The Pennsylvania State University

The Graduate School

College of Agricultural Sciences

ASSESSING THE DIVERSITY OF THE MONOMORPHIC SUBSPECIES *BIFIDOBACTERIUM ANIMALIS* SUBSP. *LACTIS*

A Dissertation in

Food Science

by

Joseph R. Loquasto

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2013

The dissertation of Joseph R. Loquasto was reviewed and approved* by the following

Robert F. Roberts
Department Head of Food Science
Professor of Food Science
Co-Dissertation Adviser
Chair of Committee

Edward G. Dudley
Casida Development Professor of Food Science
Associate Professor of Food Science
Co- Dissertation Adviser

Rodolphe Barrangou
Adjunct Professor of Food Science
Associate Professor of Food, Bioprocessing and Nutrition Sciences
North Carolina State University

Eric Harvill
Professor of Microbiology and Infectious Disease

*Signatures are on file in the Graduate School

Abstract

Bifidobacterium animalis subsp. *lactis* is a widely consumed probiotic microorganism commonly added to fermented dairy products such as yogurt. Health benefits associated with this subspecies are considered to be strain-specific, thus proper identification of these strains is critical. However, the identification and differentiation of strains has remained difficult using both phenotypic and molecular methods. Prior to this work very little genomic sequence existed for strains of this subspecies and no completed genomes were publically available. To develop a better understanding of the subspecies *B. animalis* subsp. *lactis*, the genome of a commercial strain Bl-04 and the Type strain DSM 10140 were sequenced and compared.

Comparison of the DSM 10140 and Bl-04 genome sequences (99.975% identical) revealed a high degree of identity and synteny as well as a total of 47 single nucleotide polymorphism (SNPs) and 4 insertion/deletions (INDELs). One non-synonymous SNP was identified in a putative glucose uptake protein (*glcU*). The two strains were shown to have differential ability to grow on glucose as the sole carbohydrate source. The high degree of similarity observed between these two strains explains the difficulties encountered in differentiation of strains of this subspecies.

A collection of 24 strains of *B. animalis* subsp. *lactis* were screened using the SNPs/INDELs identified between DSM 10140 and Bl-04 as possible targets. Results obtained from this analysis revealed a combination of nine SNPs/INDELs could be used to differentiate strains into 14 distinct genotypic groups. The method reported here is the first available for *B. animalis* subsp. *lactis* and can be used in clinical, regulatory, and commercial applications requiring identification of *B. animalis* subsp. *lactis* at the strain

level.

The genome of the closely related *B. animalis* subsp. *animalis* ATCC 25527 was sequenced in an effort to understand what genomic features differentiate the two subspecies. Comparative genomic analysis revealed 156 and 182 genes that were unique to and absent in the *B. animalis* subsp. *animalis* genome when compared to *B. animalis* subsp. *lactis* genome, respectively. Among the differential content was a set of unique clustered regularly interspaced short palindromic repeats (CRISPR)-associated genes and a novel CRISPR locus containing 30 spacers in the genome of *B. animalis* subsp. *animalis*. Although, previous research had suggested the ability to grow in milk was a differential phenotype between these two subspecies. Analysis revealed there is no significant difference between the subspecies in their ability to grow in milk and no differential gene content that would lead to such a result. Supplementation with a protein hydrolysate allowed growth in milk by both subspecies suggesting lack of the ability to utilize milk proteins as the factor limiting growth in milk.

Subsequent to publication of the complete genome of *B. animalis* subsp. *lactis* DSM 10140 and Bl-04, seven additional strains of the subspecies have been reported and deposited in GenBank. These genomes reveal remarkably little diversity leading to the term "monomorphic" being applied to this group. *B. animalis* subsp. *lactis* ATCC 27673 was selected for sequencing based on the report of a unique MLST profile when compared to other strains of the subspecies.

The complete genome of ATCC 27673 was 1,963,012 bp and contained 1,616 genes, 4 rRNA operons, and had a G+C content of 61.55%. Comparative analyses of the typical *B. animalis* subsp. *lactis* revealed the genome of ATCC 27673 contained six

distinct genomic islands, encoding 83 ORFs that are absent from other strains of the same subspecies. Phage or mobile genetic elements were identified in four of these genomic islands. In island 6, the largest island, a novel type I-E CRISPR-cas system was identified, which contains 81 novel spacers. It is noteworthy that all other sequenced strains of this subspecies contain type I-C systems. This study revealed ATCC 27673 is a strain of *B. animalis* subsp. *lactis* with novel genetic content, suggesting the lack of genetic variability observed to date is potentially due to the repeated sequencing of a limited number of widely distributed commercial strains.

In conclusion, the 47 SNPs and 4 INDELs identified between the sequenced strains DSM 10140 and Bl-04 separated the collection into 14 distinct genomic clusters. *B. animalis* subsp. *animalis* ATCC 25527 was sequenced and compared to sequenced strains of *B. animalis* subsp. *lactis* and the previously identified defining phenotype that separates the two subspecies was examined and no differences were observed when both subspecies were grown in milk. *B. animalis* subsp. *lactis* ATCC 27673 was identified as a potentially novel strain of the subspecies and was sequenced. Six genomic islands containing novel lactis genomic sequence were identified. The identification of this unique strain suggests more unique strains exist and the full diversity of the subspecies has not yet been explored.

Table of Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Figures

Chapter 2

Chapter 3

List of Abbreviations

%G+C	Percent guanine and cytosine
ABC	ATP-binding cassette
ANOVA	Analysis of variance
ATCC	American Type Culture Collection
BLAST	Basic local alignment search tool
bp	Base pair
CFu	Colony forming units
COG	Clusters of orthologous groups
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkultrun GMbH
F6PPK	Fructose-6-phosphate phosphoketolase
FAO/WHO	Food and Agriculture Organization/World Health Organization
GIT	Gastrointestinal tract
INDEL	Insertion/deletion
h	Hour
HGT	Horizontal gene transfer
ITS	Internally transcribed spacer
LG	Liver glucose
LL	Liver lactose
M+	RSM + peptone + casamino acids

Mb	Megabase
MC	RSM + casamino acids
mL	Milliliter
Min	Minute
MRS	de Man Rogosa and Sharpe
MRSC	MRS with cysteine hydrochloride
ORF	Open reading frame
NCBI	National Center for Biotechnology Information
PCR	Polymerase Chain Reaction
PCR-DGGE	PCR-Denaturing Gel Gradient Electrophoresis
PFGE	Pulsed-field gel electrophoresis
PGAAP	Prokaryotic Genomes Automatic Annotation Pipeline
RAPD-PCR	Randomly amplified polymorphic DNA-PCR
RAST	Rapid Annotation using Subsystems Technology
RB	Roberts-Briczinski culture collection
rDNA	Ribosomal DNA
rRNA	Ribosomal RNA
RSM	Reconstituted skim milk
SNP	Single nucleotide polymorphism
tRNA	Transfer RNA

Acknowledgements

I would first like to thank Dr. Roberts for his many years of support and guidance starting in my undergraduate and continuing throughout graduate school. I first started working for Dr. Roberts as a junior washing dishes in his lab. Since then, I have come a long way, learning and experiencing a great deal which is in no small order due to his guidance. Dr. Roberts has always challenged me and has positioned me to succeed in my endeavors.

I would also like to thank my committee, Dr. Ed Dudley, Dr. Rodolphe Barrangou, and Dr. Eric Harvill. Dr. Dudley has been incredibly helpful to my research and development as a scientist. He has provided me with great direction during this time of growth. Rodolphe has also been an invaluable resource. Despite not being on the Penn State campus, Rodolphe has always been willing to help, answer questions, and provide guidance when needed. Finally, I must express appreciation for Dr. Harvill's willingness to answer questions and provide insight.

I must thank Beth Briczinski, who provided me with an immeasurable amount of knowledge and mentoring as an undergraduate and young graduate student. Additionally, our lab simply would not run without the help of Emily Furumoto, who always takes care of lab issues and makes sure that we are always well fed.

In addition I would like to express gratitude to the entire Food Science Department faculty, staff, and graduate students, who have been incredibly helpful throughout my time here and has made it feel like a home.

Finally, I would like to thank Meriel and my family for their love and support.

Chapter 1
Literature Review

The genus *Bifidobacterium*

Many of the microorganisms added to food as probiotics belong to the genus *Bifidobacterium*. Members of the genus *Bifidobacterium* have a high G+C content, around 60%, have genomes ranging in size from 1.9 to 2.8 Mb, and have been associated with the intestine of mammals, insects, or isolated from milk (Lee and O'Sullivan, 2010, Turroni et al., 2011). Generally, bifidobacteria are added to fermented dairy products, or consumed in freeze-dried pill form, for their perceived health benefits. Several species of *Bifidobacterium* are used as and considered probiotic including, *B. adolescentis, B. bifidum, B. breve, B. longum* subsp. *longum, B. longum* subsp. *infantis,* and *B. animalis* subsp. *lactis* (Champagne et al., 2005).

Bifidobacteria were discovered by Tissier (1899) in the feces of a healthy infant and he named the organism *Bacillus bifidus*. This organism was later named *Lactobacillus bifidus* because of related morphological characteristics (gram-positive, non-sporeforming) and similar biochemical traits to *Lactobacillus acidophilus* (Poupard et al., 1973*)*. In 1924 Orla-Jensen suggested a new genus, *Bifidobacterium*, however the genus name was not widely accepted. What is now *Bifidobacterium bifidum* (originally *Bacillus bifidus*) endured 13 recognized name changes, belonging to ten different genera before the present, accepted taxonomy, which was introduced in the 8th edition of Bergey's Manual (Poupard et al., 1973). The majority of the members of the genus *Bifidobacterium* were originally isolated from the feces or intestines of humans or animals.

A characteristic feature of the genus *Bifidobacterium* is the activity of the enzyme fructose-6-phosphate phosphoketolase (F6PPK; EC 4.1.2.22). Detection of F6PPK activity is used to identify potential new members of the genus. The F6PPK pathway has been labeled the 'bifid shunt' because of its specificity to *Bifidobacterium.* The bifid shunt was once considered exclusive to the genus *Bifidobacterium*, however reclassification of certain bifidobacterial species, and creation of new genera, has resulted in the bifid shunt serving only as an indicator of the family *Bifidobacteriaceae* (Felis and Dellaglio, 2007). The end products of hexose fermentation via the bifid shunt result in a theoretical ratio of 3:2 of acetic acid to lactic acid (Biavati and Mattarelli, 2006). Phenotypically, an isolate is considered a member of the genus *Bifidobacterium* if it is a Gram positive, obligate anaerobe, non-motile, non-sporeforming, positive for F6PPK activity, end products of glucose fermentation are acetic and lactic acids, and catalase negative, as recently conducted by (Kim et al., 2010). Until recently, F6PPK activity and bifid shunt end products, were exploited in the majority of schemes designed to identify putative bifidobacterial isolates to the genus level. However, the inability of these characteristics to provide clear separation among genera, species, and strains, and the development of DNA-based techniques has made these methods obsolete. DNA-based method such as genus specific primers targeting the 16S rDNA-ITS region designed by Kaufmann et al. (1997) have gained in popularity. Additionally, the relative ease and rapidity of DNA-based methods, and abundance of 16S rDNA sequence data has allowed for accurate and reliable identification of bifidobacterial species.

Bifidobacterium animalis* subsp. *lactis

B. lactis, as it was first identified, was isolated and described by Meile et al. (1997) from a commercial yogurt sample. The *B. lactis* strain UR1 (later named DSM 10140), was selected from a sample of 60 bifidobacterial isolates as exhibiting the most oxygen-tolerance when grown in media with 50 ml (10% of total volume) of oxygen added to a 500 ml flask containing 400 ml of brain heart infusion containing 5 g/L of yeast extract, 0.5 g/L of L-cysteine hydrochloride and 0.001 g/L of resazurine. In an effort to assign the oxygen-tolerant isolate to a species, 16S rDNA sequence was amplified and compared to known *Bifidobacterium* sequences. The 16S rDNA *B. lactis* UR1 closely matched the 16S rDNA sequence from *B. animalis* with a similarity of 98.6%. Since the 16S rDNA sequence of UR1 did not perfectly match any of the known sequences Meile *et al.* classified UR1 as a new species. To further investigate the species status issue, the authors conducted DNA-DNA hybridization assay with *B. animalis* and *B. longum*. The authors observed less than 10% identity with *B. longum* and 27% similarity with *B. animalis*. The authors state the low DNA hybridization level did not meet the criterion proposed by Stackebrandt and Goebel (1994), and the phenotypic data supported UR1 as a separate species.

For several years *B. lactis* remained its own species until Cai et al. (2000) reclassified *B. lactis* as a subjective synonym of *B. animalis* based on DNA-DNA hybridizations. Two separate techniques were used to evaluate DNA relatedness. The microplate method revealed 92.3% DNA relatedness and the membrane filter method indicated 85.7% DNA relatedness between the type strains of *B. lactis* and *B. animalis* (DSM 10140 and ATCC 25527). *B. lactis* Bb-12 was also compared to *B. lactis* DSM

10140 for DNA relatedness. The microplate method revealed 95.9% DNA similarity, while the membrane filter method indicated 90.3% relatedness. Based on these findings it was suggested that *B. lactis* be considered a junior subjective synonym of *B. animalis.*

In 2001, Ventura et al. (2001) developed a set of PCR primers for the identification of *B. lactis*, Bflact2/Bflact5. This primer set targeted the 16S rDNA and the 16S-23S ITS (internally transcribed sequence) region. These primers were specific in the differentiation of *B. lactis* from *B. animalis*, a PCR amplicon would be observed with *B. lactis* strains but not with *B. animalis* strains. In a subsequent study, Ventura and Zink (2002), developed a primer set, Ban2 and 23Si, that targeted two 8 bp insertions in the 16S-23S ITS and 23S of *B. animalis* strains for the specific identification of the species. Based on the sequence variation observed by the authors, it was proposed *B. animalis* and *B. lactis* should not be considered subjective synonyms but rather *B. lactis* should be considered a subspecies of *B. animalis*.

Zhu et al. (2003) evaluated partial HSP60 sequence similarity among five *B. animalis* and two *B. lactis* strains. HSP60 sequence segregated the seven strains into two groups. The first group contained *B. lactis* DSM 10140, *B. lactis* JB-1, and *B. animalis* B83, the second group contained four strains of *B. animalis*. Based on this sequence analysis Zhu et al. (2003) concluded *B. lactis* strains should be considered a subspecies of *B. animalis* as previously proposed by Ventura and Zink (2002).

The *B. lactis* species status was further evaluated using a multiple technique approach to assess relatedness to *B. animalis* (Masco et al., 2004). In this study, the authors compared the DNA sequence of *groEL* and *aptD*, conducted SDS-PAGE protein profiles, assessed fluorescent amplified fragment length polymorphism (FAFLP) banding

patterns, and evaluation of growth in milk. The authors note each of these methods provided unambiguous separation of *B. animalis* and *B. lactis*. Strains of *B. lactis* exhibited 1.5-2.5 log growth in milk compared to less than 0.5 log of growth by strains of *B. animalis*. DNA based methods showed that within a (sub-) species the degree of differences were similar to the degree of differences within other *Bifidobacterium* species (Masco et al., 2004), and that growth in milk was only observed for strains of *B. lactis* and not for strains of *B. animalis*. The authors conclude since strains of both can clearly be distinguished, *B. lactis* and *B. animalis* should be differentiated on the subspecies level. At the present time, the species *B. animalis* consists of two subspecies *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis*.

Recently doubt has been cast on the original source of the strain DSM 10140 (Klein, 2009). Debate exists as to if DSM 10140 represents an original isolation or was actually the re-isolation of an existing commercial strain previously isolated from a different source. Even though this strain was isolated from a commercial yogurt product, dairy is most likely not the natural reservoir for this subspecies due to the likelihood of heat processing treatments involved in yogurt manufacture that would destroy the microorganisms present in milk prior to yogurt production. With the original source of *B. animalis* subsp. *lactis* in doubt and the fact that increased growth is seen in milk when supplemented with additional protein and some carbohydrates casts some doubt to whether *B. animalis* subsp. *lactis* has specifically evolved for growth in milk (Gomes et al., 1998).

After establishment of the taxonomic status of *B. animalis* subsp. *lactis*, the subspecies was described by Masco et al. (2004). The optimum growth temperature is

39-42°C with no growth on agar plates exposed to air but growth does occur when 10% oxygen is included in the headspace above liquid media. Strains exhibit growth in milk and milk-based media (more recent data would suggest this is not a defining phenotype). Under anaerobic conditions acetate is formed in the molar ratio of 10:1 to lactate. DNA G+C content is 61 $\pm$ 0.5% (Masco *et al*, 2004).

Felis and Dellaglio (2007) evaluated the phylogenetic relationship between all species contained in the genus *Bifidobacterium* based on 16S rDNA sequence alignment. Phylogenetic groups were established, with the oldest species name providing the name of the group. The *B. animalis* species was categorized as a member of the *B. pseudolongum* group with *B. choerinum, B. cuniculi, B. gallicum, B. pseudolongum*. *B. choerinum* was found to be the most closely related species to *B. animalis* (Ventura et al., 2006). Further analysis employing a concatenated gene tree based on *clpC, dnaB, dnaG, dnaJ1, purF, rpoC,* and *xfp* sequences confirmed *B. animalis* as a member of the *B. pseudolongum* group. This analysis revealed *B. choerinum* to be the most closely related species to *B. animalis*. *B. choerinum* is a species that was originally isolated from pig feces and it is suggested that it may be well adapted to the gut of pre-weaned piglets (Maxwell et al., 2004).

B. animalis subsp. *lactis* as a probiotic

Probiotics are defined as "Live microorganisms which, when administered in adequate amounts confer a health benefit on the host" (FAO/WHO, 2002). The majority of microorganisms believed to have probiotic properties belong to the genera *Bifidobacterium* and *Lactobacillus.* The term "probiotic" was first coined by Lilly and

Stillwell (1965), referring to growth-promoting factors produced by microorganisms and literally means "for life" in direct contrast to antibiotics. However, the term has now come to refer to health-promoting microorganisms that are intentionally consumed. The potential for replacing harmful organisms in the gastrointestinal tract with those that are "good" was first postulated by Metchnikoff in his book "The Prolongation of Life" (Metchinkoff, 1907), where he observed the beneficial aspects of yogurt consumption.

Putative Health benefits associated with *B. animalis* subsp. *lactis*

B. animalis subsp. *lactis* is commonly added to dairy products due to its perceived health benefits and its technological advantages over species of the same genus (Mättö et al., 2006). A variety of health benefits have been associated with strains of *B. animalis* subsp. *lactis* (Sanders, 2006). The entire breadth of evidence supporting *B. animalis* subsp. *lactis* as a probiotic is beyond the scope of this work and will not be reviewed. Rather, the purpose of this section is to introduce the range of health benefits and provide an appreciation for potential health promoting activities of this organism.

Putative Gastrointestinal Benefits

Several studies have evaluated strains of *B. animalis* subsp. *lactis* for its ability to reduce intestinal transit time and/or increase defecation frequency. Marteau et al. (2002) evaluated the ability of a fermented dairy product containing *B. animalis* subsp. *lactis* DN- 173 010 to increase defecation frequency in women. Female subjects consumed 3 cups/day of fermented milk containing the probiotic microorganism. Total colonic transit time was significantly reduced from 60.7 h in the control group to 51.5 h in the group

consuming the probiotic. This study is of particular interest due to the author's claim that this was the first time a probiotic had been shown to significantly reduce colonic transit time.

Waller et al. (2011) investigated the effect of *B. animalis* subsp. *lactis* HN019, at two different levels, on whole gut transit time in adults. The high-dose group received 1.72×10^{10} CFu/day and the low-dose group received 1.8×10^{9} CFu/day both in capsule form. Decreases in whole gut transit time were observed in both the high-dose and low-dose treatments, but no decrease was observed in the control group. The high-dose group reduced from of 49 ± 30 to 21 ± 32 h, low-dose from 60 ± 33 to 41 ± 39 h, whereas placebo increased slightly from 43 ± 31 to 44 ± 33 h. Due to the larger percentage decrease observed in the high-dose group the authors conclude that the *B. animalis* subsp. *lactis* HN019 reduced whole gut transit time in a dose-dependent manner.

However, not all studies evaluating *B. animalis* subsp. *lactis* have shown significant outcomes. Tabbers et al. (2011) evaluated *B. animalis* subsp. *lactis* DN- 173 010 for its ability to increase defecation frequency in children. Although there was an increase in frequency, the change was comparable to the control group, which was given a non-fermented milk product.

Chouraqui et al. (2004) investigated *B. animalis* subsp. *lactis* Bb-12 and its ability to reduce the risk of diarrhea in infants. Ninety infants were either fed conventional formula or formula supplemented with *B. animalis* subsp. *lactis* Bb-12. In total, 28.3% of infants fed formula with Bb-12 experienced diarrhea, whereas 38.7% of infants had diarrhea in the control group. This difference was not statistically significant, however the total number of days with diarrhea was 1.15 ± 2.5 days in the Bb-12 group and 2.3 ±

4.5 days in the control group, which was statistically significant. The authors conclude that the consumption of an acidified formula supplemented with *B. animalis* subsp. *lactis* Bb-12 can provide a protective effect against acute diarrhea in infants.

Guyonnet et al. (2009) evaluated *B. animalis* subsp. *lactis* DN- 173 010 in 197 women reporting minor digestive disorders but without diagnosed gastrointestinal (GI) disorders. *B. animalis* subsp. *lactis* DN- 173 010 was added to a fermented milk (Activia®, Danone) and overall GI well-being was assessed. The fermented milk product was consumed for a 4-week period. Statistically significant improvements were observed in borborygmi (a rumbling noise caused by gas moving through the intestines), stool consistency, bloating, and overall frequency of digestive symptoms, however no significant change was observed in defecation frequency.

Immune Related Benefits

Veiga et al. (2010) investigated the consumption of *B. animalis* subsp. *lactis* DN-173 010 in a fermented milk in mice with intestinal inflammation. Investigators found statistically significant decreased inflammation in mice fed fermented milk containing *B. animalis* subsp. *lactis,* compared with non-fermented milk diet and a control diet. Significantly more *B. animalis* subsp. *lactis* cells were recovered from feces of test mice with lower colitis scores (0-3) than those with higher scores suggesting certain mice may be more responsive to the treatment than other mice. Consumption of fermented milk containing *B. animalis* subsp. *lactis* also increased the levels of lactate-producing and butyrate-producing microorganisms resulting in a decrease in cecal pH (Veiga et al., 2010).

Healthy children between ages 3 and 5 were given a probiotic supplement of *B. animalis* subsp. *lactis* Bi-07 and/or *L. acidophilus* NCFM twice a day for six months (Leyer et al., 2009). When compared to placebo consumption of the probiotic mixture exhibited significantly less frequent flu-like symptoms (cough, fever), prescription of antibiotics, and duration of symptoms. When comparing the *L. acidophilus* NCFM vs. *L. acidophilus* NCFM and *B. animalis* subsp. *lactis* Bi-07 protection against symptoms was more pronounced, suggesting a mixture of probiotic strains may be more effective.

B. animalis subsp. *lactis* Bb-12 supplemented yogurt beverage was given to 172 children ages 2-4 years and monitored for missed days of school as the primary outcome. Participants drank 4 oz. of active or control (without Bb-12) yogurt for 90 consecutive days (Merenstein et al., 2011). Children in the active group missed an average of 2.54 days/100 days and children in control group missed 2.42 days/100 days, thus exhibiting no protective effect. The authors hypothesize the no effect observed may be due to the control group consuming yogurt with active starter cultures, where previously other studies have used a placebo with no active cultures.

The results from studies regarding the health-promoting effects of probiotic supplementation reviewed here are not conclusive. This is possibly due to the target populations, health concerns, delivery vehicle, and doses administered were not equivalent across studies. Thus it is difficult to draw definitive conclusions about the consumption of *B. animalis* subsp. *lactis* on overall health.

Presence of *B. animalis* subsp. *lactis* in dairy products

Due to the perceived and documented health benefit associated with *B. animalis* subsp. *lactis* the organism is often added to fermented dairy products. *B. animalis* subsp. *lactis* is able to remain viable during the shelf-life (7 weeks at 4°C) of fermented dairy products (Kailasapathy, 2006). *Bifidobacterium animalis* subsp. *lactis* was found in the majority of commercial dairy products claiming to contain probiotic microorganisms by Fasoli et al. (2003) even when *Bifidobacterium* are not claimed in the label.

In the Fasoli et al. (2003) study, 14 commercial products from the Italian marketplace were evaluated for their probiotic content. Both species identification and quantification via plate counts were investigated in this study. Species identification was carried out by species-specific PCR and PCR-DGGE (PCR-Denaturing Gel Gradient Electrophoresis). Of the 7 commercial yogurts evaluated only one claimed to contain *B. animalis* subsp. *lactis*, five claimed to contain *Bifidobacterium* or *Bifidobacterium* spp., while one claimed to contain *B. bifidus*. In all of the commercial yogurts *B. animalis* subsp. *lactis* was detected by species-specific PCR and PCR-DGGE, with a viable plate count ranging from 4×10^5 to 8×10^7. Seven commercial lyophilized probiotic supplements were also evaluated. Of these products, none claimed to contain *B. animalis* subsp. *lactis* however, in five of the products *B. animalis* subsp. *lactis* was detected by both detection methods.

In another study, 14 commercial fermented milks were analyzed for their probiotic content (Gueimonde et al., 2004). Of the 14 fermented milks, eight claimed to contain *Bifidobacterium* as the probiotic microorganism. Identification of species was determined by partial sequence of 16S rDNA. All eight of the commercial products

claiming to contain *Bifidobacterium* were found to contain *B. animalis* subsp. *lactis*. In addition, all isolates showed 100% similarity in the partial 16S rRNA sequence and showed identical PFGE profiles using *Spe*I.

Grand et al. (2003) evaluated three sour milks and a soft cheese and compared PFGE profiles of the isolated strains to three widely used commercialized strains of *B. animalis* subsp. *lactis,* Bb-12, DN- 173 010, and B420 using *Spe*I. In their PFGE analysis, all three of the reference strains and all of the isolated *B. animalis* subsp. *lactis* strains isolated from commercial products gave the exact same profiles. Interestingly, the authors comment that most of the bifidobacteria isolated from a dairy origin belong to the same species, *B. animalis* subsp. *lactis,* even when this specific subspecies is not claimed on the label.

Mayer et al. (2003) evaluated 36 strains, 27 isolated from their own work, and 9 strains obtained from the work of Bonaparte (1997). All food isolates showed similarities to the type strain of one of three species, *B. longum, B. animalis,* or *B. lactis*. However, the authors state they are unable to distinguish isolated strains between *B. animalis* and *B. lactis.* In all food items tested, *B. animalis*/*lactis* was isolated from commercial products from Hungary, Poland, and Germany. Strains isolated from food and analyzed by FT-IR spectroscopy clustered into 4 groups. FT-IR spectroscopy produces an infrared spectra patterns from the total cellular composition that are considered species-specific (Mayer et al., 2003). Clusters 1 and 2 were comprised of *B. longum* and *B. infantis*, Clusters 3 and 4 contain 36 strains of *B. animalis*/*lactis*, including type strains and strains isolated from food products. Sub-cluster 3a contained the type strains of both *B. animalis* subsp.

animalis and *B. animalis* subsp. *lactis.* Sub-clusters 3b, 3c, and 3d are all comprised of strains isolated from food products.

Masco et al. (2005) investigated the claims of 58 commercial products stating to contain bifidobacteria. From the 58 products, 626 isolates were tested by genus specific PCR, 434 of those isolates were confirmed as bifidobacteria. Of the 434 confirmed bifidobacteria, 154 were further evaluated and identified to the species level by BOX-PCR fingerprinting. *B. breve* was identified in 20% of the products while, *B. animalis* subsp. *lactis* was identified in 80% of all the commercial products tested. All but one commercial yogurt contained *B. animalis* subsp. *lactis.* A subset of *B. animalis* subsp. *lactis* isolates were tested further by PFGE using *Spe*I. PFGE analysis resulted in four distinguishable patterns. The majority of isolates gave a single PFGE pattern and three isolates gave three different patterns. Further PFGE analysis was conducted with *Xba*I, however the four patterns were not easily distinguishable. This result suggests the majority of commercial products are supplemented with a single strain or a group of highly similar strains, which are genetically similar but can be differentiated using certain methods (e.g. PFGE with *Spe*I). The authors suggest the reason why *B. animalis* subsp. *lactis* is isolated from the majority of commercial products is the organism's relatively higher level of oxygen-tolerance compared to other bifidobacteria (Masco et al., 2005). Additionally, Temmerman et al. (2003) analyzed the commercial yogurts Activia and Proflora which claim to contain bifidobacteria and found these yogurts to contain *B. animalis* subsp. *lactis* using DGGE.

Historically it has been difficult to differentiate strains from commercial products at the species level within the genus *Bifidobacterium* (Tmanova et al., 2012). Many of

products claiming to contain other species of bifidobacteria actually contain *B. animalis* subsp. *lactis* most likely due to enhanced survival under processing conditions. With the greater use of DNA-based identification methods it can be expected this issue will decrease in occurrence.

Survival of *B. animalis* subsp. *lactis* in dairy products

Although a number of studies have demonstrated a dose-dependent effect by *B. animalis* subsp. *lactis* (Larsen et al., 2006, Waller et al., 2011) the minimum effective dose has not been established. Regardless of the effective dose, the ability of a probiotic microorganism added to a fermented dairy product, to survive throughout the shelf-life is a critical parameter for assessing the suitability of that organism. The probiotic microorganisms will encounter several challenges while being added to and surviving in a fermented product, mainly oxygen and the acidic environment of the fermented product. Survival of the microorganisms is important in order to ensure an effective dose of the probiotic is delivered to the consumer. The dose delivered in a commercial product must be high enough to be effective but not so concentrated that the sensory properties of the product becomes undesirable (Sanders et al., 2012).

B. animalis subsp. *lactis* has been shown to survive significantly better than other bifidobacteria when stored at 4°C in skim milk (Jayamanne and Adams, 2006). *B. animalis* subsp. *lactis* survival was compared to *B. longum*, *B. breve*, *B. adolescentis*, *B. longum* biotype *infantis*, and *B. bifidum* over 9 days, with pH adjusted to 4.25. Levels of *B. animalis* subsp. *lactis* viable cells remained constant, while all other organisms fell below 10^6 CFu/ml after 7 days of storage with an initial population greater than 10^8

CFu/ml. These results indicate *B. animalis* subsp. *lactis* is potentially suitable for survival in a fermented dairy product and that other species of bifidobacteria may not be able to survive at a level that is acceptable over the entire shelf-life or would have to be added at high levels resulting in possible off-flavors. It was shown that viable counts of strains identified as *B. animalis* subsp. *lactis* remain relatively stable after 30 days of refrigerated storage in commercial fermented milks (Gueimonde et al., 2004). Initial bifidobacterial populations ranged from 10^6-10^8 CFu/ml and decreased less than one log in commercial fermented milks stored at 4°C.

Survival of *B. animalis* subsp. *lactis* B94 was evaluated in a yogurt product over of 35 days. Yogurt was inoculated with 10^8 CFu/g of freeze-dried culture and after 35 days of storage the population of *B. animalis* subsp. *lactis* was 5.45×10^7 CFu/g. Further analysis showed a direct correlation between pH of the yogurt and viable *B. animalis* subsp. *lactis* cells. Yogurt pH at day 0 was approximately pH 4.45 and decreased to pH 4.3 by the day 35 (Kailasapathy et al., 2008). Despite the direct correlation with viable cells and pH, only pH was monitored during the shelf-life, so the possibility remains that other factors (e.g. starter culture metabolites) may have been responsible for a decrease in viable cells.

In a separate study, survival of *B. animalis* subsp. *lactis* B94 was evaluated over 28 days in yogurt at four different pHs (4.6, 4.55, 4.5, 4.45). A statistically significant decrease in pH was only seen in the yogurt when the fermentation was terminated at pH 4.6 (Donkor et al., 2006). It should be noted that in this sample the pH decline by 0.17 pH units by day 28. Additionally, the yogurt batch terminated at pH 4.45 decreased by 0.23 pH units at day 28, but there was not significant decrease in *B. animalis* subsp. *lactis*

despite the greater difference and overall lower pH. These results suggest factors other than pH also play a significant role in *B. animalis* subsp. *lactis* survival in a yogurt product.

B. animalis subsp. *lactis* has been shown to survive well when stored in milk and dairy products and has been shown to survive better than other members of the genus *Bifidobacterium* (Jayamanne and Adams, 2006). Other than survival in probiotic-containing products, survival in the human gastrointestinal tract can be considered a critical trait for an effective probiotic microorganism. Microorganisms in a consumed probiotic fermented dairy product will also encounter bile in the gastrointestinal tract.

Challenges to potential probiotic microorganisms

In the product and following consumption, three significant challenges face a probiotic organism, namely acid, oxygen, and bile. *B. animalis* subsp. *lactis* and other probiotic microorganisms are often added to fermented dairy products with a low pH (~4.6) and stored for an extended period of time. Also, the organisms will be challenged by the low pH of the human stomach. Organisms will additionally encounter oxygen during addition to the food as well as dissolved oxygen present in the product. Ingested organisms will be subjected to the effects of bile after leaving the stomach in the duodenum (Begley et al., 2005).

Acid Tolerance

Generally, American style yogurt is fermented to a terminal pH of 4.6, with European style yogurt having a slightly lower pH. Thus, microorganisms contained in a

fermented product, must be able to survive in this acidic environment over an extended period of time (~60 days). This ability is considered to be a critical criterion for selecting a probiotic microorganism (Saarela et al., 2000). In addition, microorganisms will encounter the low pH of the human stomach. The pH of human gastric juice can vary between individuals based on different times of the day, and meal status but is on average a pH of 2.0 while fasting but can be higher (pH 3-4) while consuming food (McLauchlan et al., 1989). While the time exposed in the stomach is much shorter the environment is much harsher.

Several studies have shown that *B. animalis* subsp. *lactis* is the most acid resistant (sub)-species among species of *Bifidobacterium* (Matsumoto et al., 2004, Mättö et al., 2004, Vernazza et al., 2006). Matsumoto et al. (2004), investigated two strains of *B. animalis* subsp. *lactis,* four strains of *B. animalis* subsp. *animalis*, four strains of *B. bifidum*, two strains of *B. catenulatum* and *B. longum*, and one strain of each *B. infantis*, *B. adolescentis, B. breve*, and *B. pseudocatenulatum*. The two strains of *B. animalis* subsp. *lactis* evaluated in this study were LKM 512 (Bb-12) and JCM 10602 (DSM 10140) and the four *B. animalis* subsp. *animalis* strains were JCM 1190 (presumably ATCC 25527), JCM 1253, JCM 7117, and JCM 7124. Each strain was evaluated for survival at four different pH levels (pH 2, 3, 4, and 5) at five time points (0, 0.5, 1.0, 1.5, 2.0, and 3.0 h). Based on these experiments the authors concluded bifidobacteria were generally non-acid tolerant except for *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis*. Little to no decrease (less than 1 log) was observed in the viable cell counts when cells were incubated at pH 3 for up to three hours in both subspecies. The authors also suggest acid tolerance is species specific rather than strain-specific because all the

acid tolerant strains belonging to *B. animalis* subsp. *lactis* or *animalis* and only minor differences noted within a sub-species (Matsumoto et al., 2004). However, only a limited sample size of strains was evaluated in this study.

Mättö et al. (2006) evaluated two strains of *B. animalis* subsp. *lactis* (E2010 and E701), two strains of *B. adolescentis*, and one strain of *B. bifidum*, *B. breve*, and *B. longum*. Cells were subjected to low pH (2.0, 2.5, 3.0) for 2 hours and with or without the addition of pepsin. No decrease in the population of viable cells was observed in *B. animalis* subsp. *lactis* cells when treated at pH 3.0 (with or without pepsin), but for each other species < 3.0 log CFu/ml was detected. Interestingly, at pH 2.5 less than 3.0 log CFu/ml viable cell were recovered when *B. animalis* subsp. *lactis* E701 was treated, in contrast 5.5 log CFu/ml *B. animalis* subsp. *lactis* E2010 were recovered following the same treatment. The difference in recovered cells at pH 2.5 within the subspecies indicates that acid tolerance in *B. animalis* subsp. *lactis* can be strain-specific, contradicting previous findings (Matsumoto et al., 2004). Additionally, differences were observed in recovery of viable cells treated at pH 2.0 and 2.5 in *B. animalis* subsp. *lactis* in the presence or absence of 3 g/L pepsin. It was hypothesized addition of pepsin provided a protective effect by sustaining H^+-ATPase activity by the addition of ATP in impurities of the pepsin preparation (Mättö et al., 2006).

Vernazza et al. (2006) assessed survival of *B. animalis* subsp. *lactis* Bb-12, *B. adolescentis*, *B. infantis*, and two strains of *B. longum* incubated at pH 2, 3, and 4 for up to 20 minutes in Wilkin-Chalgren broth with pH adjusted using HCl. The population of *B. animalis* subsp. *lactis* Bb-12 was not reduced, even after 20 minutes of exposure to the lowest pH. All other organisms evaluated decreased to 0% recovery at all pH levels

evaluated (except for *B. infantis* at pH 4) after twenty minutes of exposure. These results indicate that *B. animalis* subsp. *lactis* Bb-12 is significantly more resistant to acid than other species tested. *B. animalis* subsp. *lactis* Bb-12 showed excellent survival at low pH, although cells were only exposed for 20 minutes, which is a relatively short period of time (Vernazza et al., 2006).

While some investigators have focused on the differing survival of different species and in some cases different strains, others have attempted to explain the mechanism by which *B. animalis* subsp. *lactis* tolerates acidic environments and why it is generally the most acid tolerant species of bifidobacteria. In an attempt to determine a physiological explanation for the observed differences, the activity of H^+-ATPase in acid tolerant and non-acid tolerant strains was evaluated. For the acid tolerant strains, H^+-ATPase peaked at pH of 4, when compared to pH 3 and 5. Non-acid tolerant strains showed peak activity H^+-ATPase at pH of 5, and lower H^+-ATPase at pH of 4 when compared to acid tolerant strains. It was suggested the acid tolerant strains are better able to discharge H^+ via H^+-ATPase activity and the ability to synthesize H^+-ATPase in response to low pH is important in the determining acid tolerance in bifidobacteria (Matsumoto et al., 2004).

The *B. animalis* subsp. *lactis atpBEFHAGDC* operon was analyzed and compared to similar operons in other bifidobacterial species (Ventura et al., 2004). These eight genes encode the F_1-F_0-ATPase, which is responsible for creating a proton gradient by expelling H^+ from the cytoplasm driven by ATP hydrolysis (Ventura et al., 2004). Protein sequences from *B. animalis* subsp. *lactis* DSM 10140 showed a high degree of similarity with the same operons in *B. breve* NCIMB 8808 and *B. longum* NCC 2705. A

15-fold increase in transcription level of *atpD* at pH 3.5 was detected after 100 minutes, suggesting adaptation to acidic conditions (Ventura et al., 2004). Despite the protein level similarity observed in this study, other reports have observed a significant difference in the ability of *B. animalis* subsp. *lactis* and *B. longum* to survive in acidic conditions (Mättö et al., 2006). This would indicate the response to high acidity in *B. animalis* subsp. *lactis* is more complex, potentially involving multiple genes or operons.

It has been shown that pre-exposure to mild acidic conditions can improve the survival of those pre-treated when subsequently exposed to more harsh conditions (Foster, 1993). However, a pretreatment at pH 3-4 did not provide protection against exposure to pH 2.5 in *B. animalis* subsp. *lactis* E2010 (Saarela et al., 2004). Only a modest protective effect was observed in other bifidobacterial species when exposed at pH 4.5 (Waddington et al., 2010), suggesting a mild acidic pre-treatment of bifidobacteria will not significantly improve culture viability when added to a fermented dairy product. Though adding cells in stationary phase may increase survival, as stationary phase cells were shown to survive better at low pH than exponential phase cells (Waddington et al., 2010).

At this time, the response in *B. animalis* subsp. *lactis* to acidic environments is not thoroughly understood. Although it has been shown that F_1-F_0-ATPase is, at least, partially responsible for the expelling of H^+ out of the cell (Matsumoto et al., 2004) significant differences have been observed in species that have similar genetic structure of the genes encoding this ATPase, suggesting more components are involved. No studies have thoroughly investigated variation in acid tolerance among strains of *B. animalis* subsp. *lactis*. One study showed acid-tolerance is strain-specific in *B. animalis* subsp.

lactis (Mättö et al., 2006), while one study showed there was no difference between strains in regards to acid tolerance (Matsumoto et al., 2004). Additional studies should be conducted, with a large set of genetically diverse strains, to assess differences between strains in resistance as well as in transcription using whole genome microarray.

Oxygen Tolerance

Members of the genus *Bifidobacterium* are obligate anaerobes and thus oxygen is toxic to these organisms. Thus, the ability to tolerate environments with oxidative stresses is another critical criterion to assess when selecting a probiotic strain. It would be practically impossible to design a processing system and fermented product, in which, the microorganism is not exposed to some level of oxygen. When an organism is in direct contact with oxygen, the organism is likely killed due to intracellular production of hydrogen peroxide (Champagne et al., 2005). In addition to direct contact with oxygen, strains of *L. delbrueckii* are known to produce peroxide in the medium (Villegas and Gilliland, 1998, Marty-Teysset et al., 2000) which is another potential source of oxidative stress during storage of a fermented dairy product.

When *B. animalis* subsp. *lactis* was originally isolated, it was selected for its tolerance to oxidative stress (Meile et al., 1997). Survival in the presence of air has been shown to be strain-specific in bifidobacteria (Beerens et al., 2000). Strains of *B. animalis* subsp. *lactis* did not significantly differ in growth from *B. animalis* subsp. *animalis* and all classified as "moderate" (>5% aerobic growth compared to anaerobic) in oxygen tolerance rating (Simpson et al., 2005). Only three species evaluated were observed to have a "high" (>30%) oxygen tolerance rating; *B. boum, B. minimum,* and *B.*

pyschraerophilum.

When three strains of *B. animalis* subsp. *lactis* were evaluated for resistance to H_2O_2 a large difference in the survival was observed between Bl-04 (moderate) and DSM 10140 (poor) (Oberg et al., 2011). Additionally, *B. longum* subsp. *infantis* ATCC 15697 showed the greatest resistance to H_2O_2, *B. longum* NCC2705 and *B. animalis* subsp. *lactis* RH-1 showed intermediate resistance, and *B. longum* D2957 showed poor resistance, comparatively. This study provides strong evidence for strain-specific oxidative stress response in *B. animalis* subsp. *lactis* especially due to the high degree of similarity between DSM 10140 and Bl-04.

Li et al. (2010) isolated at moderately oxygen tolerant strain from a traditional Chinese fermented milk. Several strains were isolated but the strain Qq08 showed the highest relative bacterial growth rate (RBGR) defined as dividing of absorbance of aerobic growth by anaerobic growth (Talwalkar et al., 2001). Sequencing of the 16S rRNA gene sequence revealed this isolate to be a strain of *B. animalis* subsp. *lactis* (Li et al., 2010), further providing support for the aero-tolerance of strains of the subspecies.

Aerobic bacteria are able to detoxify molecules causing oxidative stress by decomposition of these compounds, generally by enzymes such as catalase and superoxide dismutase (Li et al., 2010). The genomes of *B. animalis* subsp. *lactis* are devoid of these traditional stress response genes (Oberg et al., 2011). Some investigators have measured superoxide dismutase activity, but have observed no correlation with activity and exposure to oxygen (Shimamura et al., 1992, Talwalkar and Kailasapathy, 2004), suggesting measurement of non-specific activity or the presence of a gene with similar activity that has yet to be annotated. Despite the lack of these stress response

genes, some species and strains of bifidobacteria are oxygen tolerant, albeit to varying degrees. It has been proposed bifidobacteria use two NADH oxidative enzyme systems in order to detoxify oxygen to water (Shimamura et al., 1992). Interestingly, NAD-oxidase and NAD-peroxidase were shown to be most active at a pH of 5, with very little activity near neutral pH. These enzymes showed moderate activity at pH 4.5, which would be the pH closest to typical yogurt pH. Recently, expression of genes, which encoded for proteins that showed significant difference in production when grown in the presence of oxygen was investigated in *B. animalis* subsp. *lactis* IPLA4549 (Ruiz et al., 2012). When *B. animalis* subsp. *lactis* was grown under aerobic conditions NADH oxidase activity was shown to have increased. Additionally, coproporphyrinogen III oxidase was up-regulated in *B. animalis* subsp. *lactis*, leading to speculation this enzyme plays a role in the oxidative stress response possibly by the oxidation of protoporphyrin IX by H_2O_2 (Ruiz et al., 2012). However, how this precisely acts to reduce oxidative stress in *B. animalis* subsp. *lactis* is not well understood.

In summary, response to oxidative stress is not well understood in bifidobacteria or more specifically in *B. animalis* subsp. *lactis*. *B. animalis* subsp. *lactis* has been shown to be moderately oxygen tolerant and is able to tolerate oxygen at higher levels than other members of the genus (Meile et al., 1997, Simpson et al., 2005). Additionally, the degree of oxygen tolerance is a strain-specific, even among very closely related strains (Oberg et al., 2011).

Bile Tolerance

An ingested probiotic must also be able to survive in the presence of bile salts. Bile salts exist as natural emulsifiers of dietary lipids but also have antimicrobial characteristics (Kim and Lee, 2008). Bile salts exert antimicrobial activity primarily by disrupting cell membranes resulting in lysis, disassociation of membrane bound proteins, or allowing leakage of cellular contents (Begley et al., 2005).

In previous works assessing the ability of bifidobacteria to tolerate bile, *B. infantis* ATCC 15697 was found to be the most resistant while *B. longum* ATCC 15707 was the least resistant when subjected to bile (glycocholate) at four levels (0.0, 0.6, 1.5 and 3.0 g/L) (Ibrahim and Bezorovainy, 1993). In a subsequent study, it was found that *B. longum* was the most tolerant species and *B. infantis* was the least resistant to bile salts (Clark and Martin, 1994). In a third study, *B. longum* BL2 and *B. infantis* BI420 were shown to have comparable survival in the presence of bile salts (Chung et al., 1999). These results suggest bile tolerance is a strain-specific trait or significant differences in results were obtained due to variation in experimental parameters. In the evaluation of 36 *Bifidobacterium* strains exposed to bile, in the form of ox bile, only 6 strains were classified as "resistant", among those strains was *B. animalis* subsp. *animalis* ATCC 25527 (no *lactis* strains were evaluated). Strains of *B. breve* were resistant, moderately resistant, and susceptible, further providing evidence of strain-specificity in response to bile in bifidobacteria (Kociubinski et al., 1999). The chemical composition and source of the bile influences survivability of *B. animalis* subsp. *lactis* (Saarela et al., 2005). In this previous work bile extract (porcine source) exerted a great decrease of *B. animalis* subsp. *lactis* cells than bile acids (a mixture of bovine and ovine).

