-blind, randomized, controlled study. *Aliment. Pharmacol. Ther.* 16:587–593.
- Simpson PJ, Stanton C, Fitzgerald GF, Ross RP. 2005. Intrinsic tolerance of *Bifidobacterium* species to heat and oxygen and survival following spray drying and storage. *J. Appl. Microbiol.* 99:493–501.
- Matsumoto M, Ohishi H, Benno Y. 2004. H<sup>+</sup>-ATPase activity in *Bifidobacterium* with special reference to acid tolerance. *Int. J. Food Microbiol.* 93:109–113.
- Vernazza CL, Gibson GR, Rastall RA. 2006. Carbohydrate preference, acid tolerance and bile tolerance in five strains of *Bifidobacterium*. *J. Appl. Microbiol.* 100:846–853.
- Sanchez B, Champomier-Verges MC, Stuer-Lauridsen B, Ruas-Madiedo P, Anglade P, Baraige F, de los Reyes-Gavilán CG, Johansen E, Zagorec M, Margolles A. 2007. Adaptation and response of *Bifidobacterium animalis* subsp. *lactis* to bile: a proteomic and physiological approach. *Appl. Environ. Microbiol.* 73:6757–6767.
- Kailasapathy K, Harmstorff I, Phillips M. 2008. Survival of *Lactobacillus acidophilus* and *Bifidobacterium animalis* ssp. *lactis* in stirred fruit yogurts. *LWT Food Sci. Technol.* 41:1317–1322.
- Meile L, Ludwig W, Reuger U, Gut C, Kaufmann P, Dasen G, Wenger S, Teuber M. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. *Syst. Appl. Microbiol.* 20: 57–64.
- Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, Horvath P, Coute-Monvoisin AC, Leyer G, Rendulic S, Steele JL, Broadbent JR, Oberg T, Dudley EG, Schuster S, Romero DA, Roberts RF. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J. Bacteriol.* 191:4144–4151.
- Kim JF, Jeong H, Yu DS, Choi SH, Hur CG, Park MS, Yoon SH, Kim DW, Ji GE, Park HS, Oh TK. 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. *J. Bacteriol.* 191: 678–679.
- Garrigues C, Johansen E, Pedersen MB. 2010. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. *J. Bacteriol.* 192:2467–2468.
- Sun Z, Chen X, Wang J, Gao P, Zhou Z, Ren Y, Sun T, Wang L, Meng H, Chen W, Zhang H. 2010. Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. *J. Bacteriol.* 192:4080–4081.
- Chervaux C, Grimaldi C, Bolotin A, Quinquis B, Legrain-Raspaud S, van Hylckama Vlieg JE, Denariac G, Smokvina T. 2011. Genome sequence of the probiotic strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494. *J. Bacteriol.* 193:5560–5561.
- Bottacini F, Dal Bello F, Turrone F, Milani C, Duranti S, Foroni E, Viappiani A, Strati F, Mora D, van Sinderen D, Ventura M. 2011. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BLC1. *J. Bacteriol.* 193:6387–6388.
- Stahl B, Barrangou R. 2012. Complete genome sequences of probiotic strains of *Bifidobacterium animalis* subsp. *lactis* B420 and Bi-07. *J. Bacteriol.* 194:4131–4132.
- Milani C, Duranti S, Lugli GA, Bottacini F, Strati F, Arioli S, Foroni E, Turrone F, van Sinderen D, Ventura M. 2013. Comparative genomics of *Bifidobacterium animalis* subsp. *lactis* reveals a strict monophyletic bifidobacterial taxon. *Appl. Environ. Microbiol.* 79:4304–4315.
- Loquasto JR, Barrangou R, Dudley EG, Roberts RF. 2011. Short communication: the complete genome sequence of *Bifidobacterium animalis* subspecies *animalis* ATCC 25527(T) and comparative analysis of growth in milk with *B. animalis* subspecies *lactis* DSM 10140(T). *J. Dairy Sci.* 94:5864–5870.
- Makarova KS, Koonin EV. 2013. Evolution and classification of CRISPR-cas systems and Cas protein families, p 61–91. In Barrangou R, van der Oost J (ed), *CRISPR-cas systems*. Springer-Verlag, Berlin, Germany.

22. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV. 2011. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9:467–477.
23. Delétoile A, Passet V, Aires J, Chambaud I, Butel MJ, Smokvina T, Brisse S. 2010. Species delineation and clonal diversity in four *Bifidobacterium* species as revealed by multilocus sequencing. *Res. Microbiol.* 161: 82–90.
24. de Man JD, Rogosa M, Sharpe ME. 1960. A medium for the cultivation of lactobacilli. *J. Appl. Bacteriol.* 23:130–135.
25. Ventura M, Zink R. 2002. Rapid identification, differentiation, and proposed new taxonomic classification of *Bifidobacterium lactis*. *Appl. Environ. Microbiol.* 68:6429–6434.
26. Zhao F, Zhao F, Li T, Bryant DA. 2008. A new pheromone trail-based genetic algorithm for comparative genome assembly. *Nucleic Acids Res.* 36:3455–3462.
27. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi:10.1371/journal.pone.0011147.
28. Ning Z, Cox AJ, Mullikin JC. 2001. SSAHA: a fast search method for large DNA databases. *Genome Res.* 11:1725–1729.
29. 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.
30. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuicchio M, Federhen S, Feolo M, Fingerhman IM, Geer LY, Helmberg W, Kapustin Y, Krasnov S, Landsman D, Lipman DJ, Lu Z, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Karsch-Mizrachi I, Ostell J, Panchenko A, Phan L, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Slotta D, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Wang Y, Wilbur WJ, Yaschenko E, Ye J. 2012. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 40:D13–D25. doi:10.1093/nar/gkr1184.
31. Sonnhammer ELL, Durbin R. 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* 167:GC1–GC10.
32. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35:W52–W57. doi:10.1093/nar/gkm360.
33. Vernikoff GS, Parkhill J. 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. *Bioinformatics* 22:2196–2203.
34. Ventura M, Reniero R, Zink R. 2001. Specific identification and targeted characterization of *Bifidobacterium lactis* from different environmental isolates by a combined multiplex-PCR approach. *Appl. Environ. Microbiol.* 67:2760–2765.
35. Ventura M, Turroni F, Zomer A, Foroni E, Giubellini V, Bottacini F, Canchaya C, Claesson MJ, He F, Mantzourani M, Mulas L, Ferrarini A, Gao B, DelleDonne M, Henrissat B, Coutinho P, Oggioni M, Gupta RS, Zhang Z, Beighton D, Fitzgerald GF, O'Toole PW, van Sinderen D. 2009. The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. *PLoS Genet.* 5:e1000785. doi:10.1371/journal.pgen.1000785.
36. Lee PS, Grossman AD. 2006. The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol. Microbiol.* 60:853–869.
37. Bhattacharjee MK, Meyer R. 1993. Specific binding of MobA, a plasmid-encoded protein involved in the initiation and termination of conjugal DNA transfer, to single stranded oriT DNA. *Nucleic Acids Res.* 21:4536–4568.
38. Zhang S, Meyer R. 1997. The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. *Mol. Microbiol.* 25:509–516.
39. Turroni F, Bottacini F, Foroni E, Mulder I, Kim J-H, Zomer A, Sanchez B, Bidossi A, Ferrarini A, Giubellini V, DelleDonne M, Henrissat B, Coutinho P, Oggioni M, Fitzgerald GF, Mills D, Margolles A, Kelly D, van Sinderen D, Ventura M. 2010. Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc. Natl. Acad. Sci. U. S. A.* 107:19514–19519.
40. Kim TH, Sebastian S, Pinkham JT, Ross RA, Blalock LT, Kasper DL. 2010. Characterization of the O-antigen polymerase (Wzy) of *Francisella tularensis*. *J. Biol. Chem.* 285:27839–27849.
41. Rasko DA, Ravel J, Oekstad OA, Helgason E, Cer RZ, Jiang L, Shores KA, Fouts DE, Tourasse NJ, Angiuoli SV, Kolonay J, Nelson WC, Kolstü A-B, Fraser CM, Read TD. 2004. The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. *Nucleic Acids Res.* 32:977–988.