Bile resistant derivatives have been developed in bifidobacteria by increasingly exposing to higher concentrations of bile salts. Bile resistant derivatives displayed two to sixteen fold increase in resistance to oxgall or four to eight fold increase to sodium cholate, including *B. animalis* subsp. *lactis* 4549dOx, which was originally misidentified as *B. bifidum* (Noriega et al., 2004, Margolles and Sanchez, 2012). Some of the bile resistant strains also showed greater resistance to some antibiotics, including tetracycline and greater resistance to low pH (Noriega et al., 2004, Luis Noriega et al., 2005). In addition to altered antibiotic resistance, it has been shown that the derivative strain *B. animalis* subsp. *lactis* 4549dOx differed from the parental strain in terms of growth on certain carbohydrate sources (Ruas-Madiedo et al., 2005).

The F_1F_0-ATPase system has also been shown to play a role in the response to bile in *B. animalis* subsp. *lactis* (Sanchez et al., 2006). Activity of this ATPase was shown to increase in the presence of bile and was most active at a pH of 5.0. To further support their claim the F_1F_0-ATPase system is involved in bile resistance it was shown decreasing amounts of intracellular ATP was detected in cells exposed to increasing concentrations of bile. In response to bile *B. animalis* subsp. *lactis* has been shown to modify its membrane fatty acid composition in a bile resistant mutant, noting a shift to more saturated fatty acids in the resistant derivative (Ruiz et al., 2007).

Another mechanism bifidobacteria can utilize for bile resistance is bile salt hydrolase (BSH) activity. Bile salt hydrolase hydrolyzes glycine or taurine conjugated bile acids into free bile acids and amino acids (Kim and Lee, 2008). A bile salt hydrolase was identified in *B. animalis* subsp. *lactis* KL612 and was found to be transcribed with two other genes, potentially suggesting it is part of an operon of genes involved with bile

resistance, although no function was ascribed to the other two genes in the putative operon (Kim and Lee, 2008). Increased BSH activity was observed in *B. animalis* subsp. *lactis* Bi30, with maximum catalytic activity at pH 5.1 (Jarocki, 2011). Up-regulation of DnaK was shown in *B. animalis* subsp. *lactis* Bi-07 in the presence of bile salts, which interacts with human plasminogen, potentially increasing its ability to colonize the human GIT (Candela et al., 2010). A bile efflux transporter was identified in *B. longum* and was shown to confer bile resistance, but no homolog was found in *B. animalis* subsp. *lactis*, possibly highlighting the importance of BSH in *B. animalis* subsp. *lactis* (Gueimonde et al., 2009).

Multiple elements, including BSH and a bile efflux transporter, have been shown to be active in the bile stress response of bifidobacteria. It has been suggested species which are considered to be autochthonous members of the GIT confer bile resistance by different means than species that are not, including *B. animalis* subsp. *lactis* (Gueimonde et al., 2009). BSH activity has been shown to be an important factor in bile resistance in *B. animalis* subsp. *lactis* and has been found to be constitutively expressed in Bb-12 (Garrigues et al., 2005).

Several studies have noted increased resistance to one challenge (e.g. bile) leads to increased resistance to another in bifidobacteria (e.g. acid) (Noriega et al., 2004, Sanchez et al., 2006, Sanchez et al., 2007). A possible explanation is for the same or similar systems to involved in protection against more than one challenge. For example, the F_1F_0-ATPase system was found to be involved in both acid and bile resistance in *B. animalis* subsp. *lactis*.

Dairy adaptation of *B. animalis* subsp. *lactis*

The original isolation of *B. animalis* subsp. *lactis* was made from a commercial yogurt sample (Meile et al., 1997). Because the original isolate came from a dairy product, the species was assigned the name *lactis*. In addition, because the original isolate came from a dairy source, it was speculated *B. animalis* subsp. *lactis* was adapted for the dairy environment (Lee and O'Sullivan, 2010). Several reports have identified growth in milk and/or milk-based media as the defining phenotypic difference between *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* (Ventura and Zink, 2002, Masco et al., 2004). Masco et al. (2004) observed between 1.5-2.5 log growth of 10 strains of *B. animalis* subsp. *lactis*, and less than 0.5 log growth in two strains of *B. animalis* subsp. *animalis* in reconstituted skim milk measured by the change in conductivity of lactose to lactic acid. Li et al. (2010) report isolation of new strain of *B. animalis* subsp. *lactis* identified based on 16S rRNA sequence and also report positive growth in unsupplemented milk, but do not state the level of growth observed.

Despite observations of growth in milk and the perceived dairy adaptation of *B. animalis* subsp. *lactis*, it has been shown growth of this organism can be enhanced in milk or milk-based media. Gomes et al. (1998) found the addition of milk protein hydrolyzates to a milk medium significantly increased the growth of *B. animalis* subsp. *lactis* Bo. Additionally, Gomes et al. found this organism did not clot pure milk at 37°C for 24 h with a 5% inoculum. *B. animalis* subsp. *lactis* Bo was also co-cultured with *L. acidophilus*, whose growth was not influenced by milk protein hydrolyzates, and observed increased growth under this condition. This is likely because *L. acidophilus* is able to hydrolyze milk proteins for its own growth and produce peptides in the media

stimulating bifidobacterial growth (Gomes et al., 1998). Several other species of bifidobacteria including *B. longum, B. infantis, B. breve, B. adolescentis* and *B. pseudocatenulatum* exhibited increased growth when co-cultured with a *Lactococcus lactis* with proteinase (PrtP) activity (Yonezawa et al., 2010).

B. animalis subsp. *lactis* DSM 10140 was evaluated for growth in skim milk and skim milk supplemented with caseinmacropeptide, either from bovine whey or combined caprine and ovine whey, or whey protein concentrate (Janer et al., 2004). After 24 hours of incubation significantly increased growth was observed in supplemented milks when compared to skim milk alone. Additionally, significant growth was observed in the skim milk supplemented with whey protein concentrate compared to either milk supplemented with bovine or combined caseinmacropeptide (Janer *et al*, 2004).

To further understand the ability or lack of ability of *B. animalis* subsp. *lactis* to grow in milk, its proteolytic activity was evaluated (Janer et al., 2005). Whole cells did not grow in milk and cell-wall bound fractions were unable to hydrolyze casein and whey proteins. Some intracellular fractions did possess the ability to hydrolyze casein, but due to it intracellular location it is not believed to play a role in growth in milk (Janer et al., 2005).

In addition to added proteins, addition of alternative carbohydrate sources has been observed to result in increased growth in milk. Shin et al. (2000) evaluated *Bifidobacterium* spp. Bf-1 and Bf-6 (Bf-6 is a strain of *B. animalis* subsp. *lactis*) in milk and milk supplemented with fructooligosaccharide, galactooligosaccharide, or inulin, each at four levels. Both *Bifidobacterium* Bf-1 and Bf-6 exhibited a significant decrease in doubling time with 5% supplementation of each carbohydrate source. Both organisms

also exhibited a significant decrease in doubling time in the presence of 3% galactooligosaccharide, and 3 and 1% fructooligosaccharide added. Increased viability at the same levels of each carbohydrate that decreased doubling time was observed after four weeks of storage at 4°C (Shin et al., 2000). Bruno et al. (2002) evaluated growth of five bifidobacteria in skim milk and supplemented skim milk containing various prebiotic carbohydrate sources. *B. animalis* Bb-5 exhibited a significant decrease in doubling time in skim milk supplemented with lactulose, raftilose, and inulin but not with Hi-maize (high amylose corn starch). *B. infantis* did not exhibit a significant decrease in doubling time when any supplement was added. Both *B. longum* strains and *B. pseudolongum* had significantly decreased doubling times when lactulose or inulin was added. The ratio of acetic acid to lactic acid production did not significantly change for *B. animalis* Bb-5 in the presence of any of the carbohydrates (Bruno et al., 2002).

Differentiating strains of a monomorphic subspecies

Achtman (2008) used the term "genetically monomorphic" to refer to a group of organisms that "have such low levels of sequence diversity that only few polymorphisms, or even none at all, are found upon sequencing a few genes" and used the term to describe certain pathogenic organisms. One of the main concerns when studying organisms with genetically monomorphic genomes is examining and understanding diversity among these strains (Comas et al., 2009). One subspecies that can be considered to have this high level of homogeneity is *B. animalis* subsp. *lactis.* Due to the high degree of genetic similarity it is difficult to distinguish strains (Briczinski and Roberts, 2006). Because health-promoting effects are considered to be strain-specific

and cannot be projected to other strains of the same species, reliable strain identification is critical (WHO/FAO, 2002). In addition, the European Food Safety Authority states "For microorganisms (e.g. bacteria, yeast), as well as species identification, there should be sufficient characterization (genetic typing) at strain level by internationally accepted molecular methods…" (EFSA, 2009). While not currently required in the United States, strain identification of commercial probiotics may soon become necessary as consumers and regulatory officials take a greater interest in this growing market.

Identifying a bacterial isolate to the strain level, especially among closely related organisms, can be a difficult task. One main reason for this is the term "strain" is not well-defined. One proposed definition is "any culture knowingly derived from the original strain" (De Vos and Truper, 2000). However, this definition does not account for the identification and classification of new or unknown strains. Another definition states "descendants from a single isolation in pure culture" (Brenner et al., 2001). However, this definition infers knowledge of original isolation source and does not allow for genetic changes in an organism that may occur over many years leading to changes in genotype and phenotype. Another, more useful definition of strain for closely related organisms and probiotic research is "a strain is an isolate the can be differentiated from other isolates of the same genus, species, and subspecies, by at least one phenotypic or genotypic characteristic" (Briczinski, 2007).

Pulsed-Field Gel Electrophoresis

Pulsed-Field Gel Electrophoresis (PFGE) is considered to be the gold standard for differentiation among strains of microorganisms (FAO/WHO, 2002). This method is

based on the pattern produced by a rare-cutting restriction enzyme digest of total DNA. PFGE can be used for several purposes (e.g. isolate comparisons, genome size estimation) and is often discriminatory enough to differentiate between strains (Basim and Basim, 2001). Choice of restriction enzyme can greatly influence the ability of this method to differentiate between strains or isolates.

When one hundred sixty bifidobacterial pig cecum isolates were evaluated by PFGE using the restriction enzyme *Xba*I (Simpson et al., 2003) seven PFGE types and eight subtypes were identified based on 15 distinct restriction patterns. This result showed different species of *Bifidobacterium* will exhibit different restriction patterns digested with *Xba*I.

While different species may give a different pattern the same may not be true of different strains of the same species. When comparing three strains of *B. animalis* subsp. *lactis* (DSM 10140, Bb-12, and B94), it was found that all three exhibited the same exact restriction pattern when using *Xba*I (Crittenden et al., 2001). Mayer et al. (2007) used the restriction enzyme *Spe*I and found nine strains of *B. animalis* subsp. *lactis* gave the same fragment pattern while two strains (Bf2 and Bf45) exhibited an additional band from PFGE analysis. All strains exhibited identical patterns with *Apa*I, *Not*I, and *Sma*I. In comparison, two *B. longum* strains gave two unique PFGE patterns using *Spe*I. Similar results were obtained by Mättö *et al,* (Mättö et al., 2004) when examining intestinal bifidobacterial isolates. Four isolates identified as *B. animalis* subsp. *lactis* gave the same pattern after restriction with *Xba*I. Ten *B. adolescentis* isolates gave nine different restriction patterns with the same method and ten *B. longum* isolates gave ten distinct patterns also with this method. In another study that included 20 commercial strains,

DSM 10140, and ATCC 27536, all strains exhibited exactly the same pattern with *Xba*I, and with *Spe*I all commercial strains exhibited to same pattern as DSM 10140 although ATCC 27536 showed a two-band difference (Briczinski, 2007). *B. animalis* subsp. *animalis* ATCC 25527 and ATCC 27674 gave the same PFGE pattern following digestion with *Spe*I or *Xba*I (Roy and Sirois, 2000).

PFGE is generally considered highly discriminatory and is generally able to differentiate strains of most species of *Bifidobacterium* (Simpson et al., 2003, Mättö et al., 2004). However, PFGE is unable to differentiate *B. animalis* subsp. *lactis*, likely due to the high genome homogeneity among strains of *B. animalis* subsp. *lactis*.

Randomly Amplified Polymorphic DNA-PCR

Randomly Amplified Polymorphic DNA PCR is a PCR based technique utilizing random primers used to generate an amplicon profile (Ward and Roy, 2005). A benefit of this method is that no prior knowledge of the genomic make-up of the organism is required. Additionally, this method has been found useful when attempting to discriminate between closely related strains (Williams et al., 1990).

Vincent et al. (1998) used five different RAPD primers and was able to group *B. adolescentis, B. bifidum, B. breve* and *B. animalis* into separate clusters. Strains classified as *B. animalis* isolated from commercial products separated into two sub-clusters most likely as the separate subspecies. In another study seven different RAPD primers were used to generate profiles. All of the commercial strains of *B. animalis* subsp. *lactis* gave identical profiles but the technique was able to differentiate between strains of *B. infantis* and *B. adolescentis* (Briczinski, 2007). Mayer et al. (2007) used 50

different RAPD primers and found identical patterns for all of the commercial strains tested except for one. The strain *B. animalis* subsp. *lactis* Bf47 (LAFTI B94) exhibited one additional amplicon compared to the other strains of the subspecies when using two of the primers. One reference strain, one commercial strain and two intestinal isolates of *B. animalis* subsp. *lactis* gave the same RAPD-PCR profile (Mättö et al., 2004). These data suggest RAPD-PCR is not discriminatory enough to distinguish between strains of *B. animalis* subsp. *lactis.*

Other methods

Ventura and Zink (2002) evaluated sequence of the 16S-23S ITS among six commercial strains of *B. animalis* subsp. *lactis* and the Type strain of *B. animalis* subsp. *animalis* ATCC 25527. In this analysis *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* clearly formed two clusters. Interestingly, strains of *B. animalis* subsp. *lactis* were found to have ~93-97% similar 16S-23S ITS, meaning strains could be differentiated from one another using the sequence from this region. However, recently the sequence of the 16S-23S ITS of these strains was re-evaluated and no differences were observed (unpublished data). A key differentiating feature of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* is an 8-bp insertions in the 16S-23S ITS which is used as a target for the design of the primer Ban2, which in conjunction with the primer 23Si is able to differentiate between the subspecies (Ventura and Zink, 2002).

Ventura and Zink (2003) evaluated *tuf*, *recA* and polymorphism in the 16S-23S ITS to differentiate *B. animalis* subsp. *lactis* from *B. animalis* subsp. *animalis*. Five strains originally identified as *animalis* and six strains of *lactis* were investigated.

Evaluation of both the *tuf* and *recA* genes allowed discrimination between the two subspecies. Three strains originally identified as *B. animalis* subsp. *animalis* exhibited the exact same sequence as strains of *B. animalis* subsp. *lactis* (which all have the same exact sequence) for both genes. Analysis of restriction length fragment polymorphism of the 16S-23S ITS exhibited the same results.

In an effort to differentiate and assess isolates of various species of bifidobacteria, Delétoile et al. (2010) subjected a number of species of bifidobacteria to a multi-locus sequence type (MLST) typing scheme based on sequences of *clpC, fusA, gyrB, ileS, purF, rplB,* and *rpoB.* This analysis adequately distinguished and clustered different species of bifidobacteria and showed high discriminatory power. *B. animalis* clustered into two subgroups, with two strains of *B. animalis* subsp. *animalis* in one group, nine strains of *B. animalis* subsp. *lactis* in another group and one strain of *B. animalis* subsp. *lactis,* ATCC 27673, which was differentiated from the rest of the *lactis* subgroup. This MLST scheme was able to differentiate various strains of *B. longum* subsp. *infantis, B. breve, B. bifidum* and *B. longum* subsp. *longum* but was only able to differentiate *B. animalis* subsp. *lactis* into 2 groups (one with 9 strains and one containing ATCC 27673).

While nucleic-acid based method were unable distinguish two strains of *B. animalis* subsp. *lactis* (DSM 10140 and Bl-04), Briczinski et al. (2008) was able to differentiate these strains based on media containing glucose as its sole carbohydrate source (Briczinski et al., 2008), which could be used to distinguish these two strains which is correlated to a specific *glcU* allele.

A novel, recently employed technique, matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) was used to differentiate strains of

B. animalis subsp. *lactis* (Ruiz-Moyano et al., 2012). Using this method the authors were able to separate 20 strains into 9 clusters based on peaks identified form MALDI-TOF MS spectra. This method shows good discrimination and with a rapid analysis time (~20 minutes) it is possibly a useful method to screen strains of this subspecies (Ruiz-Moyano et al., 2012). However, the equipment necessary to conduct this type of analysis may not be available in many microbiology laboratories.

The general lack of genomic sequences available for strains of *B. animalis* subsp. *lactis* hinders the ability to develop methods to differentiate strains or identify methods that would provide greater discrimination. The addition of genome sequence can allow for the identification of targets for strain-level differentiation.

Bifidobacterium genomics

Genome sequencing of health-promoting bacteria, such as bifidobacteria, has increased in popularity in recent years. This has lead to the development of a new field of study which some have called probiogenomics, (Ventura et al., 2009a) however the term has yet to gain popular use. Due to the general lack of genetic manipulation tools for bifidobacteria, full genome sequencing has been critical for developing a better understanding of the genetic make-up of these beneficial organisms. Additionally, genomic sequencing provides a means to gain a better appreciation for the genetic diversity that exists between species of *Bifidobacterium* and within species at the strain level. Generating a large amount of genomic information may be beneficial in leading to identification of species and/or strains that possess the best ability to survive the harsh environments of the human GIT and have the most beneficial effects on human health.

Bifidobacteria belong to the phylum Actinobacteria and the family *Bifidobacteriaceae*, which contains the eight genera *Bifidobacterium*, *Gardnerella*, *Scardovia*, *Parascardovia, Aeriscardovia, Alloscardovia, Metascardovia,* and *Pseudoscardovia* (Ventura et al., 2007, Bonaparte and Klein, 2012, Killer et al., 2013). *Bifidobacterium* has been divided into six phylogenetic groups based on sequence of 16S rRNA gene; *B. boum* group, *B. pseudolongum* group (*B. animalis* sp. belong to this group), *B. asteroides* group, *B.* pullorum group, B. *adolescentis* group, *B. longum* group (Ventura et al., 2006, Turroni et al., 2011). *Bifidobacterium* have high %G+C content ranging from 54% to 67.5% (Lee and O'Sullivan, 2010), genomes that vary in size from 1.9 to 2.8 Mb, and have between 3 and 5 rRNA operons (Turroni et al., 2009).

Non-*B. animalis* sequenced genomes

A number of completely sequenced and publically available genomes are publically available including *B. adolescentis* (1), *B. asteroides* (1), *B. bifidum* (4), *B. breve* (3), *B. dentium* (1), *B. longum* subsp. *longum* (7), *B. longum* subsp. *infantis* (2), and *B. thermophilum* (1). To date a total of 20 strains have fully sequenced (Table 1). Genomes of sequenced bifidobacteria range in size from *B. adolescentis* ATCC 17703 at 2.09 Mb to *B. longum* subsp. *infantis* ATCC 15697 2.83 Mb. The number of genes annotated in these genomes range from 1,702 (ATCC 15703) to 2,588 genes (ATCC 15697). *B. dentium* Bd1 has the lowest %GC with 58.5% and *B. bifidum* S17 has the highest %GC with 62.8%. *B. longum* subsp. *longum* NCC2705 was the first bifidobacterial strain to be sequenced (Schell et al., 2002). The non-*lactis* subspecies with the most sequenced genomes is *B. longum* subsp. *longum* with seven fully sequences

genomes. These genomes range in size from 2.27 Mb (BBMN68) to 2.48 Mb (JDM301), 59.8 to 60.3% GC, and 1,744 (F8) to 2,074 genes (DJO10A).

B. animalis subsp. *lactis* sequenced genomes

In 2009, the first sequenced genome of *B. animalis* subsp. *lactis*, strain AD011 became publically available (Kim et al., 2009). Including the genomes completed for this thesis, there are currently ten *B. animalis* subsp. *lactis* genomes available as well as one complete genome for *B. animalis* subsp. *animalis*. Most sequenced strains of *B. animalis* subsp. *lactis* are commercial strains available for use in food as probiotic additives. Additionally, draft sequences are available for two strains (HN019 and BS 01) are publically available.

B. animalis subsp. *lactis* AD011

Strain AD011 was originally isolated from the feces of healthy breast-fed infant (Kim et al., 2009). The authors claim this strain has shown a high degree of tolerance to gastric acid and bile acids, however the data is not shown. One paper has shown a health-promoting activity of this strain which was shown to reduce allergic response when orally administered to mice who have been sensitized to ovalbumin (Kim et al., 2008).

The genome of *B. animalis* subsp. *lactis* AD011 is 1,933,695 bp, with a 60.49% G+C, making *B. animalis* subsp. *lactis* the smallest of the sequenced bifidobacteria. Two rRNA operons were identified with 52 tRNAs and 1,528 coding sequences (Kim et al., 2009). *B. animalis* subsp. *lactis* AD011 is the only strain of *B. animalis* subsp. *lactis* that

reportedly contains only two rRNA operons. The presence of only two rRNA operons as well as anomalies in genome architecture suggest the genome of AD011 is incomplete and misassembled (Garrigues et al., 2010). The genome sequence for of *B. animalis* subsp. *lactis* AD011 is available from GenBank with the accession number CP001213.

B. animalis subsp. *lactis* Bb-12

The genome of commercial strain *B. animalis* subsp. *lactis* Bb-12, available from Chr. Hansen A/S, was published in 2010 (Garrigues et al., 2010) however the majority of genomic sequence was known since 2005 but not publically available (Garrigues et al., 2005). Bb-12 is one of the most studied probiotic-bacteria; the authors claim over 200 publications related to Bb-12. Studies have shown that *B. animalis* subsp. *lactis* Bb-12 posses the ability to impact fecal consistency (Larsen et al., 2006), improve immune function (Rizzardini et al., 2012), tolerate acidic environments (Vernazza et al., 2006), and survive in dairy products (Desfosses-Foucault et al., 2012). The commercial relevance of this organism made it a very good candidate for full genomic sequencing. The *B. animalis* subsp. *lactis* Bb-12 genome is 1,942,198 bp in length, contains four rRNA operons and 53 tRNAs and 1,642 open reading frames are present (Garrigues et al., 2010). The original isolation source of this strain is not disclosed. The genome of *B. animalis* subsp. *lactis* Bb-12 is available in GenBank with the accession number CP001853.

B. animalis subsp. *lactis* V9

B. animalis subsp. *lactis* V9 was first isolated from the feces of a healthy child in China and further assessed for its suitability as a probiotic (Gao et al., 2009). This isolate exhibited 99.7% survival when incubated anaerobically at pH 2.0 for 3 h. This strain also tolerated 0.3% oxgall, suggesting it had survival characteristics that made it a good candidate for use as a probiotic. This isolate was identified as *B. animalis* subsp. *lactis* by 16S rDNA analysis and exhibited a very high degree of identity with the sequence of *B. animalis* subsp. *lactis* Bb-12 (Gao et al., 2009). The genome of *B. animalis* subsp. *lactis* V9 is 1,944,050 bp in length with a %G+C content of 60.5%, contains 1,636 genes, 52 tRNAs and 4 rRNA operons (Sun et al., 2010). Comparative analysis of *B. animalis* subsp. *lactis* V9 showed 13 SNPs between V9 and Bl-04 and 44 SNPs between V9 and DSM 10140 (Sun et al., 2010). Genome sequence for *B. animalis* subsp. *lactis* V9 is available with the accession number CP001892.

B. animalis subsp. *lactis* CNCM I-2494

B. animalis subsp. *lactis* CNCM I-2494 (also known as DN-173 010 and *Bifidus regularis*® was sequenced and made publically available in 2011 and by Danone Research (Chervaux et al., 2011). CNCM I-2494 is the culture collection designation for the strain DN-173 010 in the French National Collection of Cultures and Microorganisms (Agostini et al., 2012). It is unclear why the sequence was published as *B. animalis* subsp. *lactis* CNCM I-2494 and not as DN-173 010, which is the probiotic strain in Activia® yogurt. Fermented milk containing DN-173 010 (or CNCM I-2494) has been shown to reduce inflammation in mice (Veiga et al., 2010), improve gastrointestinal

health (Guyonnet et al., 2009), and shorten colonic transit time (Marteau et al., 2002). In 2010 sales of Activia reached over $444 million in the United States (Starling, 2011), demonstrating its commercial importance. The genome of CNCM I-2494 1,943,113 bp in length and 1,660 predicted ORFs with a %G+C 0f 60.5%, 4 rRNA operons. Comparative analysis revealed between 69-319 SNPs between CNCM I-2494 and AD011, DSM 10140, Bl-04 and Bb-12 (Chervaux et al., 2011). The original isolation source of this strain is not disclosed. The genome of *B. animalis* subsp. *lactis* CNCM I-2494 is available in GenBank with the accession number CP002915.

B. animalis subsp. *lactis* BLC1

B. animalis subsp. *lactis* BLC1 was sequenced and made publically available in 2011, from a collaborative group in Italy and Ireland (Bottacini et al., 2011). The strain BLC1 has been commercially available in Italy since about 2005 and in Italy the strain is sold by Clerici Sacco International (Bottacini et al., 2011). *B. animalis* subsp. *lactis* BLC1 has been shown to inhibit growth of *Clostridium difficile* and *C. perfringens* (Schoster et al., 2013) and produce conjugated linoleic acid from linoleic acid (Rodríguez-Alcalá et al., 2011). The BLC1 genome is 1,943,990 bp long and has a G+C content of 60.5%. This genome contains 1,622 genes 52 tRNAs and 4 rRNA operons (Bottacini et al., 2011). The BLC1 genome showed a very high degree of conservation when compared to all other sequenced strains of *B. animalis* subsp. *lactis*, and was especially similar to Bb-12 (Bottacini et al., 2011). This strain claims to have a human origin (BLC1 fact sheet, obtained via personal correspondence). The genome sequence

of *B. animalis* subsp. *lactis* BLC1 is available from GenBank with the accession number CP003039.

B. animalis subsp. *lactis* Bi-07 and B420

The genomic sequences of *B. animalis* subsp. *lactis* Bi07 and B420 became publically available in 2012, and were sequenced by DuPont Nutrition and Health (Stahl and Barrangou, 2012). *B. animalis* subsp. *lactis* Bi-07 has been shown to prevent bloating in patients with bowel disorders (Ringel-Kulka et al., 2011) and has exhibited immune modulation properties (Paineau et al., 2008). *B. animalis* subsp. *lactis* B420 has also exhibited immunomodulatory effects (Roessler et al., 2008) and help reduce nasal colonization by pathogenic bacteria (Gluck and Gebbers, 2003). Both strains are sold commercially as probiotics from DuPont Nutrition and Health. Comparative analysis identified nine previously unreported SNPs in Bi-07 and one previously unreported SNP in B420 previously when the sequences were aligned with other sequenced genomes of *B. animalis* subsp. *lactis* (Stahl and Barrangou, 2012). The origin for these strains was not disclosed. The genome of *B. animalis* subsp. *lactis* Bi-07 is 1,938,822 bp in length with 1,661 genes and B420 is 1,938,595 bp in length with 1,625 genes. Both genomes have a %G+C of 60.5% and 4 rRNA operons. *B. animalis* subsp. *lactis* Bi-07 is available from GenBank with the accession number CP003498 and B420 has the accession number CP003497.

B. animalis subsp. *lactis* Bl12

B. animalis subsp. *lactis* Bl12 is a bifidobacterial strain, which was isolated from a colon sample. This strain was one of 76 isolates of *B. animalis* subsp. *lactis* fecal and colonoscopic samples subjected to minimal inhibitor concentration (MIC) assay to assess resistance against tetracycline (Milani et al., 2013). This strain showed the greatest susceptibility to tetracycline (MIC = 16 μg/ml) and was significantly different than that of other sequenced strains (Bb-12 and BLC1 MIC = 32 μg/ml). Based on the difference observed in tetracycline resistance this strain was selected for sequencing to assess diversity within the *lactis* subspecies (Milani et al., 2013). The complete genome of Bl12 is 1,938,605 bp in length, with a %G+C of 60.5%, contains 1518 ORFs, 52 tRNAs and 4 rRNA loci. Alignment of Bl12 versus other sequenced strains exhibited >99.8% similarity, confirming the high degree of similarity within the subspecies even with the addition of new genomic sequence. The genome sequence is available from GenBank with the accession number CP004053.

Stress Response

The F_1-F_0-ATPase system has been suggested to be integral to acid resistance in *B. animalis* subsp. *lactis* (Matsumoto et al., 2004). The *atp* operon has been identified and analyzed in *B. animalis* subsp. *lactis* (Ventura et al., 2004). Additionally, analogs have been identified in sequenced genomes of other species (Lee and O'Sullivan, 2010). In the genome of *B. dentium* Bd1 an alternative acid resistance system was identified. This genome contains genes for the glutamate-dependent acid resistance system, which encodes glutamate decarboxylase (GadB) and glutamate/gamma-aminobutyrate anti-

porter GadC (Ventura et al., 2009c). The F_1-F_0-ATPase operon was not upregulated in *B. dentium* Bd1, when exposed to acid, highlighting the importance of the glutamate-dependent system (Ventura et al., 2009c)

NADH oxidase activity has been shown to significantly increase when *B. animalis* subsp. *lactis* is grown in the presence of oxygen suggesting a role in detoxification (Ruiz et al., 2012). While superoxide dismutase activity has been reported in bifidobacteria (Shimamura et al., 1992, Talwalkar and Kailasapathy, 2003), genome analysis of sequenced bifidobacteria has not identified the presence of superoxide dismutase genes (Lee and O'Sullivan, 2010). Despite the identification and analysis of a putative NADH oxidase gene (Ruiz et al., 2012), NADH peroxidase gene has not been identified in sequenced genomes (Lee and O'Sullivan, 2010) even though activity has been detected. This may suggest that false-positive activity has been measured, a novel NADH peroxidase is present bifidobacteria, or bifidobacteria use an alternative mechanism to rid the cell of H_2O_2. Lee and Sullivan propose an alternative method is the reduction of H_2O_2 to H_2O via the reduction of thioredoxin by thioredoxin reductase and NADH. This system has been identified in bifidobacteria and was evaluated by Ruiz et al. (2012) but did not show a significant increase in expression when cells were exposed to oxygen.

A bile salt hydrolase gene was identified in *B. animalis* subsp. *lactis* KL612 (Kim and Lee, 2008). A bile salt hydrolase-encoding gene in *B. dentium* (BDP_1106) was shown to increase 97-fold in expression with an extended exposure to acid, it was speculated that exposure to acid prepares the cell for exposure to bile salts, mimicking human consumption (Ventura et al., 2009c). Two bile efflux pumps have been described

in bifidobacteria. A cholate transporter, *ctr*, was identified in the genome *B. longum* NCIMB 702259, which was shown to provide resistance against cholate (Price et al., 2006). A bile-induced efflux transporter, BL0920, was identified in the genome of *B. longum* NCC2705 and the expression of this gene correlated to resistance to bile when inserted in *E. coli.* Homologs of BL0920, Bbr_0838, BAD_1491, Blon_1650 were also identified in *B. breve* UCC2003, *B. adolescentis* ATCC 15703, and *B. longum* subsp. *infantis* ATCC 15697, respectively (Gueimonde et al., 2009). This suggests *B. animalis* subsp. *lactis* may have a different mechanism for bile resistance than other members of bifidobacteria.

CRISPR in *Bifidobacterium*

Clustered regularly interspaced short palindromic repeats (CRISPR) were first recognized as early as 1987 (Ishino et al., 1987) but the term wasn't coined until 2002 (Jansen et al., 2002). A CRISPR locus consists of alternating repeat sequence separated by variable (spacer) sequences, with repeat sequence ranging from 21-48 bp and spacer sequence varying from 21-73 bp, and generally adjacent to *cas* genes (CRISPR-associated genes) (Barrangou et al., 2007, Deveau et al., 2010, Barrangou and Horvath, 2012). *In silico* analysis revealed spacer sequence exhibited identity to foreign DNA such as prophage and plasmid DNA (Pourcel et al., 2005, Barrangou et al., 2007). The CRISPR/Cas system has been shown to provide immunity against bacteriophage infection and cleave plasmid DNA (Barrangou et al., 2007, Garneau et al., 2010). Due to the fact that spacer sequences are added, when exposed to bacteriophage or plasmid DNA, sequentially from the leader end and deleted from the trailer end, the spacer

content and order can provide a historical record of exposure to bacteriophage, which can be exploited for strain identification and differentiation (Horvath et al., 2008, Horvath et al., 2009, Barrangou and Horvath, 2012). Typing methods, at least partially based on CRISPR sequence, have been developed for *Salmonella* (Liu et al., 2011, Shariat et al., 2013), and the potential exists for additional typing methods in other organisms (Barrangou and Horvath, 2012). Due to the high homogeneity of *B. animalis* subsp. *lactis*, and the potential for hypervariability of the CRISPR loci, this locus may provide a target for strain-level typing.

Analysis of bifidobacterial genomes revealed a general lack of phage-resistance machinery. Thus it is possible the CRISPR/Cas system plays an important role as a defense mechanism (Ventura et al., 2009b). Thus, genetic variability may exist in different species/strains of bifidobacteria making it useful to analyze for strain identification purposes. CRISPR/Cas systems have been identified in the majority of sequenced bifidobacterial genomes. Most sequenced genomes contain one locus however *B. dentium* Bd1 contains two loci (Ventura et al., 2009c), Barrangou and Horvath, 2012). The two CRISPR loci contained in Bd1 are from unrelated families, one with 80 spacers and the second with 16 spacers suggesting the occurrence of two separate horizontal gene transfer (HGT) events (Ventura et al., 2009c).

CRISPR loci can be strain-specific, as has been observed in *B. bifidum*. *B. bifidum* S17 contains one CRISPR locus containing 44 repeats and 43 spacers, while *B. bifidum* PRL 2010 is devoid of a CRISPR locus (Turroni et al., 2010, Zhurina et al., 2011, Barrangou and Horvath, 2012). Targeting this locus could make for a rapid, easy method to differentiate these strains. Additionally, *B. longum* subsp. *longum* DJO10A

contains a CRISPR locus with 43 repeats and 42 spacers, whereas *B. longum* subsp. *longum* NCC2705 does not contain a CRISPR locus (Schell et al., 2002, Lee et al., 2008, Barrangou and Horvath, 2012). These examples suggest CRISPR content can be considered strain-specific in bifidobacteria.

Comparative *Bifidobacterium* genomics

Recently, the core and pan-genome of the genus *Bifidobacterium* was determined using 14 completed or nearly complete genomes (Bottacini et al., 2010). The core genome is defined as "the pool of genes shared by all the strains of the same bacterial species" and the pan-genome is defined as "the global gene repertoire of a bacterial species: core genome plus dispensable genome" (Medini et al., 2005). The *Bifidobacterium* pan-genome contained 5,125 genes with a core genome of approximately 967 genes (Bottacini et al., 2010). From this analysis, each genome contributed, on average, 152 new genes to the pan-genome. In the case of *B. animalis* subsp. *lactis* the number of new unique genes became zero with the addition of new sequenced strains, reflective of the high relatedness of these strains. Functional analysis of the core genes reveals mostly housekeeping genes involved in replication, carbohydrate metabolism, transcription and translation (Bottacini et al., 2010). With the addition of new genomic sequence the pan-genome is expected to grow by addition of unique genes.

More recently, the core and pan-genomes was assessed of the genus *Bifidobacterium* using 19 complete genomes across nine species (Lukjancenko et al., 2012). This analysis revealed a bifidobacterial open-genome of 6,980 genes and a core

genome of 724 genes. Additionally, the average size of sequenced genomes was found to be 2,209 kbp, 59.5% G+C, and mean gene number of 1,796 (Lukjancenko et al., 2011). COG groups of each the core and pan-genome were also analyzed. The bifidobacterial core genome contains 33.9% information storage, 20.2% cellular processes and signaling, and 45.9% metabolism. The pan-genome contains 30% information storage, 21.9% cellular processes and signaling and 48.1% metabolism, with no significant differences between core and pan-genome in percent composition of COG categories. When comparing proteomes by BLAST analysis among sequenced bifidobacterial genomes the closest relative to *B. animalis* subsp. *lactis* DSM 10140 was found to be *B. gallicum* DSM 20083 with a proteome similarity 45.3% proteome similarity (Lukjancenko et al., 2011). Previous phylogenetic analysis using 16S rRNA gene sequences also indicated that *B. gallicum* is closest relative of analyzed species (Turroni et al., 2011). Interestingly, *B. adolescentis* ATCC 15703 and *B. animalis* subsp. *lactis* DSM 10140 share 41.4% proteome similarity, however these two species are distantly related based on 16S rRNA gene sequence and are present in different *Bifidobacterium* groups (Turroni et al., 2011, Lukjancenko et al., 2011). *B. animalis* subsp. *lactis* DSM 10140 shares between 30.7-41.3% proteome similarity with the remaining strains analyzed (Lukjancenko et al., 2011).

Recently, the pan-genome of *B. animalis* subsp. *lactis* was determined. It was determined the number of unique genes added to the pan-genome decreased significantly after the addition of the third genome (Milani et al., 2013). This suggested a closed pan-genome and the limited variability existing in this subspecies. Even the presence of a strain that was isolated from a human who purportedly did not consume probiotics or

probiotic containing food and a phenotypic difference (a small difference in tetracycline resistance) did not add new genomic information to the subspecies (Milani et al., 2013). With a greater amount of genomic information becoming available comparative genomic studies will gain in value.

Conclusion

B. animalis subsp. *lactis* is a microorganism commonly added to fermented dairy product due to its perceived health benefits. This organism has been shown to be commonly isolated from commercial probiotic dairy products even when the product does not claim to contain this subspecies. This is most likely due to the ability of *B. animalis* subsp. *lactis* to withstand the conditions of processing and storage in a fermented product. Despite its wide use little is known about the genetic features that allow for these phenotypic characteristics. In efforts to better understand this organism several sequencing projects have occurred. Sequence data from several strains has revealed a very high degree of similarity between strains, which has traditionally made differentiation of strains difficult. Methods to accurately differentiate strains and assess diversity within the subspecies are needed.

Table 1. General features of fully sequenced non-*B. animalis Bifidobacterium* strains

Organism	Size (Mb)	% G+C	Accession Number	Genes	Publication
B. adolescentis					
ATCC 15703	2.09	59.2	AP009256	1,702	unpublished
B. asteroides					
PRL 2011	2.17	60.1	CP003325	1,731	Bottacini et al. (2012)
B. bifidum					
BGN4	2.22	62.6	CP001361	1,902	Yu et al. (2012)
LMG 13195	2.28	62.5	AMPL01000000	1,966	Gueimonde et al. (2012)
PRL2010	2.21	62.7	CP001840	1,791	Turroni et al. (2010)
S17	2.28	62.8	CP002220	1,845	Zhurina et al. (2011)
B. breve					
ACS-071-V-Sch8b	2.33	58.7	CP002743	2,011	unpublished
CECT 7263	2.31	58.9	AFVV00000000	1,868	Jimenez et al. (2012)
UCC2003	2.42	58.7	CP000303	1,985	O'Connell Motherway et al. (2011)
B. dentium					
Bd1	2.64	58.5	CP001750	2,197	Ventura et al. (2009c)
B. longum subsp. *infantis*					
ATCC 15697	2.83	59.9	CP001905	2,588	Sela et al. (2008)
157F	2.41	60.1	AP010889	2,070	Fukuda et al. (2011)
B. longum subsp. *longum*					
NCC2705	2.26	60.1	AF139129	1,800	Schell et al. (2002)
DJO10A	2.39	60.1	CP000605	2,074	Lee et al. (2008)
KACC 91563	2.4	59.8	CP002794	2,050	Ham et al. (2011)
BBMN68	2.27	59.9	CP002286	1,878	Hao et al. (2011)
F8	2.38	59.9	FP929034	1,744	unpublished
JDM301	2.48	59.8	CP001095	2,035	Wei et al. (2012)
JCM 1217	2.39	60.3	AP010888	2,009	Fukuda et al. (2011)
B. thermophilum					
RBL67	2.29	60.1	CP004346	1,904	unpublished

Scope of this work

The first objective of this project was to sequence the complete genome of *B. animalis* subsp. *lactis* DSM 10140, the type strain. The second objective was to use genome sequence information to develop a typing scheme to differentiate strains of *B. animalis* subsp. *lactis*. The third objective was to sequence the genome of the type strain of *B. animalis* subsp. *animalis* ATCC 25527, the other subspecies of the *animalis* species to gain insight into the genetic variability of the subspecies and to compare growth in milk of *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis*. The fourth and final objective was to sequence *B. animalis* subsp. *lactis* ATCC 27673, a strain with a unique MLST type to examine genomic diversity in the subspecies *B. animalis* subsp. *lactis*.

Works Cited

Achtman, M. 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol 62:53-70.

Agostini, S., M. Goubern, V. Tondereau, C. Salvador-Cartier, V. Bezirard, M. Leveque, H. Keranen, V. Theodorou, S. Bourdu-Naturel, N. Goupil-Feuillerat, S. Legrain-Raspaud, and H. Eutamene. 2012. A marketed fermented dairy product containing *Bifidobacterium lactis* CNCM I-2494 suppresses gut hypersensitivity and colonic barrier disruption induced by acute stress in rats. Neurogastroenterology and Motility: The Official Journal of the European Gastrointestinal Motility Society 24(4):376-e172.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315(5819):1709-1712.

Barrangou, R. and P. Horvath. 2012. CRISPR: new horizons in phage resistance and strain identification. Annual Review of Food Science and Technology 3:143-162.

Basim, E. and H. Basim. 2001. Pulsed-Field Electrophoresis (PFGE) Technique and its use in Molecular Biology. Turkish Journal of Biology 25:405-418.

Beerens, H., F. Gavini, and C. Neut. 2000. Effect of exposure to air on 84 strains of bifidobacteria. Anaerobe 6:65-67.

Begley, M., C. G. Gahan, and C. Hill. 2005. The interaction between bacteria and bile. FEMS Microbiology Reviews 29(4):625-651.

Biavati, B. and P. Mattarelli. 2006. The family *Bifidobacteriaceae*. Pages 322-382 in The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community. Vol. 3. M. Dworkin, ed. Springer-Verlag, New York, NY.

Bonaparte, C. 1997. Selective isolation and taxonomic position of bifidobacteria isolated from commercial fermented dairy products in central Europe. Page 199. Berlin Technical University Dissertation.

Bonaparte, C. and G. Klein. 2012. International Committee on Systematics of Prokaryotes. Subcommittee on the taxonomy of *Bifidobacterium, Lactobacillus* and related organisms. Minutes of the meetings, 31 August 2011, Egmond aan Zee, The Netherlands. Int J Syst Evol Microbiol 62(Pt 10):2546-2548.

Bottacini, F., F. Dal Bello, F. Turroni, C. Milani, S. Duranti, E. Foroni, A. Viappiani, F. Strati, D. Mora, D. van Sinderen, and M. Ventura. 2011. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BLC1. Journal of Bacteriology 193(22):6387-6388.

Bottacini, F., D. Medini, A. Pavesi, F. Turroni, E. Foroni, D. Riley, V. Giubellini, H. Tettelin, D. van Sinderen, and M. Ventura. 2010. Comparative genomics of the genus *Bifidobacterium*. Microbiology 156(Pt 11):3243-3254.

Bottacini, F., C. Milani, F. Turroni, B. Sanchez, E. Foroni, S. Duranti, F. Serafini, A. Viappiani, F. Strati, A. Ferrarini, M. Delledonne, B. Henrissat, P. Coutinho, G. F. Fitzgerald, A. Margolles, D. van Sinderen, and M. Ventura. 2012. *Bifidobacterium asteroides* PRL2011 genome analysis reveals clues for colonization of the insect gut. PLoS One 7(9):e44229.

Brenner, D. J., J. T. Staley, and N. R. Krieg. 2001. Classification of procaryotic organisms and the concept of bacterial speciation. Pages 27-31 in Bergey's Manual of Systematic Bacteriology. D. R. Boone and R. W. Castenholz, ed. Srpinger-Verlag, New York, NY.

Briczinski, E. P. 2007. Characterization of *Bifidobacterium animalis* ssp. *lactis* from commericial culture manufacturers- a study of glucose transport. Page 252 in Food Science. Vol. Ph. D. Pennsylvania State University.

Briczinski, E. P., A. T. Phillips, and R. F. Roberts. 2008. Transport of glucose by *Bifidobacterium animalis* subsp. *lactis* occurs via facilitated diffusion. Applied and Environmental Microbiology 74(22):6941-6948.

Briczinski, E. P. and R. F. Roberts. 2006. Technical Note: A Rapid Pulsed-Field gel electrophoresis method for analysis of Bifidobacteria. Journal of Diary Science 89:2424-2427.

Bruno, F. A., W. E. V. Lankaputhra, and N. P. Shah. 2002. Growth, Viability and Activity of *Bifidobacterium* spp. in Skim Milk Containing Prebiotics. Journal of Food Science 67(7):2740-2744.

Cai, Y., M. Matsumoto, and Y. Benno. 2000. *Bifidobacterium lactis* Meile et al. 1997 Is a Subjective Synonym of *Bifidobacterium animalis* (Mitsuoka 1969) Scardovi and Trovatelli 1974. Microbiology and Immunology 44(10):815-820.

Candela, M., M. Centanni, J. Fiori, E. Biagi, S. Turroni, C. Orrico, S. Bergmann, S. Hammerschmidt, and P. Brigidi. 2010. DnaK from *Bifidobacterium animalis* subsp. *lactis* is a surface-exposed human plasminogen receptor upregulated in response to bile salts. Microbiology 156(Pt 6):1609-1618.

Champagne, C. P., N. J. Gardner, and D. Roy. 2005. Challenges in the addition of probiotic cultures to foods. Critical Reviews in Food Science and Nutrition 45(1):61-84.

Chervaux, C., C. Grimaldi, A. Bolotin, B. Quinquis, S. Legrain-Raspaud, J. E. van Hylckama Vlieg, G. Denariaz, and T. Smokvina. 2011. Genome sequence of the probiotic strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494. Journal of Bacteriology 193(19):5560-5561.

Chouraqui, J.-P., L.-D. Van Egroo, and M.-C. Fichot. 2004. Acidified milk formula supplemented with *Bifidobacterium lactis*: Impact on infant diarrhea in residential care settings. Journal of Pediatric Gastroenterology and Nutrition 38:288-292.

Chung, H. S., Y. B. Kim, S. L. Chun, and G. E. Ji. 1999. Screening and selection of acid and bile resistant bifidobacteria. International Journal of Food Microbiology 47:25-32.

Clark, P. A. and J. H. Martin. 1994. Selection of bifidobaceria for use of dietary adjuncts in cultured dairy foods. III. Tolerance to simulated bile concentrations of human small intestines. Cultured Dairy Products Journal 29(3):18-21.

Comas, I., S. Homolka, S. Niemann, and S. Gangneux. 2009. Genotyping of Genetically Monomorphic Bacteria: DNA Sequencing in *Mycobacterium tuberculosis* Highlights the Limitations of Current Methodologies. PLoS ONE 4(11):e7815.

Crittenden, R. G., L. F. Morris, M. L. Harvey, L. T. Tran, H. L. Mitchell, and M. J. Playne. 2001. Selection of a *Bifidobacterium* strain to complement resistant starch in synbiotic yogurt. Journal of Applied Microbiology 90:266-278.

De Vos, P. and H. G. Truper. 2000. Judicial Commission of the International Committee on Systematic Bacteriology. IXth International (IUMS) Congress of Bacteriology and Applied Microbiology. International Journal of Systematic and Evolutionary Microbiology 50:2239-2244.

Delétoile, A., V. Passet, J. Aires, I. Chambaud, M. J. Butel, T. Smokvina, and S. Brisse. 2010. Species delineation and clonal diversity in four *Bifidobacterium* species as revealed by multilocus sequencing. Research in Microbiology 161(2):82-90.

Desfosses-Foucault, E., V. Dussault-Lepage, C. Le Boucher, P. Savard, G. Lapointe, and D. Roy. 2012. Assessment of Probiotic Viability during Cheddar Cheese Manufacture and Ripening Using Propidium Monoazide-PCR Quantification. Frontiers in Microbiology 3:350.

Deveau, H., J. E. Garneau, and S. Moineau. 2010. CRISPR/Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol 64:475-493.

Donkor, O. N., A. Henriksson, T. Vasiljevic, and N. P. Shah. 2006. Effect of acidification on the activity of probiotics in yoghurt during cold storage. International Dairy Journal 16(10):1181-1189.

EFSA. 2009. Frequently asked questions (FAQ) related to the EFSA assessment of Article 14 and 13.5 health cliams application. EFSA Journal 7(9):1-8.

FAO/WHO. 2002. Report of a joint FAO/WHO working group on drafting guidelines for the evaluation of probiotics in food.

Fasoli, S., M. Marzotto, L. Rizzotti, F. Rossi, F. Dellaglio, and S. Torriani. 2003. Bacterial composition of commercial probiotic products as evaluated by PCR-DGGE analysis. International Journal of Food Microbiology 82:59-70.

Felis, G. E. and F. Dellaglio. 2007. Taxonomy of Lactobacilli and Bifidobacteria. Current Issues in Intestinal Microbiology 8:44-61.

Foster, J. W. 1993. The acid tolerance response to *Salmonella typhimurium* involved transient sythesis of key acid shock proteins. Journal of Bacteriology 175(7):1981-1987.

Fukuda, S., H. Toh, K. Hase, K. Oshima, Y. Nakanishi, K. Yoshimura, T. Tobe, J. M. Clarke, D. L. Topping, T. Suzuki, T. D. Taylor, K. Itoh, J. Kikuchi, H. Morita, M. Hattori, and H. Ohno. 2011. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature 469(7331):543-547.

Gao, P., Z. Sun, S. Ma, Q. Wang, J. Gao, C. Deng, and H. Zhang. 2009. Screening and identification of probiotic bifidobacterium from Mongolian children. Acta Microbiologica Sinica 49(2):210-216.

Garneau, J. E., M. E. Dupuis, M. Villion, D. A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A. H. Magadan, and S. Moineau. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468(7320):67-71.

Garrigues, C., E. Johansen, and M. B. Pedersen. 2010. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. Journal of Bacteriology 192(9):2467-2468.

Garrigues, C., B. Steur-Lauridsen, and E. Johansen. 2005. Characterisation of *Bifidobacterium animalis* subsp. *lactis* BB-12 and other probiotic bacteria using genomics, transriptomics and proteomics. Australian Journal of Dairy Technology 60(2):84-92.

Gluck, U. and J.-O. Gebbers. 2003. Ingested probiotics reduce nasal colonization with pathogenic bacteria (*Staphylococcus aureus*, *Streptococcus pneumoniae*, and -hemolytic streptococci). American Journal of Clinical Nutrition 77:517-520.

Gomes, A. M. P., F. X. Malcata, and F. A. M. Klaver. 1998. Growth Enhancement of *Bifidobacterium lactis* Bo and *Lactobacillus acidophilus* Ki by Milk Hydrolyzates. Journal of Dairy Science 81(11):2817-2825.

Grand, M., M. Kuffer, and A. Baumgartner. 2003. Quantitative analysis and molecular identification of bifidobacteria strains in probiotic milk products. European Food Research and Technology 217(1):90-92.

Gueimonde, M., S. Delgado, B. Mayo, P. Ruas-Madiedo, A. Margolles, and C. G. de los Reyes-Gavilán. 2004. Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. Food Research International 37(9):839-850.

Gueimonde, M., C. Garrigues, D. van Sinderen, C. G. de los Reyes-Gavilan, and A. Margolles. 2009. Bile-inducible efflux transporter from *Bifidobacterium longum* NCC2705, conferring bile resistance. Applied and Environmental Microbiology 75(10):3153-3160.

Gueimonde, M., M. Ventura, A. Margolles, and B. Sanchez. 2012. Genome sequence of the immunomodulatory strain *Bifidobacterium bifidum* LMG 13195. Journal of Bacteriology 194(24):6998.

Guyonnet, D., A. Schlumberger, L. Mhamdi, S. Jakob, and O. Chassany. 2009. Fermented milk containing *Bifidobacterium lactis* DN-173 010 improves gastrointestinal well-being and digestive symptoms in women reporting minor digestive symptoms: a randomised, double-blind, parallel, controlled study. The British Journal of Nutrition 102(11):1654-1662.

Ham, J. S., T. Lee, M. J. Byun, K. T. Lee, M. K. Kim, G. S. Han, S. G. Jeong, M. H. Oh, D. H. Kim, and H. Kim. 2011. Complete genome sequence of *Bifidobacterium longum* subsp. *longum* KACC 91563. Journal of Bacteriology 193(18):5044.

Hao, Y., D. Huang, H. Guo, M. Xiao, H. An, L. Zhao, F. Zuo, B. Zhang, S. Hu, S. Song, S. Chen, and F. Ren. 2011. Complete genome sequence of *Bifidobacterium longum* subsp. *longum* BBMN68, a new strain from a healthy chinese centenarian. Journal of Bacteriology 193(3):787-788.

Horvath, P., A. C. Coute-Monvoisin, D. A. Romero, P. Boyaval, C. Fremaux, and R. Barrangou. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol 131(1):62-70.

Horvath, P., D. A. Romero, A. C. Coute-Monvoisin, M. Richards, H. Deveau, S. Moineau, P. Boyaval, C. Fremaux, and R. Barrangou. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. Journal of Bacteriology 190(4):1401-1412.

Ibrahim, S. A. and A. Bezorovainy. 1993. Survival of Bifidobacteria in the presence of bile salt. Journal of Agricultural and Food Chemistry 62:351-354.

Ishino, Y., H. Shinagawa, K. Makino, M. Amemura, and A. Nakata. 1987. Nucleotide Sequence of the iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia coli, and Identification of the Gene Product. Journal of Bacteriology 169(12):5429-5433.

Janer, C., F. Arigoni, B. H. Lee, C. Pelaez, and T. Requena. 2005. Enzymatic ability of *Bifidobacterium animalis* subsp. *lactis* to hydrolyze milk proteins: identification and characterization of endopeptidase O. Applied and Environmental Microbiology 71(12):8460-8465.

Janer, C., C. Peláez, and T. Requena. 2004. Caseinomacropeptide and whey protein concentrate enhance *Bifidobacterium lactis* growth in milk. Food Chemistry 86(2):263-267.

Jansen, R., J. D. A. van Embden, W. Gaastra, and L. M. Schouls. 2002. Identification of novel family sequence repeats among prokaryotes. OMICS 6(1):22-33.

Jarocki, P. 2011. Molecular Characterization of Bile Salt Hydrolase from *Bifidobacterium animalis* subsp. *lactis* Bi30. Journal of Microbiology and Biotechnology 21(8):838-845.

Jayamanne, V. S. and M. R. Adams. 2006. Determination of survival, identity and stress resistance of probiotic bifidobacteria in bio-yoghurts. Letters in Applied Microbiology 42(3):189-194.

Jimenez, E., M. A. Villar-Tajadura, M. Marin, J. Fontecha, T. Requena, R. Arroyo, L. Fernandez, and J. M. Rodriguez. 2012. Complete genome sequence of *Bifidobacterium breve* CECT 7263, a strain isolated from human milk. Journal of Bacteriology 194(14):3762-3763.

Kailasapathy, K., I. Harmstorf, and M. Phillips. 2008. Survival of *Lactobacillus acidophilus* and *Bifidobacterium animalis* ssp. *lactis* in stirred fruit yogurts. LWT - Food Science and Technology 41(7):1317-1322.

Kaufmann, P., A. Pfefferkorn, M. Teuber, and L. Meile. 1997. Identification and quantification of *Bifidobacterium* species isolated from food with genus-specific 16S rRNA-targeted probes by colony hybridization and PCR. Applied and Environmental Microbiology 63(4):1268-1273.

Killer, J., J. Mrazek, V. Bunesova, J. Havlik, I. Koppova, O. Benada, V. Rada, J. Kopecny, and E. Vlkova. 2013. Pseudoscardovia suis gen. nov., sp. nov., a new member of the family *Bifidobacteriaceae* isolated from the digestive tract of wild pigs (*Sus scrofa*). Systematic and Applied Microbiology 36(1):11-16.

Kim, G. B. and B. H. Lee. 2008. Genetic analysis of a bile salt hydrolase in *Bifidobacterium animalis* subsp. *lactis* KL612. J Appl Microbiol 105(3):778-790.

Kim, J. F., H. Jeong, D. S. Yu, S. H. Choi, C. G. Hur, M. S. Park, S. H. Yoon, D. W. Kim, G. E. Ji, H. S. Park, and T. K. Oh. 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. Journal of Bacteriology 191(2):678-679.

Kim, J. Y., Y. O. Choi, and G. E. Ji. 2008. Effect of Oral Probiotics (*Bifidobacterium lactis* AD011 and *Lactobacillus acidophilus* AD031) Administration on Ovalbumin-Induced Food Allergy Mouse Model. Journal of Microbiology and Biotechnology 18(8):1393-1400.

Kim, M. S., S. W. Roh, and J. W. Bae. 2010. *Bifidobacterium stercoris* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 60(Pt 12):2823-2827.

Klein, G. 2009. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Bifidobacterium*, *Lactobacillus* and related organisms: Minutes of the meetings, 3 September 2008, Egmond aan Zee, The Netherlands. International Journal of Systematic and Evolutionary Microbiology 59(12):3181-3183.

Kociubinski, G., P. Perez, and G. De Antoni. 1999. Screening of Bile Resistance and Bile Precipitation in Lactic Acid Bacteria and Bifidobacteria. Journal of Food Protetction 62(8):905-912.

Larsen, C. N., S. Nielsen, P. Kaestel, E. Brockmann, M. Bennedsen, H. R. Christensen, D. C. Eskesen, B. L. Jacobsen, and K. F. Michaelsen. 2006. Dose-response study of probiotic bacteria *Bifidobacterium animalis* subsp *lactis* BB-12 and *Lactobacillus paracasei* subsp *paracasei* CRL-341 in healthy young adults. European Journal of Clinical Nutrition 60(11):1284-1293.

Lee, J. H., V. N. Karamychev, S. A. Kozyavkin, D. Mills, A. R. Pavlov, N. V. Pavlova, N. N. Polouchine, P. M. Richardson, V. V. Shakhova, A. I. Slesarev, B. Weimer, and D. J. O'Sullivan. 2008. Comparative genomic analysis of the gut bacterium *Bifidobacterium longum* reveals loci susceptible to deletion during pure culture growth. BMC Genomics 9:247.

Lee, J. H. and D. J. O'Sullivan. 2010. Genomic insights into bifidobacteria. Microbiology and Molecular Biology Reviews : MMBR 74(3):378-416.

Leyer, G. J., S. Li, M. E. Mubasher, C. Reifer, and A. C. Ouwehand. 2009. Probiotic effects on cold and influenza-like symptom incidence and duration in children. Pediatrics 124(2):e172-179.

Li, Q., Q. Chen, H. Ruan, D. Zhu, and G. He. 2010. Isolation and characterisation of an oxygen, acid and bile resistant *Bifidobacterium animalis* subsp. *lactis* Qq08. Journal of the Science of Food and Agriculture 90(8):1340-1346.

Lilly, D. and R. H. Stillwell. 1965. Probiotics: growth promoting factors produced by microorganisms. Science 47:747-748.

Liu, F., S. Kariyawasam, B. M. Jayarao, R. Barrangou, P. Gerner-Smidt, E. M. Ribot, S. J. Knabel, and E. G. Dudley. 2011. Subtyping *Salmonella enteric*a serovar enteritidis isolates from different sources by using sequence typing based on virulence genes and clustered regularly interspaced short palindromic repeats (CRISPRs). Applied and Environmental Microbiology 77(13):4520-4526.

Luis Noriega, C. G., C. G. de Los Reyes-Gavilan, and A. Margolles. 2005. Acquisition of bile salt resistance promots antibiotic susceptibility changes in *Bifidobacterium*. Journal of Food Protetction 68(9):1916-1919.

Lukjancenko, O., D. W. Ussery, and T. M. Wassenaar. 2012. Comparative genomics of *Bifidobacterium, Lactobacillus* and related probiotic genera. Microbial Ecology 63(3):651-673.

Margolles, A. and B. Sanchez. 2012. Selection of a *Bifidobacterium animalis* subsp. *lactis* strain with a decreased ability to produce acetic acid. Applied and Environmental Microbiology 78(9):3338-3342.

Marteau, P., E. Cuillerier, S. Meanace , M. F. Gerhardt, A. Myara, M. Bouvier, C. Bouley, F. Tondu, G. Bommelaer, and J. C. Grimaud. 2002. *Bifdobacterium animalis* strain DN-173 010 shortens the colonic transit time in healthy women: a double-blind, randomized, controlled study. Alimentary Pharmocology and Therapeutics 16:587-593.

Marty-Teysset, C., F. de la Torre, and J. R. Garel. 2000. Increased Production of Hydrogen Peroxide by *Lactobacillus delbrueckii* subsp. *bulgaricus* upon Aeration: Involvement of an NADH Oxidase in Oxidative Stress. Applied and Environmental Microbiology 66(1):262-267.

Masco, L., G. Huys, E. De Brandt, R. Temmerman, and J. Swings. 2005. Culture-dependent and culture-independent qualitative analysis of probiotic products claimed to contain bifidobacteria. Int J Food Microbiol 102(2):221-230.

Masco, L., M. Ventura, R. Zink, G. Huys, and J. Swings. 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(Pt 4):1137-1143.

Matsumoto, M., H. Ohishi, and Y. Benno. 2004. H+-ATPase activity in *Bifidobacterium* with special reference to acid tolerance. Int J Food Microbiol 93(1):109-113.

Mättö, J., H.-L. Alakomi, A. Vaari, I. Virkajärvi, and M. Saarela. 2006. Influence of processing conditions on *Bifidobacterium animalis* subsp. *lactis* functionality with a special focus on acid tolerance and factors affecting it. International Dairy Journal 16(9):1029-1037.

Mättö, J., E. Malinen, M. L. Suihko, M. Alander, A. Palva, and M. Saarela. 2004. Genetic heterogeneity and functional properties of intestinal bifidobacteria. J Appl Microbiol 97(3):459-470.

Maxwell, F. J., S. H. Duncan, G. Hold, and C. S. Stewart. 2004. Isolation, growth on prebiotics and probiotic potential of novel bifidobacteria from pigs. Anaerobe 10(1):33-39.

Mayer, A., H. Seiler, and S. Scherer. 2003. Isolation of bifidobacteria from food and human faeces and rapid identification by Fourier transform infrared spectroscopy. Pages 299-313 in Annals of Microbiology. Vol. 53.

Mayer, H. K., E. Amtmann, E. Philippi, G. Steinegger, S. Mayrhofer, and W. Kneifel. 2007. Molecular discrimination of new isolates of *Bifidobacterium animalis* subsp. *lactis* from reference strains and commercial probiotic strains. International Dairy Journal 17(5):565-573.

McLauchlan, G., G. M. Fullarton, G. P. Crean, and K. E. L. McColl. 1989. Comparison of gastric body and antral pH: a 24 hour ambulatory study in healthy volunteers. Gut 30:573-578.

Medini, D., C. Donati, H. Tettelin, V. Masignani, and R. Rappuoli. 2005. The microbial pan-genome. Current opinion in Genetics & Development 15(6):589-594.

Meile, L., W. Ludwig, U. Reuger, C. Gut, P. Kaufmann, G. Dasen, S. Wenger, and M. Teuber. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. Systematic and Applied Microbiology 20:57-64.

Merenstein, D., J. Gonzalez, A. G. Young, R. F. Roberts, M. E. Sanders, and S. Petterson. 2011. Study to investigate the potential of probiotics in children attending school. European Journal of Clinical Nutrition 65(4):447-453.

Metchinkoff, E. 1907. The prolongation of life. Pages 1-100 in Optimistic Studies. W. Heinemann, ed. G. P. Putnam & Sons, London, UK.

Milani, C., S. Duranti, G. A. Lugli, F. Bottacini, F. Strati, S. Arioli, E. Foroni, F. Turroni, D. van Sinderen, and M. Ventura. 2013. Comparative genomics of *Bifidobacterium animalis* subsp. *lactis* reveals a strict monophyletic bifidobacterial taxon. Applied and Environmental Microbiology.

Noriega, L., M. Gueimonde, B. Sanchez, A. Margolles, and C. G. de los Reyes-Gavilan. 2004. Effect of the adaptation to high bile salts concentrations on glycosidic activity, survival at low PH and cross-resistance to bile salts in *Bifidobacterium*. Int J Food Microbiol 94(1):79-86.

O'Connell Motherway, M., A. Zomer, S. C. Leahy, J. Reunanen, F. Bottacini, M. J. Claesson, F. O'Brien, K. Flynn, P. G. Casey, J. A. Munoz, B. Kearney, A. M. Houston, C. O'Mahony, D. G. Higgins, F. Shanahan, A. Palva, W. M. de Vos, G. F. Fitzgerald, M. Ventura, P. W. O'Toole, and D. van Sinderen. 2011. Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. Proceedings of the National Academy of Sciences of the United States of America 108(27):11217-11222.

Oberg, T. S., J. L. Steele, S. C. Ingham, V. V. Smeianov, E. P. Briczinski, A. Abdalla, and J. R. Broadbent. 2011. Intrinsic and inducible resistance to hydrogen peroxide in *Bifidobacterium* species. Journal of Industrial Microbiology & Biotechnology 38(12):1947-1953.

Paineau, D., D. Carcano, G. Leyer, S. Darquy, M. A. Alyanakian, G. Simoneau, J. F. Bergmann, D. Brassart, F. Bornet, and A. C. Ouwehand. 2008. Effects of seven potential

probiotic strains on specific immune responses in healthy adults: a double-blind, randomized, controlled trial. FEMS Immunology and Medical Microbiology 53(1):107-113.

Poupard, J. A., I. Husain, and R. F. Norris. 1973. Biology of bifidobacteria. Bacteriological Reviews:136-165.

Pourcel, C., G. Salvignol, and G. Vergnaud. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 151(Pt 3):653-663.

Price, C. E., S. J. Reid, A. J. Driessen, and V. R. Abratt. 2006. The Bifidobacterium longum NCIMB 702259T ctr gene codes for a novel cholate transporter. Applied and Environmental Microbiology 72(1):923-926.

Ringel-Kulka, T., O. S. Palsson, D. Maier, I. Carrol, J. A. Galanko, G. Leyer, and Y. Ringel. 2011. Probiotic Bacteria *Lactobacillus acidophilu*s NCFM and *Bifidobacterium lactis* Bi-07 Versus Placebo for the Symptoms of Bloating in Patients With Functional Bowel Disorders A Double-blind Study. Journal of Clinical Gastroenterology 45:518-525.

Rizzardini, G., D. Eskesen, P. C. Calder, A. Capetti, L. Jespersen, and M. Clerici. 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 randomised, double-blind, placebo-controlled study. The British Journal of Nutrition 107(6):876-884.

Rodríguez-Alcalá, L. M., T. Braga, F. Xavier Malcata, A. Gomes, and J. Fontecha. 2011. Quantitative and qualitative determination of CLA produced by *Bifidobacterium* and lactic acid bacteria by combining spectrophotometric and Ag+-HPLC techniques. Food Chemistry 125(4):1373-1378.

Roessler, A., U. Friedrich, H. Vogelsang, A. Bauer, M. Kaatz, U. C. Hipler, I. Schmidt, and G. Jahreis. 2008. The immune system in healthy adults and patients with atopic dermatitis seems to be affected differently by a probiotic intervention. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 38(1):93-102.

Roy, D. and S. Sirois. 2000. Molecular differentiation of *Bifidobacterium* species with amplified ribosomal DNA restriction analysis and alignment of short regions of the ldh gene. FeMS Microbiolgy Letters 191(1):17-24.

Ruas-Madiedo, P., A. Hernandez-Barranco, A. Margolles, and C. G. de los Reyes-Gavilan. 2005. A bile salt-resistant derivative of *Bifidobacterium animalis* has an altered fermentation pattern when grown on glucose and maltose. Applied and Environmental Microbiology 71(11):6564-6570.

Ruiz, L., M. Gueimonde, P. Ruas-Madiedo, A. Ribbera, C. G. de Los Reyes-Gavilan, M. Ventura, A. Margolles, and B. Sanchez. 2012. Molecular clues to understand the aerotolerance phenotype of *Bifidobacterium animalis* subsp. *lactis*. Applied and Environmental Microbiology 78(3):644-650.

Ruiz, L., B. Sanchez, P. Ruas-Madiedo, C. G. de Los Reyes-Gavilan, and A. Margolles. 2007. Cell envelope changes in *Bifidobacterium animalis* ssp. *lactis* as a response to bile. FEMS Microbiol Lett 274(2):316-322.

Ruiz-Moyano, S., N. Tao, M. A. Underwood, and D. A. Mills. 2012. Rapid discrimination of *Bifidobacterium animalis* subspecies by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Food Microbiology 30(2):432-437.

Saarela, M., G. Mogensen, R. Fonden, J. Matto, and T. Mattila-Sandholm. 2000. Probiotic bacteria: safety, functional, and technological properties. Journal of Biotechnology 84:197-215.

Saarela, M., M. Rantala, K. Hallamaa, L. Nohynek, I. Virkajarvi, and J. Matto. 2004. Stationary-phase acid and heat treatments for improvement of the viability of probiotic lactobacilli and bifidobacteria. J Appl Microbiol 96(6):1205-1214.

Saarela, M., I. Virkajarvi, H. L. Alakomi, T. Mattila-Sandholm, A. Vaari, T. Suomalainen, and J. Matto. 2005. Influence of fermentation time, cryoprotectant and neutralization of cell concentrate on freeze-drying survival, storage stability, and acid and bile exposure of *Bifidobacterium animalis* ssp. *lactis* cells produced without milk-based ingredients. J Appl Microbiol 99(6):1330-1339.

Sanchez, B., M. C. Champomier-Verges, B. Stuer-Lauridsen, P. Ruas-Madiedo, P. Anglade, F. Baraige, C. G. de los Reyes-Gavilan, E. Johansen, M. Zagorec, and A. Margolles. 2007. Adaptation and response of *Bifidobacterium animalis* subsp. *lactis* to bile: a proteomic and physiological approach. Applied and Environmental Microbiology 73(21):6757-6767.

Sanchez, B., C. G. de los Reyes-Gavilan, and A. Margolles. 2006. The F1F0-ATPase of *Bifidobacterium animalis* is involved in bile tolerance. Environ Microbiol 8(10):1825-1833.

Sanders, M. E. 2006. Summary of probiotic activities of *Bifidobacterium lactis* HN019. Journal of Clinical Gastroenterology 40(9):776-783.

Sanders, M. E., F. Guarner, R. Guerrant, P. R. Holt, E. M. M. Quigley, R. B. Sartor, P. M. Sherman, and E. A. Mayer. 2012. An update on the use and investigation of probiotics in health and disease. Gut

Schell, M. A., M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M. C. Zwahlen, F. Desiere, P. Bork, M. Delley, R. D. Pridmore, and F. Arigoni. 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human

gastrointestinal tract. Proceedings of the National Academy of Sciences of the United States of America 99(22):14422-14427.

Schoster, A., B. Kokotovic, A. Permin, P. D. Pedersen, F. D. Bello, and L. Guardabassi. 2013. In vitro inhibition of *Clostridium difficile* and *Clostridium perfringens* by commercial probiotic strains. Anaerobe 20:36-41.

Sela, D. A., J. Chapman, A. Adeuya, J. H. Kim, F. Chen, T. R. Whitehead, A. Lapidus, D. S. Rokhsar, C. B. Lebrilla, J. B. German, N. P. Price, P. M. Richardson, and D. A. Mills. 2008. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. Proceedings of the National Academy of Sciences of the United States of America 105(48):18964-18969.

Shariat, N., M. J. Dimarzio, S. Yin, L. Dettinger, C. H. Sandt, J. R. Lute, R. Barrangou, and E. G. Dudley. 2013. The combination of CRISPR-MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of Salmonella enterica subsp. enterica serovar Enteritidis. Food Microbiology 34(1):164-173.

Shimamura, S., F. Abe, N. Ishibashi, H. Miyakawa, T. Yaeshima, T. Araya, and M. Tomita. 1992. Relationshlp Between Oxygen Sensitivity and Oxygen Metabolism of Blfldobacterium Species. Journal of Diary Science 75(12):3296-3306.

Shin, H. S., J. H. Lee, J. J. Pestka, and Z. Ustunol. 2000. Growth and Viability of Commerical *Bifidobacterium* spp in Skim Milk Containing Oligosaccharides and Inulin. Journal of Food Science 65(5):884-887.

Simpson, P. J., C. Stanton, G. F. Fitzgerald, and R. P. Ross. 2003. Genomic Diversity and Relatedness of Bifidobacteria Isolated from a Porcine Cecum. Journal of Bacteriology 185(8):2571-2581.

Simpson, P. J., C. Stanton, G. F. Fitzgerald, and R. P. Ross. 2005. Intrinsic tolerance of *Bifidobacterium* species to heat and oxygen and survival following spray drying and storage. J Appl Microbiol 99(3):493-501.

Stahl, B. and R. Barrangou. 2012. Complete genome sequences of probiotic strains of *Bifidobacterium animalis* subsp. *lactis* B420 and Bi-07. Journal of Bacteriology 194(15):4131-4132.

Starling, S. 2011. Dannon probiotic yogurt sales up 16% despite claims crackdown. Nutra Ingredients-usa.com.

Sun, Z., X. Chen, J. Wang, P. Gao, Z. Zhou, Y. Ren, T. Sun, L. Wang, H. Meng, W. Chen, and H. Zhang. 2010. Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. Journal of Bacteriology 192(15):4080-4081.

Tabbers, M. M., A. Chmielewska, M. G. Roseboom, N. Crastes, C. Perrin, J. B. Reitsma, O. Norbruis, H. Szajewska, and M. A. Benninga. 2011. Fermented milk containing

Bifidobacterium lactis DN-173 010 in childhood constipation: a randomized, double-blind, controlled trial. Pediatrics 127(6):e1392-1399.

Talwalkar, A. and K. Kailasapathy. 2003. Metabolic and Biochemical Responses of Probiotic Bacteria to Oxygen. Journal of Diary Science 86:2537-2546.

Talwalkar, A. and K. Kailasapathy. 2004. A review of oxygen toxicity in yogurts: Influence on the survival of probiotic bacteria and protective techniques. Comprehensive Reviews in Food Science and Food Safety 3:117-124.

Talwalkar, A., K. Kailasapathy, P. Peiris, and R. Arumugaswamy. 2001. Application of RBGR- a simple way for screening of oxygen tolerance in probiotic bacteria. International Journal of Food Microbiology 71:245-248.

Temmerman, R., I. Scheirlinck, G. Huys, and J. Swings. 2003. Culture-Independent Analysis of Probiotic Products by Denaturing Gradient Gel Electrophoresis. Applied and Environmental Microbiology 69(1):220-226.

Tissier, H. 1899. La reaction chromphile d'Escherich et le bacterium coli. Comptes Rendus Hebdomadaires des Seances et Memoires de la Societe de Biologie 51:943-945.

Tmanova, L. L., A. Onyenwoke, and R. F. Roberts. 2012. Short communication: Identification and differentiation of bifidobacteria obtained from Ukraine. Journal of Dairy Science 95(1):91-97.

Turroni, F., F. Bottacini, E. Foroni, I. Mulder, J.-H. Kim, A. Zomer, B. Sanchez, A. Bidossi, A. Ferrarini, V. Giubellini, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, G. F. Fitzgerald, D. Mills, A. Margolles, D. Kelly, D. van Sinderen, and M. Ventura. 2010. Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-dervied glycan foraging PNAS 107(45):19514-19519.

Turroni, F., D. van Sinderen, and M. Ventura. 2009. Bifidobacteria: from ecology to genomics. Frontiers in Bioscience 14:4673-4684.

Turroni, F., D. van Sinderen, and M. Ventura. 2011. Genomics and ecological overview of the genus *Bifidobacterium*. Int J Food Microbiol 149(1):37-44.

Veiga, P., C. A. Gallini, C. Beal, M. Michaud, M. L. Delaney, A. DuBois, A. Khlebnikov, J. E. T. van Hylckama Vlieg, S. Punit, J. N. Glickman, A. Onderdonk, L. H. Glimcher, and W. S. Garrett. 2010. *Bifidobacterium animalis* subsp. *lactis* fermented milk product reduces inflammation by altering a niche for colitogenic microbes. Proceedings of the National Academy of Sciences 107(42):18132-18137.

Ventura, M., C. Canchaya, A. Del Casale, F. Dellaglio, E. Neviani, G. F. Fitzgerald, and D. van Sinderen. 2006. Analysis of bifidobacterial evolution using a multilocus approach. Int J Syst Evol Microbiol 56(Pt 12):2783-2792.

Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater, and D. van Sinderen. 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiology and molecular biology reviews : MMBR 71(3):495-548.

Ventura, M., C. Canchaya, D. van Sinderen, G. F. Fitzgerald, and R. Zink. 2004. Bifidobacterium lactis DSM 10140: Identification of the *atp* (*atpBEFHAGDC*) Operon and Analysis of Its Genetic Structure, Characteristics, and Phylogeny. Applied and Environmental Microbiology 70(5):3110-3121.

Ventura, M., S. O'Flaherty, M. J. Claesson, F. Turroni, T. R. Klaenhammer, D. van Sinderen, and P. W. O'Toole. 2009a. Genome-scale analyses of health-promoting bacteria: probiogenomics. Nature reviews. Microbiology 7(1):61-71.

Ventura, M., R. Reniero, and R. Zink. 2001. Specific identification and targeted characterization of *Bifidobacterium lactis* from different environmental isolates by a combined multiplex-PCR approach. Applied and Environmental Microbiology 67(6):2760-2765.

Ventura, M., F. Turroni, G. Lima-Mendez, E. Foroni, A. Zomer, S. Duranti, V. Giubellini, F. Bottacini, P. Horvath, R. Barrangou, D. A. Sela, D. A. Mills, and D. van Sinderen. 2009b. Comparative analyses of prophage-like elements present in bifidobacterial genomes. Applied and Environmental Microbiology 75(21):6929-6936.

Ventura, M., F. Turroni, A. Zomer, E. Foroni, V. Giubellini, F. Bottacini, C. Canchaya, M. J. Claesson, F. He, M. Mantzourani, L. Mulas, A. Ferrarini, B. Gao, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, R. S. Gupta, Z. Zhang, D. Beighton, G. F. Fitzgerald, P. W. O'Toole, and D. van Sinderen. 2009c. The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. PLoS Genetics 5(12):e1000785.

Ventura, M. and R. Zink. 2002. Rapid Identification, Differentiation, and Proposed New Taxonomic Classification of *Bifidobacterium lactis*. Applied and Environmental Microbiology 68(12):6429-6434.

Ventura, M. and R. Zink. 2003. Comparative Sequence Analysis of the tuf and recA Genes and Restriction Fragment Length Polymorphism of the Internal Transcribed Spacer Region Sequences Supply Additional Tools for Discriminating Bifidobacterium lactis from *Bifidobacterium animalis*. Applied and Environmental Microbiology 69(12):7517-7522.

Vernazza, C. L., G. R. Gibson, and R. A. Rastall. 2006. Carbohydrate preference, acid tolerance and bile tolerance in five strains of *Bifidobacterium*. J Appl Microbiol 100(4):846-853.

Villegas, E. and S. E. Gilliland. 1998. Hydrogen peroxide production by *Lactobacillus delbrueckii* subsp. *lactis* at 5C. Journal of Food Science 63(6):1070-1074.

Vincent, D., D. Roy, F. Mondou, and C. Dery. 1998. Characterization of bifidobacteria by random DNA amplification. International Journal of Food Microbiology 43:185-193.

Waddington, L., T. Cyr, M. Hefford, L. T. Hansen, and M. Kalmokoff. 2010. Understanding the acid tolerance response of bifidobacteria. J Appl Microbiol 108(4):1408-1420.

Waller, P. A., P. K. Gopal, G. J. Leyer, A. C. Ouwehand, C. Reifer, M. E. Stewart, and L. E. Miller. 2011. Dose-response effect of *Bifidobacterium lactis* HN019 on whole gut transit time and functional gastrointestinal symptoms in adults. Scandinavian Journal of Gastroenterology 46(9):1057-1064.

Ward, P. and D. Roy. 2005. Review of molecular methods for identification, characterization and detection of bifidobacteria. Le Lait 85(1-2):23-32.

Wei, Y. X., Z. Y. Zhang, C. Liu, P. K. Malakar, and X. K. Guo. 2012. Safety assessment of *Bifidobacterium longum* JDM301 based on complete genome sequences. World Journal of Gastroenterology : WJG 18(5):479-488.

Williams, J. G. K., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18(22):6531-6535.

Yonezawa, S., J. Z. Xiao, T. Odamaki, T. Ishida, K. Miyaji, A. Yamada, T. Yaeshima, and K. Iwatsuki. 2010. Improved growth of bifidobacteria by cocultivation with *Lactococcus lactis* subspecies *lactis*. Journal of Dairy Science 93(5):1815-1823.

Yu, D. S., H. Jeong, D. H. Lee, S. K. Kwon, J. Y. Song, B. K. Kim, M. S. Park, G. E. Ji, T. K. Oh, and J. F. Kim. 2012. Complete genome sequence of the probiotic bacterium *Bifidobacterium bifidum* strain BGN4. Journal of Bacteriology 194(17):4757-4758.

Zhu, L., W. Li, and D. X. 2003. Species identification of genus Bifidobacterium based on partial HSP60 gene sequences and proposal of *Bifidobacterium thermacidophilum* subsp. *porcinum* subsp. nov. International Journal of Systematic and Evolutionary Microbiology 53(5):1619-1623.

Zhurina, D., A. Zomer, M. Gleinser, V. F. Brancaccio, M. Auchter, M. S. Waidmann, C. Westermann, D. van Sinderen, and C. U. Riedel. 2011. Complete genome sequence of *Bifidobacterium bifidum* S17. Journal of Bacteriology 193(1):301-302.

Problem Statement

B. animalis subsp. *lactis* is a widely consumed probiotic. Prior to this work the differentiation of strains of this subspecies was not possible by current methods. Because health-promoting effects are considered to be strain specific the differentiation of strains is important. Identification of strains is also important to identify what strains are present in commercial probiotic containing products. Additionally, because of the difficulty to differentiate strains an understanding of the diversity within the subspecies in not well understood.

Chapter 2

The complete genome sequence of *Bifidobacterium animalis* subsp. *animalis* ATCC 25527^T and comparative analysis of growth in milk with *Bifidobacterium animalis* subsp. *lactis* DSM 10140^T

J. R. Loquasto, R. Barrangou, E. G. Dudley, R. F. Roberts

Published as: **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). Journal of dairy science **94:**5864-5870.

Statement of Contribution: The candidate was responsible for design, implementation, analysis, and manuscript preparation.

Abstract

The objective of this work was to sequence the genome of *B. animalis* subsp. *animalis* ATCC 25527^T the subspecies most closely related to the widely used *B. animalis* subsp. *lactis* probiotic strains. The complete 1,932,963 bp genome was determined by a combination of 454 shotgun sequencing and PCR gap closing and the completed assembly was verified by comparison to a *Kpn*I optical map. Comparative analysis of the *B. animalis* subsp. *animalis* ATCC 25527^T and *B. animalis* subsp. *lactis* DSM 10140^T genomes revealed high degrees of both synteny and sequence homology. Comparative genomic analysis revealed 156 and 182 genes that were unique to and absent in, the *B. animalis* subsp. *animalis* genome, respectively. Among these was a set of unique CRISPR-associated genes and a novel CRISPR locus containing 30 spacers in the genome of *B. animalis* subsp. *animalis*. Although previous research has suggested one of the defining phenotypic differences between *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* is the ability of the later to grow in milk and milk-based media, the differential gene content did not provide insights to explain these differences. Furthermore, growth and acid production in milk and milk-based media did not differ significantly between *B. animalis* subsp. *lactis* (DSM 10140^T and Bl04) and *B. animalis* subsp. *animalis* (ATCC 25527^T). Growth of these strains in supplemented milk suggested growth was limited by lack of available low molecular weight nitrogen in the three strains examined.

Key Words: *B. animalis* subsp. *animalis*, genome, growth in milk

Introduction

Members of the genus *Bifidobacterium* are considered to be natural inhabitants of the human and other mammalian gastrointestinal tract (GIT) and numerous fermented dairy products are supplemented with bifidobacteria as a probiotic adjunct (Turroni et al., 2009). One widely used bifidobacterial species is *B. animalis*, which contains the two subspecies *lactis* and *animalis*. Generally, many strains within the *lactis* subspecies are regarded as technologically suitable for use as probiotics because they have been reported to be more acid-, bile-, and oxygen-tolerant than other members of the genus (Mainville et al., 2005, Jayamanne and Adams, 2006). Furthermore, *B. animalis* subsp. *lactis* has been reported to grow in milk and milk-based media, whereas *B. animalis* subsp. *animalis* has been reported to lack this ability (Meile et al., 1997, Masco et al., 2004). Some authors have speculated *B. animalis* subsp. *lactis* diverged from *B. animalis* subsp. *animalis* and adapted specifically for growth in milk and that its genome has been streamlined for this specific niche (Lee and O'Sullivan, 2010). It has also been suggested the ability to grow in milk differentiates the two subspecies (Masco et al., 2004).

Although multiple fully sequenced genomes are available for *B. animalis* subsp. *lactis*, (Barrangou et al., 2009, Kim et al., 2009, Garrigues et al., 2010, Sun et al., 2010), the genome of the most closely related subspecies, *B. animalis* subsp. *animalis*, has not been sequenced. Thus, the objective of this work was to completely sequence the genome of *B. animalis* subsp. *animalis* ATCC 25527^T and compare it to that of *B. animalis* subsp. *lactis*. Additionally, the ability of both subspecies to grow in milk and milk-based media was assessed.

Materials and Methods

Bacterial Strains and Culture Conditions

The type strain of *B. animalis* subsp. *lactis* DSM 10140^T was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM; The German Collection of Microorganisms and Cell Cultures; Braunshweig, Germany) and grown on MRS (de Man et al., 1960) supplemented with 0.05% (w/v) cysteine hydrochloride (MRSC). The commercial strain *B. animalis* subsp. *lactis* Bl04 was obtained from Danisco USA, Inc. (Madison, WI). *B. animalis* subsp. *animalis* ATCC 25527^T was obtained from the American Type Culture Collection (ATCC; Manassas, VA). *Bifidobacterium*-specific primers developed by Kaufmann *et al* (1997) were used with the following modified PCR protocol: 20 cycles of 94°C for 1 min, 57°C for 0.5 min, 72°C for 1.5 min. Cultures of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* were verified to the subspecies level using PCR primers and conditions designed by Ventura and Zink (2002).

Sequencing and assembly of the *B. animalis* subsp. *animalis* genome

Genomic DNA was isolated from 10 mL overnight culture of *B. animalis* subsp. *animalis* ATCC 25527^T grown in MRS + cysteine hydrochloride (0.05%) broth using the Wizard Genomic DNA Purification Kit (Promega; Madison, WI). Genomic DNA was submitted for 454 pyrosequencing on a GS-FLX sequencer at The Pennsylvania State University. Sequencing was followed by *de novo* assembly in Newbler (Roche; Branford, CT). Following assembly contigs were aligned using PGA (pheromone trail-

based genetic algorithm (Zhao et al., 2008)) using the *B. animalis* subsp. *lactis* DSM 10140^T genome as a reference (Barrangou et al., 2009). To close gaps, primers were designed on the end of each contig and PCR was conducted with appropriate primers. PCR products were sequenced at the Huck Institute Genomic Core Facility at The Pennsylvania State University using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators; Applied Biosystems; Foster City, CA) and a ABI 3730XL DNA analyzer with ABI sequence analysis software (version 5.1.1). Sequences were aligned and added to contigs to generate a completed circular genome using SeqMan (DNASTAR; Madison, WI).

A *Kpn*I optical map of the *B. animalis* subsp. *animalis* ATCC 25527^T genome was generated at OpGen (Gaithersburg, MD). The assembled genome was aligned with the optical map using MapSolver (OpGen) to confirm genome assembly and also with an *in silico Kpn*I optical map of *B. animalis* subsp. *lactis* DSM 10140^T. The genome sequence of *B. animalis* subsp. *animalis* ATCC 25527^T was submitted for automated annotation at NCBI and deposited in GenBank (accession number CP002567).

Comparative Genomic Analysis

The genomic sequences of *B. animalis* subsp. *animalis* ATCC 25527^T and *B. animalis* subsp. *lactis* DSM 10140^T were aligned using the Mauve alignment tool with default settings (Darling *et al,* 2004). Differential gene content was identified using RAST (Aziz *et al,* 2008) as having less than 60% amino acid sequence identity (Chaudhuri et al., 2010) between sequences in the two genomes. Clustered regularly interspaced short palindromic repeats (CRISPR) were identified in the *B. animalis* subsp.

animalis ATCC 25527^T genome using Dotter and CRISPRFinder (Grissa et al., 2007). The total number of SNPs existing between the ATCC 25527^T and DSM 10140^T genomes was determined following alignment using Mauve. When protein sequences were aligned, BLASTP was used (Sayers et al., 2012).

Growth in Milk

Growth of *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* was evaluated in 10% reconstituted skim milk (RSM, Becton Dickinson; Franklin Lakes, NJ), RSM supplemented with 0.5% peptone and 1% yeast extract (M+), and RSM supplemented with 0.5% casamino acids (MC). Prior to inoculation of milk cultures were prepared from frozen stock by growth in MRSC containing 2% (w/v) lactose at 37°C in an anaerobic incubator (85% nitrogen, 10% carbon dioxide, and 5% hydrogen) for 16-18 h. One mL of turbid culture was then transferred to autoclaved 10% RSM and allowed to incubate anaerobically for 24 h. After activation, one mL of the 24 h culture was transferred to each treatment, mixed and then incubated anaerobically at 37°C. Samples were taken initially, and then every 6 hours, for 24 hours, for determination of pH and of viable counts. Counts were made by pour-plating on MRSC containing 2% (w/v) lactose agar and incubating anaerobically for 48-72 h. Two replications were performed. Mean increase in populations and decrease in pH were compared using ANOVA in Minitab (State College, PA, version 15).

Results and Discussion

Genome Sequencing and Comparison

The complete genome of *B. animalis* subsp. *animalis* ATCC 25527^T was 1,932,693 bp in length, 5,790 bp shorter than that of *B. animalis* subsp. *lactis* DSM 10140^T (1,938,483 bp, Table 1). The assembly was verified by comparison to a *Kpn*I optical map (Figure 1A) and only minor differences were observed between the *in vivo* and *in silico* optical maps, confirming the assembly of the *B. animalis* subsp. *animalis* ATCC 25527^T genome was accurate. Automated annotation from NCBI predicted the genome of *B. animalis* subsp. *animalis* ATCC 25527^T contains 1,597 genes, 34 less than the 1,629 genes annotated in the genome of *B. animalis* subsp. *lactis* DSM 10140^T. The general characteristics of these two genomes along with the other sequenced *B. animalis* subsp. *lactis* strains appear in Table 1. Analysis of the two genomes for differential content revealed 156 and 182 genes in the *animalis* subspecies that are unique and absent, respectively, as compared to the *lactis* subspecies. Unique genes were identified as having less than 60% amino acid identity as detected by RAST. The Mauve alignment shows a high degree of both homology and synteny between the two genomes (Figure 1B) but also identified 73,021 SNPs between the two genomes indicating sequence diversity.

Differential content analysis did not reveal obvious genes or loci that may play a role in the ability of either organism to grow in milk. Genes thought to potentially be important for growth in milk were examined by protein-protein alignment in BLAST. When genes predicted to play a role in lactose transport and utilization were evaluated

((Balat_0475 (*lacS*) and Balat_0476 (*lacZ*)) they were found to share 99% amino acid identity with their respective homologs (Banan_2470 and Banan_2475). Balat_1174, annotated as *pepO* and previously evaluated by Janer et al. (2005), was aligned with Banan_5790, exhibiting 98% amino acid sequence identity. Likewise, all other genes with predicted peptidase activity examined showed greater than 96% amino acid sequence identity, again suggesting a minimal difference in predicted function.

CRISPR

A novel CRISPR locus, Bana1, was identified in the genome of *B. animalis* subsp. *animalis* ATCC 25527^{T}, at 1,656-1,668 Mb, which contains 32 repeats and 30 spacers, and 8 *cas* (CRISPR-associated) genes. The novel 29 bp CRISPR repeat, 5'-GTTTGCCCCGCACAGGCGGGGATGATCCG-3' is homologous to repeats from the Ldbu1 CRISPR family (Horvath et al., 2008), and belongs to the Type IE CRISPR/Cas system (previously known as *E. coli* type), with universal *cas1* and *cas2* genes, as well as the signature Type I gene *cas3* (COG1203) (Makarova et al., 2011). Although this CRISPR repeat sequence is novel, homologous 29 bp CRISPR repeats have been identified in the genomes of *Bifidobacterium gallicum* DSM20093, *Bifidobacterium catenulatum* DSM16992 and *Bifidobacterium angulatum* DSM20098. The repeat-spacer array is interrupted by an independently confirmed 77 bp random sequence located in-between the 10th and 11th repeat. Notwithstanding the paucity of sequence information available for bifidobacterial phages and plasmids, spacer S3 showed high homology to a tape measure protein from a *Bifidobacterium dentium* prophage sequence, and spacers S4, S19 and S28 showed homology to plasmid sequences from *Burkholderia*, *Ralstonia*

and *Streptomyces*, respectively (Figure 2). Additionally, several spacers showed homology to sequences from environmental genomic surveys, likely containing high levels of viral particles

Although genome content, architecture and synteny is overall high between *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis*, several polymorphic islands provide insight with regards to speciation. Interestingly, CRISPR content was hypervariable between the two subspecies, with the presence of a unique locus in each subspecies. Indeed, the Bana1 and Bala (Barrangou et al., 2009) loci are unique to *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis*, respectively. This is consistent with CRISPR locus hypervariability previously observed in bifidobacteria (Briczinski et al., 2009, Horvath et al., 2009), and their susceptibility to horizontal gene transfer. Together with spacer sequence homology to foreign genetic elements, notably a *Bifidobacterium* prophage sequence (Ventura et al., 2009) this is consistent with the involvement of CRISPR/Cas systems in providing defense against invasive genetic elements (Barrangou et al., 2007, Horvath and Barrangou, 2010). The occurrence of different CRISPR/Cas systems in closely related organisms is likely an indicator of the selective pressure foreign genetic elements, such as phages, apply on these bacterial hosts (Andersson and Banfield, 2008, Tyson and Banfield, 2008, Horvath and Barrangou, 2010). The critical role CRISPR/Cas systems play in bifidobacteria is also highlighted by the notable absence of other phage resistance mechanisms such as restriction modification and abortive infection (Ventura et al., 2009).