42. Briczinski EP, Loquasto JR, Barrangou R, Dudley EG, Roberts AM, Roberts RF. 2009. Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* by using single-nucleotide polymorphisms, insertions, and deletions. *Appl. Environ. Microbiol.* 75:7501–7508.
43. Ventura M, Turroni F, Lima-Mendez G, Foroni E, Zomer A, Duranti S, Giubellini V, Bottacini F, Horvath P, Barrangou R, Sela DA, Mills DA, van Sinderen D. 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. *Appl. Environ. Microbiol.* 75:6929–6936.
44. Jorth P, Whiteley M. 2012. An evolutionary link between natural transformation and CRISPR adaptive immunity. *mBio* 3(5):e00309-12. doi:10.1128/mBio.00309-12.
45. Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. 2012. CRISPR interference can prevent natural transformation and virulence acquisition during *in vivo* bacterial infection. *Cell Host Microbe* 12:177–186.
46. Achtman M. 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu. Rev. Microbiol.* 62:53–70.
47. Achtman M. 2012. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367:860–867.
48. Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, Feldkamp M, Kusecek B, Vogler AJ, Li Y, Cui Y, Thomson NR, Jombart T, Leblois R, Lichtner P, Rahalison L, Petersen JM, Balloux F, Keim P, Wirth T, Ravel J, Yang R, Carniel E, Achtman M. 2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat. Genet.* 42:1140–1143.
49. Monot M, Honore N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, Matsuoka M, Taylor GM, Donoghue HD, Bouwman A, Mays S, Watson C, Lockwood D, Khamesipour A, Dowlati Y, Jianping S, Rea TH, Vera-Cabrera L, Stefani MM, Banu S, Macdonald M, Sapkota BR, Spencer JS, Thomas J, Harshman K, Singh P, Busso P, Gattiker A, Rougemont J, Brennan PJ, Cole ST. 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat. Genet.* 41:1282–1289.
50. Losada L, Ronning CM, DeShazer D, Woods D, Fedorova N, Kim HS, Shabalina SA, Pearson TR, Brinkac L, Tan P, Nandi T, Crabtree J, Badger J, Beckstrom-Sternberg S, Saqib M, Schutzer SE, Keim P, Nierman WC. 2010. Continuing evolution of *Burkholderia mallei* through genome reduction and large-scale rearrangements. *Genome Biol. Evol.* 2:102–116.
51. Bart MJ, van Gent M, van der Heide HG, Boekhorst J, Hermans P, Parkhill J, Mooi FR. 2010. Comparative genomics of prevaccination and modern *Bordetella pertussis* strains. *BMC Genomics* 11:627. doi:10.1186/1471-2164-11-627.
52. Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, Dolecek C, Achtman M, Dougan G. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat. Genet.* 40:987–993.
53. Kuroda M, Serizawa M, Okutani A, Sekizuka T, Banno S, Inoue S. 2010. Genome-wide single nucleotide polymorphism typing method for identification of *Bacillus anthracis* species and strains among *B. cereus* group species. *J. Clin. Microbiol.* 48:2821–2829.
54. Comas I, Chakravarti J, Small PM, Galagan J, Niemann S, Kremer K, Ernst JD, Gagneux S. 2010. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat. Genet.* 42:498–503.
55. Diez-Villasenor C, Almendros C, Garcia-Martinez J, Mojica FJ. 2010. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 156:1351–1361.
56. Touchon M, Rocha EP. 2010. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS One* 5:e11126. doi:10.1371/journal.pone.0011126.
57. Oberg TS, Steele JL, Ingham SC, Smeianov VV, Briczinski EP, Abdalla A, Broadbent JR. 2011. Intrinsically and inducible resistance to hydrogen peroxide in *Bifidobacterium* species. *J. Ind. Microbiol. Biotechnol.* 38: 1947–1953.
58. Scardovi V, Trovatielli LD. 1974. *Bifidobacterium animalis* (Mitsuoka) comb. nov. and the “minimum” and “subtile” groups of new bifidobacteria found in sewage. *Int. J. Syst. Bacteriol.* 24:21–28.