Comparison of Growth in Milk

Growth of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* ATCC 25527^T in milk and supplemented milk was evaluated over 24 h (Figure 3). The increase in population was not statistically different when compared across each of the milk-based media evaluated. Populations of *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* increased by approximately 1 log after 24 h incubation in un-supplemented milk. Populations in the supplemented milks (for all three organisms) were similar and increased by ~2-2.5 logs (10-fold greater), indicating a statistically significant increase when compared to those grown in un-supplemented milk ($p < 0.001$).

Acid production (pH) was also monitored over 24 hours to assess the ability of *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* to acidify milk-based media. Organisms grown in milk reduced the pH approximately 0.5 pH units after 24 hours (no statistical difference between organisms, terminal pH ranging from 5.87 to 6.09). Organisms grown in milk supplemented with 0.5% casamino acids or 0.5% peptone and 1% yeast extract decreased the pH 2.0-2.5 units after 24 hours, (terminal pH ranging from 4.01 to 4.46) except for DSM 10140^T grown in 0.5% casamino acids which decreased the pH approximately 1.5 units. Although DSM 10140^T grew to similar populations as ATCC 25527^T and Bl04 in milk supplemented with 0.5% casamino acids, the growth resulted in a significantly smaller reduction of pH ($p < 0.001$).

Supplementation of RSM with casamino acids, a source of amino acids and peptides from casein, stimulated growth of both *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* suggesting proteolytic activity is limiting for growth in milk. This finding differs from the data reported previously by Masco et al. (2004) who observed an

increase of 2.0-2.5 log CFu/ml after 24 hours for *B. animalis* subsp. *lactis* strains grown in RSM.

Conclusions

The study provides insight into the genetic diversity that exists in the *B. animalis* species. Differential gene content did not provide a clear explanation for phenotypic differences alleged between the two subspecies. A novel CRISPR locus, Bana1, containing 32 repeats and 30 spacers was identified in the genome of *B. animalis* subsp. *animalis* ATCC 25527^T. In addition and in contrast to previous reports, growth of *B. animalis* subsp. *animalis* ATCC 25527^T in milk did not statistically differ from growth of *B. animalis* subsp. *lactis* DSM 10140^T or Bl04. It was also determined that addition of casamino acids improved growth after 24 hours, suggesting that *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* may both lack the necessary proteolytic activity needed for good growth in milk.

Acknowledgements

J.R.L is funded, in part, by the Penn State University Microbial Functional Genomics USDA-AFRI Training Grant (Contract 2010-65110-20488). The authors would also like to thank Dr. Stephan Schuster for assistance with 454-sequencing and assembly and the Penn State Genomics Core Facility, University Park, PA for Sanger sequencing.

Works Cited

Andersson, A. F. and J. F. Banfield. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. Science 320(5879):1047-1050.

Barrangou, R., E. P. Briczinski, L. L. Traeger, J. R. Loquasto, M. Richards, P. Horvath, A. C. Coute-Monvoisin, G. Leyer, S. Rendulic, J. L. Steele, J. R. Broadbent, T. Oberg, E. G. Dudley, S. Schuster, D. A. Romero, and R. F. Roberts. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. Journal of Bacteriology 191(13):4144-4151.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315(5819):1709-1712.

Briczinski, E. P., J. R. Loquasto, R. Barrangou, E. G. Dudley, A. M. Roberts, and R. F. Roberts. 2009. Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* by using single-nucleotide polymorphisms, insertions, and deletions. Applied and Environmental Microbiology 75(23):7501-7508.

Chaudhuri, R. R., M. Sebaihia, J. L. Hobman, M. A. Webber, D. L. Leyton, M. D. Goldberg, A. F. Cunningham, A. Scott-Tucker, P. R. Ferguson, C. M. Thomas, G. Frankel, C. M. Tang, E. G. Dudley, I. S. Roberts, D. A. Rasko, M. J. Pallen, J. Parkhill, J. P. Nataro, N. R. Thomson, and I. R. Henderson. 2010. Complete Genome Sequence and Comparative Metabolic Profiling of the Prototypical Enteroaggregative *Escherichia coli* Strain 042. PLoS ONE 5(1):e8801.

de Man, J. D., M. Rogosa, and M. E. Sharpe. 1960. A medium for the cultivation of lactobacilli. Journal of Applied Bacteriology (23):130-135.

Garrigues, C., E. Johansen, and M. B. Pedersen. 2010. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. Journal of Bacteriology 192(9):2467-2468.

Grissa, I., G. Vergnaud, and C. Pourcel. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Research 35(Web Server issue):W52-57.

Horvath, P. and R. Barrangou. 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327(5962):167-170.

Horvath, P., A. C. Coute-Monvoisin, D. A. Romero, P. Boyaval, C. Fremaux, and R. Barrangou. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol 131(1):62-70.

Horvath, P., D. A. Romero, A. C. Coute-Monvoisin, M. Richards, H. Deveau, S. Moineau, P. Boyaval, C. Fremaux, and R. Barrangou. 2008. Diversity, activity, and

evolution of CRISPR loci in *Streptococcus thermophilus*. Journal of Bacteriology 190(4):1401-1412.

Janer, C., F. Arigoni, B. H. Lee, C. Pelaez, and T. Requena. 2005. Enzymatic ability of *Bifidobacterium animalis* subsp. *lactis* to hydrolyze milk proteins: identification and characterization of endopeptidase O. Applied and Environmental Microbiology 71(12):8460-8465.

Jayamanne, V. S. and M. R. Adams. 2006. Determination of survival, identity and stress resistance of probiotic bifidobacteria in bio-yoghurts. Letters in Applied Microbiology 42(3):189-194.

Kim, J. F., H. Jeong, D. S. Yu, S. H. Choi, C. G. Hur, M. S. Park, S. H. Yoon, D. W. Kim, G. E. Ji, H. S. Park, and T. K. Oh. 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. Journal of Bacteriology 191(2):678-679.

Lee, J. H. and D. J. O'Sullivan. 2010. Genomic insights into bifidobacteria. Microbiology and Molecular Biology Reviews : MMBR 74(3):378-416.

Mainville, I., Y. Arcand, and E. R. Farnworth. 2005. A dynamic model that simulates the human upper gastrointestinal tract for the study of probiotics. International Journal of Food Microbiology 99(3):287-296.

Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost, and E. V. Koonin. 2011. Evolution and classification of the CRISPR-Cas systems. Nature Reviews. Microbiology 9(6):467-477.

Masco, L., M. Ventura, R. Zink, G. Huys, and J. Swings. 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(Pt 4):1137-1143.

Meile, L., W. Ludwig, U. Reuger, C. Gut, P. Kaufmann, G. Dasen, S. Wenger, and M. Teuber. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. Systematic and Applied Microbiology 20:57-64.

Sayers, E. W., T. Barrett, D. A. Benson, E. Bolton, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. Dicuccio, S. Federhen, M. Feolo, I. M. Fingerman, L. Y. Geer, W. Helmberg, Y. Kapustin, S. Krasnov, D. Landsman, D. J. Lipman, Z. Lu, T. L. Madden, T. Madej, D. R. Maglott, A. Marchler-Bauer, V. Miller, I. Karsch-Mizrachi, J. Ostell, A. Panchenko, L. Phan, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, M. Shumway, K. Sirotkin, D. Slotta, A. Souvorov, G. Starchenko, T. A. Tatusova, L. Wagner, Y. Wang, W. J. Wilbur, E. Yaschenko, and J. Ye. 2012. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 40(Database issue):D13-25.

Sun, Z., X. Chen, J. Wang, P. Gao, Z. Zhou, Y. Ren, T. Sun, L. Wang, H. Meng, W. Chen, and H. Zhang. 2010. Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. Journal of Bacteriology 192(15):4080-4081.

Turroni, F., E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G. L. de'Angelis, F. Shanahan, D. van Sinderen, and M. Ventura. 2009. Exploring the diversity of the bifidobacterial population in the human intestinal tract. Applied and Environmental Microbiology 75(6):1534-1545.

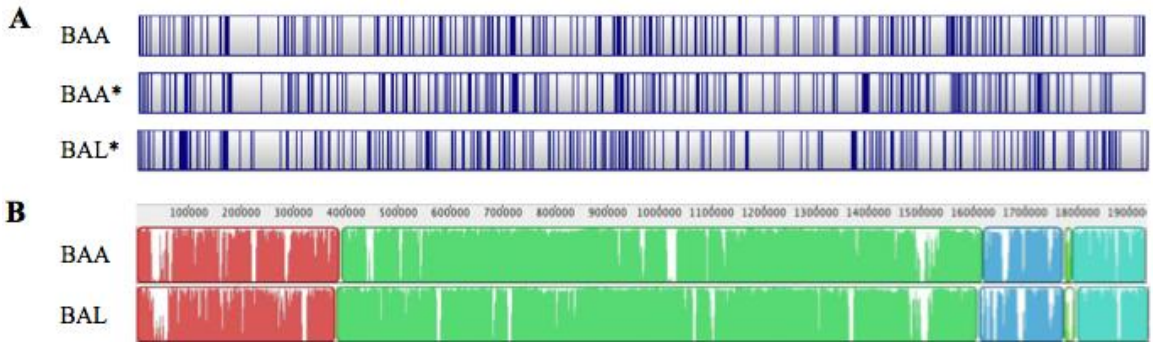
Tyson, G. W. and J. F. Banfield. 2008. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Environ Microbiol 10(1):200-207.

Ventura, M., F. Turroni, G. Lima-Mendez, E. Foroni, A. Zomer, S. Duranti, V. Giubellini, F. Bottacini, P. Horvath, R. Barrangou, D. A. Sela, D. A. Mills, and D. van Sinderen. 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. Applied and Environmental Microbiology 75(21):6929-6936.

Ventura, M. and R. Zink. 2002. Rapid Identification, Differentiation, and Proposed New Taxonomic Classification of *Bifidobacterium lactis*. Applied and Environmental Microbiology 68(12):6429-6434.

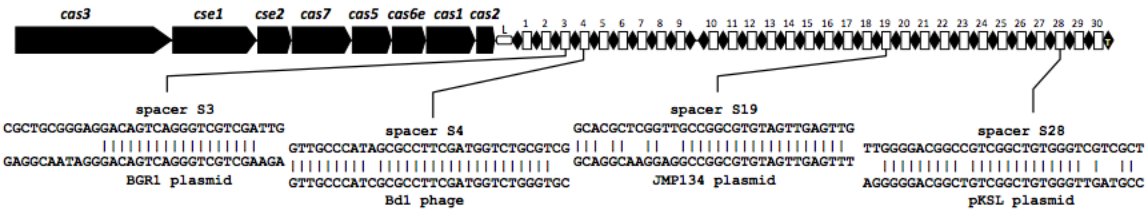
Zhao, F., F. Zhao, T. Li, and D. A. Bryant. 2008. A new pheromone trail-based genetic algorithm for comparative genome assembly. Nucleic Acids Research 36(10):3455-3462.

Figure 1. Comparisons of the complete genomes of *B. animalis* subsp. *animalis* ATCC 25527^{T} (BAA) and *B. animalis* subsp. *lactis* DSM 10140^{T} (BAL)



Panel A) Experimental and *in silico* (*) *Kpn*I optical maps of BAA and BAL. Panel B) Alignment of the complete genomes of BAA and BAL constructed using the Mauve alignment tool (Darling *et al*, 2004). Both figures begin at the same position on the BAA genome.

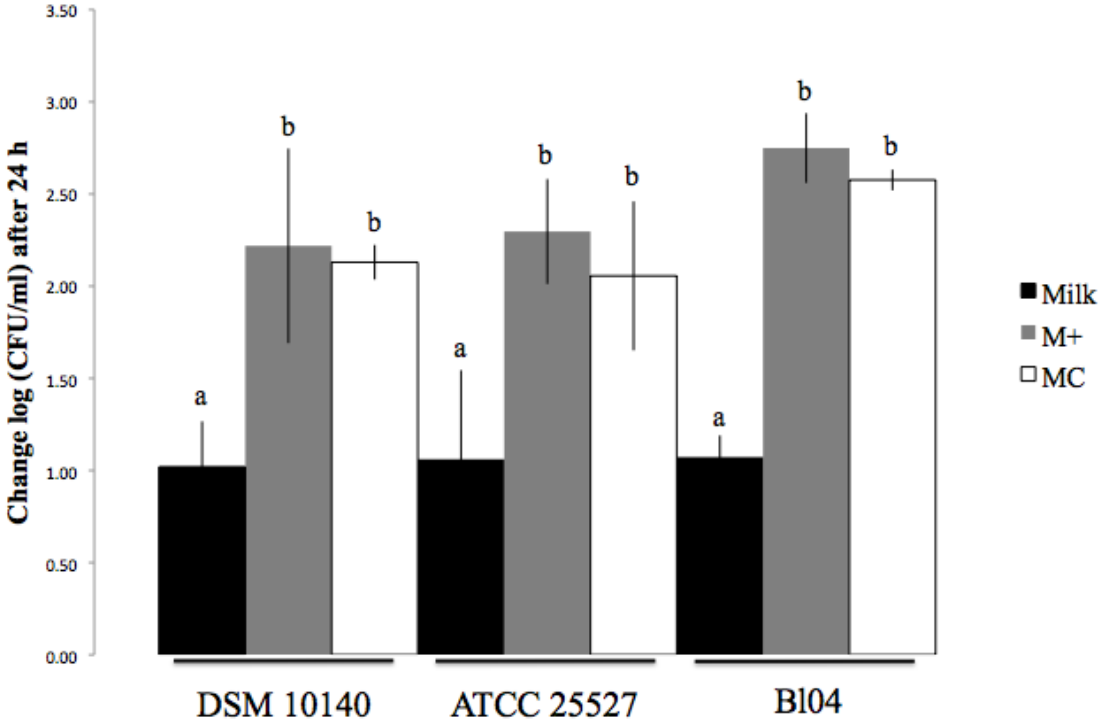
Figure 2. The *B. animalis* subsp. *animalis* ATCC 25527^{T} CRISPR/Cas system locus (Bana1) as it appears in the genome



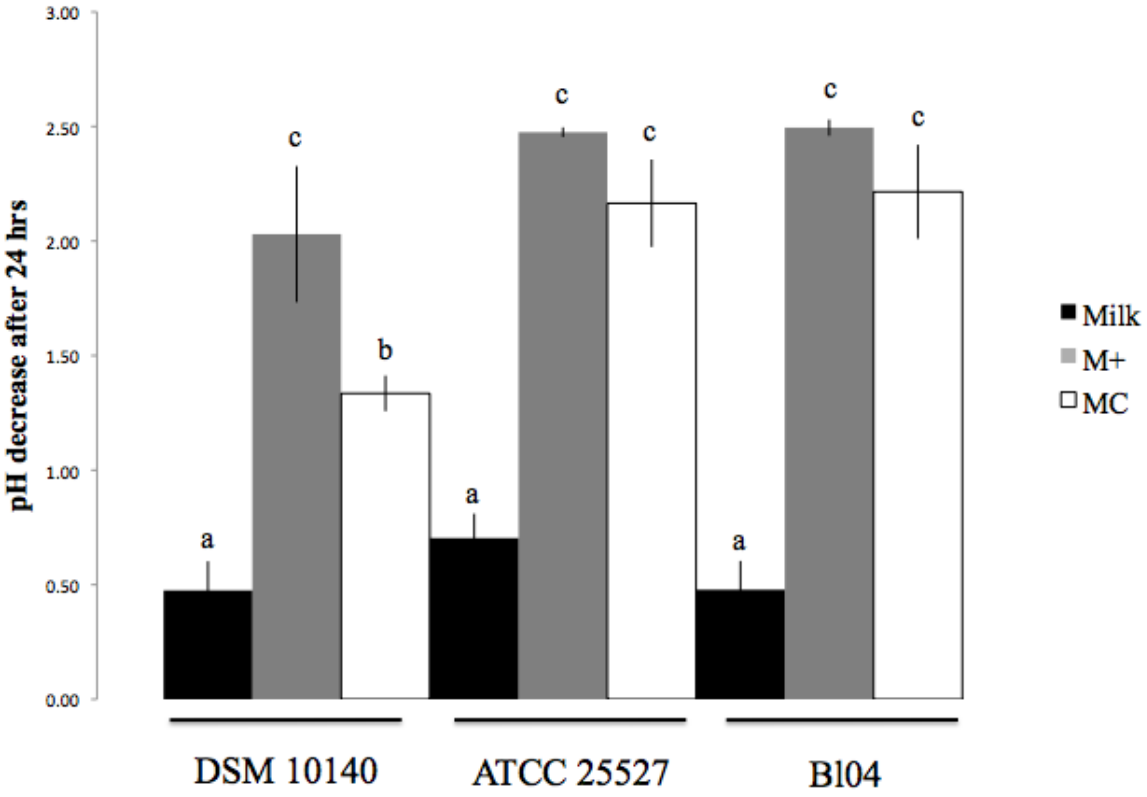
The *cas* genes are located on the left followed by the leader sequence (designated L) and the repeat-spacer region. Repeats and spacers are designated as black diamonds and white boxes, respectively. The terminal repeat is indicated by a "T". The sequence of spacers S3, S4, S19 and S28 were found to have significant homology with the indicated element.

Figure 3. Evaluation of growth in milk and milk supplemented with 0.5% casamino acids or supplement with 0.5% peptone and 1% yeast extract

A



B



Panel A) Change in population (log CFu/mL) of each organism after anaerobic incubation for 24h at 37C in the indicated medium. Panel B) pH decrease after 24 hours. Letters represent statistically significant differences after 24 hours ($P < 0.05$). Milk = 10% RSM, M+ = 10% RSM + 1% peptone + 0.5% casamino acids, MC = 10% RSM + 0.5% casamino acids. Initial levels for each inoculation were approximately 3.0×10^6 CFu/mL.

1

Table 1. General Characteristics

Organism	GenBank Accession #	Length (bp)	Coding %	% G+C	Genes	rRNA operons	tRNAs	IS elements
B. animalis subsp. *animalis* ATCC 25527^T	CP002567	1,932,963	85.4	60.47	1,597	4	52	4
B. animalis subsp. *lactis* DSM 10140^T	CP001606	1,938,483	90.3	60.48	1,629	4	51	6
B. animalis subsp. *lactis* Bl04	CP001515	1,938,709	90.5	60.48	1,631	4	52	6
B. animalis subsp. *lactis* Bb-12	CP001853	1,942,198	89.9	60.48	1,642	4	52	6
B. animalis subsp. *lactis* V9	CP001892	1,944,050	86.0	60.50	1,636	4	52	7
B. animalis subsp. *lactis* HN019[1]	ABOT00000000	1,915,892	86.0	60.00	1,632	1	52	4
B. animalis subsp. *lactis* AD011[2]	CP001213	1,933,695	84.6	60.49	1,528	2	52	7

[1]Currently exists as 28 contigs

[2] Assembly and annotation issues might explain differences in size and number of genes

Chapter 3

Bifidobacterium animalis subsp. *lactis* ATCC 27673 is a genomically unique strain within this conserved subspecies

Joseph R. Loquasto, Rodolphe Barrangou, Edward G. Dudley, Buffy Stahl, Chun Chen, Robert F. Roberts

Currently under review at Applied and Environmental Microbiology

Statement of Contribution: The candidate proposed the idea, developed and implemented the project, analyzed the data, and prepared the manuscript.

Abstract

Many strains of *Bifidobacterium animalis* subsp. *lactis* are considered health-promoting probiotic microorganisms and are commonly formulated into fermented dairy foods such as yogurt. Analyses of previously sequenced genomes of *B. animalis* subsp. *lactis* has revealed very little genetic diversity, suggesting it is a monomorphic subspecies. However, during an MLST survey of the genus *Bifidobacterium* it was revealed that *B. animalis* subsp. *lactis* ATCC 27673 gave a distinct profile from 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 and contained 1,616 genes, 4 rRNA operons, and had a G+C content of 61.55%. Comparative analyses revealed the genome of ATCC 27673 contained six distinct genomic islands, encoding 83 ORFs not found in other strains of the same subspecies. In four islands, either phage or mobile genetic elements were identified. In island 6, the largest island, a novel CRISPR locus was identified, which contains 81 novel spacers. 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 ATCC 27673 is a strain of *B. animalis* subsp. *lactis* with novel genetic content and suggests the lack of genetic variability observed to date is likely due to the repeated sequencing of a limited number of widely distributed commercial strains

Introduction

Many *Bifidobacterium animalis* subsp. *lactis* strains are considered to be probiotic microorganisms, and are commonly added to a variety of fermented and non-fermented foods (Gueimonde et al., 2004, Masco et al., 2004). 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 (Rizzardini et al., 2012), increased digestive comfort (Guyonnet et al., 2009), and reduction of colonic transit time (Marteau et al., 2002). Technological advantages claimed for this subspecies include tolerance to oxygen (Simpson et al., 2005), resistance to acid (Matsumoto et al., 2004, Vernazza et al., 2006) and bile (Sanchez et al., 2007), as well as viability during extended refrigerated storage (Kailasapathy et al., 2008). *B. animalis* subsp. *lactis* DSM 10140 (originally identified as *B. lactis*) was first described in 1997 as unique species of *Bifidobacterium* and was identified as an oxygen tolerant isolate from a fermented milk sample (Meile et al., 1997).

In previous work, the genome of the Type strain of the subspecies DSM 10140 and a commercial strain Bl-04 were completely sequenced and subjected to comparative analysis (Barrangou et al., 2009). This analysis revealed the two genomes were highly conserved as only 47 SNPs and 4 INDELs were identified between the two strains. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) 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 Bl-04 (Barrangou et al, 2009).

Sequencing of 8 additional genomes has confirmed the highly monomorphic nature of the *B. animalis* subsp. *lactis* genome (Kim et al., 2009, Garrigues et al., 2010, Sun et al., 2010, Bottacini et al., 2011, Chervaux et al., 2011, Stahl and Barrangou, 2012, Milani et al., 2013). 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 isolates of *B. animalis* subsp. *lactis* have been completely sequenced it is not clear 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* (Loquasto et al., 2011) was sequenced. The genome of *B. animalis* subsp. *animalis* ATCC 25527 exhibited approximately 96.23% similarity with the genome of *B. animalis* subsp. *lactis* (Loquasto et al., 2011). Of note, *B. animalis* subsp. *animalis* had a CRISPR locus that differed in terms of spacer and repeat sequences and overall size compared to *B. animalis* subsp. *lactis* and classified as a Type I-E CRISPR-*cas* system (Makarova et al., 2011, Makarova and Koonin, 2013).

When evaluating speciation within the genus *Bifidobacterium*, Delétoile et al. (2010) subjected a number of species of bifidobacteria to a multi-locus 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* based on 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 it with 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 (de Man et al., 1960) supplemented with 0.05% (wt/vol) cysteine hydrochloride (MRSc). All cultures were grown anaerobically (85% N, 10% CO_2, 5% H_2) at 37°C. Genomic DNA was isolated from 10 mL of overnight culture using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) according to manufacturer's instructions. Subspecies specific primers and conditions developed by Ventura and Zink (2002) were used to verify the culture as *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 was generated. Sequencing was followed by *de novo* assembly in Newbler (Roche, Branford, CT) and the contigs generated were aligned with the genome of *B. animalis* subsp. *lactis* DSM 10140 (Barrangou et al., 2009) using the pheromone trail-based genetic algorithm (Zhao et al., 2008). To close remaining gaps in

the sequence, PCR primers were designed on the end of each ordered contig and PCR was conducted with the corresponding contig PCR primer. PCR products were sequenced at The Pennsylvania State University using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators; Applied Biosystems, Foster City, CA) and ABI 3730XL DNA analyzer with ABI sequence analysis software (version 5.1.1.)

A *de novo Kpn*I optical map of *B. animalis* subsp. *lactis* ATCC 27673 DNA was constructed by OpGen (Gaithersburg, MD) and was compared to a *Kpn*I *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) and deposited in GenBank with the accession number CP003941.

Comparative Genomic Analysis

Genomic alignments were conducted using the progressiveMauve tool (Darling et al., 2010) with default settings. Alignments for tree construction were conducted using SSAHA: Sequence Search and Alignment by Hashing Algorithm (Ning et al., 2001), and the tree was visualized using MEGA 5.1 (Tamura et al., 2011). 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 (Barrangou et al., 2009) and *B. animalis* subsp. *animalis* ATCC 25527 (Loquasto et al., 2011). Protein sequences identified as unique genes present in a specific strain were searched for

homology using BLASTP in the non-redundant database (Sayers et al., 2012). Clustered regularly interspaced short palindromic repeats (CRISPR) were identified using Dotter (Sonnhammer and Durbin, 1996) and CRISPRFinder (Grissa et al., 2007). Genomic islands were defined as four consecutive differential ORFs (open reading frames) 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 (Vernikos and Parkhill, 2006).

SNPs were identified from alignments conducted using progressive Mauve (Darling et al., 2010) and "export SNPs" feature with default settings. Each genome was aligned pairwise with each other strain and SNPs identified. Synonymous, non-synonymous, and intergenic SNPs were identified using a script generated in our laboratory and publically available annotations from NCBI.

Results

General Features

The complete genome sequence of *B. animalis* subsp. *lactis* ATCC 27673 is 1,963,012 bp long 24,529 bp larger than the genome of the Type strain *B. animalis* subsp. *lactis* DSM 10140 and 24,303 larger than the commercial strain Bl-04 (Table 1). Annotation using the PGAAP at NCBI revealed the ATCC 27673 genome contained 1,616 genes, which was thirteen fewer than the type strain and fifteen fewer than Bl-04. The genome also contained 4 rRNA operons (Table 1). The %G+C content was found to be 61.55%, slightly higher than the other sequenced genomes of *B. animalis* subsp. *lactis*.

The genome of ATCC 27673 exhibited a coding percentage of 83.4%, lower than Bl-04, with coding percentage of 90.5%. *B. animalis* subsp. *lactis* ATCC 27673 was verified to the subspecies level using subspecies-specific primers for *lactis* and *animalis*, as a negative control (Supp. Figure 1) (Ventura et al., 2001, Ventura and Zink, 2002). BLAST analysis showed 100% identity of ATCC 27673 16S DNA with 16S DNA of other *lactis* strains. SNPs in ATCC 27673 identified by the MLST scheme developed by Delétoile *et al.* were confirmed in the genome sequence.

Assembly Verification

Whole genome alignments of fully sequenced genomes of *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* verified genetic homogeneity and synteny across strains (Figure 1). Comparison of *in vivo* and *in silico Kpn*I optical maps of ATCC 27673 revealed only minor differences and confirmed genome assembly. Additionally, *in silico Kpn*I optical maps of *B. animalis* subsp. *lactis* ATCC 27673 and Bl-04 and *B. animalis* subsp. *animalis* ATCC 25527 were also aligned (Supp. Figure 2). 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 SSAHA alignment provided insight into the relationship between strains of *B. animalis* (Figure 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 the *animalis* subspecies than other subspecies *lactis* strains. Alignment of *B. animalis* subsp. *lactis* strains show very high degree of synteny and content among previously sequenced strains. Alignment of previously sequenced strains with ATCC 27673 show a high degree of synteny and content despite novel gene content. Whole genome SNP tree analysis indicated the strains Bi-07 and Bl-04 are more closely related to ATCC 27673 then the other strains of *B. animalis* subsp. *lactis*.

Because of the high degree of similarity between the previously sequenced *B. animalis* subsp. *lactis* strains comparison 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 (Figure 2). The core genome included 83.29% of the genes in the genome of ATCC 27673. The genome of ATCC 27673 contains 96 unique genes when compared to both Bl-04 and ATCC 25527, which is 5.94% of the genes present, and differs significantly from previous numbers of unique genes reported within this subspecies. When comparing strains of the *lactis* subspecies 1,430 shared genes were identified. Recently, it was determined the core genome of *B. animalis* subsp. *lactis* was comprised of 1,518 genes (Milani et al., 2013). As expected, addition of a new strain with 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 when compared to two strains of *B. animalis* subsp.

lactis, thus showing a closer relationship with Bl-04 based on the number of unique genes identified.

When compared to *B. animalis* subsp. *lactis* Bl-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 Supplemental Table S1. Most of the genes found in the genomic islands have yet to be assigned functions. Island 3 (Figure 3A) contains four genes (BLAC_00780, BLAC_00785, BLAC_00790, 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 in the non-redundant database (BDP_125, BDP_126, and BDP_127), as BLAC_00780 and BLAC_00785 share a high degree of similarity to BDP_125 (Ventura et al., 2009b).

Island 5 (Figure 3B) contains 12 putative genes (Blac_05725-Blac_05780), half with unassigned function. In addition, Island 5 contains the putative genes, Blac_05745 (partioning protein ParA (Lee and Grossman, 2006)), Blac_05760 and Blac_05765 (mobilization proteins, MobA (Bhattacharjee and Meyer, 1993) and MobC (Zhang and Meyer, 1997)). Interestingly, best BLAST matches of Blac_05760 and Blac_05765, were hits against *B. bifidum* PRL2010 MobA and MobC (E = 0.0), which are also located on chromosomal DNA (Turroni et al., 2010).

Island 6 (Figure 3C), the largest island, is comprised of 29 ORFs. Eight of these are CRISPR-associated genes not found in other *B. animalis* subsp. *lactis* genomes but with similar predicted function to those found in *B. animalis* subsp. *animalis* ATCC 25527. This locus contains 81 CRISPR spacers, which exhibit no homology with CRISPR spacer/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*) (Kim et al., 2010). BLASTP analysis shows best hit is to an O-antigen polymerase (*wzy*) identified in *B. cereus* ATCC 10987 (E= 1e-14) (Rasko et al., 2004).

Four of the islands, islands 2, 4, 5 and 6, contain either insertion sequence (IS) elements, genes putatively assigned functions from bacteriophage or other putative mobile elements. Islands 2 and 4 contain (Supp. Figure 3B and C) genes predicted to encode for phage integrase proteins (BLAC_00465 and BLAC_03510, respectively). Island 6 (Figure 3C) contains four transposases (BLAC_07020, BLAC_07060, BLAC_07065, 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 Bl-04 83 (66.4%) were found to be in genomic islands.

In an effort to better understand gene differences between ATCC 27673 and other *lactis* strains, Bl-04 was chosen as a representative genome for unique gene content analysis. Six genomic islands were also identified in Bl-04 (Table S2). 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 Bl-04. This is a lower percentage of genes present in genomic islands than what was identified in ATCC 27673. Island 1 is the largest island of 25 ORFs in the genome of Bl-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 contain mobile genetic elements, in Island 3 an integrase (Balac_1179) and a putative phage prohead protease (Balac_1191) were identified and Island 6 contains a putative plasmid transfer protein (Balac_1443)

and a phage integrase (Balac_1448). Island 4 contains 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. For comparison, *B. animalis* subsp. *animalis* ATCC 25527 was reported to contain 73,021 SNPs when compared to DSM 10140 (Loquasto et al., 2011), whereas ATCC 27673 contains 12,053 SNPs when compared to 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 340-407 SNPs when compared to all other previously sequenced *B. animalis* subsp. *lactis* strains. Further analysis of the SNPs identified between Bb-12 and the other sequenced strains revealed the majority SNPs reside in regions of relatively high density of suggesting regions of hypervariablility or regions of imperfect sequence. The high degree of similarity between strains of *B. animalis* subsp. *lactis* is highlighted by analysis of Bi07 and Bl-04, which revealed only 11 SNPs and B420 and Bl-04, which revealed only 12 SNPs. The recently sequenced stain Bl12 showed the most similarity with B420 exhibiting only 19 SNPs between the two strains.

SNPs were also evaluated to determine whether they were synonymous or non-synonymous using the publicly available annotations. The majority of SNPs identified were within open reading frames and were non-synonymous. This is in accordance with

the previous comparison between DSM 10140 and Bl-04 (Barrangou et al., 2009, Briczinski et al., 2009), where a greater number of SNPs were found to be non-synonymous.

CRISPR-*cas* system

A novel CRISPR locus was identified in the genome of *B. animalis* subsp. *lactis* ATCC 27673. This locus contains 81 spacers/82 repeats and 8 CRISPR-associated (*cas*) genes belonging to the Type I-E CRISPR-*cas* system, and the signature Type I *cas3* gene (Barrangou and Horvath, 2012) with repeat sequence 5'-GTGTTCCCCGCAAGCGCGGGGATGATCCC-3'. Some degeneracy exists in the repeat sequence, with the alternative repeat sequence 5'-GTGTTCCCCGCAAGCGCGGGGATGATCCT -3' existing as 26 repeats towards the terminal end of the locus and an additional 8 distinct repeats that exist once as the last eight repeats of the locus (Figure 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 all other *B. animalis* subsp. *lactis* strains that 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 the *lactis* subspecies. To date, three versions of the CRISPR loci have been identified in the subspecies *lactis.* Previously identified CRISPR loci in *B. animalis* subsp. *lactis* were Type I-C and either contained 23 repeats, 20 repeats or 19 repeats (Barrangou et al., 2009, Briczinski et al., 2009, Milani et al., 2013). The CRISPR locus in ATCC 27673 contains 82 repeats and 81 spacers, none that match any of the repeat/spacers present in any other strain of *lactis*. This locus is

interrupted by 280-bp of seemingly random sequence, which is possibly degenerative repeat/spacer sequence. BLAST analysis of the ATCC 27673 *cas*3 DNA sequence resulted in 263 nucleotides of perfect homology in an intergenic region present in the other *lactis* genomes. Also of note, this region of perfect homology is immediately adjacent to a transposase (Balac_1413), at the same relative position in the genome of Bl-04 (and other *lactis*) strains possibly suggesting a deletion event.

Although there are no known phage 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 (Ventura et al., 2009a, Milani et al., 2013). The number of spacers suggests 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 phage 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 non-redundant (nr/nt), whole-genome shotgun contigs (wgs), and genomic survey sequences (gss) databases. One spacer (S55) with the sequence 5'-GCCCACGCGTGAAATCGATGCGTGTGGCCGTG-3', was a perfect hit against a genomic 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 genomic sequence of a closely related organism and not that of bacteriophage or 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 (Bikard et al., 2012, Jorth and Whiteley, 2012).

Discussion

The overall architecture of the genome of ATCC 27673 was similar to the genomes of other sequenced strains of the *lactis* subspecies but also exhibited significant sequence diversity 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" (Milani et al., 2013). The monomorphic nature observed in the previously sequenced genomes may exist for biological or non-biological reasons. One possible non-biological 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 based on similar criteria from environmental samples. This may have artificially reduced the diversity detected within the subspecies because of selection bias. It also a possibility that new strains simply represent the re-isolation and re-naming of existing strains. ATCC 27673 has not been associated with the use as a probiotic food additive which lends creditability to this suggestion. Some authors have argued that isolates of *B. animalis* subsp. *lactis* recovered from human subjects represent unique strains (Sun et al., 2010, Milani et al., 2013). While this is potentially true, the possibility that human isolates simply represent re-isolation of widely-distributed commercial strains consumed by these individuals 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 and only allowed for the survival of 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 (Achtman, 2008).

Monomorphic lineages have been previously defined as "have such low levels of sequence diversity that only few polymorphisms, or even none at all, are found upon sequencing a few genes" (Achtman, 2008). 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* (Achtman, 2012). SNPs within the core genomes of each lineage were identified and ranged from 7-226 SNPs per 100 kb (Holt et al., 2008, Monot et al., 2009, Bart et al., 2010, Comas et al., 2010, Kuroda et al., 2010, Losada et al., 2010, Morelli et al., 2010, Achtman, 2012). In comparison, across nine full genomes of *B. animalis* subsp. *lactis* (not including ATCC 27673) 29 SNPs per 100 kb were identified. The addition of all *B. animalis* subsp. *lactis* strain ATCC 27673 to this analysis results in 685 SNPs per 100 kb, highlighting the highly monomorphic nature of the previously sequenced strains. Even with the addition of the ATCC 27673 the group would still be considered monomorphic.

Unique genes have been identified between ATCC 27673, Bl-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 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. This suggests much of the diversity observed in ATCC 27673 has been acquired via horizontal evolution (Vernikos and Parkhill, 2006).

One region identified as a strong candidate for HGT is Genome Island 6 containing 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 *E. coli* CRISPR1/2 and CRISPR3/4 (Diez-Villasenor et al., 2010, Touchon and Rocha, 2010). Alignment of ATCC 27673 CRISPR repeats revealed 4 SNPs with ATCC 25527 and 8 SNPs with DSM 20098, whereas alignment with either *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 (Milani et al., 2013). This strain, reportedly isolated from a person who did not consume probiotics, was chosen for sequencing based on increased susceptibility to tetracycline (minimum inhibitory concentration (MIC) of 16 μg/ml, half of the observed MIC of Bb-12). Despite selection based on phenotypic characteristics, the addition of this new genomic sequence did not add to the genetic variability of the subspecies (Milani et al., 2013).

Despite the lack of genetic variability, strains of *B. animalis* subsp. *lactis* have been shown to differ in resistance to oxidative stress (Oberg et al., 2011). This suggests small differences in phenotypic characteristics may not be sufficient 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 (Scardovi and Trovatelli, 1974), 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 for 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 HGT as a possible driver of diversity in this subspecies. A novel CRISPR locus was identified, which differs from other *lactis* loci in terms of CRISPR-*cas* systems type, as well as spacer number and content. Analysis of ATCC 27673 provides evidence the true diversity within this subspecies has yet to be explored and potentially unattributed phenotypes in this subspecies exist which may be beneficial for industrial or human health purposes. We propose the most probable

explanation for the lack of diversity within the *lactis* subspecies is the intense focus on commercially relevant strains and the likely re-isolation of these strains and assigning them as new strains.

Acknowledgements

The authors would like to thank the Penn State University Genomics Core Facility for assistance with Sanger sequencing. Also, we would like to thank Stephan Schuster, for assistance with 454-sequencing and assembly and Jihye Park for bioinformatics assistance.

Works Cited

Achtman, M. 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol 62:53-70.

Achtman, M. 2012. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences 367(1590):860-867.

Barrangou, R., E. P. Briczinski, L. L. Traeger, J. R. Loquasto, M. Richards, P. Horvath, A. C. Coute-Monvoisin, G. Leyer, S. Rendulic, J. L. Steele, J. R. Broadbent, T. Oberg, E. G. Dudley, S. Schuster, D. A. Romero, and R. F. Roberts. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. Journal of Bacteriology 191(13):4144-4151.

Barrangou, R. and P. Horvath. 2012. CRISPR: new horizons in phage resistance and strain identification. Annual Review of Food Science and Technology 3:143-162.

Bart, M. J., M. van Gent, H. G. van der Heide, J. Boekhorst, P. Hermans, J. Parkhill, and F. R. Mooi. 2010. Comparative genomics of prevaccination and modern *Bordetella pertussis* strains. BMC Genomics 11:627.

Bhattacharjee, M. K. and R. Meyer. 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 Research 21(19):4536-4568.

Bikard, D., A. Hatoum-Aslan, D. Mucida, and L. A. Marraffini. 2012. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. Cell Host & Microbe 12(2):177-186.

Bottacini, F., F. Dal Bello, F. Turroni, C. Milani, S. Duranti, E. Foroni, A. Viappiani, F. Strati, D. Mora, D. van Sinderen, and M. Ventura. 2011. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BLC1. Journal of Bacteriology 193(22):6387-6388.

Briczinski, E. P., J. R. Loquasto, R. Barrangou, E. G. Dudley, A. M. Roberts, and R. F. Roberts. 2009. Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* by using single-nucleotide polymorphisms, insertions, and deletions. Applied and Environmental Microbiology 75(23):7501-7508.

Chervaux, C., C. Grimaldi, A. Bolotin, B. Quinquis, S. Legrain-Raspaud, J. E. van Hylckama Vlieg, G. Denariaz, and T. Smokvina. 2011. Genome sequence of the probiotic strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494. Journal of Bacteriology 193(19):5560-5561.

Comas, I., J. Chakravartti, P. M. Small, J. Galagan, S. Niemann, K. Kremer, J. D. Ernst, and S. Gagneux. 2010. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nature Genetics 42(6):498-503.

Darling, A. E., B. Mau, and N. T. Perna. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5(6):e11147.

de Man, J. D., M. Rogosa, and M. E. Sharpe. 1960. A medium for the cultivation of lactobacilli. Journal of Applied Bacteriology (23):130-135.

Delétoile, A., V. Passet, J. Aires, I. Chambaud, M. J. Butel, T. Smokvina, and S. Brisse. 2010. Species delineation and clonal diversity in four *Bifidobacterium* species as revealed by multilocus sequencing. Research in Microbiology 161(2):82-90.

Diez-Villasenor, C., C. Almendros, J. Garcia-Martinez, and F. J. Mojica. 2010. Diversity of CRISPR loci in *Escherichia coli*. Microbiology 156(Pt 5):1351-1361.

Garrigues, C., E. Johansen, and M. B. Pedersen. 2010. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. Journal of Bacteriology 192(9):2467-2468.

Grissa, I., G. Vergnaud, and C. Pourcel. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Research 35(Web Server issue):W52-57.

Gueimonde, M., S. Delgado, B. Mayo, P. Ruas-Madiedo, A. Margolles, and C. G. de los Reyes-Gavilán. 2004. Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. Food Research International 37(9):839-850.

Guyonnet, D., A. Woodcock, B. Stefani, C. Trevisan, and C. Hall. 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. Journal of Digestive Diseases 10(1):61-70.

Holt, K. E., J. Parkhill, C. J. Mazzoni, P. Roumagnac, F. X. Weill, I. Goodhead, R. Rance, S. Baker, D. J. Maskell, J. Wain, C. Dolecek, M. Achtman, and G. Dougan. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. Nature Genetics 40(8):987-993.

Jorth, P. and M. Whiteley. 2012. An evolutionary link between natural transformation and CRISPR adaptive immunity. mBio 3(5).

Kailasapathy, K., I. Harmstorf, and M. Phillips. 2008. Survival of *Lactobacillus acidophilus* and *Bifidobacterium animalis* ssp. *lactis* in stirred fruit yogurts. LWT - Food Science and Technology 41(7):1317-1322.

Kim, J. F., H. Jeong, D. S. Yu, S. H. Choi, C. G. Hur, M. S. Park, S. H. Yoon, D. W. Kim, G. E. Ji, H. S. Park, and T. K. Oh. 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. Journal of Bacteriology 191(2):678-679.

Kim, T. H., S. Sebastian, J. T. Pinkham, R. A. Ross, L. T. Blalock, and D. L. Kasper. 2010. Characterization of the O-antigen polymerase (Wzy) of *Francisella tularensis*. The Journal of Biological Chemistry 285(36):27839-27849.

Kuroda, M., M. Serizawa, A. Okutani, T. Sekizuka, S. Banno, and S. Inoue. 2010. Genome-wide single nucleotide polymorphism typing method for identification of Bacillus anthracis species and strains among *B. cereus* group species. Journal of Clinical Microbiology 48(8):2821-2829.

Lee, P. S. and A. D. Grossman. 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(4):853-869.

Loquasto, J. R., R. Barrangou, E. G. Dudley, and R. F. Roberts. 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). Journal of Dairy Science 94(12):5864-5870.

Losada, L., C. M. Ronning, D. DeShazer, D. Woods, N. Fedorova, H. S. Kim, S. A. Shabalina, T. R. Pearson, L. Brinkac, P. Tan, T. Nandi, J. Crabtree, J. Badger, S. Beckstrom-Sternberg, M. Saqib, S. E. Schutzer, P. Keim, and W. C. Nierman. 2010. Continuing evolution of *Burkholderia* mallei through genome reduction and large-scale rearrangements. Genome Biology and Evolution 2:102-116.

Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost, and E. V. Koonin. 2011. Evolution and classification of the CRISPR-Cas systems. Nature Reviews. Microbiology 9(6):467-477.

Makarova, K. S. and E. V. Koonin. 2013. Evolution and Classification of CRISPR-Cas Systems and Cas Protein Families. Pages 61-91 in CRISPR-cas systems. R. Barrangou and J. van der Oost, ed. Springer-Verlag Berlin Heidelberg.

Marteau, P., E. Cuillerier, S. Meanace , M. F. Gerhardt, A. Myara, M. Bouvier, C. Bouley, F. Tondu, G. Bommelaer, and J. C. Grimaud. 2002. *Bifdobacterium animalis* strain DN-173 010 shortens the colonic transit time in healthy women: a double-blind, randomized, controlled study. Alimentary Pharmocology and Therapeutics 16:587-593.

Masco, L., M. Ventura, R. Zink, G. Huys, and J. Swings. 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(Pt 4):1137-1143.

Matsumoto, M., H. Ohishi, and Y. Benno. 2004. H+-ATPase activity in *Bifidobacterium* with special reference to acid tolerance. Int J Food Microbiol 93(1):109-113.

Meile, L., W. Ludwig, U. Reuger, C. Gut, P. Kaufmann, G. Dasen, S. Wenger, and M. Teuber. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. Systematic and Applied Microbiology 20:57-64.

Milani, C., S. Duranti, G. A. Lugli, F. Bottacini, F. Strati, S. Arioli, E. Foroni, F. Turroni, D. van Sinderen, and M. Ventura. 2013. Comparative genomics of *Bifidobacterium animalis* subsp. *lactis* reveals a strict monophyletic bifidobacterial taxon. Applied and Environmental Microbiology.

Monot, M., N. Honore, T. Garnier, N. Zidane, D. Sherafi, A. Paniz-Mondolfi, M. Matsuoka, G. M. Taylor, H. D. Donoghue, A. Bouwman, S. Mays, C. Watson, D. Lockwood, A. Khamesipour, Y. Dowlati, S. Jianping, T. H. Rea, L. Vera-Cabrera, M. M. Stefani, S. Banu, M. Macdonald, B. R. Sapkota, J. S. Spencer, J. Thomas, K. Harshman, P. Singh, P. Busso, A. Gattiker, J. Rougemont, P. J. Brennan, and S. T. Cole. 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. Nature Genetics 41(12):1282-1289.

Morelli, G., Y. Song, C. J. Mazzoni, M. Eppinger, P. Roumagnac, D. M. Wagner, M. Feldkamp, B. Kusecek, A. J. Vogler, Y. Li, Y. Cui, N. R. Thomson, T. Jombart, R. Leblois, P. Lichtner, L. Rahalison, J. M. Petersen, F. Balloux, P. Keim, T. Wirth, J. Ravel, R. Yang, E. Carniel, and M. Achtman. 2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. Nature Genetics 42(12):1140-1143.

Ning, Z., A. J. Cox, and J. C. Mullikin. 2001. SSAHA: A fast search method for large DNA databases. Genome Research 11(1):1725-1729.

Oberg, T. S., J. L. Steele, S. C. Ingham, V. V. Smeianov, E. P. Briczinski, A. Abdalla, and J. R. Broadbent. 2011. Intrinsic and inducible resistance to hydrogen peroxide in *Bifidobacterium* species. Journal of Industrial Microbiology & Biotechnology 38(12):1947-1953.

Rasko, D. A., J. Ravel, O. A. Oékstad, E. Helgason, R. Z. Cer, L. Jiang, K. A. Shores, D. E. Fouts, N. J. Tourasse, S. V. Angiuoli, J. Kolonay, W. C. Nelson, A.-B. Kolstù, C. M. Fraser, and T. D. Read. 2004. The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. Nucleic Acids Research 32(3):977-988.

Rizzardini, G., D. Eskesen, P. C. Calder, A. Capetti, L. Jespersen, and M. Clerici. 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 randomised, double-blind, placebo-controlled study. The British Journal of Nutrition 107(6):876-884.

Sanchez, B., M. C. Champomier-Verges, B. Stuer-Lauridsen, P. Ruas-Madiedo, P. Anglade, F. Baraige, C. G. de los Reyes-Gavilan, E. Johansen, M. Zagorec, and A. Margolles. 2007. Adaptation and response of *Bifidobacterium animalis* subsp. *lactis* to bile: a proteomic and physiological approach. Applied and Environmental Microbiology 73(21):6757-6767.

Sayers, E. W., T. Barrett, D. A. Benson, E. Bolton, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. Dicuccio, S. Federhen, M. Feolo, I. M. Fingerman, L. Y. Geer, W. Helmberg, Y. Kapustin, S. Krasnov, D. Landsman, D. J. Lipman, Z. Lu, T. L. Madden, T. Madej, D. R. Maglott, A. Marchler-Bauer, V. Miller, I. Karsch-Mizrachi, J. Ostell, A. Panchenko, L. Phan, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, M. Shumway, K. Sirotkin, D. Slotta, A. Souvorov, G. Starchenko, T. A. Tatusova, L. Wagner, Y. Wang, W. J. Wilbur, E. Yaschenko, and J. Ye. 2012. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 40(Database issue):D13-25.

Scardovi, V. and L. D. Trovatelli. 1974. *Bifidobacterium animalis* (Mitsuoka) comb. nov. and the "*minimum*" and "*subtile*" groups of new Bifidobacteria found in sewage. International Journal of Systematic Bacteriology 24(1):21-28.

Simpson, P. J., C. Stanton, G. F. Fitzgerald, and R. P. Ross. 2005. Intrinsic tolerance of *Bifidobacterium* species to heat and oxygen and survival following spray drying and storage. J Appl Microbiol 99(3):493-501.

Sonnhammer, E. L. L. and R. Durbin. 1996. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene:GC1-10.

Stahl, B. and R. Barrangou. 2012. Complete genome sequences of probiotic strains of *Bifidobacterium animalis* subsp. *lactis* B420 and Bi-07. Journal of Bacteriology 194(15):4131-4132.

Sun, Z., X. Chen, J. Wang, P. Gao, Z. Zhou, Y. Ren, T. Sun, L. Wang, H. Meng, W. Chen, and H. Zhang. 2010. Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. Journal of Bacteriology 192(15):4080-4081.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28(10):2731-2739.

Touchon, M. and E. P. Rocha. 2010. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. PLoS One 5(6):e11126.

Turroni, F., F. Bottacini, E. Foroni, I. Mulder, J.-H. Kim, A. Zomer, B. Sanchez, A. Bidossi, A. Ferrarini, V. Giubellini, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, G. F. Fitzgerald, D. Mills, A. Margolles, D. Kelly, D. van Sinderen, and M. Ventura. 2010. Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-dervied glycan foraging PNAS 107(45):19514-19519.

Ventura, M., R. Reniero, and R. Zink. 2001. Specific identification and targeted characterization of *Bifidobacterium lactis* from different environmental isolates by a combined multiplex-PCR approach. Applied and Environmental Microbiology 67(6):2760-2765.

Ventura, M., F. Turroni, G. Lima-Mendez, E. Foroni, A. Zomer, S. Duranti, V. Giubellini, F. Bottacini, P. Horvath, R. Barrangou, D. A. Sela, D. A. Mills, and D. van Sinderen. 2009a. Comparative analyses of prophage-like elements present in bifidobacterial genomes. Applied and Environmental Microbiology 75(21):6929-6936.

Ventura, M., F. Turroni, A. Zomer, E. Foroni, V. Giubellini, F. Bottacini, C. Canchaya, M. J. Claesson, F. He, M. Mantzourani, L. Mulas, A. Ferrarini, B. Gao, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, R. S. Gupta, Z. Zhang, D. Beighton, G. F. Fitzgerald, P. W. O'Toole, and D. van Sinderen. 2009b. The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. PLoS Genetics 5(12):e1000785.

Ventura, M. and R. Zink. 2002. Rapid Identification, Differentiation, and Proposed New Taxonomic Classification of *Bifidobacterium lactis*. Applied and Environmental Microbiology 68(12):6429-6434.

Vernazza, C. L., G. R. Gibson, and R. A. Rastall. 2006. Carbohydrate preference, acid tolerance and bile tolerance in five strains of *Bifidobacterium*. J Appl Microbiol 100(4):846-853.

Vernikos, G. S. and J. Parkhill. 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. Bioinformatics 22(18):2196-2203.

Zhang, S. and R. Meyer. 1997. The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transer. Molecular Microbiology 25(3):509-516.

Zhao, F., F. Zhao, T. Li, and D. A. Bryant. 2008. A new pheromone trail-based genetic algorithm for comparative genome assembly. Nucleic Acids Research 36(10):3455-3462.

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

Organism	GenBank Accession #	Length (bp)	Coding %	G+C %	Genes	rRNA operons	tRNAs	Reference
B. animalis subsp. *lactis*								
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	(Barrangou et al., 2009)
Bl-04	CP001515	1,938,709	90.5	60.5	1,631	4	52	(Barrangou et al., 2009)
Bb-12	CP001853	1,942,198	89.9	60.5	1,642	4	52	(Garrigues et al., 2010)
V9	CP001892	1,944,050	86.0	60.5	1,636	4	52	(Sun et al., 2010)
AD011[1]	CP001213	1,933,695	84.6	60.5	1,528	2	52	(Kim et al., 2009)
CNCM I-2494	CP002915	1,943,113	90.9	60.5	1,724	4	52	(Chervaux et al., 2011)
BLC1	CP003039	1,943,990	86.4	60.5	1,622	4	52	(Bottacini et al., 2011)
Bi-07	CP003498	1,938,822	86.5	60.5	1,661	4	52	(Stahl and Barrangou, 2012)
B420	CP003497	1,938,595	86.1	60.5	1,625	4	52	(Stahl and Barrangou, 2012)
Bl12	CP004053	1,938,605	86.1	60.5	1,607	4	52	(Milani et al., 2013)
B. animalis subsp. *animalis*								
ATCC 25527	CP002567	1,932,963	85.4	60.5	1,597	4	52	(Loquasto et al., 2011)

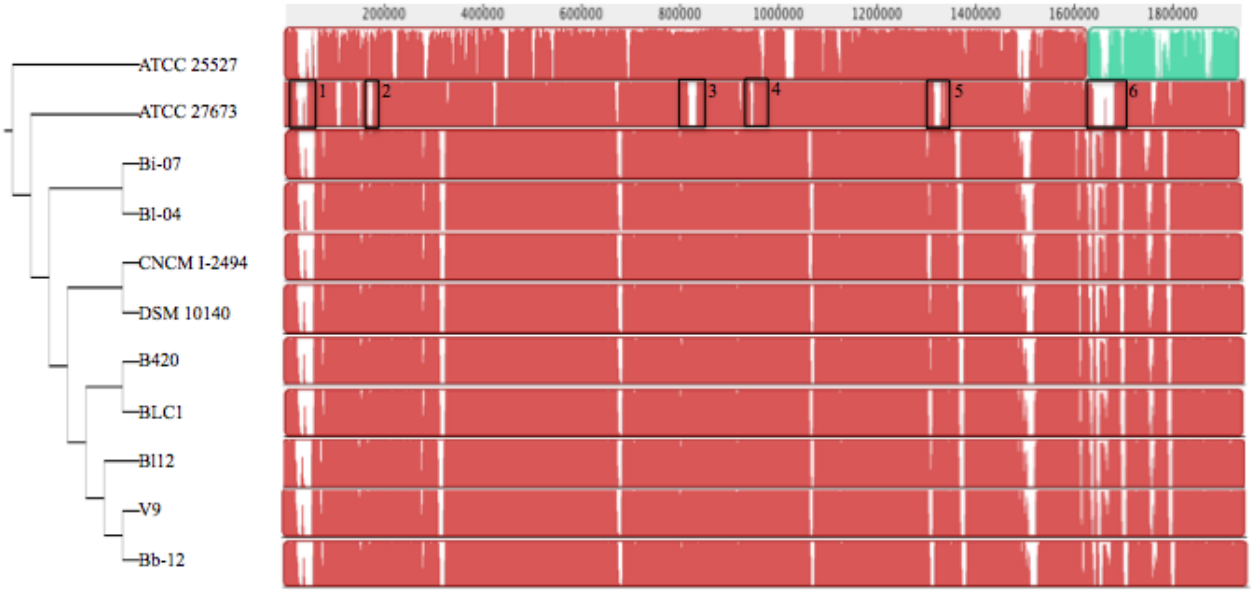
[1]Excluded from bioinformatic analyses because of issues associated with assembly (Kim et al., 2009).

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

	ATCC 27673	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
Bl12										

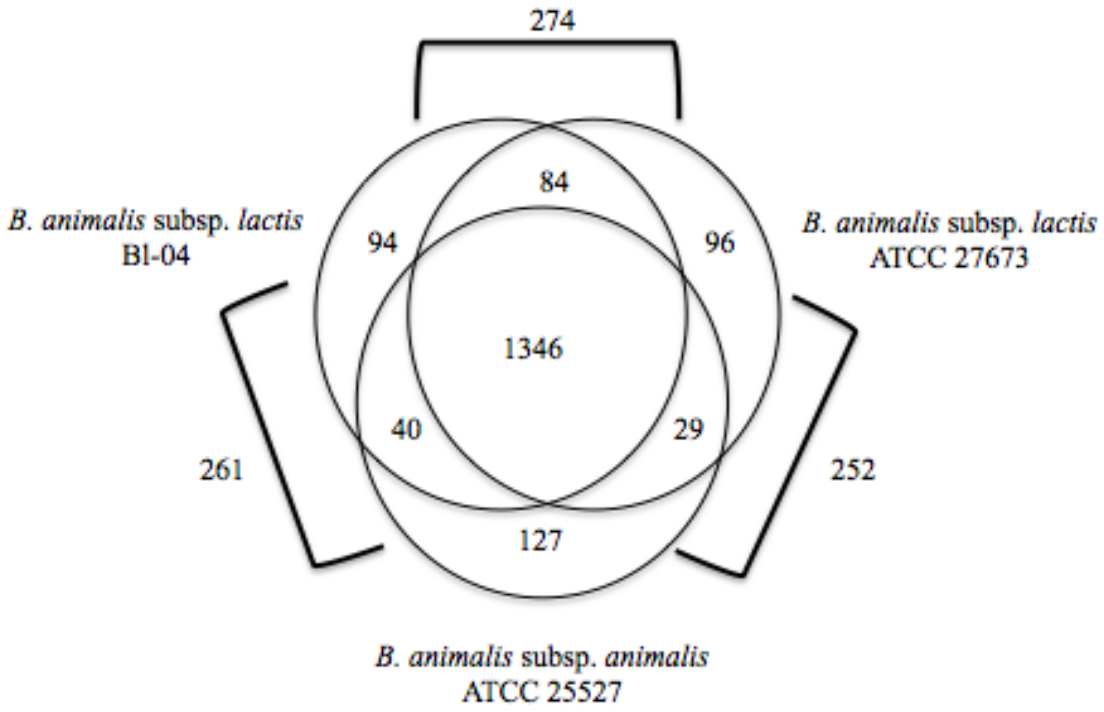
Note- Numbers represent the total number of SNPs identified between pairs genomes identified in Mauve.

Figure 1. Mauve alignments of sequenced stains of *B. animalis*



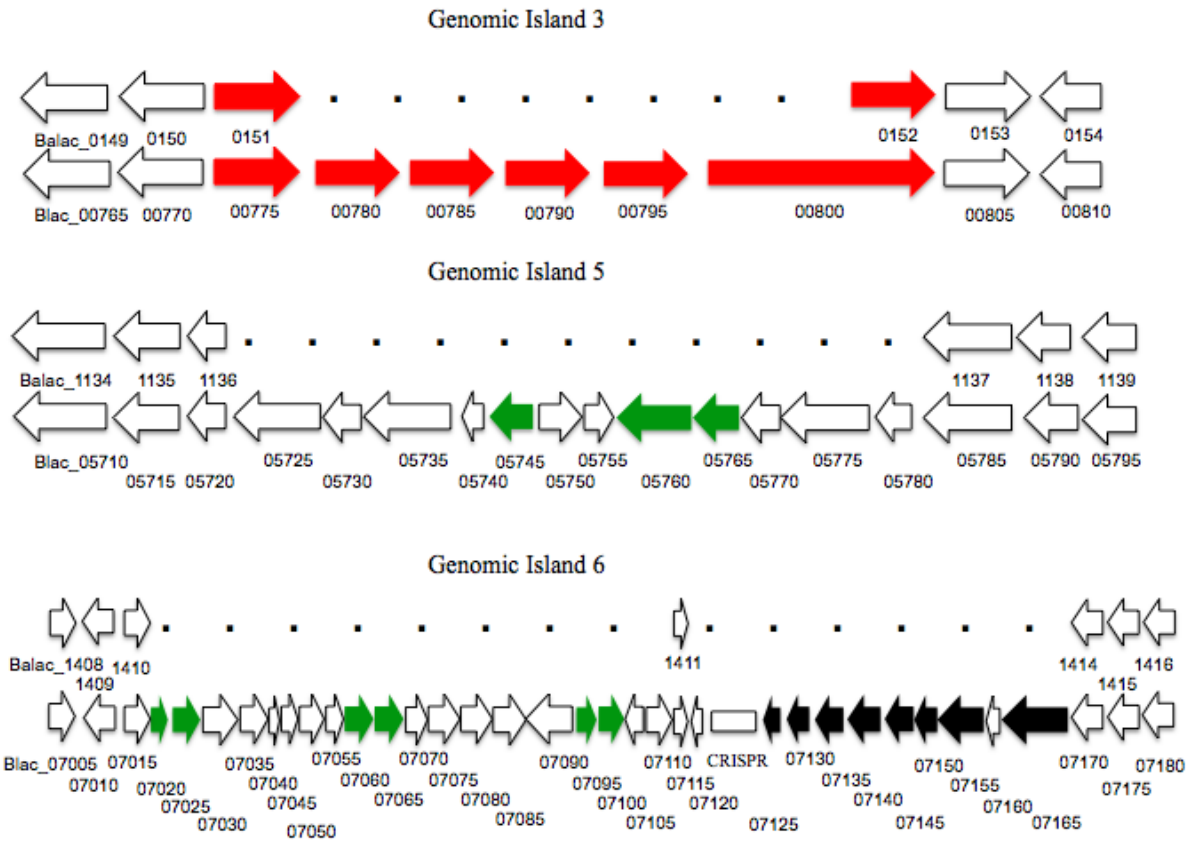
Pairwise alignments were made with the genome of *B. animalis* subsp. *lactis* ATCC 27673 using progressiveMauve. All strains are *B. animalis* subsp. *lactis* except for *B. animalis* subsp. *animalis* ATCC 25527.

Figure 2. Shared and unique genes between three groups of *B. animalis*



ATCC 27673 = *B. animalis* subsp. *lactis* ATCC 27673, Bl-04 = *B. animalis* subsp. *lactis* Bl-04, and ATCC 25527 = *B. animalis* subsp. *animalis* ATCC 25527. Unique genes were identified as having less than 60% amino acid similarity between any pair of strains. Note- Based on the similarity of the sequenced *B. animalis* subsp. lactis genomes, Bl-04 was chosen as a representative of the group.

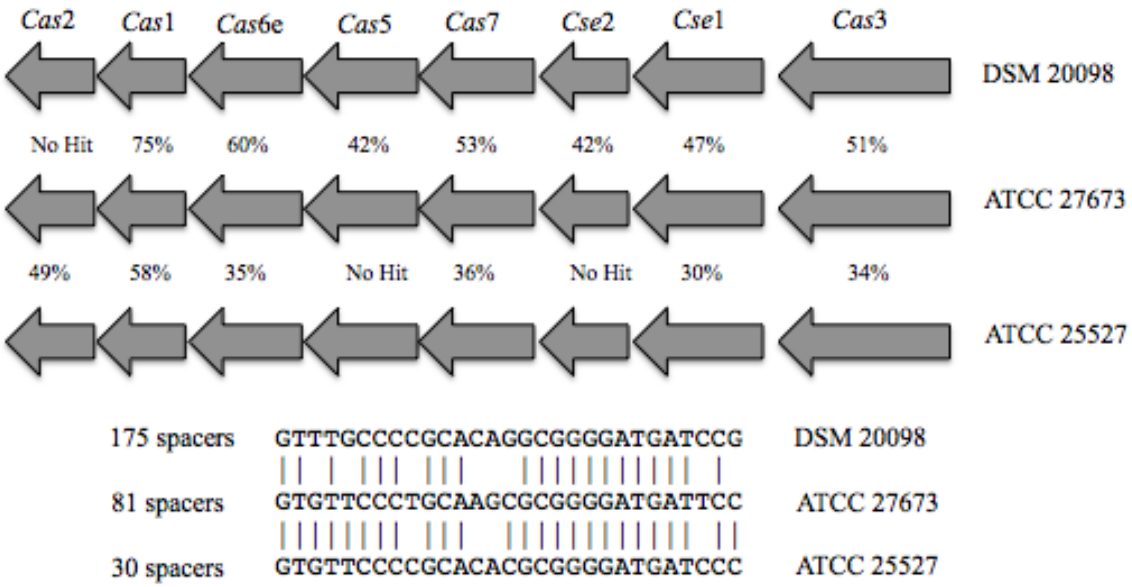
Figure 3. Genomic Islands present in *B. animalis* subsp. *lactis* ATCC 27673



Genomic islands identified in *B. animalis* subsp. *lactis* ATCC 27673 compared to in *B. animalis* subsp. *lactis* Bl-04, A. Genomic Island 3, B. Genomic Island 5, and C. Genomic Island 6. Red arrows correspond to sugar transport genes, green to mobile genetic elements, black to CRISPR-associated genes and white to hypothetical or not-specified genes. Aligned genes

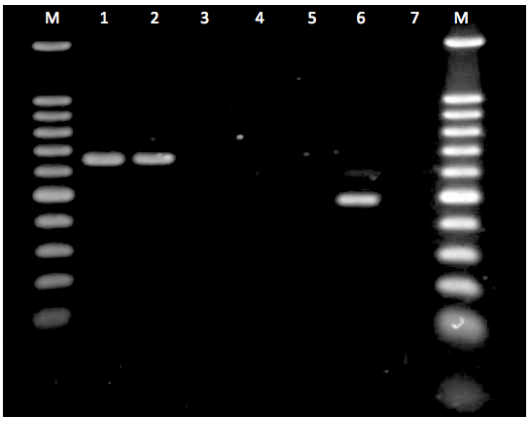
from Bl-04 (top) and ATCC 27673 (bottom) represent gene homology flanking Genomic Island. Data for other genomic islands is presented in Supplemental Figure 3.

Figure 4. CRISPR-*cas* system of *B. animalis* subsp. *lactis* ATCC 27673 and comparison to similar systems



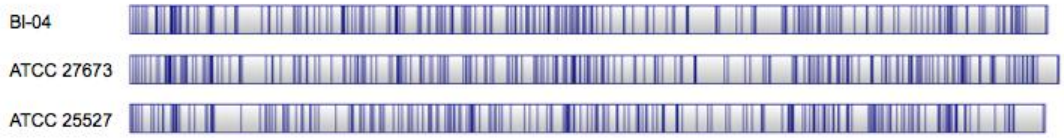
Gene similarity represents percent amino acid similarity between genes of the same assigned function. In addition, repeats from the CRISPR array are aligned with matches to ATCC 27673 highlighted by connecting line. ATCC 27673 = *B. animalis* subsp. *lactis* ATCC 27673, ATCC 25527 = *B. animalis* subsp. *animalis* ATCC 25527, DSM 20098 = *B. angulatum* DSM 20098.

Supplementary Figure 1. PCR identification of *B. animalis* subsp. *lactis* ATCC 25527 using subspecies specific primers Bflact2/5



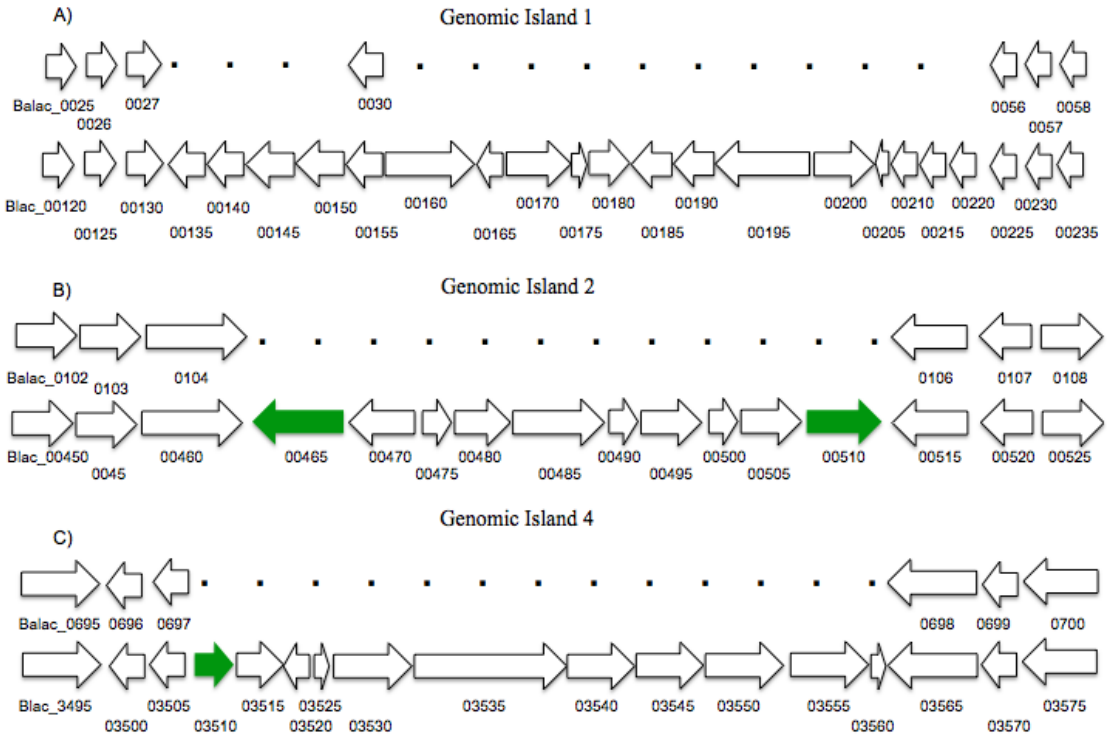
M- 100 bp ladder; Lane 1- DSM 10140 DNA + Bflact2/5; Lane 2- ATCC 27673 DNA + Bflact2/5; Lane 3- ATCC 25527 DNA + Bflact2/5; Lane 4- DSM 10140 DNA + Ban2/23Si; Lane 5- ATCC 27673 DNA + Ban2/23Si; Lane 6- ATCC 25527 DNA + Ban2/23Si; Lane 7- Negative Control

Supplementary Figure 2. *In silico Kpn*I optical maps of three *B. animalis* strains



In silico comparison of Bl-04 (*B. animalis* subsp. *lactis)*, ATCC 27673 (*B. animalis* subsp. *lactis*), and ATCC 25527 (*B. animalis* subsp. *animalis*).

Supplementary Figure 3. Additional Genomic Islands present in *B. animalis* subsp. *lactis* ATCC 27673



Supplemental Figure 3 Legend. Genomic islands identified in *B. animalis* subsp. *lactis* ATCC 27673 compared to in *B. animalis* subsp. *lactis* Bl-04, A. Genomic Island 1, B. Genomic Island 2, and C. Genomic Island 4. Green arrow correspond to mobile genetic elements, and white to hypothetical or not-specified genes. Aligned genes from Bl-04 (top) and ATCC 27673 (bottom) represent gene homology flanking Genomic Island.

Supplementary Table 1. Unique genes in *B. animalis* subsp. *lactis* ATCC 27673

BLAC_Number	Amino Acid Length	Putative Gene Product	BLAC_Number	Amino Acid Length	Putative Gene Product
Island 1			**Island 5**		
BLAC_00135	134	Histone acetyltransferase HPA2 and related acetyltransferases	BLAC_05725	593	Site-specific recombinases, DNA invertase
BLAC_00140	134	hypothetical protein	BLAC_05730	148	hypothetical protein
BLAC_00145	236	hypothetical protein	BLAC_05735	768	hypothetical protein
BLAC_00150	262	hypothetical protein	BLAC_05740	77	hypothetical protein
BLAC_00160	681	Uncharacterized conserved protein	BLAC_05745	186	atpase chromosome partitioning protein ParA
BLAC_00165	101	hypothetical protein	BLAC_05750	118	cAMP-binding domain-catabolite gene activator
BLAC_00170	402	hypothetical protein	BLAC_05755	93	hypothetical protein
BLAC_00175	83	hypothetical protein	BLAC_05760	509	mobilization protein
BLAC_00180	188	hypothetical protein	BLAC_05765	173	mobilization protein
BLAC_00185	212	hypothetical protein	BLAC_05770	102	hypothetical protein
BLAC_00190	191	hypothetical protein	BLAC_05775	622	hypothetical protein
BLAC_00195	702	hypothetical protein	BLAC_05780	74	transcriptional regulator
BLAC_00200	437	hypothetical protein	BLAC_05990	32	hypothetical protein
BLAC_00205	62	hypothetical protein	BLAC_06025	533	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
BLAC_00210	101	hypothetical protein	BLAC_06485	97	hypothetical protein
BLAC_00215	94	hypothetical protein	BLAC_06490	87	hypothetical protein

BLAC_00220	157	hypothetical protein
BLAC_00335	63	hypothetical protein
Island 2		
BLAC_00465	382	phage integrase
BLAC_00470	251	hypothetical protein
BLAC_00475	98	hypothetical protein
BLAC_00480	130	hypothetical protein
BLAC_00485	429	hypothetical protein
BLAC_00490	89	hypothetical protein
BLAC_00495	208	hypothetical protein
BLAC_00500	62	hypothetical protein
BLAC_00505	220	hypothetical protein
BLAC_00510	289	phage major capsid protein, HK97
BLAC_00690	70	hypothetical protein
BLAC_00695	471	hypothetical protein
BLAC_00700	711	hypothetical protein
BLAC_00705	45	hypothetical protein
BLAC_00725	145	hypothetical protein
Island 3		
BLAC_00780	463	sugar ABC transport permease- binding protein
BLAC_00785	461	sugar ABC transport permease- binding protein
BLAC_00790	356	sugar ABC transport
BLAC_06495	279	Protein tyrosine phosphatase (EC 3.1.3.48)
BLAC_06500	268	hypothetical protein
BLAC_06505	90	hypothetical protein
BLAC_06525	200	Transcriptional regulator, MarR family
BLAC_06535	130	Alcohol dehydrogenase (EC 1.1.1.1)
BLAC_06540	63	hypothetical protein
BLAC_06650	155	hypothetical protein
BLAC_06760	76	hypothetical protein
BLAC_06940	311	hypothetical protein
Island 6		
BLAC_07020	107	IS3 family transposase
BLAC_07025	311	transposase subunit B
BLAC_07030	635	hypothetical protein
BLAC_07035	268	hypothetical protein
BLAC_07040	63	hypothetical protein
BLAC_07045	140	hypothetical protein
BLAC_07050	276	hypothetical protein
BLAC_07055	307	hypothetical protein
BLAC_07060	439	transposase
BLAC_07065	428	transposase
BLAC_07070	281	glycosyl transferase family protein
BLAC_07075	462	glycosyl transferase, group

		permease			1
BLAC_00795	313	binding protein dependent transport systems- inner membrane component	BLAC_07080	414	glycosyltransferase
BLAC_2100	30	hypothetical protein	BLAC_07085	433	O-antigen polymerase (wzy)
BLAC_2265	70	hypothetical protein	BLAC_07090	867	endo-1,4-beta-xylanase
BLAC_2400	111	hypothetical protein	BLAC_07095	107	IS3 family transposase
BLAC_2880	76	hypothetical protein	BLAC_07100	311	transposase subunit B
BLAC_2885	48	hypothetical protein	BLAC_07105	150	hypothetical protein
BLAC_2900	141	Predicted acetyltransferase	BLAC_07110	329	RelA/SpoT domain-containing protein
BLAC_2910	491	Type II restriction endonuclease, MutH family	BLAC_07120	74	hypothetical protein
BLAC_2915	435	DNA-cytosine methyltransferase (EC 2.1.1.37)	BLAC_07125	120	CRISPR-associated protein, Cas2
BLAC_2920	257	possible restriction endonuclease	BLAC_07130	110	CRISPR-associated protein Cas1
Island 4			BLAC_07135	235	CRISPR-associated protein, Cas6e
BLAC_03510	237	phage integrase	BLAC_07140	244	CRISPR-associated protein, Cas5
BLAC_03515	300	hypothetical protein	BLAC_07145	338	CRISPR-associated protein, Cas7
BLAC_03520	149	hypothetical protein	BLAC_07150	209	CRISPR-associated protein, Cse2
BLAC_03525	71	hypothetical protein	BLAC_07155	581	CRISPR-associated protein Cse1

BLAC_03530	695	type II restriction endonuclease
BLAC_03535	1128	hypothetical protein
BLAC_03540	438	hypothetical protein
BLAC_03545	404	hypothetical protein
BLAC_03550	648	helicase domain-containing protein
BLAC_03555	568	hypothetical protein
BLAC_03560	86	hypothetical protein
BLAC_03730	189	hypothetical protein
BLAC_03745	34	RNA binding methyltransferase FtsJ like
BLAC_04005	233	hypothetical protein
BLAC_04010	274	xylose isomerase
BLAC_04125	109	Transposase
BLAC_04130	288	hypothetical protein
BLAC_04135	265	Putative helicase
BLAC_04145	69	hypothetical protein
BLAC_05220	79	hypothetical protein
BLAC_05665	419	ABC-type multidrug transport system, permease component
BLAC_07160	61	hypothetical protein
BLAC_07165	1035	CRISPR-associated helicase Cas3
BLAC_07450	519	possible cell surface protein
BLAC_07455	853	hypothetical protein
BLAC_07590	141	hypothetical protein
BLAC_07995	462	Pyridine nucleotide-disulfide oxidoreductase

Supplementary Table 2. Unique genes in *B. animalis* subsp. *lactis* Bl-04

BALAC_Number	Amino Acid Length	Putative Gene Product
Balac_0011	441	multidrug transport protein
Island 1		
Balac_0028	92	hypothetical protein
Balac_0029	90	hypothetical protein
Balac_0030	117	hypothetical protein
Balac_0031	361	hypothetical protein
Balac_0032	70	hypothetical protein
Balac_0033	454	DNA-cytosine methyltransferase (EC 2.1.1.37)
Balac_0034	282	hypothetical protein
Balac_0035	808	hypothetical protein
Balac_0036	194	hypothetical protein
Balac_0037	712	hypothetical protein
Balac_0038	123	hypothetical protein
Balac_0039	38	hypothetical protein
Balac_0040	68	hypothetical protein
Balac_0041	77	hypothetical protein
Balac_0042	50	hypothetical protein
Balac_0043	87	hypothetical protein
Island 3		
Balac_1179	367	Integrase
Balac_1180	206	hypothetical protein
Balac_1181	69	hypothetical protein
Balac_1182	71	hypothetical protein
Balac_1183	79	hypothetical protein
Balac_1184	64	hypothetical protein
Balac_1185	64	hypothetical protein
Balac_1186	401	hypothetical protein
Balac_1187	91	hypothetical protein
Balac_1188	63	hypothetical protein
Balac_1189	79	hypothetical protein
Balac_1190	106	hypothetical protein
Balac_1191	588	Putative phage prohead protease
Balac_1192	74	hypothetical protein
Balac_1204	599	Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
Balac_1258	494	phosphotransferase
Balac_1285	73	hypothetical protein

Balac_0044	128	transcriptional modulator of MazE/toxin, MazF
Balac_0045	297	Transcriptional regulator
Balac_0046	88	Aldo/keto reductase of diketogulonate reductase family(EC:1.1.1.274)
Balac_0047	217	oxidoreductase, aldo/keto reductase family
Balac_0048	750	Beta-glucosidase (EC 3.2.1.21)
Balac_0049	146	hypothetical protein
Balac_0052	509	lysosomal glucosyl ceramidase-like protein
Balac_0053	702	Beta-galactosidase (EC 3.2.1.23)
Balac_0054	385	Glycoside-Pentoside-Hexuronide (GPH):Cation symporter family protein
Balac_0055	214	Transcriptional regulator, TetR family
Balac_0122	250	Transcriptional regulator, MerR family
Balac_0177	733	hypothetical protein
Balac_0261/*fbp*	2697	hypothetical protein
Balac_1292	56	hypothetical protein
Balac_1293	257	protein tyrosine/serine phosphatase
Balac_1294	97	hypothetical protein
Balac_1298	184	possible MarR-type transcriptional regulator
Island 4		
Balac_1301	63	hypothetical protein
Balac_1302	87	hypothetical protein
Balac_1303	349	CRISPR-associated Csb3
Balac_1304	1018	CRISPR-associated Cas3
Balac_1305	546	CRISPR-associated Csb2
Balac_1306	398	CRISPR-associated Csb1
Balac_1307	98	CRISPR-associated Cas2
Balac_1308	497	CRISPR-associated RecB family exonuclease Cas4b / CRISPR-associated protein Cas1

Balac_0262	42	ATP-dependent DNA helicase RecG (EC 3.6.1.-)
Balac_0299	314	putative membrane protein
Balac_0435	66	O-acetylhomoserine sulfhydrylase (EC 2.5.1.49)
Balac_0447	709	putative 67 kDa myosin-crossreactive streptococcal antigen
Balac_0460	337	UDP-glucose 4-epimerase (EC 5.1.3.2)
Balac_0472	1166	hypothetical protein
Balac_0476/*lacZ*	1068	Beta-galactosidase (EC 3.2.1.23)
Balac_0481	58	hypothetical protein
Balac_0502	303	hypothetical protein
Balac_0518	642	Sialic acid-specific 9-O-acetylesterase
Balac_0537	268	Methionine aminopeptidase (EC 3.4.11.18)
Balac_0547	80	hypothetical protein
Balac_0565	653	ATP-dependent DNA helicase RecQ
Balac_1329	299	Dihydrodipicolinate synthase (EC 4.2.1.52)
Balac_1333	1518	ATP-dependent DNA helicase
Balac_1350	31	hypothetical protein
Balac_1371	537	UDP-phosphate sugar phosphotransferase (EC 2.7.8.6)
Balac_1387	311	polysaccharide biosynthesis protein
Island 5		
Balac_1389	445	flipppase wzx
Balac_1390	335	hypothetical protein
Balac_1391	288	Group 2 family glycosyl transferase
Balac_1399	379	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
Balac_1400	155	polysaccharide biosynthesis protein CpsF
Balac_1401	167	Beta-1,4-galactosyltransferase CpsIVG
Balac_1402	318	Glycosyl transferase

Balac_0572	128	hypothetical protein
Balac_0574	1042	DNA/RNA helicase of DEAD/DEAH box family
Balac_0575	119	hypothetical protein
Balac_0576	656	hypothetical protein
Balac_0603	59	hypothetical protein
Balac_0616	45	hypothetical protein
Balac_0713	446	ABC transporter, permease protein
Balac_0744/*glmU*	461	N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23) / Glucosamine-1-phosphate N-acetyltransferase (EC 2.3.1.157)
Balac_0767	505	hypothetical protein
Balac_0779	190	hypothetical protein
Balac_0813	97	hypothetical protein
Balac_0821	208	Holliday junction DNA helicase RuvA
Island 2		
Balac_1403	279	Glycosyl transferase
Balac_1404	378	hypothetical protein
Balac_1405	1194	multidrug ABC transporter ATPase and permease
Balac_1412	112	hypothetical protein
Island 6		
Balac_1440	373	hypothetical protein
Balac_1441	62	hypothetical protein
Balac_1442	132	hypothetical protein
Balac_1443	254	putative plasmid transfer protein
Balac_1444	306	replication initiation protein
Balac_1445	76	hypothetical protein
Balac_1446	373	HNH endonuclease
Balac_1447	364	modification methylase EcoRI (Adenine-specificmethyltransferase EcoRI) (M.EcoRI) (EC:2.1.1.72)

Balac_0924	327	Pullulanase (EC 3.2.1.41)
Balac_0925	57	hypothetical protein
Balac_0927	230	hypothetical protein
Balac_0928	401	hypothetical protein
Balac_0929	390	hypothetical protein
Balac_0930	142	hypothetical protein
Balac_0931	151	hypothetical protein
Balac_0932	322	hypothetical protein
Balac_0956	293	ABC transporter, ATP-binding protein
Balac_1005	550	DNA segregation ATPase
Balac_1132	214	galactoside O-acetyltransferase (EC:2.3.1.79, EC:2.3.1.18)

Balac_1448	407	phage integrase
Balac_1451	44	hypothetical protein
Balac_1463	32	hypothetical protein
Balac_1484/*capB*	561	possible cell surface protein
Balac_1485	672	hypothetical protein
Balac_1510	36	hypothetical protein
Balac_1517	536	1,4-beta-N-acetylmuramidade
Balac_1519	186	N-acetylmuramoyl-L-alanine amidase
Balac_1520	176	Tpx-1 family protein
Balac_1521	234	hypothetical protein
Balac_1522	443	hypothetical protein
Balac_1541	35	hypothetical protein
Balac_1578	197	hypothetical protein
Balac_1599	438	Multiple sugar ABC transporter, substrate-binding protein
Balac_1607	997	cation-transporting ATPase PacL

Chapter 4

Conclusions and Future Directions

Conclusions

Comparison of the genomes of strains DSM 10140 and Bl-04 revealed an extremely high degree of synteny and similarity and a total of only 47 SNPs and 4 INDELs leading to the term monomorphic being applied to the genomes of this subspecies. SNPs and INDELs identified between the genomes of *B. animalis* subsp. *lactis* strains DSM 10140 and Bl-04 were used to design a typing scheme that was able to successfully differentiate a collection of 24 strains of *B. animalis* subsp. *lactis* into 14 groups that were previously unable to be differentiated.

In an effort to develop a better understanding of the *B. animalis* group, the genome of the closely related subspecies *B. animalis* subsp. *animalis* was sequenced. Comparison of the *B. animalis* subsp. *animalis* genome with that of *B. animalis* subsp. *lactis* revealed 156 unique genes including a novel CRISPR locus. Although the literature had previously indicated growth in milk was a defining characteristic between *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* no difference was observed when this was assessed phenotypically and no obvious genomic differences related to this phenotype were detected (Chapter 2).

Subsequent to publication of the DSM 10140 and Bl-04 genomes, the genomes of a number of other commercial strains of *B. animalis* subsp. *animalis* were sequenced by various groups. Analysis of these genomes confirmed the monomorphic nature of the *B. animalis* subsp. *lactis* genome of the strains sequenced to date. In a report by Delétoile et al. (2010) analyzing speciation of the genus *Bifidobacterium* by MLST it was observed

that although most strains of *B. animalis* subsp. *lactis* grouped together as expected, strain ATCC 27673 was located between the *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* clusters. Because of the lack of variability observed in sequenced strains of *B. animalis* subsp. *lactis* the genome of ATCC 27673 was chosen for sequencing. Comparison of the ATCC 27673 genome with the genomes of other sequenced strains of *B. animalis* subsp. *lactis* revealed six genomic islands of unique sequence when compared to other strains of this subspecies. Among the unique gene sequence was a novel set of CRISPR spacers, which show no similarity to spacers identified in other strains of *B. animalis* subsp. *lactis*. In addition, a Type I-E CRISPR-*cas* system was identified in the genome of ATCC 27673, whereas Type I-C systems have been identified in all other *B. animalis* subsp. *lactis* genomes. It is hypothesized that the general lack of diversity observed in *B. animalis* subsp. *lactis* may be due to the re-isolation and sequencing a small subset of widely utilized commercial strains (Chapter 3).

Future Directions

Currently, ten non-ATCC 27673 strains of *B. animalis* subsp. *lactis* have been fully sequenced. This set of strains, if truly "different", represents a potentially useful group for understanding the functionality of the *B. animalis* subsp. *lactis*. No comprehensive studies have been conducted to assess if and what phenotypic differences exist between these strains. Previous work has shown several phenotypes important for organisms used as probiotics are strain specific and health benefits should be considered strain-specific as well (FAO/WHO, 2002). It would be interesting to investigate differences in acid-tolerance, bile-tolerance, resistance to oxidative stress, and survival in

fermented dairy products across all sequenced strains. Comparison of important phenotypes could identify strains that are best suited for use in fermented foods as a probiotic. In addition, because these strains are so closely related they can serve as a good model for identifying genomic regions of interest that control specific phenotypes. Microarray expression analysis could be utilized to identify regions that are up-regulated during stressed conditions. In a similar manner it would be interesting to evaluate experimentally, and through meta-analysis of existing data, differences in these strains with respect to various probiotic functions (immune modulation, gastrointestinal transit time, etc.). Such a functional-genomics approach might allow identification of gene(s) responsible or associated with specific probiotic function (Ventura's so called "probiogenomics" (Ventura et al., 2009)).

The sequence of ATCC 27673 has demonstrated the full diversity of the *lactis* subspecies has not yet been fully explored. It is not known whether ATCC 27673 represents a unique isolate or it is part of a larger lineage of different strains. The strain of *B. animalis* subsp. *lactis* ATCC 27673 should be further characterized as genotypes and phenotypes important for probiotics organisms have yet to be fully assessed. Identification of this diverse strain gives optimism that other strains with unique genomic sequence exist but have yet to be isolated or identified. Places for isolation include but are not limited to, sewage, farm mammal feces (pig, cow, horse, dog etc…), and non-probiotic consuming humans. Understanding the genotypic differences that drive phenotypic differences can be a critical advancement in the field of probiotics. This may also lead to the identification of strains that can be best suited for inclusion in commercial dairy products.

Appendix A

Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04

Rodolphe Barrangou, Elizabeth P. Briczinski, Lindsay L. Traeger, **Joseph R. Loquasto**, Melissa Richards, Philippe Horvath, Anne-Claire Coûté-Monvoisin, Greg Leyer, Snjezana Rendulic, James L. Steele, Jeffery R. Broadbent, Taylor Oberg, Edward G. Dudley, Stephan Schuster, Dennis A. Romero, Robert F. Roberts

Published as: 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. Journal of Bacteriology **191:**4144-4151.

Statement of Contribution: The candidate was responsible for the design of closing primers, conducting PCR to obtain Amplicons for sequencing, and assembly of the DSM 10140 contigs into a closed genome. In addition the candidate was responsible for design of primers to confirm or refute more than 200 putative SNPs. The candidate also assisted with the data analysis and manuscript preparation.

Abstract

Bifidobacteria are important members of the human gut flora, especially in infants. Comparative genomic analysis of two *Bifidobacterium animalis* subsp. *lactis* strains revealed evolution by internal deletion of consecutive spacer-repeat units within a novel CRISPR locus, which represented the largest differential content between the two genomes. Additionally, 47 single nucleotide polymorphisms were identified, consisting primarily of non-synonymous mutations, indicating positive selection and/or recent divergence. A particular non-synonymous mutation in a putative glucose transporter was linked to a negative phenotypic effect on the ability of the variant to catabolize glucose, consistent with a modification in the predicted protein transmembrane topology. Comparative genome sequence analysis of three *Bifidobacterium* species provides a core genome set of 1,117 orthologs complemented by a pan-genome of 2,445 genes. The genome sequences of the intestinal bacterium *B. animalis* subsp. *lactis* provide insights into rapid genome evolution and the genetic basis for adaptation to the human gut environment, notably with regards to catabolism of dietary carbohydrates, resistance to bile and acid, and interaction with the intestinal epithelium. The high degree of genome conservation observed between the two strains in terms of size, organization, and sequence is indicative of a genomically monomorphic subspecies and explains the inability to differentiate the strains by standard techniques such as PFGE.

There are six supplementary figures and five supplementary tables included in this manuscript. The genome sequences for Bl-04 and DSM 10140 were deposited at NCBI under the temporary accession numbers CP001515 and 32893, respectively.

Introduction

Actinobacteria, *Firmicutes*, *Proteobacteria* and *Bacteroidetes* are dominant microbial phyla widely distributed in diverse ecosystems on the planet (Dos Voltos et al., 2008, Federici et al., 2004, Gopal et al., 2001, Haft et al., 2005, Let et al., 2008, Paineau et al., 2008, Tatusov et al., 2000). Metagenomic analyses of the microbial landscape inhabiting various mammalian environments, notably the human gastrointestinal tract (GIT) and skin, have specifically identified *Actinobacteria* as an important and occasionally dominant phylum (Gill et al., 2006, Grice et al., 2008, Ley et al., 2008). Among the members of the large, diverse, and dynamic microbial community residing in the human GIT, *Bifidobacterium* is a dominant genus considered beneficial to humans and includes probiotic strains (live microorganisms which, when administered in adequate amounts, confer a health benefit on the host) (FAO/WHO, 2002). The population of bifidobacteria in the human intestine varies over time. Following vaginal delivery, GIT of healthy newborns is typically colonized by bifidobacteria, especially in breast-fed infants, during the first few days of life (Favier et al., 2002). Interindividual variation, however, is remarkable in the human infant intestinal flora (Palmer et al., 2007), and dominant genera are not always consistent across metagenomic analyses of the human gut flora (Gill et al., 2006, Kurokawa et al., 2007, Palmer et al., 2007, Ley et al., 2008). Over time, the infant intestinal ecosystem becomes more complex as the diet becomes more diverse, with bifidobacteria typically remaining dominant until weaning (Kurokawa et al., 2007).

Bifidobacterium animalis subsp. *lactis* is a gram-positive lactic acid bacterium commonly found in the gut of healthy humans and has been identified in the infant gut

biota, particularly in ileal, fecal and mucosal samples (Wall et al., 2008, Turroni et al., 2009). Some strains of *B. animalis* subsp. *lactis* are able to survive in the GIT, to adhere to human epithelial cells *in vitro*, to modify fecal flora, to modulate the host immune response, or to prevent microbial gastroenteritis and colitis (Gopal et al., 2001, Bartosch et al., 2005, Foligne et al., 2007, Paineau et al., 2008, Wall et al., 2008, Turroni et al., 2009). Additionally, *B. animalis* subsp. *lactis* has been reported to utilize non-digestible oligosaccharides, which may contribute to the organism's ability to compete in the human gut. Carbohydrates resistant to enzymatic degradation and not absorbed in the upper intestinal tract are a primary source of energy for microbes residing in the large intestine. The benefits associated with probiotic strains of *B. animalis* subsp. *lactis* have resulted in their inclusion in the human diet via formulation into a large array of dietary supplements and foods including dairy products, such as yogurt. Deciphering the complete genome sequence of such microbes will provide additional insight into the genetic basis for survival and residence in the human gut, notably with regards to their ability to survive gastric passage and utilize available nutrients. Also, these genomes provide reference sequences for the on-going metagenomic analyses of the human environment, including the gut metagenome.

Bifidobacterium animalis subsp. *lactis* is the most common bifidobacteria utilized as a probiotic in commercial dairy products in North America and Europe (Gueimonde et al., 2004, Masco et al., 2005). However, despite its commercial and probiotic significance, strain-level differentiation of *B. animalis* subsp. *lactis* has been hindered by the high genetic similarity of this organism, as determined by pulsed-field gel electrophoresis and other nucleic acid-based techniques (Ventura and Zink, 2002,

Briczinski and Roberts, 2006, Wall et al., 2008), and the lack of available genomic sequence information. The genome sequence of strain BB-12 (Garrigues et al., 2005) is not currently publicly available and only a draft genome sequence in 28 contigs is available for strain HN019 (GenBank Project #28807). The complete *B. animalis* subsp. *lactis* genome for strain AD011 (Kim et al., 2009) was only recently (2009) published. While this was an important first step, a single genome does not allow identification of unique targets for strain differentiation or comparative analyses within the subspecies.

The objectives of this study were to determine the complete genome sequences of two *B. animalis* subsp. *lactis* strains, the Type strain and a widely used commercial strain, to provide insights into both the functionality of this species and speciation and strain specialization.

Materials and Methods

Bacterial strains

Bifidobacterium animalis subsp. *lactis* Bl-04 (also known as DGCC2908 and RB 4825) was originally isolated from a fecal sample of a healthy adult, is a widely-used commercial strain, and has been deposited at the American Type Culture Collection safe deposit as SD5219. DSM 10140, which is the *B. animalis* subsp. *lactis* Type strain, was originally isolated from a commercial product (Meile et al., 1997) and was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Braunschweig, Germany). Additionally, *B. animalis* subsp. *lactis* strains Bi-07 (ATCC SD5220), HN019 and B420 from the Danisco Global Culture Collection were used. DSM 10140 was

propagated at 37°C anaerobically in LL (Lapierre et al., 1992). All other cultures were propagated at 37°C anaerobically in MRS supplemented with 0.05% cysteine.

Genome sequencing

The *B. animalis* subsp. *lactis* Bl-04 genome draft sequence was obtained by whole genome shotgun sequencing carried out at Macrogen (Rockville, MD). A DNA library was prepared in pC31 and shotgun sequenced using an ABI 3730XL (Roche, Nutley, NJ), targeting 6× coverage. To complement the Bl-04 draft genome, a 454 pyrosequencing run was performed, targeting 15× coverage using a GS-20.

Independently, the *B. animalis* subsp. *lactis* DSM 10140 genome draft was sequenced at The Pennsylvania State University using 454 pyrosequencing targeting 30× coverage, followed by *de novo* assembly in Newbler (Roche).

Concurrently, an optical map was generated for the Bl-04 strain at OpGen (Madison, WI) using *Not*I. Assembled contigs were aligned with the optical map and remaining gaps were closed by walking across PCR products. PCR amplicons were sequenced at the Huck Institute's Nucleic Acid Facility at The Pennsylvania State University using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators; Applied Biosystems, Foster City, CA) and run on an ABI 3730XL DNA Analyzer using ABI Data Collection Program (v 2.0), or at Davis Sequencing (Davis, CA) Facilities. Data were analyzed with ABI Sequencing Analysis software (Version 5.1.1). Sequencing reactions were performed using protocol #4303237 from Applied Biosystems. Complete genome sequences were annotated at Integrated Genomics using

the ERGO package (Integrated Genomics, Chicago, IL) without manual curation and at NCBI.

In silico analyses

The two *B. animalis* subsp. *lactis* genomes were aligned using BioEdit (Isis Biosciences) and SeqMan (DNASTAR, Madison, WI). Single nucleotide polymorphisms (SNP) reports were generated to identify insertions, deletions, and mutations. Comparative genomic analyses for Clusters of Orthologous Groups (COGs, Tatusov et al., 2000) were carried out using the ERGO package by similarity clustering and visualized using MeV (JCVI, MD, Saeed et al., 2003).

SNP re-sequencing

After identification of putative SNPs between the two genome drafts, primers were designed to amplify regions containing SNPs in both genomes. PCR amplicons were purified using the QIAquick kit (Qiagen, Valencia, CA) and sequenced at The Pennsylvania State University or the Davis Sequencing facilities.

CRISPR analyses

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) loci were identified in the Bl-04 and DSM 10140 genome sequences using Dotter (Sonnhammer and Durbin, 1995). An internal segment of the putative CRISPR locus was amplified by PCR, using primers 5'-TTGGATGCAAGCCCTCAATGAAGC-3' and 5'-TGAGGGAAGCCGAACTCAATCACA-3'. The PCR amplicons were subsequently

sequenced (Davis Sequencing) from both ends with each primer. The CRISPR spacers were visualized as two-tone color combinations as previously described (Barrangou et al., 2007, Horvath et al., 2008b).

Data deposit

Both genome sequences were deposited at NCBI under temporary accession numbers CP001515 and CP001606 for Bl-04 and DSM 10140, respectively.

Results

B. animalis subsp. *lactis* genome

The Bl-04 shotgun sequencing generated 1,824,258 bp assembled in 121 contigs. Complementation by pyrosequencing generated 1,948,770 bp assembled in 16 contigs. Concurrently and independently, the DSM 10140 pyrosequencing generated 33× coverage of the 1.9-Mbp genome, assembled in 30 contigs. Following a few rounds of gap closing (by contig-end pairing), the remaining contigs (3 for Bl-04 and 10 for DSM 10140) from both genome drafts were aligned using the Bl-04 optical map as a template. Remaining gaps were closed subsequently and independently in both genomes.

The distinct sequencing strategies used for the two genomes generated comparable assemblies, with the exception of one misassembly in the pyrosequencing-only genome draft. Upon comparison of the two genomes, a total of 211 putative SNPs were identified. Fourteen of the putative SNPs were located in rRNA operons or IS

elements and were not resequenced. After resequencing the remaining 197 SNPs, 150 were found to be the result of sequencing errors, while 47 SNP sites were confirmed.

Overall, the Bl-04 and DSM 10140 genomes are 1,938,709 bp and 1,938,483 bp, respectively. Using the largest genome as a reference (Bl-04), 1,655 open reading frames (ORFs) were identified and putative functions were assigned to 1,072 of them, which is typical across sequenced *Bifidobacterium* genomes (Table 1). The features of the *B. animalis* subsp. *lactis* genomes are detailed in Table 1 and Supplementary Figure A.

The origin and terminus of replication in the *B. animalis* subsp. *lactis* genome were predicted (Supplementary Figure A and Supplementary Figure B) based on *dnaA* (Balac_0001), DnaA boxes, and base composition asymmetry between the leading and lagging strands (Frank and Lobry, 2000, Mackiewicz et al., 2004, Hansen et al., 2007). The putative *oriC* origin of replication was identified in an AT-rich intergenic region upstream of *dnaA* in the vicinity of a cluster of hypothetical DnaA boxes (Supplementary Figure B). Similar putative *oriC* with comparable DnaA boxes were also identified in other bifidobacteria genomes (Supplementary Figure B), were consistent with previously observed degeneracy at certain sites and reflected the high G+C content of the genome (Mackiewicz et al., 2004, Hansen et al., 2007). These conserved elements might be used to locate the origin of replication in *Bifidobacterium* species and other *Actinobacteria* genomes, since GC-skew analysis alone does not necessarily identify an unequivocal *oriC* (Schell et al., 2002).

Comparative analysis of two *B. animalis* subsp. *lactis* genomes

Comparison of the two *B. animalis* subsp. *lactis* genomes revealed nearly perfect alignment. Of the 47 SNPs validated in the two genome sequences, 39 were in predicted coding sequences and 8 were in intergenic regions (Figure 1 and Supplementary Table 2). Of the 39 coding SNPs, 31 represent non-synonymous mutations while 8 are synonymous (Figure 1), indicating positive selection and/or recent divergence. Four distinct insertion/deletion sites (INDELs) were also identified, totaling 443 bp (Supplementary Table 3). INDEL1 is a 121-bp sequence encoding tRNA-Ala-GGC, present in the Bl-04 genome (881,420-881,540 bp), which is absent in the DSM 10140 genome. INDEL2 is a 54-bp sequence within the long chain fatty acid-coA ligase gene (Balac_0771, EC 6.2.1.3, COG1022), present in the DSM 10140 genome (902,893-902,946 bp), and absent in the Bl-04 genome. INDEL2 yields an in-frame deletion of 18 amino acids in the Bl-04 protein. Interestingly, the importance of long chain fatty acyl-coA synthetases in bifidobacteria has been noted previously (Schell et al., 2002). INDEL3 is a 214-bp sequence within the CRISPR locus (Clustered Regularly Interspaced Short Palindromic Repeats), present in Bl-04 (1,512,373-1,512,586 bp) and absent in the DSM 10140 genome, which corresponds to three repeat-spacer units. INDEL4 is a 54-bp sequence present in the DSM 10140 genome (1,715,507-1,715,560 bp), and absent in the Bl-04 genome.

Overall, notwithstanding the 47 SNPs and the 443-bp INDELs, the two genomes were 99.975% identical. In addition to the two complete genome sequences, optical mapping was used to analyze the genome layouts of four *B. animalis* subsp. *lactis* strains. The highly similar optical maps of this strain set (comprised of the Type strain,

commercial strains, and isolates used in functional and clinical studies) indicate a high degree of genome conservation in terms of size, organization, and sequence (Figure 1). The lack of polymorphism observed is indicative of a genomically monomorphic subspecies.

Comparative genomic analysis of bifidobacteria

Analysis of the distribution of annotated ORFs over COG categories revealed a high overall conservation of COG representation across *B. animalis* subsp. *lactis, B. longum,* and *B. adolescentis* (Table 1) which belong to three different *Bifidobacterium* clusters (Felis and Dellaglio, 2007, Ventura et al., 2007, Turroni et al., 2009). Specifically, the top four functional categories were identical, namely general prediction, translation, carbohydrate metabolism, and amino acid metabolism, which is typical of lactic acid bacteria (Makarova et al., 2006a, Makarova and Koonin, 2007). However, fewer ORFs were associated with carbohydrate utilization in *B. animalis* subsp. *lactis* than in *B. longum,* and *B. adolescentis*. Overall, most genes (1117, 67%) were conserved in all three species, representing the "core" bifidobacterial genome (Supplementary Figure C), whereas 416 (25%), 368, and 298 genes were unique to *B. animalis* subsp. *lactis, B. longum* and *B. adolescentis,* respectively. Interestingly, a relatively small proportion of genes were shared between two genomes only, while a larger proportion of genes were unique to each of the three genomes examined, suggesting distinct speciations (Supplementary Figure C). Alignment of the three chromosomes indicates *B. animalis* subsp. *lactis* shares a relatively high level of synteny with *B. adolescentis* but exhibits little colinearity with *B. longum* (Supplementary Figure D). This further confirms the

observed differences between the genomes of *B. longum* and *B. adolescentis* (Sela et al., 2008).

Comparative analysis of gene conservation across the three *Bifidobacterium* species through BLASTp (Supplementary Table 1) revealed eight areas containing consecutive sets of ORFs present in *B. animalis* subsp. *lactis* and absent in both *B. adolescentis* and *B. longum* (Figure 2). Notably, a prophage remnant is located in area 5, including a phage-related integrase (Balac_1179, COG0582) and a phage-related prohead protein (Balac_1191, COG3740). Although there is only anecdotal evidence of bacteriophage in bifidobacteria (Sgorbati et al., 1983), a previous analysis of bifidobacterial genome sequences revealed the presence of prophages integrated in a $tRNA^{Met}$ gene (Ventura et al., 2005, Ventura et al., 2007). The observed remnants of a prophage-like element in the *B. animalis* subsp. *lactis* genome do not seem to be adjacent to a $tRNA^{Met}$ sequence. An *eps* cluster with a low GC content was identified in area 7, (Balac_1383-Balac_1391), which may be involved in the synthesis of membrane-associated exopolysaccharides. Two genes potentially involved in oxalic acid catabolism (Federici et al., 2004) were identified in area 8, *oxc* (Balac_1453) and *frc* (Balac_1449), which encode an oxalyl-coenzymeA decarboxylase (EC 4.1.1.8) and a formyl-coenzymeA transferase (EC 2.8.3.16), respectively.

Analysis of the *B. animalis* subsp. *lactis* CRISPR locus

A novel CRISPR-Cas system, Bala1, was identified in the *B. animalis* subsp. *lactis* genome in area 6 (Figure 2), which is different from other CRISPR loci previously identified in bifidobacteria. It is the ninth CRISPR family identified in lactic acid bacteria

genomes and the fourth CRISPR family identified in bifidobacteria, in addition to Blon1 (in *B. longum*), Lhel1 (in *B. adolescentis*), and Ldbu1 (present in *Bifidobacterium catenulatum*) (Horvath et al., 2008a). Novelty was observed both in terms of CRISPR repeat sequence and *cas* (CRISPR-associated) gene content. In Bl-04, the typical 36-bp CRISPR repeat, 5'-ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG-3', which is partially palindromic, is present 23 times and separated by 22 unique spacer sequences (Figure 3, Supplementary Figure F). The typical repeat sequence is conserved in the first 21 repeats, while the last 2 have SNPs (Supplementary Figure F), notably at the 3' end, as previously shown in other CRISPR systems (Horvath et al., 2008b). Two remnant CRISPR repeats were also identified between *cas2* and *csb1* (Figure 3). The co-occurrence of remnant CRISPR repeats and a typical CRISPR repeat-spacer array has been observed previously, notably in *Streptococcus thermophilus* CRISPR3 (Horvath et al., 2008b) and *B. adolescentis* (Horvath et al., 2008a). Six *cas* genes were identified downstream of the repeat-spacer region, including the universal nuclease *cas1* (Balac_1308, COG1518, TIGR00287), the endonuclease *cas2* (Balac_1307, COG1343, TIGR01573) and the helicase *cas3* (Balac_1304, COG1203, TIGR02621). Additionally, three novel putative *cas* genes, *csb1-3* (Cas subtype Bifidobacterium Bala1, with a nomenclature derived from Haft et al., 2005) were also identified in this CRISPR locus. Two of them, *csb1* (Balac_1306) and *csb2* (Balac_1305), contain *cas*-type conserved elements, cas_GSU0053 and cas_GSU0054, respectively. The spacer size varied between 34 and 39 bp. Some Bl-04 CRISPR spacers showed similarity to phage sequences (S3 homology with *Streptomyces* phage phi-BT1 AJ550940; S17 homology with frog virus AY548484; S20 homology with a phage capsid protein in *Chromohalobacter salexigens*

CP000285) and metagenomic-derived sequences (S19 homology with human gut metagenome BABA01032251; S20 homology with marine metagenome AACY021620797). However, the analysis of CRISPR spacers is limited by the absence of bifidobacteria phage sequences in public databases.

While the Bala1 CRISPR locus was identified in both Bl-04 and DSM 10140 (Figure 3), polymorphism was observed in terms of spacer content. Although the spacer content was identical both at the leader and the trailer end of the locus, three consecutive internal repeat-spacer units were unique to Bl-04 (Figure 3). The Bala1 CRISPR locus was sequenced in several *B. animalis* subsp. *lactis* strains and spacer content analysis revealed only two versions of the CRISPR locus, as seen in the two genomes, with presence/absence of three consecutive internal repeat-spacer units (Figure 3). The Bala1 CRISPR locus was subjected to amplification by PCR in several bifidobacterial species, including the *B. animalis* subsp. *animalis* Type strain. This locus was present exclusively in *B. animalis* subsp. *lactis* strains and absent in *B. animalis* subsp. *animalis*, *B. longum, B. adolescentis, B. dentium, B. catenulatum, Bifidobacterium breve,* and *Bifidobacterium bifidum,* suggesting it is sub-species specific. Interestingly, the GC content of the Bala1 CRISPR locus was approximately 49.74%, while that of the genome is 60.19%, suggesting that it may have been acquired laterally from a low-GC microbe as previously discussed for CRISPR loci (Godde and Bickerton, 2006, Horvath et al., 2008a).

Discussion

The relatively small size of the *B. animalis* subsp. *lactis* genome, as compared to other bifidobacteria, is consistent with a genome simplification process that reduces

biosynthetic capabilities and favors the retention and acquisition of genes involved in the utilization of a broad repertoire of carbon and nitrogen sources (Makarova et al., 2006a, Makarova and Koonin, 2007). This is typical of genomic evolution of microbes that live in nutritionally-rich environments such as the human gut that rely on the host and other members of the intestinal flora for energy sources and metabolic intermediates (Makarova et al., 2006a, Makarova and Koonin, 2007). This likely explains the lower number of hydrolases in *B. animalis* subsp. *lactis,* as compared to *B. longum* and *B. adolescentis* (Supplementary Table 4). Notably, the genes involved in the catabolism of human milk oligosaccharides in *B. longum* subsp. *infantis* (Sela et al., 2008) and degradation and utilization of mucin (endo-α-*N*-acetylgalactosaminidase) were not identified. Also, unlike most bacteria, ABC transporters do not appear in the typical operon configuration and only two copies of a carbohydrate-specific ATP-binding protein, Balac_0062 (COG1129) and Balac_1610 (COG3839) were identified in *Bifidobacterium animalis* subsp. *lactis.* No phosphotransferase system (PTS) component was identified, contradicting a previous report of fructose-PTS activity in the DSM 10140 strain (Parche et al., 2007), and in contrast to the presence of PTS in *B. longum* and *B. adolescentis.*

While the number of carbohydrate hydrolases is fewer in *B. animalis* subsp. *lactis,* the variety of hydrolases present suggests the ability to utilize a wide range of complex carbohydrates including milk galactosides and undigestible plant-derived oligosaccharides (Supplementary Table 4). This is consistent with a previous report that genes in the carbohydrate transport and metabolism COG are over-represented in the

human gut microbiome (Kurokawa et al., 2007). This may also reflect the strong influence of the host diet on bacterial gut communities in mammals (Ley et al., 2008).

Additionally, genes involved in adaptation to and survival of this organism in the human GIT were identified. Specifically, two paralogs encoding putative N-acetylmuramidases (COG3757, Balac_1516 and Balac_1517) were identified, which may be involved in the degradation of bacterial cell wall components available in the intestinal environment. Four genes encoding putative cell surface proteins that could be involved in interactions with human epithelial cells were identified in the *B. animalis* subsp. *lactis* genomes including two putative collagen-adhesion proteins (*capA*, Balac_1456, COG4932 and *capB*, Balac_1484), an elastin-binding protein (*ebpS*) and a fibronectin-binding protein (*fbp*, Balac_0271). A typical LPxTG anchor motif was only identified in CapB (LPLTG). An alternative putative anchor motif, VAATG, was identified *in silico* in the vicinity of a poly-R sequence at the C-terminus of CapA and Fbp. The presence of two distinct sortase-encoding genes in the genome of *B. animalis* subsp. *lactis*, *srtB* (Balac_1349) and *srtA* (Balac_1485) is consistent with the presence of two distinct anchor motifs.

Numerous genes commonly associated with stress response were identified in the *B. animalis* subsp. *lactis* genome (Supplementary Table 5). Notably, a bile salt hydrolase was identified, *bsh* (Balac_0863, EC 3.5.1.24), encoding a member of the Ntn-PVA family of enzymes (COG3049). This gene family is enriched in the human gut microbiome (Jones et al., 2008) and is likely involved in the ability of *B. animalis* subsp. *lactis* to tolerate bile and survive in the human gut environment.

Genomic content was compared across members of three distinct *Bifidobacterium* phylogenetic clusters: *B. adolescentis, B. longum,* and *B. animalis* subsp. *lactis* (a member of the *Bifidobacterium pseudolongum* group) (Felis and Dellaglio, 2007, Ventura et al., 2007). Comparative analysis revealed a high degree of conservation and synteny overall (Figure 2 and Supplementary Figure D), as previously observed in *Bifidobacterium* genomes (Sela et al., 2008). However, there are significant, functionally relevant differences, notably CRISPR content, catabolism of dietary compounds, cell surface proteins and polysaccharides, and prophages. Of note, cell surface proteins identified in the *B. animalis* subsp. *lactis* genome may be involved in its ability to adhere to human intestinal epithelial cells (Gopal et al., 2001) and possess immunomodulation properties.

Overall, it appears the genomes of the *B. animalis* subsp. *lactis* strains included in this study are highly monomorphic, consistent with previous reports indicating high similarity among *B. animalis* subsp. *lactis* strains isolated from infant samples and various commercial products, as determined by pulsed-field gel electrophoresis (Ventura and Zink, 2002, Briczinski and Roberts, 2006, Wall et al., 2008) which may suggest clonal ancestry.

Comparison of the Bl-04 genome to the recently published genome sequence of *B. animalis* subsp. *lactis* AD011 (Kim et al., 2009) revealed the AD011 genome to be 5,014 bp shorter than Bl-04 and indicated a marked difference in assembly between the two genomes. An advantage of the approach taken in the current work – independently sequencing two related strains – is the increased confidence and ability to identify strain-level differences, which has been a challenge for this particular sub-species. In this

particular case, the number of SNPs initially identified between Bl-04 and DSM 10140 was decreased approximately 5-fold by re-sequencing. Thus, by sequencing two genomes independently, using both traditional and pyrosequencing methods, and re-sequencing SNPs regardless quality score, a low error rate (approximately 0.00035%) was obtained leading to increased confidence in the genome sequences. Comparison of the Bl-04 genome to the HN109 draft genome (currently in 28 contigs) did not reveal any novel content.

Despite the genetic homogeneity among *B. animalis* subsp. *lactis* strains, inter-strain functional differences have been observed, such as variability in immunogenic properties (Foligne et al., 2007) and the ability to catabolize glucose (Briczinski et al., 2008). Interestingly, a SNP was identified in *glcU* (Balac_1097, COG4975), a putative glucose uptake protein. Bl-04 has a reduced ability to grow using glucose as the primary carbohydrate source, which is linked to a significant reduction in its ability to transport glucose into the cell (as RB 4825, Briczinski et al., 2008). *In silico* analysis of the putative transmembrane domains of the two *glcU* variants indicates the non-synonymous CGT (Arg in Bl-04) to GGT (Gly in DSM 10140) mutation impacts the predicted structural arrangement of the resulting proteins in the cell membrane (Supplementary Figure E). Perhaps this SNP resulting in the loss of the ability to transport glucose is indicative of rapid evolution driven either by the lack of selective pressure to maintain a functional glucose transporter in the human gut environment, where glucose is likely absent due to its absorption in the upper GIT, or selective pressure against maintaining an arguably useless gene.

While it is difficult to study the evolutionary changes of such monomorphic bacterial genomes, SNPs and INDELs provide insights into the evolution and diversity of clonal bacteria (Dos Voltos et al., 2008). Based on the size and the number of SNPs identified in these two *B. animalis* subsp. *lactis* genomes (Lenski et al., 2003, Zhang et al., 2006, Perfeito et al., 2007), these strains are separated by 1,268 to 304,414 generations, or 6-1,522 years (based on 200 generations per year). The primary observed genomic difference between the two strains consists of three CRISPR repeat-spacer units present in the Bl-04 genome and absent in the DSM 10140 genome. CRISPR represents a family of repeated DNA elements that provides acquired immunity against foreign genetic elements (Makarova et al., 2006b, Barrangou et al., 2007, Brouns et al., 2008, Sorek et al., 2008). Since it has been previously shown that CRISPR loci evolve primarily by polarized addition of novel spacers at the leader end of the locus and by internal deletion of contiguous vestigial spacers (Barrangou et al., 2007, Deveau et al., 2008, Horvath et al., 2008b), the observed difference between the two *B. animalis* subsp. *lactis* genomes is likely the result of an internal deletion within the Bala1 CRISPR locus. This suggests Bl-04 is the ancestral strain from which a variety of *B. animalis* subsp. *lactis* clonal offspring may have recently derived.

The determination of two *B. animalis* subsp. *lactis* genome sequences provides insights into the genetic basis for survival in the human GIT. The ability of microbes to survive in the human environment has been associated with resistance to acid and bile, as well as their opportunistic capacity to utilize undigested dietary compounds. This can be achieved by encoding the enzymatic machinery necessary to catabolize either a wide range of carbohydrates or niche-specific nutrients, as documented for the cariogenic

Streptococcus mutans in the oral cavity (Ajdić et al., 2002), the adaptation of *Lactobacillus plantarum* to a variety of environmental niches (Kleerebezem et al., 2003), the presence of *Lactobacillus johnsonii* and *L. acidophilus* in the GIT (Pridmore et al., 2004, Altermann et al., 2005) and the residence of *B. longum* in the colon (Schell et al., 2002). The comparison of two nearly identical genome sequences revealed rapid evolution via internal deletion of three consecutive repeat-spacer units within a CRISPR locus, three small INDELs and 47 SNPs, including a non-synonymous mutation resulting in the loss of the ability to utilize glucose efficiently, perhaps resulting from directed evolutionary pressure. In addition, the availability and analysis of these genome sequences may allow for development of molecular methods for strain differentiation and will aid in metagenomic analyses of the human microbiome.

Acknowledgements

The authors would like to acknowledge Theresa Walunas for assistance with annotation at Integrated Genomics, and Emily Zentz and Buffy Stahl for optical mapping at OpGen. The work at The Pennsylvania State University was sponsored in part by grants from NutriCorp North East and The Penn State Ice Cream Short Course. Work at Danisco was sponsored by Danisco USA Inc.

Works Cited

Ajdić, D., W. M. McShan, R. E. McLaughlin, G. Savić, J. Chang, M. B. Carson, C. Primeaux, R. Tian, S. Kenton, H. Jia, S. Lin, Y. Qian, S. Li, H. Zhu, F. Najar, H. Lai, J. White, B. A. Roe, and J. J. Ferretti. 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc. Natl. Acad. Sci. U. S. A. 99(22):14434-14439.

Altermann, E., W. M. Russell, M. A. Azcarate-Peril, R. Barrangou, B. L. Buck, O. McAuliffe, N. Souther, A. D. W. Dobson, T. Duong, M. Callanan, S. Lick, A. Hamrick, R. Cano, and T. R. Klaenhammer. 2005. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. Proc. Natl. Acad. Sci. U. S. A. 102(11):3906-3912.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315(5819):1709-1712.

Bartosch, S., E. J. Woodmansey, J. C. M. Paterson, M. E. T. McMurdo, and G. T. Macfarlane. 2005. Microbiological effects of consuming a synbiotic containing *Bifidobacterium bifidum, Bifidobacterium lactis,* and oligofructose in elderly persons, determined by real-time polymerase chain reaction and counting of viable bacteria. Clin. Infect. Dis. 40(1):28-37.

Briczinski, E. P., A. T. Phillips, and R. F. Roberts. 2008. Transport of glucose by *Bifidobacterium animalis* subsp. *lactis* occurs via facilitated diffusion. Applied Environmental Microbiology 74(22):6941-6948.

Briczinski, E. P. and R. F. Roberts. 2006. Technical note: A rapid pulsed-field gel electrophoresis method for analysis of bifidobacteria. Journal of Dairy Science 89(7):2424-2427.

Brouns, S. J. J., M. M. Jore, M. Lundgren, E. R. Westra, R. J. H. Slijkhuis, A. P. L. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin, and J. van der Oost. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321(5891):960-964.

Crooks, G. E., G. Hon, J.-M. Chandonia, and S. E. Brenner. 2004. WebLogo: A sequence logo generator. Genome Research 14(6):1188-1190.

Deveau, H., R. Barrangou, J. E. Garneau, J. Labonté, C. Fremaux, P. Boyaval, D. A. Romero, P. Horvath, and S. Moineau. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. Journal of Bacteriology 190(4):1390-1400.

Dos Voltos, T., O. Mestre, J. Rauzier, M. Golec, N. Rastogi, V. Rasolofo, T. Tonjum, C. Sola, I. Matic, and B. Gicquel. 2008. Evolution and diversity of clonal bacteria: The paradigm of *Mycobacterium tuberculosis*. PLoS ONE 3(2):e1538.

FAO/WHO. 2002. Joint FAO/WHO Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food. London, Ontario.

Favier, C. F., E. E. Vaughan, W. M. De Vos, and A. D. L. Akkermans. 2002. Molecular Monitoring of Succession of Bacterial Communities in Human Neonates. Applied and Environmental Microbiology 68(1):219-226.

Federici, F., B. Vitali, R. Gotti, M. R. Pasca, S. Gobbi, A. B. Peck, and P. Brigidi. 2004. Characterization and heterologous expression of the oxalyl coenzyme A decarboxylase gene from *Bifidobacterium lactis*. Appl. Environ. Microbiol. 70(9):5066-5073.

Felis, G. E. and F. Dellaglio. 2007. Taxonomy of lactobacilli and bifidobacteria. Curr. Issues Intest. Microbiol. 8(2):44-61.

Foligne, B., S. Nutten, C. Grangette, V. Dennin, D. Goudercourt, S. Poiret, J. Dewulf, D. Brassart, A. Mercenier, and B. Pot. 2007. Correlation between *in vitro* and *in vivo* immunomodulatory properties of lactic acid bacteria. World J. Gastroenterol. 13(2):236-243.

Frank, A. C. and J. R. Lobry. 2000. Oriloc: Prediction of replication boundaries in unannotated bacterial chromosomes. Bioinformatics 16(6):560-561.

Garrigues, C., B. Stuer-Lauridsen, and E. Johansen. 2005. Characterisation of *Bifidobacterium animalis* subsp. *lactis* BB-12 and other probiotic bacteria using genomics, transcriptomics and proteomics. Aust. J. Dairy Technol. 60(2):84-92.

Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. Science 312(5778):1355-1359.

Godde, J. S. and A. Bickerton. 2006. The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. J. Mol. Evol. 62(6):718-729.

Gopal, P. K., J. Prasad, J. Smart, and H. S. Gill. 2001. In vitro adherence properties of *Lactobacillus rhamnosus* DR20 and *Bifidobacterium lactis* DR10 strains and their antagonistic activity against an enterotoxigenic *Escherichia coli*. International Journal of Food Microbiology 67:207-216.

Grice, E. A., H. H. Kong, G. Renaud, A. C. Young, NISC Comparative Sequencing Program, G. G. Bouffard, R. W. Blakesley, T. G. Wolfsberg, M. L. Turner, and J. A. Segre. 2008. A diversity profile of the human skin microbiota. Genome Research 18(7):1043-1050.

Gueimonde, M., S. Delgado, B. Mayo, P. Ruas-Madiedo, A. Margolles, and C. G. de los Reyes-Gavilán. 2004. Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. Food Research International. 37:839-850.

Haft, D. H., J. Selengut, E. F. Mongodin, and K. E. Nelson. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. 1(6):e60.

Hansen, F. G., B. B. Christensen, and T. Atlung. 2007. Sequence characteristics required for cooperative binding and efficient *in vivo* titration of the replication initiator protein DnaA in *E. coli*. J. Mol. Biol. 367(4):942-952.

Horvath, P., A.-C. Coûté-Monvoisin, D. A. Romero, P. Boyaval, C. Fremaux, and R. Barrangou. 2008a. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int. J. Food Microbiol. doi:10.1016.

Horvath, P., D. A. Romero, A.-C. Coûté-Monvoisin, M. Richards, H. Deveau, S. Moineau, P. Boyaval, C. Fremaux, and R. Barrangou. 2008b. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. Journal of Bacteriology 190(4):1401-1412.

Jones, B. V., M. Begley, C. Hill, C. G. M. Gahan, and J. R. Marchesi. 2008. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. Proc. Natl. Acad. Sci. U. S. A. 105(36):13580-13585.

Kim, J. F., H. Jeong, D. S. Yu, S.-H. Choi, C.-G. Hur, M.-S. Park, S. H. Yoon, D.-W. Kim, G. E. Ji, H.-S. Park, and T. K. Oh. 2009. Genome sequence of the probiotic bacterium Bifidobacterium animalis subsp. lactis AD011. Journal of Bacteriology 191(2):678-679.

Kleerebezem, M., J. Boekhorst, R. van Kranenburg, D. Molenaar, O. P. Kuipers, R. Leer, R. Tarchini, S. A. Peters, H. M. Sandbrink, M. W. E. J. Fiers, W. Stiekema, R. M. K. Lankhorst, P. A. Bron, S. M. Hoffer, M. N. N. Groop, R. Kerkhoven, M. de Vries, B. Ursing, W. M. de Vos, and R. J. Siezen. 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc. Natl. Acad. Sci. U. S. A. 100(4):1990-1995.

Kurokawa, K., T. Itoh, T. Kuwahara, K. Oshima, H. Toh, A. Toyoda, H. Takami, H. Morita, V. K. Sharma, T. P. Srivastava, T. D. Taylor, H. Noguchi, H. Mori, Y. Ogura, D. S. Ehrlich, K. Itoh, T. Takagi, Y. Sakaki, T. Hayashi, and M. Hattori. 2007. Comparative Metagenomics Revealed Commonly Enriched Gene Sets in Human Gut Microbiomes. DNA Research 14(4):169-181.

Lapierre, L., P. Undeland, and L. J. Cox. 1992. Lithium chloride-sodium propionate agar for the enumeration of bifidobacteria in fermented dairy products. Journal of Dairy Science. 75:1192-1196.

Lenski, R. E., C. L. Winkworth, and M. A. Riley. 2003. Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. J. Mol. Evol. 56(4):498-508.

Ley, R. E., M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight, and J. I. Gordon. 2008. Evolution of mammals and their gut microbes. Science 320(5883):1647-1651.

Mackiewicz, P., J. Zakrzewska-Czerwińska, A. Zawilak, M. R. Dudek, and S. Cebrat. 2004. Where does bacterial replication start? Rules for predicting the *oriC* region. Nucleic Acids Res. 32(13):3781-3791.

Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J.-H. Lee, I. Díaz-Muñiz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer, and D. Mills. 2006a. Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. U. S. A. 103(42):15611-15616.

Makarova, K. S., N. V. Grishin, S. A. Shabalina, Y. I. Wolf, and E. V. Koonin. 2006b. A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct 1:7.

Makarova, K. S. and E. V. Koonin. 2007. Evolutionary genomics of lactic acid bacteria. Journal of Bacteriology 189(4):1199-1208.

Masco, L., G. Huys, E. De Brandt, R. Temmerman, and J. Swings. 2005. Culture-dependent and culture-independent qualitative analysis of probiotic products claimed to contain bifidobacteria. Int. J. Food Microbiol. 102(2):221-230.

Meile, L., W. Ludwig, U. Rueger, C. Gut, P. Kaufmann, G. Dasen, S. Wenger, and M. Teuber. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. Syst. Appl. Microbiol. 20:57-64.

Paineau, D., D. Carcano, G. Leyer, S. Darquy, M.-A. Alyanakian, G. Simoneau, J.-F. Bergmann, D. Brassart, F. Bornet, and A. C. Ouwehand. 2008. Effects of seven potential probiotic strains on specific immune responses in healthy adults: A double-blind, randomized, controlled trial. FEMS Immunol. Med. Microbiol. 53(1):107-113.

Palmer, C., E. M. Bik, D. B. DiGiulio, D. A. Relman, and P. O. Brown. 2007. Development of the human infant intestinal microbiota. PLoS Biol. 5(7):e177.

Parche, S., J. Amon, I. Jankovic, E. Rezzonico, M. Beleut, H. Barutçu, I. Schendel, M. P. Eddy, A. Burkovski, F. Arigoni, and F. Titgemeyer. 2007. Sugar transport systems of *Bifidobacterium longum* NCC 2705. J. Mol. Microbiol. Biotechnol. 12:9-19.

Perfeito, L., L. Fernandes, C. Mota, and I. Gordo. 2007. Adaptive mutations in bacteria: high rate and small effects. Science 317(5839):813-815.

Pridmore, R. D., B. Berger, F. Desiere, D. Vilanova, C. Barretto, A.-C. Pittet, M.-C. Zwahlen, M. Rouvet, E. Altermann, R. Barrangou, B. Mollet, A. Mercenier, T. Klaenhammer, F. Arigoni, and M. A. Schell. 2004. The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. Proc. Natl. Acad. Sci. U. S. A. 101(8):2512-2517.

Saeed, A. I., V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Currier, M. Thiagarajan, A. Sturn, M. Snuffin, A. Rezantsev, D. Popov, A. Ryltsov, E. Kostukovich, I. Borisovsky, Z. Liu, A. Vinsavich, V. Trush, and J. Quackenbush. 2003. TM4: A free, open-source system for microarray data management and analysis. BioTechniques 34(2):374-378.

Schell, M. A., M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M.-C. Zwahlen, F. Desiere, P. Bork, M. Delley, R. D. Pridmore, and F. Arigoni. 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. Proc. Natl. Acad. Sci. U. S. A. 99(22):14422-14427.

Sela, D. A., J. Chapman, A. Adeuya, J. H. Kim, F. Chen, T. R. Whitehead, A. Lapidus, D. S. Rokhsar, C. B. Lebrilla, J. B. German, N. P. Price, P. M. Richardson, and D. A. Mills. 2008. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. Proc. Natl. Acad. Sci. U. S. A. 105(48):18964-18969.

Sgorbati, B., M. B. Smiley, and T. Sozzi. 1983. Plasmids and phages in *Bifidobacterium longum*. Microbiologica 6(2):169-173.

Sonnhammer, E. L. and R. Durbin. 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167(1-2):GC1-10.

Sorek, R., V. Kunin, and P. Hugenholtz. 2008. CRISPR- A widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. 6(3):181-186.

Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin. 2000. The COG database: A tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28(1):33-36.

Turroni, F., E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G. L. de'Angelis, F. Shanahan, D. van Sinderen, and M. Ventura. 2009. Exploring the diversity of the bifidobacterial population in the human intestinal tract. Applied and Environmental Microbiology. 75(6):1534-1545.

Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater, and D. van Sinderen. 2007. Genomics of Actinobacteria: Tracing the evolutionary history of an ancient phylum. Microbiol. Mol. Biol. Rev. 71(3):495-548.

Ventura, M., J.-H. Lee, C. Canchaya, R. Zink, S. Leahy, J. A. Moreno-Munoz, M. O'Connell-Motherway, D. Higgins, G. F. Fitzgerald, D. J. O'Sullivan, and D. van

Sinderen. 2005. Prophage-like elements in bifidobacteria: Insights from genomics, transcription, integration, distribution, and phylogenetic analysis. Applied Environmental Microbiology. 71(12):8692-8705.

Ventura, M. and R. Zink. 2002. Rapid Identification, Differentiation, and Proposed New Taxonomic Classification of *Bifidobacterium lactis*. Applied and Environmental Microbiology 68(12):6429-6434.

Wall, R., S. G. Hussey, C. A. Ryan, M. O'Neill, G. Fitzgerald, C. Stanton, and R. P. Ross. 2008. Presence of two *Lactobacillus* and *Bifidobacterium* probiotic strains in the neonatal ileum. ISME J. 2(1):83-91.

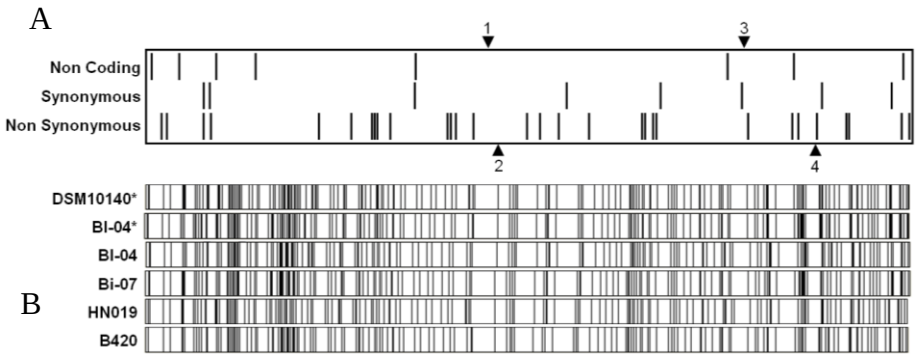
Zhang, W., W. Qi, T. J. Albert, A. S. Motiwala, D. Alland, E. K. Hyytia-Trees, E. M. Ribot, P. I. Fields, T. S. Whittam, and B. Swaminathan. 2006. Probing genomic diversity and evolution of *Escherichia coli* O157 by single nucleotide polymorphisms. Genome Res. 16(6):757-767.

Table 1. General characteristics of selected bifidobacteria

	B. animalis* subsp. *lactis		***B. adolescentis***	***B. longum***
Data Category	**Bl-04**	**DSM 10140**	**ATCC 15703**	**NCC2705**
Genome Size (bp)	1,938,709	1,938,483	2,089,645	2,260,266
Coding Percentage	90.45%	90.34%	87.57%	86.07%
G+C	60.48%	60.48%	59.18%	60.13%
Average ORF Length (bp)	1053	1056	1122	1503
Predicted ORFs	1,655	1,658	1,631	1,729
ORFs with Assigned Function	1,072	1,071	1,165	1,305
COG Matches	773	773	791	839
Pfam Domains	890	890	936	969
rRNA Operons	4	4	5	4
tRNA[2]	52	51	55	54
Transposases	9	9	15	1

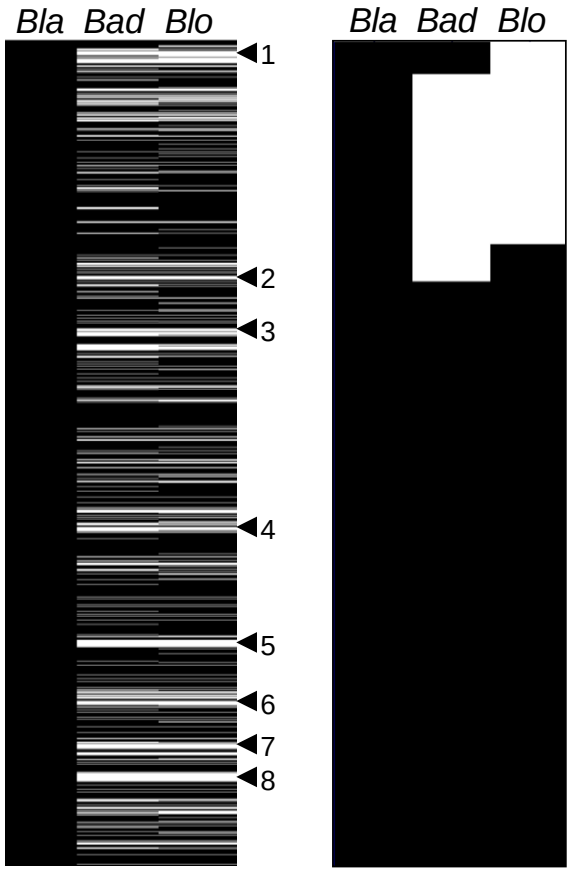
Numbers extracted from the ERGO Database at the time of analysis

Figure 1. Alignment of the *B. animalis* subsp. *lactis* genomes



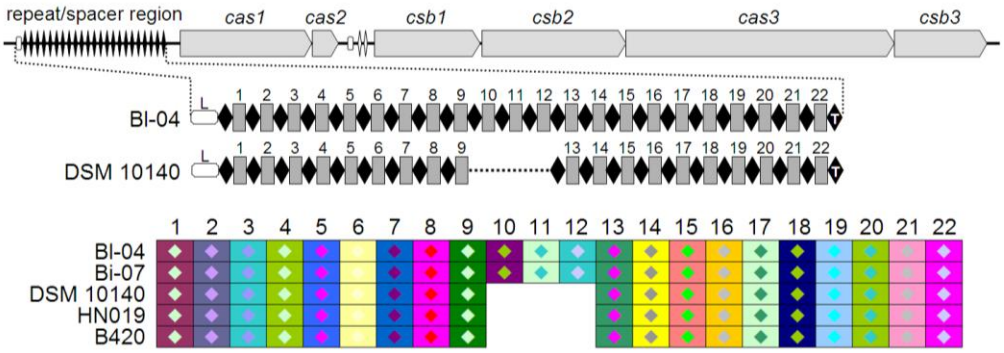
Differences observed between the two *B. animalis* subsp. *lactis* genomes: Panel A: SNPs are indicated using bars, with non-coding SNPs at the top, synonymous SNPs in the middle and non-synonymous SNPs at the bottom. INDELs are shown as arrows, with insertions in the Bl-04 genome shown on top and insertions in the DSM 10140 genome shown on the bottom. Panel B: Optical maps of various *B. animalis* subsp. *lactis* strains, visualized using the OpGen MapViewer, based on a NotI digest (bottom). The first two optical maps were generated *in silico* (indicated by the *), based on the genome sequences. The other optical maps represent experimental data.

Figure 2. Comparative genomic analysis of bifidobacteria



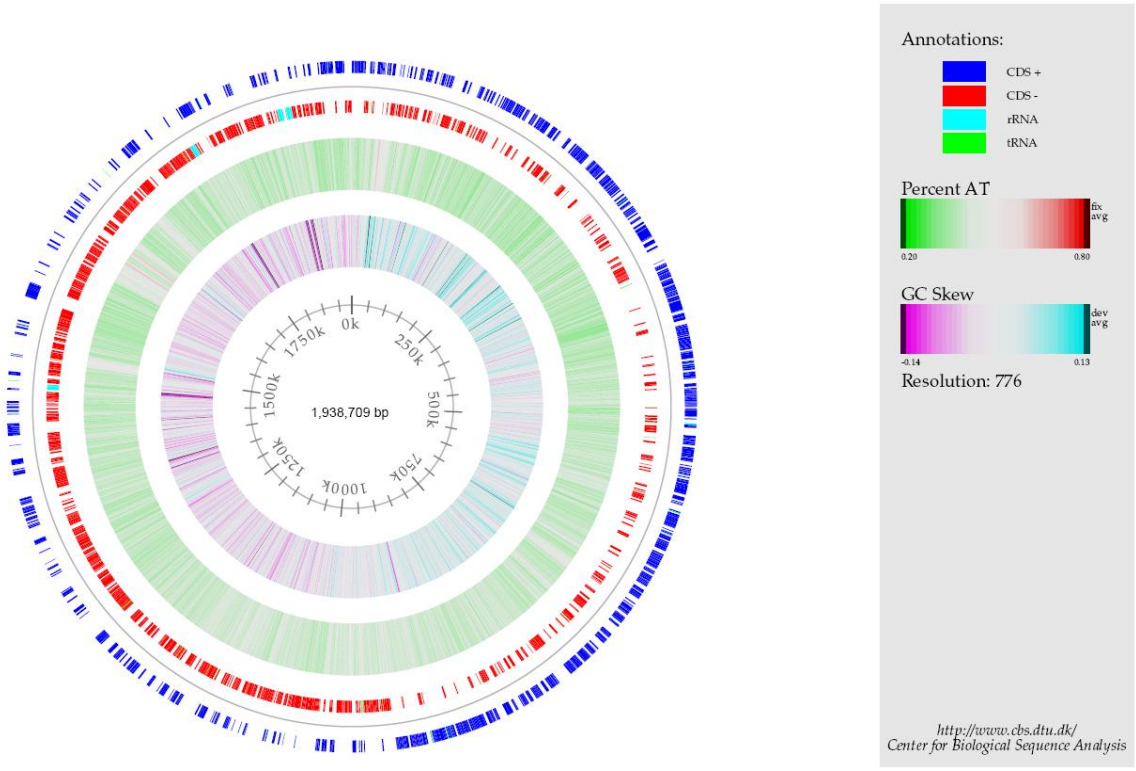
Sequences homologous to *B. animalis* subsp. *lactis* (*Bla*) proteins are shown in black, whereas absence of a homolog in the *B. adolescentis* ATCC 15703 (*Bad*) or in the *B. longum* NCC2705 (*Blo*) genomes appear in white, represented using MeV. Data represented in the chromosomal ORF order with the origin of replication at the top and numbers indicating location of areas of unique content in *B. animalis* subsp. *lactis* (left panel). Hierarchical representation of homologs across the three genomes (right panel).

Figure 3. Overview of the *B. animalis* subsp. *lactis* CRISPR locus



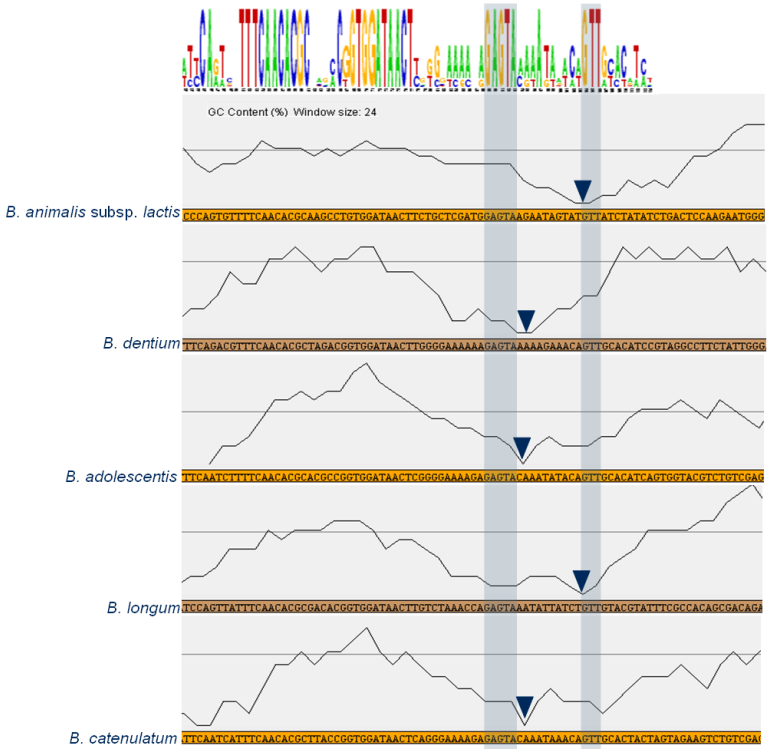
The Bala1 CRISPR locus as it appears in the genome is shown at the top with the repeat/spacer region on the left (repeats are represented as black diamonds and spacers as gray boxes) and *cas* genes on the right (top). The repeat-spacer region is shown for Bl-04 and DSMZ 10140 with repeats shown as black diamonds, spacers shown as numbered gray boxes, with the leader shown as a white box and the terminal repeat show as a black diamond annotated with a "T" (middle). CRISPR spacer overview where repeats are not included and only spacers are represented using a two-color combination (bottom).

Supplementary Figure A. *B. animalis* subsp. *lactis* genome



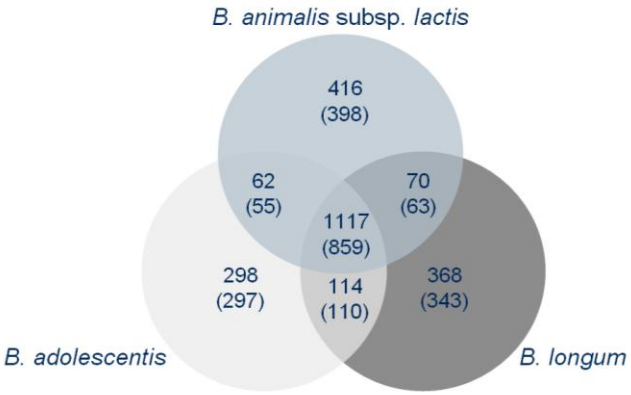
The atlas represents a circular view of the complete genome sequence of the Bl-04 chromosome. The key (right) describes circles from the outermost to innermost circle. Circle 1 (outermost circle): open reading frames on the positive strand (red) and negative strand (blue). Circle 2: GC skew. Circle 3: AT content. Circle 4: rRNA (aqua) and tRNA (green).

Supplementary Figure B. Putative origin of replication



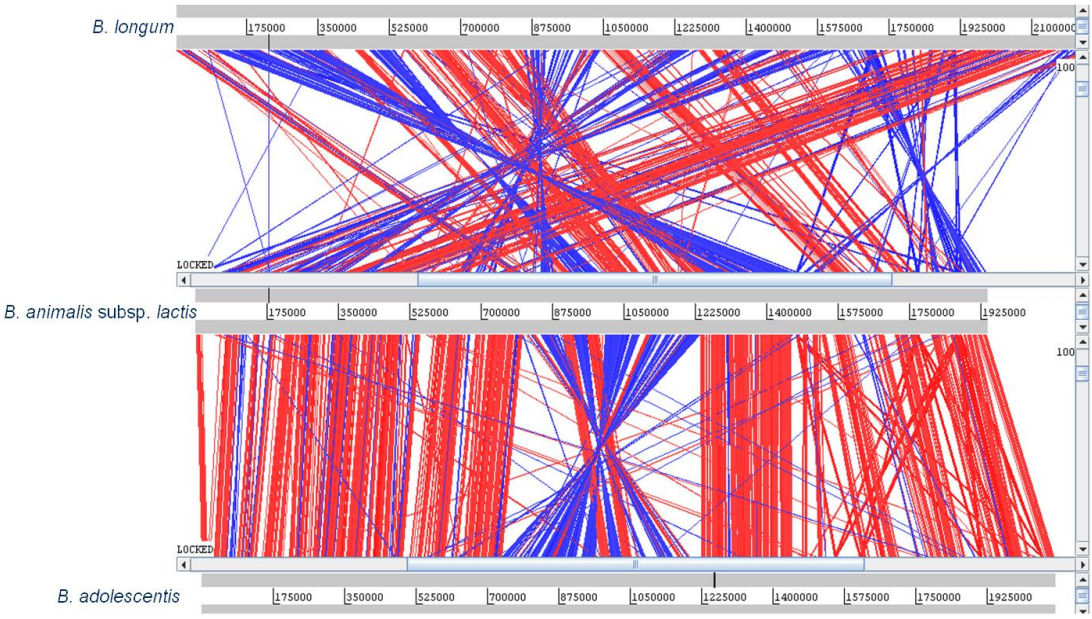
Sequences conserved across several *Bifidobacterium* genomes are shown at the top using WebLogo (Crooks et al., 2004) and the putative origins of replication are indicated as black arrows on the genome sequence of *B. animalis* subsp. *lactis, B. dentium, B. adolescentis, B. longum* and *B. catenulatum.* The GC content (%GC) is also represented graphically using Artemis (Sanger Institute).

Supplementary Figure C. Comparative analysis of genome content across *Bifidobacterium* species



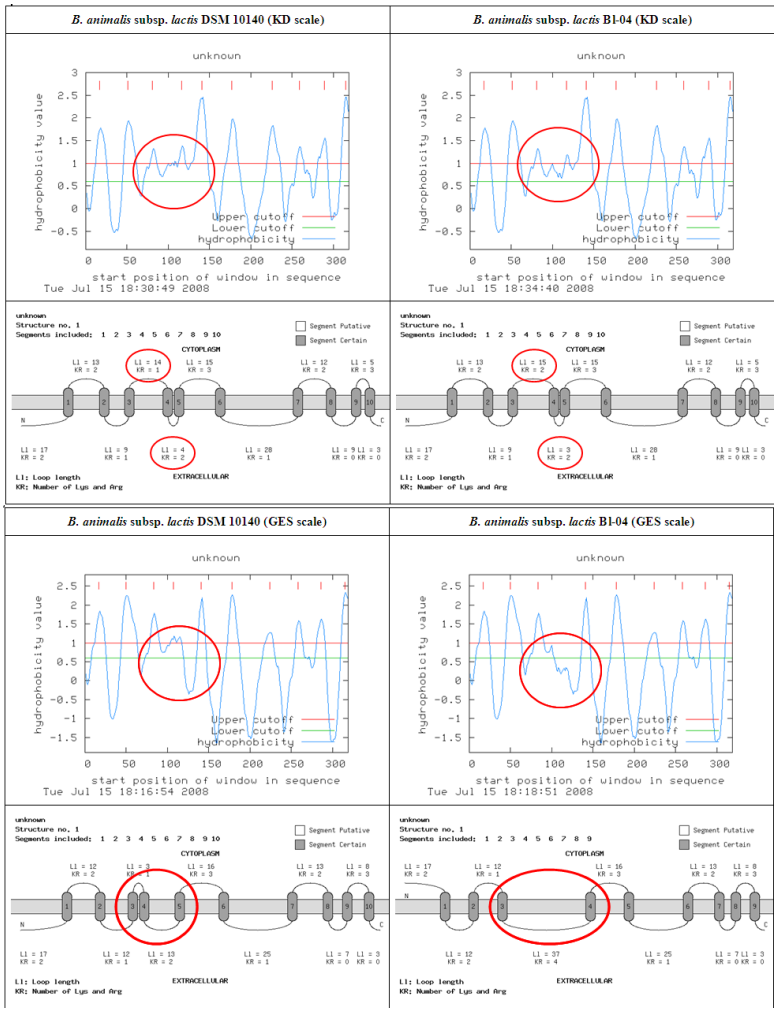
The number of orthologous sequences (top number) and orthologous groups (bottom number) are represented as a Venn diagram. Numbers were generated using the ERGO package (Integrated Genomics). While 416 putative CDS were predicted to be unique in the *B. animalis* subsp. *lactis* genome according to the ERGO *in silico* settings, 324 CDS appear valid given their size, orientation and genetic context.

Supplementary Figure D. Comparative alignment of *Bifidobacterium* genomes



Lines represent DNA-DNA similarities as defined by BLASTn matches, between *B. animalis* subsp. *lactis* and *B. longum* NCC2705 (top) and between *B. animalis* subsp. *lactis* and *B. adolescentis* ATCC 15703 (bottom) using the Artemis Comparison Tool (ACT, Sanger Institute). Matches on co-directional DNA strands are shown in red while matches on opposite DNA strands are shown in blue.

Supplementary Figure E. Comparison of putative GlcU topologies



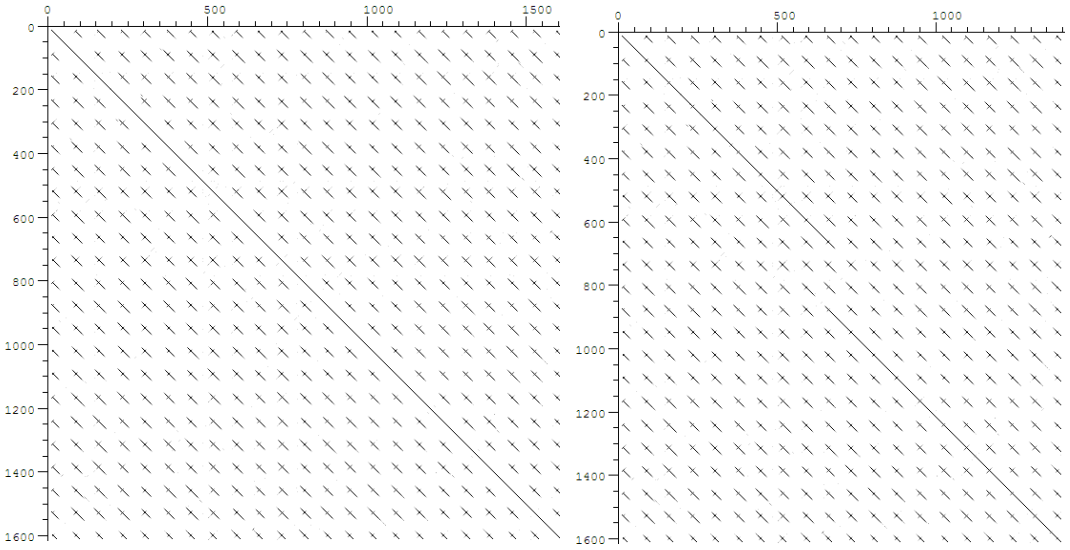
A SNP was identified in a putative glucose permease gene, *glcU* (Balac_1097), which results in a change in amino acid sequence from a glycine in DSM 10140 to an arginine in Bl-04 at position 121 in the protein. Hydrophobicity plots and predicted protein topologies were predicted using both the Kyte & Doolittle (KD) and Goldman-Engelman-Steitz (GES) scales, obtained through analysis at the Mobyle site (http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=toppred) using Toppred. The areas affected by the change in amino acid sequence are circled in red. While there are minor differences between the strains with the KD scale (top), a different number of

transmembrane segments is identified between the two strains using the GES scale (bottom).

Supplementary Figure F. Bala1 CRISPR locus in *Bifidobacterium animalis* subsp. *lactis*

```
R1  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R2  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R3  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R4  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R5  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R6  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R7  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R8  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R9  ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R10 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R11 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R12 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R13 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R14 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R15 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R16 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R17 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R18 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R19 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R20 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R21 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
R22 ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGA
R23 ATCTCCGAAGTTTTGGCTTCGGAGCTTCATTGAGGA

S1  GACGATATGGCGCTCAGCGTGGCGGAGTGGGAGGCGG
S2  AAGACCGGCACCGAACGCGACTTCACCATGACCTC
S3  GCCCACCACAACGGCAACGGCGGAGGAACACGCGCCGAA
S4  AAGCCGAACTCAATCACACGCATCAAAGCGAACA
S5  GTATTCGCCGTTCGAGAGGAATGAGAGGATGCTGTCAG
S6  TCGCCATTGGAGACGCGACGCAGGATACTATGGC
S7  ACGACAAGCCGCCGCCACCGATATTCACCTGCGA
S8  GGCCGCTTCGGTGACGGGCTGGTTTTTCCACCACACGC
S9  AATCCCAGCCGCAAGGTCTGATGCCGCCTGAAAT
S10 CACTGGTGGTGCGAATACGCCGAAACGGTGGAATGG
S11 ATTGAGATTGATACCCGTGGCGCCGCTGATGAGAC
S12 AATCCCTCGGCCCATGATTCGTCACGTGGGATCAC
S13 AAACAGGTCAATCAGCGGCGCAGGGAGGAGACGAA
S14 GAGTGAACAACTCACTGTGCGAACCATCGAACCGTT
S15 CGGTTGAGCAGCCACGTGGTGATACTGCTCGCGCCA
S16 CTTGCATCCAACGCGCACAGCATTGCATACGGGTATAG
S17 ATCATCCTCACGGAAATAGTGAGCATCCTCGAGAACCTG
S18 GGCCGCGATAGTCCACGAGGCGAACGAAGGCGTTGC
S19 GCTCAAGACACTCACCGACCAGCTCAAGAAGACCGA
S20 CGCGATCGTCACCGACTGCACTGTGTTCGCACTGTC
S21 GCGACACCGAACGCCGCCGCCACAGTCGGGATGGC
S22 AGGGCCAGCAACGTCGTGGAGATCCATCAGGAGGC
```



(top) Sequences for repeats R1 through R23 (left) and spacers S1 through S22 (right). (bottom) Self alignment of the Bl-04 Bala1 repeat-spacer region (left) and comparison between the Bl-04 (Y axis) and the DSM 10140 (X axis) Bala1 (right) using Dotter (Sonnhammer and Durbin, 1995).

Supplementary Table 1. Comparative genomics

Bla	Bad	Blo	ORF	Product
1	1	1	Balac_0001	dnaA
1	1	1	Balac_0002	DNA polymerase III subunit beta
1	1	1	Balac_0003	recombination protein RecF
1	1	1	Balac_0004	hypothetical protein
1	1	1	Balac_0005	gyrB
1	1	1	Balac_0006	DNA gyrase subunit A
1	1	1	Balac_0007	hypothetical protein
1	1	1	Balac_0008	hypothetical protein
1	1	1	Balac_0009	Putative hemolysin III-like membrane protein
1	1	1	Balac_0010	glutamate dehydrogenase
1	1	0	Balac_0011	multidrug transport protein
1	1	0	Balac_0012	hypothetical protein
1	1	0	Balac_0013	mannan endo-1,4-beta-mannosidase precursor
1	1	1	Balac_0014	large-conductance mechanosensitive channel
1	1	1	Balac_0015	hypothetical protein
1	0	0	Balac_0018	hypothetical protein
1	0	1	Balac_0019	membrane-anchored glycerophosphoryl diesterphosphodiesterase-like protein
1	0	1	Balac_0020	putative transcriptional regulator
1	0	0	Balac_0021	hypothetical protein

For the complete 38-page comparative genomics table refer to Journal of Bacteriology Supplemental Material (http://jb.asm.org/content/191/13/4144/suppl/DC1)

Supplementary Table 1. SNP sites between the *B. animalis* subsp. *lactis* Bl-04 and DSM 10140 genomes

Genome Position		Nucleotide				Bl-04	Amino Acid Change		
Bl-04	**10140**	**Bl-04**	**10140**	**HN019[1]**	**AD011[1]**	**ORF Number[2]**	**Bl-04**	**DSM 10140**	**Gene ID[3]**
SNPs in Non-coding Regions									
19098	19098	T	C	T	T	Balac_0014/Balac_0015	--	--	
87537	87537	T	C	T	T	Balac_0078/ Balac_0079			
278276	278277	T	G	T	T	Balac_0233/Balac_0232	--	--	
281373	281374	C	T	T	T	Balac_0235/Balac_0234	--	--	
686997	687000	A	G	G	G	Balac_0581/Balac_0582	--	--	
1456395	1456329	G	-	-	-	Balac_1262/Balac_1261	--	--	
1636887	1636606	G	A	A	A	Balac_1395/Balac_1396	--	--	
1916727	1916502	C	T	C	C	Balac_1613/Balac_1614	--	--	
Synonymous SNPs in Protein-coding Regions									
158448	158448	C	T	C	C	Balac_0140	L	L	ABC-type branched-chain amino acid transport systems, periplasmic component
175867	175868	G	T	G	G	Balac_0153	A	A	LacI family transcriptional regulator
685549	685552	G	A	G	G	Balac_0580	I	I	Hypothetical protein
1059607	1059543	A	T	A	A	Balac_0922	V	V	CTP synthase
1299818	1299752	G	A	G	G	Balac_1126	D	D	ABC-type multidrug transport system permease component
1496404	1496337	G	A	G	G	Balac_1296	S	S	ABC transporter
1716326	1716099	G	A	G	G	Balac_1460	G	G	Hypothetical protein
1881197	1880972	C	T	T	T	Balac_1590	I	I	LacI-type transcriptional regulator
Non-Synonymous SNPs in Protein-coding Regions									
41264	41264	G	A	G	G	Balac_0039	Q	STOP	Hypothetical protein
53321	53321	A	G	A	A	Balac_0051	L	P	Transposase (only one copy of transposase is in the HN019 genome)
159526	159526	-	G	G	G	Balac_0141	frameshift	frameshift	Acetyl-/propionyl-coenzyme A carboxylase alpha chain
177291	177292	A	G	A	A	Balac_0156	V	A	Hypothetical protein
459191	459192	C	T	C	C	Balac_0375	A	V	SSU ribosomal protein S9P
525385	525386	G	A	G	G	Balac_0443	A	T	Carboxypeptidase S1
581338	581339	C	T	C	C	Balac_0487	R	W	LacI-type translational regulator
588991	588993	-	G	G	G	Balac_0496	frameshift	frameshift	D-alanyl-D-alanine carboxypeptidase
594225	594228	G	A	G	G	Balac_0500	R	H	Dihydroneopterin aldolase
629421	629424	C	T	C	C	Balac_0529	H	Y	Oligoribonuclease
769449	769452	-	C	-	-	Balac_0653	frameshift	frameshift	Uracil-DNA glycosylase
776136	776140	T	C	T	T	Balac_0660	L	P	Sensor protein
791497	791501	A	C	A	A	Balac_0671	Y	S	ABC-type amino acid transport system periplasmic component
839425	839429	C	T	T	T	Balac_0710	A	V	Hypothetical membrane protein
959997	959934	A	G	G	G	Balac_0837	T	A	Hypothetical protein
993071	993008	C	T	T	T	Balac_0864	P	S	Hypothetical protein

Genome Position		Nucleotide				Bl-04	Amino Acid Change		
Bl-04	**10140**	**Bl-04**	**10140**	**HN019[1]**	**AD011[1]**	**ORF Number[2]**	**Bl-04**	**DSM 10140**	**Gene ID[3]**
1040110	1040047	A	-	A	A	Balac_0903	frameshift	frameshift	Long-chain-fatty-acid-CoA ligase
1108705	1108641	C	T	C	C	Balac_0971	A	V	Putative phosphoketolase
1247760	1247696	G	A	G	G	Balac_1088	Q	STOP	ATP-binding protein of ABC transporter
1255138	1255074	T	C	T	T	Balac_1094	V	A	Hypothetical protein
1260073	1260009	G	C	C	C	Balac_1097	R	G	Putative glucose uptake permease
1287673	1287609	C	-	C	C	Balac_1116	frameshift	frameshift	DNA binding protein
1278674	1287609	T	-	T	T	Balac_1116	frameshift	frameshift	DNA binding protein
1510941	1510874	C	T	C	C	Balac_1308	D	N	CRISPR-associated Cas1/Cas4 family protein
1635976	1635695	G	A	G	G	Balac_1395	A	V	H(+)-stimulated manganese uptake system protein
1652309	1652028	C	A	C	C	Balac_1407	A	S	Hypothetical protein
1702923	1702641	-	C	C	C	Balac_1453	frameshift	frameshift	Putative oxalyl-CoA decarboxylase
1792444	1792217	C	A	C	C	Balac_1524	G	W	Polysaccharide ABC transporter permease
1799512	1799285	G	A	G	G	Balac_1529	V	I	Hypothetical protein
1910534	1910309	G	A	G	G	Balac_1610	S	F	ATP binding protein of ABC transporter for sugars
1934933	1934708	G	A	G	G	Balac_1626	S	F	Glucose-inhibited division protein B
Putative SNPs Not Verified									
1304746	1304680	C	T						Transposon
1472571	1472504	G	A						rRNA
1659466	1659185	T	C						Transposon
1660202	1659921	C	G						Transposon
1660229	1659948	C	A						Transposon
1660307	1660026	C	A						Transposon
1660541	1660260	G	T						Transposon
1677062	1676781	T	-						Intergenic
1863560	1863333	G	-						rRNA
1864992	1864764	-	G						rRNA
1865023	1864796	-	T						rRNA
1865106	1864880	C	S						rRNA
1865119	1864893	C	S						rRNA
1865176	1864950	-	C						rRNA

[1] SNPs in HN019 and AD011 were not verified by resequencing and are based on analysis of the draft genome sequence found in ERGO (HN019) and accessed on 10/13/2008 and GenBank (AD011) accessed on 3/17/09.

[2] Number refers to the ORF numbering in NCBI.

[3] Gene ID refers to the annotation in ERGO.

1

Supplementary Table 3. Comparison of INDEL sites among the four *B. animalis* subsp. *lactis* genomes

INDEL[1]	Length (bp)	Identification of INDEL locus	Sequence (P for present; A = absent)			
			DSM 10140	Bl-04	HN019	AD011
INDEL1	121	tRNA-Ala-GGC	A	P	P	P
INDEL2	54	long chain fatty acid-coA ligase	P	A	P	P
INDEL3	214	CRISPR locus	A	P	A	A
INDEL4	54	Intergenic region	P	A	P	A

[1] Corresponds to numbering in Figure 1.

Supplementary Table 4. Putative carbohydrate hydrolyzing enzymes identified in the genome sequences of *Bifidobacterium*

Putative Hydrolase	**E.C. Number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
Beta-fructofuranosidase (inulinase), sucrose-6-phosphate hydrolase	3.2.1.26	1241	00105	00420 00873
Glucosylceramidase	3.2.1.45	0052 1418		00487
Galactoside O-acetyltransferase	2.3.1.18	0173	01079 01080	
Beta-mannosidase	3.2.1.25	0152		
Sucrose phosphorylase	2.4.1.7	0138	00536	01218
Glycogen phosphorylase	2.4.1.1	0076	00597	01164
Cellobiose phosphorylase	2.4.1.20	1421		01622
Beta-N-acetylhexosaminidase	3.2.1.52	1025	01031	01844
Phosphoglucomutase / Phosphomannomutase	5.4.2.2 / 5.4.2.8	0425	01630	01545
Phosphomannomutase	5.4.2.8	0999	01004	00647
Alpha-glucosidase	3.2.1.20	1566	00523 00542 01334	00627
Alpha,alpha-phosphotrehalase (Trehalose-6-phosphate hydrolase)	3.2.1.93	1573	00529	00165 01140 01637
Oligo-1,6-glucosidase	3.2.1.10	1593	01526	00592
Alpha-L-arabinofuranosidase (both debranching and exo activity)	3.2.1.55	0065	00181	00484 01093 01301 01304 01306
Alpha-L-arabinofuranosidase / Beta-xylosidase (Xylan 1,4-beta-xylosidase)	3.2.1.55 / 3.2.1.37	0512 0517	00184 00187 00544 01138 01166 01611	00482 01602 01608
Mannan endo-1,4-beta-mannosidase	3.2.1.78	1450 0012		00293

Putative Hydrolase	**E.C. Number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
Beta-galactosidase	3.2.1.23	0268	00259	00491
		0484	01168	00630
		0053	01775	00958
		0476	00978	00959
				01108
				00638
				01617
Beta-galactosidase / Transgalactosylase			00978	00604
				00638
4-alpha-glucanotransferase	2.4.1.25	0373	00527	01139
		1567	01570	01485
1,4-alpha-glucan branching enzyme	2.4.1.18	0995	0999	00651
Isoamylase	3.2.1.68	0376	00982	00667
		0977	01573	01488
				01489
Alpha-galactosidase	3.2.1.22	1537	01518	00597
		1601		01097
				01477
				01478
Endo-1,4-beta-xylanase	3.2.1.8	1574	00420	01096
		0520	00421	01620
			00682	
			01543	
			01544	
Putative xylanase /	3.2.1.8		0420	
Putative xylan esterase	3.2.1.8		0682	
Putative xylanase/chitin deacetylase	3.5.1.-	0171		
Endo-1,4-beta-xylanase / Arabinan endo-1,5-alpha-L-arabinosidase	3.2.1.8 / 3.2.1.99		00182	
			00183	
Beta-glucosidase	3.2.1.21	1551	01757	00224
		0151		00476
Thermostable b-glucosidase		0049	01763	00472
			01764	00624
Glucan 1,3-β-glucosidase / β-glucosidase	3.2.1.58 / 3.2.1.21			00017
				00025
Glucan 1,3-beta-glucosidase	3.2.1.58		01761	00481
				00483

Putative Hydrolase	**E.C. Number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
Pullulanase	3.2.1.41	1562		00694
		0924		00696
		0925		01137
Alpha-N-acetylgalactosaminidase precursor (galactomannanase)	3.2.1.49		00177	01094
Arabinogalactan endo-1,4-beta-galactosidase precursor	3.2.1.89		00257	
Sialidase (endo-α-N-acetylgalactosaminidase)	3.2.1.18		00464	
Chitinase	3.2.1.14		00895	
Cyclomaltodextrinase (neopullanase)	3.2.1.54		00907	
Glycosyl hydrolases (1,4-β-acetylmuramidase, and other cell wall-degrading enzymes)	3.2.1.17			01072
		1516		
		1517		
		1520		
Glycosyl transferases	2.4.1.-	0613	01104	01723
		0824	00826	00744
		0888	01674	01583
		1101	01676	01585
		1509	01721	01634
		0491	01422	00281
		0493	01320	00371
		0527	00215	01066
		1109	00235	00936
		1201	00045	00939
		1374	00205	00941
		1379	00213	00942
		1381	00216	01058
		1382	00217	01059
		1384	00236	01256
		1386	00438	01577
		1402	00566	01903
		1403	00672	
		1511		
Alpha-mannosidase	3.2.1.24		01327	00523
			01328	

Putative Hydrolase	**E.C. Number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
			01329	
Endo-beta-N-acetylglucosaminidase	3.2.1.96		01335	
Putative arabinoidase (hypothetical)			00146	
Putative N-acetyl-β-hexosaminidase			00056	
Putative amylase-like glucanase		1458	00388	0995

Supplementary Table 5. Putative stress response genes identified in the genome sequences of *Bifidobacterium*

Annotation in ERGO	**E.C. number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
DnaK		1557	00520	01122
GrpE		1556	00519	01121
DnaJ		1555 0840	00517 00719	01120 00829
HsrP		1554	00516	01119
GroES		0359	01558	01468
GroEL		0656	00002	01764
HrcA		0841	00718	00830
RNA polymerase ECF-type Sigma Factor		1167	01357	00352
ECF-type sigma factor negative selector		1166	01358	00351
ClpB		1481	01250	01024
ClpC		0665	00010	01772
ClpP haloenzyme	3.4.21.92			
ClpP1		1078	00945	01788
ClpP2		1077	00944	01789
ClpX		1076	00943	01790
ATP-Dependent endopeptidase Lon	3.4.21.53	1256	00094	00430
Endopeptidase DegP	3.4.21.-	0101	00555	01192
Endopeptidase HtpX	2.4.24.-	0106	00551	01197
CspA cold shock protein		0655	00001	
CspB cold shock protein		0661	00007	01769
RecN DNA repair protein		0736	01043	00096
RecA DNA repair protein		1114	01415	00286
LexA repressor		1212	00383	01310
RecX-like protein		1113	01416	00285
Thioredoxin reductase	1.8.1.9	1623 0866	00649 00614	00039
Peroxiredoxin	1.11.1.15	0865	00615	
Thioredoxin peroxidase	1.11.1.15	0889	00821	
Ferroxidase	1.16.3.1	0056	00618	

Annotation in ERGO	**E.C. number**	***B. animalis* subsp. *lactis* Bl-04**	***B. longum* subsp. *longum* NCC2705**	***B. adolescentis* ATCC 15703**
Oxygen-independent coproporphyrinogen-III oxidase	1.3.99.22	0893	00847	00740
Protoporphyrinogen oxidase	1.3.3.4	1470	00399	01013
NADH oxidase, water-forming	1.6.99.3	0244	01266	
Ferrodoxin-NADP reductase	1.18.1.2	0099	00522	01196
Chromate reductase/NADPH-dependent FMN reductase	1.5.1.-	1576	00139	00006

Appendix B

Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* using SNPs and INDELs

Elizabeth P. Briczinski, **Joseph R. Loquasto**, Rodolphe Barrangou, Edward G. Dudley, Anastasia M. Roberts, and Robert F. Roberts

Published as: 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. Applied and Environmental Microbiology **75:**7501-7508.

Statement of Contribution: The candidate assisted with design of the project and was responsible for design of PCR primers to evaluate each strain, carried out PCR, sequence analysis, analyzed and ordered clustering of allelic groups and was actively involved in preparation of the manuscript.

Abstract

Several probiotic strains of *Bifidobacterium animalis* subsp. *lactis* are widely supplemented into food products and dietary supplements due to their documented health benefits and ability to survive within the mammalian gastrointestinal tract and acidified dairy products. The strain-specificity of these characteristics demands techniques with high discriminatory power to differentiate among strains. However, to date, molecular-based approaches, such as pulsed-field gel electrophoresis and randomly amplified polymorphic DNA-PCR, have been ineffective at achieving strain separation due to the monomorphic nature of this subspecies. Previously, sequencing and comparison of two *B. animalis* subsp. *lactis* genomes (DSMZ 10140 and Bl-04) confirmed this high level of sequence similarity, identifying only 47 single nucleotide polymorphisms (SNPs) and 4 insertions and/or deletions (INDELs) between them. In this study, we hypothesized that a sequence-based typing method targeting these loci would permit greater discrimination between strains than previously attempted methods. Sequencing 50 of these loci in 24 strains of *B. animalis* subsp. *lactis* revealed a combination of 9 SNPs/INDELs could be used to differentiate strains into 14 distinct genotypic groups. In addition, the presence of a non-synonymous SNP within the gene encoding a putative glucose uptake protein was found to correlate with the ability of certain strains to transport glucose and to grow rapidly in a medium containing glucose as the sole carbon source. The method reported here can be used in clinical, regulatory, and commercial applications requiring identification of *B. animalis* subsp. *lactis* at the strain level.

Introduction

Probiotics are currently defined as live microorganisms which, when administered in adequate amounts, confer a health benefit on the host (FAO/WHO, 2002). Many of the organisms studied for their probiotic potential are members of lactic acid bacteria and the genus *Bifidobacterium*, which has resulted in their inclusion in a large variety of dietary supplements and food products. Relative to most bifidobacterial species of human origin, *Bifidobacterium animalis* subsp. *lactis* is less sensitive to stressful conditions (bile, acid, and oxygen) which might be encountered in the mammalian gastrointestinal tract or in fermented or acidified dairy products (Bonaparte and Reuter, 1997, Klein et al., 1998, Prasad et al., 1998, Mättö et al., 2004, Masco et al., 2005). *B. animalis* subsp. *lactis* is widely added to commercial products because it is better able to withstand the adverse conditions of starter culture and product manufacture and to maintain viability and stability during product shelf-life (Mättö et al., 2006). Therefore, strains of *B. animalis*, specifically *B. animalis* subsp. *lactis*, have been found in the majority of probiotic-supplemented dairy products surveyed in North America (United States, Canada) and Europe (Great Britain, France, Italy, Germany) (Biavati et al., 1992, Iwana et al., 1993, Yaeshima et al., 1996, Mattarelli et al., 2002, Fasoli et al., 2003, Grand et al., 2003, Mayer et al., 2003, Gueimonde et al., 2004, Masco et al., 2005, Jayamanne and Adams, 2006).

When selecting a probiotic microorganism to add to supplements or foods, the strain must be identified at the genus, species, and strain levels (Saarela et al., 2000). Proper characterization of a strain is important for safety and quality assurance, for identifying and differentiating putative probiotic strains, and for understanding the interactions among members of gut microbiota. In addition, proper characterization is

important to maintain consumer confidence. Product labels often list invalid names of organisms or misidentify the species the product contains, leading to consumer confusion (Biavati et al., 1992, Yaeshima et al., 1996, Reuter, 1997, Hamilton-Miller et al., 1999, O'Brien et al., 1999, Mattarelli et al., 2002, Masco et al., 2005, Ibrahim and Carr, 2006). In the case of *Bifidobacterium*, most dairy products sold in the United States do not identify species and many only refer to the invalid name "Bifid" or "Bifidus". At the very least, added microorganisms should be accurately identified to the species level on product labels.

According to the FAO/WHO guidelines for probiotic use, specific health benefits observed in research employing a specific strain cannot be extrapolated to other, closely-related strains (FAO/WHO, 2002). Although most clinical studies of probiotic strains compare strains of different genera or different species, few studies have assessed the actual variability of expected health benefits within species or subspecies. However, it is reasonable to consider that health effects, like the phenotypic traits exhibited by strains within a species, are strain-specific. Therefore, reliable techniques for identification of probiotic organisms at the strain level are required.

Characterization to the strain level has several important potential applications. Understanding the complex interactions among microorganisms in the intestinal ecosystem requires methods of differentiating a strain of interest from other strains of the same species contained in the autochthonous microbiota. Strain differentiation techniques also aid in assessing survival of a probiotic organism through the gastrointestinal system, which is particularly important for clinical trials and regulatory purposes (Hoffman et al., 2008). The ability to uniquely identify a strain also lends

credibility to statements made about the potential health benefits of consuming a particular product containing a strain with demonstrated probiotic effects and supports the licensing or intellectual property rights of the manufacturer.

The high degree of genome conservation observed between strains of *B. animalis* subsp. *lactis* in terms of size, organization, and sequence is indicative of a genomically monomorphic subspecies (Barrangou et al., 2009, Kim et al., 2009, and HN019 GenBank Project #28807). As an example, comparison of the complete genome sequences of two *B. animalis* subsp. *lactis* strains, DSMZ 10140 (the Type strain) and Bl-04 (a commercial strain, also known as RB 4825 (Barrangou et al., 2009), identified 47 single nucleotide polymorphisms (SNPs) in non-repetitive elements, as well as 443 bp distributed among 4 INDEL sites: a 121-bp tRNA-encoding sequence, a 54-bp region within the long-chain fatty acid-coenzyme A ligase gene, a 214-bp region within the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) locus, and a 54-bp intergenic sequence. Overall, this 99.975% genome identity explains the inability to differentiate these strains by techniques such as sequencing of housekeeping genes, multi-locus sequence typing (MLST), and pulsed-field gel electrophoresis (PFGE) (Roy and Sirois, 2000, Jian et al., 2001, Ventura and Zink, 2003, Zhu et al., 2003, Ventura et al., 2004, Berthoud et al., 2005, Briczinski and Roberts, 2006, Ventura et al., 2006).

The strain-specificity of reported health benefits of probiotics and the frequent use of *B. animalis* subsp. *lactis* as a probiotic in food products and supplements demands techniques with greater discriminatory power to identify and differentiate among strains within this highly homogeneous group. Unfortunately, strain-level differentiation of *B. animalis* subsp. *lactis* presents several challenges. Although Ventura and Zink were able

to differentiate strains of *B. animalis* subsp. *lactis* by sequencing the 16S-23S internal transcribed sequence (ITS) region (Ventura and Zink, 2002), analysis of the four ITS operons between DSMZ 10140 and Bl-04 indicated complete identity (Barrangou et al., 2009). However, SNPs and INDELs do have the potential for strain differentiation. According to Achtman, focusing on polymorphic SNPs is a desirable approach for the typing of monomorphic species (Achtman, 2008). Therefore, the objective of the present study was to exploit the previously identified SNP and INDEL sites to develop a technique capable of differentiating among a collection of *B. animalis* subsp. *lactis* strains obtained from culture collections and commercial starter culture companies.

Materials and Methods

Acquisition of strains

Bifidobacterium animalis subsp. *lactis* DSMZ 10140 and ATCC 27536 were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, the German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) and the American Type Culture Collection (ATCC, Manassas, VA), respectively. Twenty additional strains were obtained directly from six commercial starter culture companies (Table 1). By agreement with the suppliers, commercial strains are not identified by their trade names, but are identified by a random four-digit number preceded by "RB" (e.g., RB 0171). Isolates of *B. animalis* subsp. *lactis* Bl-04 (RB 4825), BL-01(RB 5251), Bi-07 (RB 5733), B420 (RB 7239), and HN019 were obtained from

Danisco USA, Inc. Strains Bl-04 and RB 4825 represent chronologically distinct isolates of the same strain and were analyzed separately.

Classification of strains

All strains were assayed for activity of fructose-6-phosphate phosphoketolase (F6PPK; E.C. 4.1.2.22) as described by Scardovi (Scardovi, 1981) and Biavati and Mattarelli (Biavati and Mattarelli, 2006). Genus and subspecies identification were confirmed using PCR amplification of a region of 16S rRNA gene based on the method of Kaufmann et al. (1997) and by PCR amplification of a *B. animalis* subsp. *lactis*-specific sequence in the 16S-ITS region as described by Ventura and Zink (2002), respectively. Stock cultures were prepared in Liver Lactose (LL, Lapierre et al., 1992), supplemented to 10% glycerol (v/v), and stored at -70ºC.

Phenotypic characterization

Strains were grown anaerobically in Tryptone-Phytone-Yeast Extract broth (TPY, Scardovi, 1981) until turbid, then the cells were pelleted by centrifugation. Lactic and acetic acid concentrations were determined in the supernatant by high performance liquid chromatography (HPLC) (Briczinski et al., 2008). The activities of nineteen enzymes were evaluated in strains grown on LL using api ZYM kits (bioMérieux, Inc., Durham, NC). A cell suspension was used to inoculate the test kits according to the manufacturer's instructions. Results were scored after four and a half hours of aerobic incubation at 37°C. Carbohydrate utilization was evaluated using api 50 CH kits (bioMérieux, Inc.). Commercial strains were prepared in LL or Liver Glucose (LG)

medium. ATCC and DSMZ strains were prepared in LL, LG or Reinforced Clostridial media. A cell suspension was used to inoculate api 50 CHL medium (bioMérieux, Inc.) which was then used to rehydrate the carbohydrate substrates in the test kits according to the manufacturer's instructions. Incubation was performed anaerobically at 37°C and results were scored after 48 hours. Glucose uptake assays were performed as previously described (Briczinski et al., 2008) by incubating mid-log phase (OD_{600} ~ 0.5) cells with D-[U-^{14}C]glucose.

Isolation of DNA

DNA was extracted for PCR as described by Vincent, *et al.* (1998). Briefly, overnight cultures of bifidobacteria in LL broth were harvested by centrifugation, washed, and lysed. DNA was extracted with chloroform/isoamyl alcohol (24:1) three times and precipitated with cold isopropanol. DNA concentration was estimated by measuring A_{260} with a DU-650 spectrophotometer (Beckman Coulter, Inc., Somerset, NJ).

Characterization of bifidobacteria by PFGE and RAPD-PCR

A rapid pulsed-field electrophoresis (PFGE) method (Briczinski and Roberts, 2006) was used to further characterize strains of *B. animalis* subsp. *lactis*, comparing restriction patterns of chromosomal DNA digested with *Xba*I or *Spe*I. Strain comparison by randomly amplified polymorphic DNA-PCR analyses (RAPD-PCR) was performed using seven different primers. Sequences for primers #103 (5'-GTGACGCCGC-3'), #127 (5'-ATCTGGCAGC-3'), and #173 (5'-CAGGCGGCGT-3') are from Sakata, *et al.*

(Sakata et al., 2002). Primer AB-1 (5'-GGTGCGGGAA-3') and AB-5 (5'-AACGCGCAAC-3') are from the Amersham Biosciences Ready-to-Go RAPD Analysis Beads Technical Insert. Primer sequences OPV-07 (5'-GAAGCCAGCC-3') and OPR-13 (5'-GGACGACAAG-3') are from Mayer et al. (2007).

Comparison of *B. animalis* subsp. *lactis* genome sequences

Primers were designed in the regions flanking the SNP or INDEL previously identified in the alignment of the DSMZ 10140 and Bl-04 genome sequences (Barrangou et al., 2009). Sequences surrounding the SNP or INDEL were entered into Primer3 (http://frodo.wi.mit.edu/) and primers were selected. Primer sets were evaluated using nucleotide BLAST against the DSMZ 10140 genome to ensure each primer only annealed at one position. All primers were designed with an annealing temperature of 60ºC.

Amplification mixtures (50 μL) consisted of 10 μl of 5× Colorless GoTaq® reaction buffer (Promega, Madison, WI) with a final concentration of 1.5 mM $MgCl_2$, 300 μM concentration of each deoxynucleotide triphosphate (Promega), 0.5 μM concentration of each forward and reverse primer (Supplementary Table 1; Integrated DNA Technologies, Coralville, IA), 100 ng template DNA, and GoTaq® DNA polymerase (1.5 Units). Amplifications were performed with 1 cycle of 95ºC for 5 mins, followed by 35 cycles of 95ºC for 1 min, 58ºC for 45 s, and 72ºC for 1 min, then a final cycle of 72ºC for 7 mins. Amplicons of the reaction mixtures were separated on a 1.0% agarose gel using 0.5× TBE buffer. A 100-bp DNA ladder (Promega) was included as a molecular weight marker. Electrophoresis was performed using a submerged horizontal

gel electrophoresis system at 110 V for 90 minutes. After staining in a solution of ethidium bromide for at least 1 hour, bands were visualized on a UV transilluminator and images were captured using an AlphaImager 3300 Gel Documentation System (Alpha Innotech, San Leandro, CA). Amplicons of the appropriate size were extracted from the agarose gel and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). Purified products were sequenced at the Genomic Core Facility at The Pennsylvania State University using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators; Applied Biosystems, Foster City, CA) and run on an ABI 3730XL DNA Analyzer using the ABI Data Collection Program (v 2.0). Sequences were analyzed in SeqMan version 6.0 (DNASTAR, Madison, WI) and aligned using CLUSTAL W. For quality assurance purposes, each time PCR and sequencing were performed, DSMZ 10140 and Bl-04 were included as controls.

***In silico* analyses**

All allelic sequence data resulting from the resequencing of the 47 SNP and 3 INDEL sites from 24 strains were uploaded in JMP Genomics (SAS Institute, Cary, NC). Sequences from a strain that matched the sequence from DSMZ 10140 or Bl-04 were assigned values of 1 and 3, respectively. Unique sequence from a strain that did not match either DSMZ 10140 or Bl-04 was assigned a 2. A maximum of three sequence types were observed for any locus investigated. Hierarchical two-way clustering was carried out using the fast Ward algorithm to analyze both strains and allelic profiles simultaneously and visualize polymorphism of the 24 strains across the 50 genetic loci. These analyses allow identification if the number of sequence types, clustering of strains in similar genotypes, and provide an overview of variability across genetic loci.

The discriminatory power of the method (Hunter and Gaston, 1988) with 95% confidence intervals was calculated using EpiCompare version 1.0 (Ridom GmbH, Wurzburg, Germany).

Results and Discussion

Classification of strains

All strains in the study were confirmed as *Bifidobacterium* by the presence of F6PPK activity and by PCR with genus-specific primers. Metabolism of glucose by bifidobacteria involves the action of F6PPK, producing lactic and acetic acids as the final end products in a theoretical molar ratio of 1.5:1 (Scardovi and Trovatelli, 1965). When grown in tryptone-phytone-yeast extract broth, the strains in this study exhibited production of both lactic and acetic acids (data not shown). All strains were further confirmed as *B. animalis* subsp. *lactis* by PCR using subspecies-specific primers targeting a region in the 16S-ITS (data not shown).

Phenotypic characterization

In the present study, all of the *B. animalis* subsp. *lactis* strains examined exhibited C4 esterase, leucine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-galactosidase, α-glucosidase, and β-glucosidase activities (data not shown). This is consistent with previously published work on enzymatic activities of bifidobacteria isolated from commercial products (Gueimonde et al., 2004).

With respect to carbohydrate utilization, when harvested from LL, all commercial strains of *B. animalis* subsp. *lactis* fermented D-ribose, esculin, D-maltose, D-lactose, D-melibiose, sucrose, and D-raffinose, as did DSMZ 10140 and ATCC 27536. Ten of the commercial strains, along with the two culture collection strains, also fermented glucose, while the other ten commercial strains did not (Table 1). All of the strains evaluated differed from the description of *B. animalis* subsp. *lactis* in the literature (Meile et al., 1997) because they were unable to ferment L-arabinose, D-xylose, amygdalin, gentibiose, and potassium 5-ketogluconate. According to Gueimonde, *et al*. (2004), carbohydrate fermentation patterns for *B. animalis* subsp. *lactis* strains isolated from commercial products were all positive only for glucose, esculin, and raffinose and most strains could be differentiated by utilization of at least one carbohydrate. However, this was not observed with the strains examined in this study. The results of carbohydrate assays can vary depending on the medium used to prepare the inoculum, which was also observed in this work. For example, DSMZ 10140 fermented amygdalin when cells were harvested from RC medium, but not from LL or LG media and ATCC 27536 did not ferment lactose when harvested from RC medium, but did when harvested from LL medium. It is possible differences observed among the commercial *B. animalis* subsp. *lactis* strains evaluated by Gueimonde, *et al*. and the strains evaluated in the current study are related to preparation on glucose-based (MRS) and lactose-based (LL) media, respectively.

Strains of *B. animalis* subsp. *lactis* were grown in LL and then evaluated for glucose uptake. A correlation between glucose fermentation pattern with the api CH 50 kit and glucose uptake was observed (Table 1). Strains that yielded a positive result for glucose fermentation also exhibited glucose uptake greater than 4.0 nmoles/min/mg cell

protein, while strains with a negative result for glucose fermentation exhibited glucose uptake less than 1.5 nmoles/min/mg cell protein. The results of the carbohydrate fermentation pattern and glucose uptake represented the main phenotypic difference among the strains. Based on these characteristics alone, the strains could be differentiated into two groups (Table 1).

DNA-based characterization

For differentiation of bacterial strains, PFGE has been considered the "gold standard" (O'Sullivan and Kullen, 1998, Busch and Nitschko, 1999, Biavati and Mattarelli, 2001, FAO/WHO, 2002). Therefore, PFGE was employed in an attempt to differentiate among commercial strains of *B. animalis* subsp. *lactis*. PFGE patterns (Figure 1) with *Xba*I were identical for all 22 strains examined, whereas ATCC 27536 exhibited a 2-band difference with *Spe*I (Figure 1, Lane 2), likely from the loss of a restriction site.

Homogeneity of PFGE patterns among strains of *B. animalis* subsp. *lactis* has been reported previously in the literature (Bonaparte and Reuter, 1997, Crittenden et al., 2001, Grand et al., 2003). However, different PFGE patterns were observed for two commercial strains of *B. animalis* subsp. *lactis*, HN019 and *Bifidobacterium* sp. 420, by Mayer, *et al*. (2007), although the authors acknowledged these results contradicted previous reports. One additional study reported differentiation among strains of *B. animalis* subsp. *lactis* isolated from commercial products, with four PFGE pattern types (Masco et al., 2005). The results in the current work indicate high genetic homogeneity among all commercial strains of *B. animalis* subsp. *lactis* and suggest the observation of

different patterns resulted from either instability in the *B. animalis* subsp. *lactis* genome, a more diverse strain collection, or analytical artifacts.

In an additional attempt to differentiate the strains, RAPD-PCR was performed using seven different primers. A representative gel obtained with primer 103 and the *B. animalis* subsp. *lactis* strains is shown in Figure 2. Inter- and intra-species differences were observed with RAPD-PCR primers among other species of bifidobacteria (data not shown), however, no differences were observed among the strains of *B. animalis* subsp. *lactis* with any of the seven RAPD-PCR primers evaluated. Three of the RAPD primers selected for evaluation (primers 103, 127, and 173) were included because they had been employed to differentiate among and within strains of *B. longum*, *B. infantis*, and *B. suis* (Sakata et al., 2002). Two of the RAPD-PCR primers (OPV-07 and OPR-13) were selected for evaluation because they have been shown to differentiate *B. animalis* subsp. *lactis* LAFTI™ B94 from other commercial strains of *B. animalis* subsp. *lactis* (Mayer et al., 2007). *B. animalis* subsp. *lactis* LAFTI B94 was obtained directly from the supplier and included in this study among the RB strains, however, OPV-07 and OPR-13 could not differentiate B94 from other strains in the group suggesting alternative primers are needed to differentiate these strains. It is important to note that RAPD-PCR is very sensitive and slight differences in reagents or DNA preparation may explain the results between this study and the one by Mayer, *et al*.

Differentiating among a genomically monomorphic group, such as this collection of *B. animalis* subsp. *lactis* strains, was not possible with a variety of conventional phenotypic and nucleic acid-based techniques. Although glucose uptake allowed separation of the strains into two major groups, PFGE and RAPD-PCR were ineffective.

With genome sequences for DSMZ 10140 and Bl-04 recently available (Barrangou et al., 2009), *in silico* analysis of restriction digests of DSMZ 10140 and Bl-04 with *Xba*I and *Spe*I confirmed the banding patterns observed by PFGE (data not shown). These strains did not differ in terms of the number of predicted fragments with *Xba*I and *Spe*I (29 and 24, respectively) and comparison of fragments indicated no loss or gain of restriction sites. Although the four INDELs between the two strains did result in a few fragments with different sizes, the differences were relatively too small (≤ 0.2 kb) to be discerned on the PFGE gels under the electrophoresis conditions used.

Genotypic characterization

Comparative analysis of the DSMZ 10140 and Bl-04 genomes revealed little diversity – 47 SNPs and four INDELs – between these two strains (Barrangou et al., 2009). Forty-seven SNPs and three INDELs (INDELs #1-3) were examined across all 24 of the strains in our collection, serving as a basis for development of a strain-specific typing method. Sequencing reactions for INDEL 4 failed to consistently provide unambiguous data for certain strains; therefore, this locus was omitted from the analysis.

Fourteen distinct genetic clusters were identified across the 24 strains based hierarchical clustering (Figure 3). Of those 14 strain clusters, 10 are comprised of a single strain while four contain two, three, four, or five strains that cannot be differentiated from the other strains within the group. Multistrain groups may represent different isolates of the same strain, as is known for the chronologically distinct isolates RB 4825 and Bl-04, or may differ at loci that were not examined. Perhaps not surprisingly, widely used commercial strains, such as BB-12, Bl-04, and Bi-07 (RB

5733), fall within multistrain clusters (see Figure 3, strain clusters 3, 9, 12, respectively). Other loci would need to be analyzed in order to determine if these strains are genetically different or represent different isolates of the same strain.

Ten distinct clusters of genetic loci were identified across the 50 alleles evaluated based on hierarchical clustering (Figure 3). Of these 10 clusters, 7 are comprised of a single allelic distribution across the strains. In contrast, three clusters (genetic locus clusters 6, 4, and 7 comprised of 2, 6, and 35 genetic loci, respectively) exhibit conserved allelic distributions across the strains within an individual cluster. This analysis allows selection of the minimum number and identity of genetic loci that must be evaluated for maximum genetic differentiation of this collection of strains. Accordingly, a minimum of nine genetic loci used in combination (Figure 3, one from each genetic loci clusters 1-6, 8-10) will differentiate the 24 strains in this collection into 14 strain clusters. Thirty-five of the alleles evaluated are DSMZ 10140-specific (genetic loci cluster 7 on Figure 3) and one allele, of Balat_0141, is specific for RB 0171.

In order to assess the ability of a typing scheme to separate strains, discriminatory power is often calculated (Hunter and Gaston, 1988). The discriminatory power of the allelic profiling method described was calculated to be 0.92 with a 95% confidence interval (CI) of [0.852-0.988], indicating there is a 92% chance that any two randomly selected strains will be placed into two different groups. For comparative purposes, the discriminatory power calculated for the same strain set using the glucose phenotype was a discriminatory power of 0.52 (95 % CI, 0.481-0.558).

Genotypic and phenotypic analysis of glucose uptake

A non-synonymous SNP was identified in a putative glucose uptake gene (*glcU*; DSMZ 10140 position 1,260,073). Upon sequencing, it was determined each of the strains that exhibited slow growth on glucose and low glucose uptake possess one genotype whereas the strains that exhibited normal growth on glucose and greater glucose uptake possess another genotype (Table 1). This SNP has the potential to explain the differences observed in glucose utilization and transport among these strains since the genotype is correlated to phenotype for the entire strain set. In addition, during analysis of the strain collection, a second SNP identified in the *glcU* gene (DSMZ position 1,260,380) was able to differentiate RB 1791 and RB 7239 from all other strains. Based on genotypic analysis at these two SNP sites, it is possible to break the collection into three distinct groups (Table 1), whereas it was only possible to generate two groups based on the phenotype of glucose uptake.

Conclusion

A collection of 24 strains of *B. animalis* subsp. *lactis*, including reference and commercial strains, could not be differentiated by PFGE or RAPD-PCR, which are considered to be discriminatory typing methods. The genetic dynamic range (Achtman, 2008) of this collection is defined by the distance between Bl-04 and DSMZ 10140, the two reference genomes for this subspecies. Visual analysis of the allelic distribution of the 24 strains across the 50 genetic loci clearly indicates that DSMZ (the Type strain) is the most unique strain with 35 distinct alleles (genetic locus cluster 7 on Figure 3), and

that all other strains are more similar to Bl-04. However, in addition to DSMZ 10140, it is also clear that two distinct families of strains appear to exist (Figure 3).

Overall, the SNP/INDEL analysis revealed polymorphism in the culture collection and demonstrates that a combination of only nine genetic loci needs to be analyzed to differentiate the 24 stains of this collection into 14 clusters. These nine genetic loci provide maximum genetic discrimination among the collection, which includes widely used commercial and reference strains. Although this collection of 24 isolates of *B. animalis* subsp. *lactis* exhibited a high degree of relatedness, one phenotypic difference, related to glucose uptake, was observed and correlated with the *glcU* genotype.

A functional overview of the 50 variable genetic loci (Table 2) indicates transporters and CRISPR elements are highly represented in the differential genetic content. This might indicate selective evolutionary pressure on hypervariable loci (CRISPR) and genes involved in the adaptation to the environment (transcriptional regulators, carbohydrate uptake and metabolism). This is consistent with previous reports indicating the polymorphic nature of CRISPR loci in lactic acid bacteria, notably in *Bifidobacterium* (Horvath et al., 2009).

This typing method can be used in clinical, regulatory, and commercial applications requiring identification of *B. animalis* subsp. *lactis* at the strain level. Also, it can be exploited in analyses and surveys of environmental samples.

Acknowledgements.

The authors acknowledge Allen Phillips for assistance with glucose uptake assays.

Works Cited

Achtman, M. 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62:53-70.

Barrangou, R., E. P. Briczinski, L. L. Traeger, J. R. Loquasto, M. Richards, P. Horvath, A.-C. Coûté-Monvoisin, G. Leyer, S. Rendulic, J. L. Steele, J. R. Broadbent, T. Oberg, E. G. Dudley, S. Schuster, D. A. Romero, and R. F. Roberts. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. Journal of Bacteriology 191(13):4144-4151.

Berthoud, H., F. Chavagnat, M. Haueter, and M. G. Casey. 2005. Comparison of partial gene sequences encoding a phosphoketolase for the identification of bifidobacteria. Lebensm.-Wiss. Technol. 38(1):101-105.

Biavati, B. and P. Mattarelli. 2001. The family *Bifidobacteriaceae*. in The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community. 3rd electronic ed. M. Dworkin, ed. Springer-Verlag, New York, NY.

Biavati, B. and P. Mattarelli. 2006. The family *Bifidobacteriaceae*. Pages 322-382 in The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community. Vol. 3. M. Dworkin, ed. Springer-Verlag, New York, NY.

Biavati, B., P. Mattarelli, and F. Crociani. 1992. Identification of bifidobacteria from fermented milk products. Microbiologica 15:7-13.

Bonaparte, C. and G. Reuter. 1997. Bifidobacteria in commercial dairy products: Which species are used? Microecol. Ther. 28:181-197.

Briczinski, E. P., A. T. Phillips, and R. F. Roberts. 2008. Transport of glucose by *Bifidobacterium animalis* subsp. *lactis* occurs via facilitated diffusion. Applied and Environmental Microbiology. 74(22):6941-6948.

Briczinski, E. P. and R. F. Roberts. 2006. Technical note: A rapid pulsed-field gel electrophoresis method for analysis of bifidobacteria. Journal of Dairy Science. 89(7):2424-2427.

Busch, U. and H. Nitschko. 1999. Methods for the differentiation of microorganisms. J. Chromatogr. B Biomed. Sci. Appl. 722:263-278.

Crittenden, R. G., L. F. Morris, M. L. Harvey, L. T. Tran, H. L. Mitchell, and M. J. Playne. 2001. Selection of a *Bifidobacterium* strain to complement resistant starch in a synbiotic yoghurt. J. Appl. Microbiol. 90(2):268-278.

FAO/WHO. 2002. Joint FAO/WHO Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food. London, Ontario.

Fasoli, S., M. Marzotto, L. Rizzotti, F. Rossi, F. Dellaglio, and S. Torriani. 2003. Bacterial composition of commercial probiotic products as evaluated by PCR-DGGE analysis. Int. J. Food Microbiol. 82(1):59-70.

Grand, M., M. Küffer, and A. Baumgartner. 2003. Quantitative analysis and molecular identification of bifidobacteria strains in probiotic milk products. Eur. Food Res. Technol. 217(1):90-92.

Gueimonde, M., S. Delgado, B. Mayo, P. Ruas-Madiedo, A. Margolles, and C. G. de los Reyes-Gavilán. 2004. Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. Food Res. Int. 37:839-850.

Hamilton-Miller, J. M. T., S. Shah, and J. T. Winkler. 1999. Public health issues arising from microbiological and labelling quality of foods and supplements containing probiotic microorganisms. Public Health Nutr. 2(2):223-229.

Hoffman, F. A., J. T. Heimbach, M. E. Sanders, and P. L. Hibberd. 2008. Executive summary: Scientific and regulatory challenges of development of probiotics as foods and drugs. Clin. Infect. Dis. 46(Suppl 2):S53-S57.

Horvath, P., A.-C. Coûté-Monvoisin, D. A. Romero, P. Boyaval, C. Fremaux, and R. Barrangou. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int. J. Food Microbiol. 131(1):62-70.

Hunter, P. R. and M. A. Gaston. 1988. Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. J. Clin. Microbiol. 26(11):2465-2466.

Ibrahim, S. A. and J. P. Carr. 2006. Viability of bifidobacteria in commercial yogurt products in North Carolina during refrigerated storage. Int. J. Dairy Technol. 59(4):272-277.

Iwana, H., H. Masuda, T. Fujisawa, H. Suzuki, and T. Mitsuoka. 1993. Isolation and identification of *Bifidobacterium* spp. in commercial yogurts sold in Europe. Bifidobacteria Microflora 12(1):39-45.

Jayamanne, V. S. and M. R. Adams. 2006. Determination of survival, identity and stress resistance of probiotic bifidobacteria in bio-yoghurts. Lett. Appl. Microbiol. 42:189-194.

Jian, W., L. Zhu, and X. Dong. 2001. New approach to phylogenetic analysis of the genus *Bifidobacterium* based on partial HSP60 gene sequences. Int. J. System. Evol. Microbiol. 51:1633-1638.

Kaufmann, P., A. Pfefferkorn, M. Teuber, and L. Meile. 1997. Identification and quantification of Bifidobacterium species isolated from food with genus-specific 16S rRNA-targeted probes by colony hybridization and PCR. Applied and Environmental Microbiology 63(4):1268-1273.

Kim, J. F., H. Jeong, D. S. Yu, S.-H. Choi, C.-G. Hur, M.-S. Park, S. H. Yoon, D.-W. Kim, G. E. Ji, H.-S. Park, and T. K. Oh. 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. Journal of Bacteriology. 191(2):678-679.

Klein, G., A. Pack, C. Bonaparte, and G. Reuter. 1998. Taxonomy and physiology of probiotic lactic acid bacteria. Int. J. Food Microbiol. 41:103-125.

Lapierre, L., P. Undeland, and L. J. Cox. 1992. Lithium chloride-sodium propionate agar for the enumeration of bifidobacteria in fermented dairy products. J. Dairy Sci. 75:1192-1196.

Masco, L., G. Huys, E. De Brandt, R. Temmerman, and J. Swings. 2005. Culture-dependent and culture-independent qualitative analysis of probiotic products claimed to contain bifidobacteria. Int. J. Food Microbiol. 102(2):221-230.

Mattarelli, P., G. Brandi, M. Modesto, and B. Biavati. 2002. Discrepancy between declared and recovered bifidobacteria in a human probiotic. Ann. Microbiol. 52:283-286.

Mättö, J., H.-L. Alakomi, A. Vaari, I. Virkajärvi, and M. Saarela. 2006. Influence of processing conditions on *Bifidobacterium animalis* ssp. *lactis* functionality with a special focus on acid tolerance and factors affecting it. Int. Dairy J. 16:1029-1037.

Mättö, J., E. Malinen, M.-L. Suihko, M. Alander, A. Palva, and M. Saarela. 2004. Genetic heterogeneity and functional properties of intestinal bifidobacteria. J. Appl. Microbiol. 97(3):459-470.

Mayer, A., H. Seiler, and S. Scherer. 2003. Isolation of bifidobacteria from food and human faeces and rapid identification by Fourier transform infrared spectroscopy. Ann. Microbiol. 53(3):299-213.

Mayer, H. K., E. Amtmann, E. Philippi, G. Steinegger, S. Mayrhofer, and W. Kneifel. 2007. Molecular discrimination of new isolates of *Bifidobacterium animalis* subsp. *lactis* from reference strains and commercial probiotic strains. Int. Dairy J. 17:565-573.

Meile, L., W. Ludwig, U. Rueger, C. Gut, P. Kaufmann, G. Dasen, S. Wenger, and M. Teuber. 1997. *Bifidobacterium lactis* sp. nov., a moderately oxygen tolerant species isolated from fermented milk. Syst. Appl. Microbiol. 20:57-64.

O'Brien, J., R. Crittenden, A. Ouwehand, and S. Salminen. 1999. Safety evaluation of probiotics. Trends Food Sci. Technol. 10(12):418-424.

O'Sullivan, D. J. and M. J. Kullen. 1998. Tracking of probiotic bifidobacteria in the intestine. Int. Dairy J. 8(5-6):513-525.

Prasad, J., H. Gill, J. Smart, and P. K. Gopal. 1998. Selection and characterisation of *Lactobacillus* and *Bifidobacterium* strains for use as probiotics. Int. Dairy J. 8:993-1002.

Reuter, G. 1997. Present and future of probiotics in Germany and in central Europe. Biosci. Microflora 16(2):43-51.

Roy, D. and S. Sirois. 2000. Molecular differentiation of *Bifidobacterium* species with amplified ribosomal DNA restriction analysis and alignment of short regions of the *ldh* gene. FEMS Microbiol. Lett. 191(1):17-24.

Saarela, M., G. Mogensen, R. Fondén, J. Mättö, and T. Mattila-Sandholm. 2000. Probiotic bacteria: Safety, functional and technological properties. J. Biotechnol. 84(3):197-215.

Sakata, S., M. Kitahara, M. Sakamoto, H. Hayashi, M. Fukuyama, and Y. Benno. 2002. Unification of *Bifidobacterium infantis* and *Bifidobacterium suis* as *Bifidobacterium longum*. Int. J. System. Evol. Microbiol. 52(6):1945-1951.

Scardovi, V. 1981. The genus *Bifidobacterium*. Pages 1951-1961 in The Prokaryotes. A Handbook on Habitats, Isolation, and Identification of Bacteria. Vol. 2. M. P. Starr, H. Stolip, H. G. Truper, A. Balows, and H. G. Schlegel, ed. Springer Verlag, New York, NY.

Scardovi, V. and L. D. Trovatelli. 1965. The fructose-6-phosphate shunt as peculiar pattern of hexose degradation in the genus *Bifidobacterium*. Annali di Microbiologia ed Enzimologia 15:19-29.

Ventura, M., C. Canchaya, A. Del Casale, F. Dellaglio, E. Neviani, G. F. Fitzgerald, and D. Van Sinderen. 2006. Analysis of bifidobacterial evolution using a multilocus approach. Int. J. System. Evol. Microbiol. 56(12):2783-2792.

Ventura, M., C. Canchaya, D. van Sinderen, G. F. Fitzgerald, and R. Zink. 2004. *Bifidobacterium lactis* DSM 10140: Identification of the *atp* (*atpBEFHAGDC*) Operon and Analysis of Its Genetic Structure, Characteristics, and Phylogeny. Applied and Environmental Microbiology 70(5):3110-3121.

Ventura, M. and R. Zink. 2002. Rapid Identification, Differentiation, and Proposed New Taxonomic Classification of *Bifidobacterium lactis*. Applied and Environmental Microbiology 68(12):6429-6434.

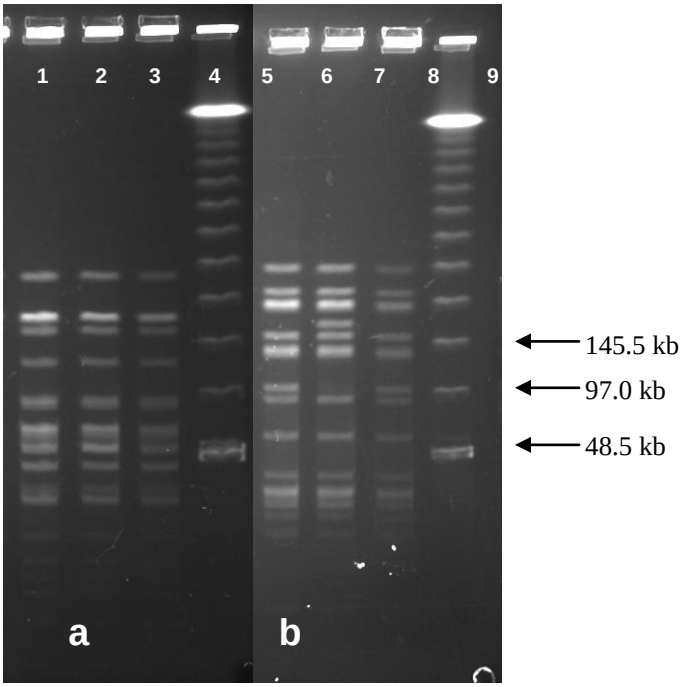
Ventura, M. and R. Zink. 2003. Comparative Sequence Analysis of the tuf and recA Genes and Restriction Fragment Length Polymorphism of the Internal Transcribed Spacer Region Sequences Supply Additional Tools for Discriminating *Bifidobacterium lactis* from *Bifidobacterium animalis*. Applied and Environmental Microbiology 69(12):7517-7522.

Vincent, D., D. Roy, F. Mondou, and C. Déry. 1998. Characterization of bifidobacteria by random DNA amplification. Int. J. Food Microbiol. 43(3):185-193.

Yaeshima, T., H. Takahashi, N. Ishibashi, and S. Shimamura. 1996. Identification of bifidobacteria from dairy products and evaluation of a microplate hybridization method. Int. J. Food Microbiol. 30(3):303-313.

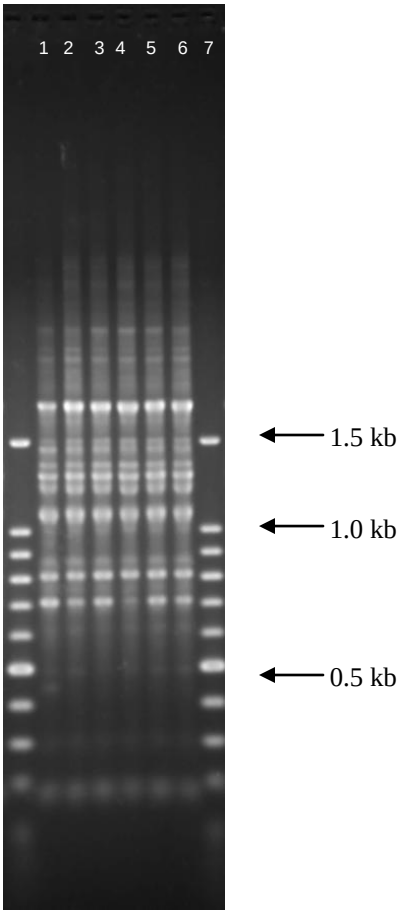
Zhu, L., W. Li, and X. Dong. 2003. Species identification of genus *Bifidobacterium* based on partial HSP60 gene sequences and proposal of *Bifidobacterium thermacidophilum* subsp. *porcinum* subsp. nov. Int. J. System. Evol. Microbiol. 53(5):1619-1623.

Figure 1. Representative PFGE of *B. animalis* subsp. *lactis* strains



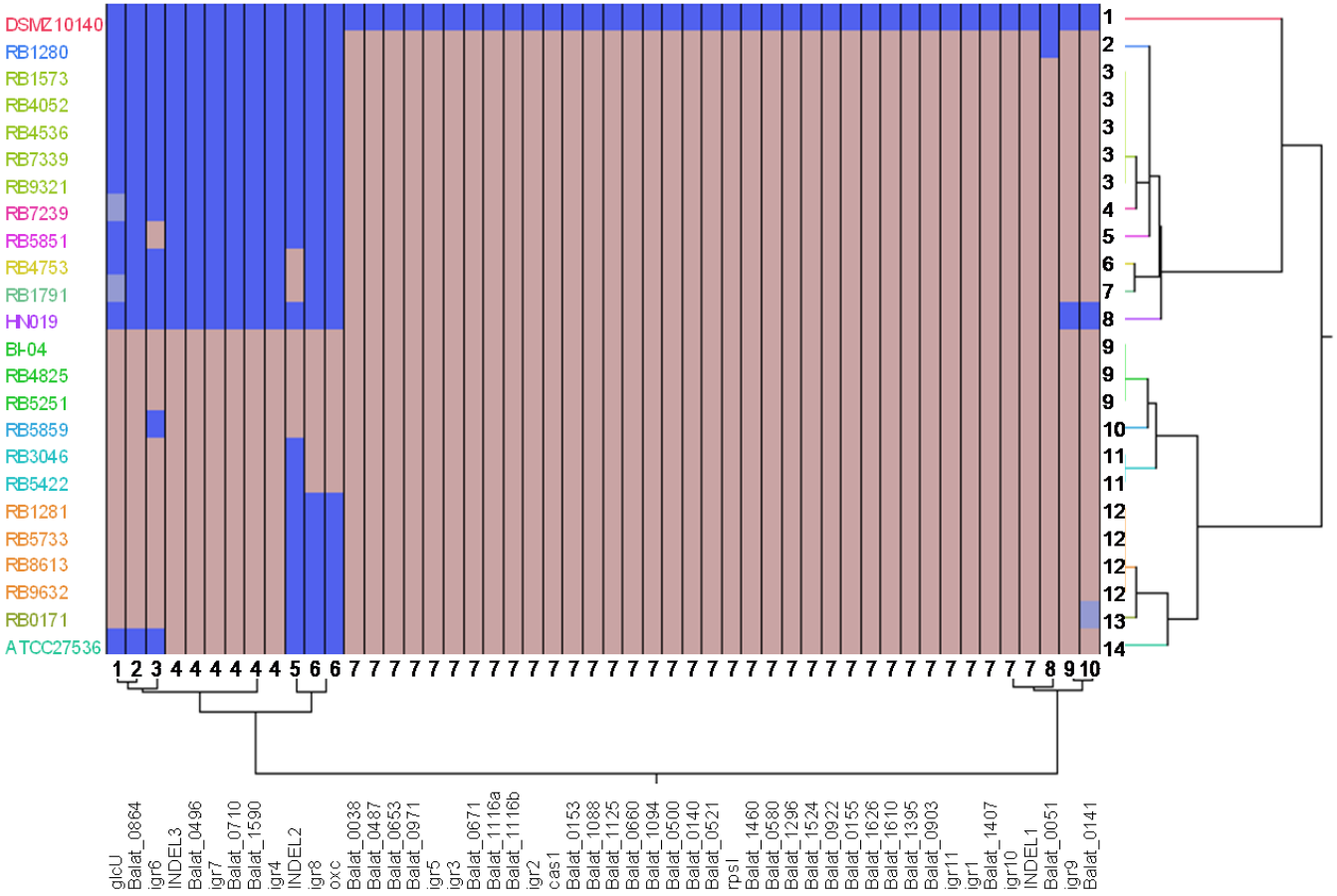
PFGE of commercial and reference bifidobacterial strains restricted using *Xba*I (panel a) and *Spe*I (panel b). Lanes 1, 8, 15: Lambda molecular weight marker. Lanes 3-7, 9-12: commercial strains of *B. animalis* subsp. *lactis*, Lane 2: *B. breve* RB 5333, Lane 13: ATCC 27536, Lane 14: DSMZ 10140.

Figure 2. Representative RAPD-PCR of *B. animalis* subsp. *lactis* strains



Representative RAPD-PCR gel with primer #103 and commercial *Bifidobacterium animalis* subsp. *lactis* strains. Lanes 1, 8, 15, 22, 28: 100-bp molecular weight marker, Lanes 2-7, 9-14, 16-21, 23-24: commercial strains, Lane 25: ATCC 27536, Lane 26: DSMZ 10140, Lane 27: negative PCR control.

Figure 3. Hierarchical cluster of *B. animalis* subsp. *lactis* strains



Hierarchical Clustering Analysis of *B. animalis* subsp. *lactis* strains across 50 genetic loci. Each row represents a strain and columns represent genetic loci. Numbers on the right indicate strain cluster and numbers on the bottom indicate genetic locus cluster.

Table 1. Comparison of glucose utilization, glucose uptake, and SNPs in the glucose uptake gene among strains of *B. animalis* subsp. *lactis.*

Strain Designation[1]	Result of glucose fermentation with commercial kits[2]	Glucose uptake[3] (nmoles/min/mg cell protein)	SNP in *glcU* (glucose uptake gene)	
			At position[4] 1,260,073	At position[4] 1,260,380
DSMZ 10140	+	4.9	C	T
ATCC 27536	+	7.3	C	T
RB 1280	+	8.3	C	T
RB 1573	+	9.0	C	T
RB 1791	+	5.1	C	C
RB 4052	+	9.7	C	T
RB 4536	+	5.8	C	T
RB 4753	+	7.0	C	T
RB 5851	+	8.1	C	T
RB 7239	+	4.1	C	C
RB 7339	+	6.9	C	T
RB 9321	+	4.8	C	T
RB 0171	-	0.8	G	T
RB 1281	-	0.5	G	T
RB 3046	-	1.0	G	T
RB 4825	-	1.4	G	T
RB 5251	-	1.0	G	T
RB 5422	-	1.1	G	T
RB 5733	-	1.0	G	T
RB 5859	-	1.1	G	T
RB 8613	-	0.7	G	T
RB 9632	-	0.5	G	T

[1] Strains beginning with an RB designation are commercial strains of *B. animalis* subsp. *lactis.*
[2] Strains were prepared on Liver Lactose (LL) media. Carbohydrate utilization was measured with api CH 50 kits after 48 hours of anaerobic incubation at 37°C. Reactions were scored numerically (0-5). Values ≥ 3 were positive (+), values < 3 were negative (-).
[3] Strains were grown in LL broth to mid-log phase. Glucose uptake assays were performed in 0.05 M potassium phosphate buffer, pH 7.5, containing 5 mM NaCl and with 1 mM glucose. Values represent the mean of duplicate assays.
[4] Position refers to numbering in the DSMZ 10140 genome sequence.

Table 2. Allelic profiles of strains of *B. animalis* subsp. *lactis* based on comparative analysis of SNPs and INDELs.

Group[1] →	**1**	**9**			**12**				**14**	**2**	**3**					**6**	**7**	**4**	**10**	**5**	**11**		**13**	**8**	**Gene Name[2]**
Genome Position in DSMZ 10140 [3] ↓	**DSMZ 10140**	**RB 4825**	**RB 5251**	**Bl-04**	**RB 1281**	**RB 5733**	**RB 8613**	**RB 9632**	**ATCC 27536**	**RB 1280**	**RB 1573**	**RB 4052**	**RB 4536**	**RB 7339**	**RB 9321**	**RB 4753**	**RB 1791**	**RB 7239**	**RB 5859**	**RB 5851**	**RB 3046**	**RB 5422**	**RB 0171**	**HN019**	
SNPs in Non-coding Regions[4]																									
19098	C	T			T				T	T	T					T	T	T	T	T	T		T	T	igr1
87537	G	A			A				A	A	A					A	A	A	A	A	A		A	A	igr2
278277	C	A			A				A	A	A					A	A	A	A	A	A		A	A	igr3
281374	T	C			C				C	*	*					*	*	*	C	*	C		C	*	igr4
525386	T	C			C				C	C	C					C	C	C	C	C	C		C	C	igr5
687000	G	A			A				*	*	*					*	*	*	*	A	A		A	*	igr6
959934	G	A			A				A	*	*					*	*	*	A	*	A		A	*	igr7
1456329	-	G			*				*	*	*					*	*	*	G	*	G		*	*	igr8
1636606	A	G			G				G	G	G					G	G	G	G	G	G		G	*	igr9
1799285	A	G			G				G	G	G					G	G	G	G	G	G		G	G	igr10
1916502	T	C			C				C	C	C					C	C	C	C	C	C		C	C	igr11
Synonymous SNPs in Protein-coding Regions [4]																									
158448	T	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_0140; ABC-type branched-chain amino acid transport systems, periplasmic component
175868	T	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_0153; LacI family transcriptional regulator
685552	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_0580; Hypothetical protein
1059543	T	A			A				A	A	A					A	A	A	A	A	A		A	A	Balat_0922; CTP synthase
1299752	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1125; ABC-type multidrug transport system permease component
1496337	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1296; ABC-type multidrug transport system, ATPase and permease components
1716099	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1460; Hypothetical protein

Group[1] →	1	9			12				14	2	3					6	7	4	10	5	11		13	8	Gene Name[2]
Genome Position in DSMZ 10140 [3] ↓	**DSMZ 10140**	**RB 4825**	**RB 5251**	**BI-04**	**RB 1281**	**RB 5733**	**RB 8613**	**RB 9632**	**ATCC 27536**	**RB 1280**	**RB 1573**	**RB 4052**	**RB 4536**	**RB 7339**	**RB 9321**	**RB 4753**	**RB 1791**	**RB 7239**	**RB 5859**	**RB 5851**	**RB 3046**	**RB 5422**	**RB 0171**	**HN019**	
1880972	T	C			C				C	*	*					*	*	*	C	*	C		C	*	Balat_1590; Transcriptional regulator
Non-Synonymous SNPs in Protein-coding Regions [4]																									
41264	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_0038; Hypothetical protein
53321	G	A			A				A	*	A					A	A	A	A	A	A		A	A	Balat_0051; Transposase
159526	G	-			-				-	-	-					-	-	-	-	-	-		--	*	Balat_0141; Acetyl-/propionyl-coenzyme A carboxylase alpha chain
177292	G	A			A				A	A	A					A	A	A	A	A	A		A	A	Balat_0155; Hypothetical protein
459192	T	C			C				C	C	C					C	C	C	C	C	C		C	C	*rpsI*; Ribosomal protein S9
581139	T	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_0487; Transcriptional regulator
588993	G	-			-				-	*	*					*	*	*	-	*	-		-	*	Balat_0496; D-alanyl-D-alanine carboxypeptidase
594228	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_0500; Dihydroneopterin aldolase
629424	T	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_0529; Oligoribonuclease
769449	C	-			-				-	-	-					-	-	-	-	-	-		-	-	Balat_0653; Uracil-DNA glycosylase
776140	C	T			T				T	T	T					T	T	T	T	T	T		T	T	Balat_0660; Sensor protein
791497	G	T			T				T	T	T					T	T	T	T	T	T		T	T	Balat_0671; ABC-type amino acid transport system periplasmic component
839429	T	C			C				C	*	*					*	*	*	C	*	C		C	*	Balat_0710; Hypothetical membrane protein
993008	T	C			C				*	*	*					*	*	*	C	*	C		C	*	Balat_0864; Hypothetical protein
1040047	-	A			A				A	A	A					A	A	A	A	A	A		A	A	Balat_0903; Long-chain-fatty-acid-CoA ligase
1108705	T	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_0971; Putative phosphoketolase
1247696	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1088; ATP-binding protein of ABC transporter
1255074	C	T			T				T	T	T					T	T	T	T	T	T		T	T	Balat_1094; Hypothetical protein
1260073	C	G			G				*	*	*					*	*	*	G	*	G		G	*	*glcU*; Putative glucose uptake permease

Group[1] →	1	9			12				14	2	3					6	7	4	10	5	11		13	8	Gene Name[2]
Genome Position in DSMZ 10140 [3] ↓	**DSMZ 10140**	**RB 4825**	**RB 5251**	**Bl-04**	**RB 1281**	**RB 5733**	**RB 8613**	**RB 9632**	**ATCC 27536**	**RB 1280**	**RB 1573**	**RB 4052**	**RB 4536**	**RB 7339**	**RB 9321**	**RB 4753**	**RB 1791**	**RB 7239**	**RB 5859**	**RB 5851**	**RB 3046**	**RB 5422**	**RB 0171**	**HN019**	
1260380	T	*			*				*	*	*					*	C	C	*	*	*		*	*	*glcU*; Putative glucose uptake permease
1287609	-	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_1116a; DNA binding protein
1287610	-	T			T				T	T	T					T	T	T	T	T	T		T	T	Balat_1116b; DNA binding protein
1510874	T	C			C				C	C	C					C	C	C	C	C	C		C	C	*cas*1; CRISPR-associated Cas1/Cas4 family protein
1635695	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1395; H(+)-stimulated manganese uptake system protein
1652028	A	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_1407; Hypothetical protein
1702641	C	-			*				*	*	*					*	*	*	-	*	-		*	*	*oxc*; Putative oxalyl-CoA decarboxylase
1792217	A	C			C				C	C	C					C	C	C	C	C	C		C	C	Balat_1524; Polysaccharide ABC transporter permease
1910309	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1610; ATP binding protein of ABC transporter for sugars
1934708	A	G			G				G	G	G					G	G	G	G	G	G		G	G	Balat_1626; Glucose-inhibited division protein B
INDELs[5]																									
881420[6]	A	P			P				P	P	P					P	P	P	P	P	P		P	P	INDEL1, tRNA-Ala-GGC
902893	P	A			*				*	*	*					A	A	*	A	*	*		*	*	INDEL2, long chain fatty acid-coA ligase
1512373[6]	A	P			P				P	*	*					*	*	*	P	*	P		P	*	INDEL3, CRISPR locus
1715507	P	A			nd				nd	nd	nd					nd	nd	nd	nd	nd	nd		nd	nd	INDEL4, Intergenic region

[1] Group refers to strain cluster number from hierarchical clustering (Figure 4).

[2] Gene Name refers to gene name from GenBank.

[3] GenBank Accession Number for DSMZ 10140 is CP001606.

[4] SNPs that are the same as DSMZ 10140 are indicated by an asterisk (*). SNPs with a different sequence are indicated by the base present (A, T, G, C) or by a gap (-). SNPs/INDELs which were not determined are indicated by "nd".

[5] The sequences for INDEL2 and INDEL4 are present (P) in the DSMZ 10140 genome. The sequences for INDEL1 and INDEL3 are absent (A) in the DSMZ genome. If an INDEL is the same as that in the DSMZ 10140 strain, it is indicated by an asterisk (*). Differences in presence or absence are indicated by P or A, respectively.

[6] Numbering corresponds to the Bl-04 (RB 4825) genome sequence; GenBank Accession Number CP001515.

Supplementary Table 1. Primers for amplification and sequencing SNPs and INDELs

Genome Position in DSMZ 10140 [1]	Gene Name[2]	Forward Primer Sequence (5'→3')	Reverse Primer Sequence (5'→3')	Product Size (bp)
SNPs in Non-coding Regions				
19098	igr1	GTTGCCACGGATACCACTG	GCGTTCGATGTCGCATTC	368
87537	igr2	CCTTGGTCTGCCTTGGTCT	GTCATTACGGGACTCGCTTT	379
278277	igr3	TAGTTGATAGCGTGCGATGC	CCCAGCCATCATCATCTCTT	432
281374	igr4	GACCTTAGTTTGTTGGATTTTGTG	ACATCCTCGCCCTCTTCC	437
525386	igr5	GGGCGATATGATTCCTCAGC	TCCACATGCGAGTAGCCAGT	879
687000	igr6	ACAGGCTCATCCGTGCTATC	AGCAGCACTGCAATCGTATG	456
959934	igr7	ATGTCCGCATGTGTTTCTGG	GCCATAGGATGCAACGTGAA	451
1456329	igr8	GTGCCTGGTGCTCATGTTC	GCATCGACAACCTCGACAC	477
1636606	igr9	CATACTGGCTGCCGACAAAC	ACGGCAATCCAAACAGCAG	751
1799285	igr10	TGGATGTCGTGGCCTAATGA	ATGCCCCAACGATTTGAGAC	535
1916502	igr11	AGAACAAGGTGGGTGATGC	GGGGGAAGACTTCTGGATCT	499
Synonymous SNPs in Protein-coding Regions				
158448	Balat_0140	ATATTTGGGACCGGCGAAC	TTCTCGCCATACACCACGTT	474
175868	Balat_0153	GGATCTTTCGATCATCTCGTG	CCACAGGAGCCCATAAACTG	391
685552	Balat_0580	TAGTCCGTACACGACCAGCA	AGATGAACCTGCCATTCACC	448
1059543	Balat_0922	GACCCATTTCACGTTGACCT	AGCATTCCGTGATGATGCT	410
1299752	Balat_1125	GCTGGTTGAACTTGGAGAGC	GGTGCGTGGCTACATCAAC	428
1496337	Balat_1296	GACGCTCTCCCCTACAAGC	AACAGGACCTCGACAAGGAA	469
1716099	Balat_1460	ATGAGCATCGGCAGGAAG	TTGCACTCTACGTGGTGCTC	454
1880972	Balat_1590	CGCAGCACAATTTGGATATG	GGAAGAAGTCCAAGCCTCCT	452
Non-Synonymous SNPs in Protein-coding Regions				
41264	Balat_0038	AGGCAGAGTTGGAGTTGCAG	CTAGCGTATTGGCGATGGTC	390
53321	Balat_0051	CCTACCGTCATCGAGTCCAG	CCTTCATCATGCCCTCCTTA	284

Genome Position in DSMZ 10140 [1]	Gene Name[2]	Forward Primer Sequence (5'→3')	Reverse Primer Sequence (5'→3')	Product Size (bp)
159526	Balat_0141	ACTACGAAAGCGCGGGAAC	CTTGTCGAAGGGCACCACA	480
177292	Balat_0155	AAGATTCCCGCGAAAGACAA	TGTCGAGTTGCTGGGTGTTC	591
459192	*rpsI*	AGGCAAAGTGAACCGAAGGA	TCCCATCGAACGTGTGGAC	703
581139	Balat_0487	TGCCGGTAGATGGAGTTGTG	GGCGTGTCGAATATGATGGA	785
588993	Balat_0496	GGTATGGCTTGGCGGTAAGA	CTGGGTTGGCAGTGTGTCTC	581
594228	Balat_0500	ATTCCGGAAGCGTTGAAGAA	GACCGTTGCACTGAGCTGTT	430
629424	Balat_0529	CACGATGAACTGTGCGAGGT	GGTTCCTTGATTTCCGTCCA	670
769449	Balat_0653	GAAGGGATGTGATCCGAAGA	CTATCCCACTCCAGGCTACG	428
776140	Balat_0660	ACAACACCAACACCGGTTCA	TGAGAGCAGATCCTCGACCA	758
791497	Balat_0671	TCGGCATCAAATACGACCAG	GAACGCACCGGGAATGTTAT	772
839429	Balat_0710	CGTGTGGTCGATATCGCTGT	AGGCGGATTTCACAGTGGTC	641
993008	Balat_0864	GGTTCGTCACTTTCGTGCAG	CCACACCACGCGTGTATTTC	537
1040047	Balat_0903	AGGAGTGCCATGTCCGTTTT	AGCCGTAGGCCTTCAACTCA	710
1108705	Balat_0971	GGGACAACGGATACCTCTCG	TGCAGCTTGATGAGGTCCAC	680
1247696	Balat_1088	CCGTCGAAGGTGTAGGTGAT	AGGTGCTTGAATTGCTCAGG	445
1255074	Balat_1094	CGAGGAACCTTCCAACGAAC	CTATGCCGCCAGAGAGAACA	715
1260073	*glcU*	ACGTAGCGGTGGAGGTCCTA	CCGGTCTTGTGCAGAAGGTT	636
1260380	*glcU*	ACGTAGCGGTGGAGGTCCTA	CCGGTCTTGTGCAGAAGGTT	636
1287609	Balat_1116a	GGTTCTGCGGTGACGTATGA	AACGTCTTGTGATCGCGTGT	790
1287610	Balat_1116b	GGTTCTGCGGTGACGTATGA	AACGTCTTGTGATCGCGTGT	790
1510874	*cas*1	ACCCGTGCGTAGGAGTATGG	CATTGGGCCCATAGGAAGAA	584
1635695	Balat_1395	TGATCGGCCTACTGCTCGTA	GCTTTGCCCATTAGGGTGAG	775
1652028	Balat_1407	ATCGCGTATGCCTTTCTCGT	CCGTGCGTTTCAACAACAGT	590
1702641	*oxc*	ACAGCAAGCGGAATGTCTTC	GGAGGATGTTTGCAATGGTT	478
1792217	Balat_1524	AGGCGAACGAAAAAGCATGT	AGCAGCACGACTGGGTAGGT	355

Genome Position in DSMZ 10140 [1]	Gene Name[2]	Forward Primer Sequence (5'→3')	Reverse Primer Sequence (5'→3')	Product Size (bp)
1910309	Balat_1610	GTTCGTCGCAGGCTTCATC	ATGCCCTACCCATCCTTGTG	654
1934708	Balat_1626	GTTTGGCAGGGTGGTTAGGA	AGACATGGCCGCTACTGCTT	782
INDELs				
8814203[3]	INDEL1, tRNA-Ala-GGC	TGATACGCAGAACACGGATT	GCCATGATGTTCCTTTCGTC	na[4]
902893	INDEL2, long chain fatty acid-coA ligase	AACCGTCTGCTGCTGTTTCT	CCCCTGAATGAAGGTGATGT	na[4]
1512373[3]	INDEL3, CRISPR locus	TTGGATGCAAGCCCTCAATGAAGC	TGAGGGAAGCCGAACTCAATCACA	na[4]

[1] GenBank Accession Number for DSMZ 10140 is CP001606.

[2] Gene Name refers to gene name from GenBank.

[3] Numbering corresponds to the Bl-04 genome sequence; GenBank Accession Number CP001515.

[4] na: Not applicable. Product size varies depending on presence or absence of the INDEL.

Vita- Joseph R. Loquasto

Education

Ph. D., Food Science- The Pennsylvania State University — December 2013
B. S., Food Science- The Pennsylvania State University — May 2007

Publications

Loquasto JR, Barrangou R, Stahl B, Dudley EG, Roberts RF. *Bifidobacterium animalis* subsp. *lactis* ATCC 27673 is a genomically unique strain within this conserved subspecies. (under review, Applied and Environmental Microbiology).

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). Journal of Dairy Science, 94 (23): 5864-70.

Harwood ML**, Loquasto JR**, Roberts RF, Ziegler GR, Hayes JE. 2013. Explaining Tolerance for Bitterness in Chocolate Ice Cream Using Solid Chocolate Preferences. Journal of Dairy Science. 2013 Aug;96(8): 4938-44

Briczinski EP, **Loquasto JR**, Barrangou R, Dudley EG, Roberts AM, Roberts RF. (2009). Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* using single-nucleotide polymorphisms, insertion and deletions. Applied and Environmental Microbiology. 75(23): 7501-7508.

Barrangou R, Briczinski EP, Traeger LL, **Loquasto JR**, Richards M, Horvath P, Coûté-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. Journal of Bacteriology. 191 (13): 4144-4151

Select Awards

USDA-AFRI Teaching in Microbial Genomics Fellowship — 2009
National Milk Producers Federation National Dairy Leadership Scholarship — 2010
Star Kay White Scholarship in Food Science — 2011

Select Experience

Food Science 414 (Dairy Science and Technology), Laboratory Instructor — 2012
Food Science 414 (Dairy Science and Technology), Teaching Assistant — 2008, 2009
Ice Cream Short Course, Teaching Assistant, Lecturer — 2006-2013
Cultured Products Short Course, Teaching Assistant — 2005-2012
Ice Cream 101, Teaching Assistant — 2006- 2013
Graduate Programs Committee — September 2010- present
Department of Food Science Head Search Committee — August 2012 – February 2013
Graduate Student Representative — 2008-2009