Genomics as a means to understand bacterial phylogeny and ecological adaptation: The case of bifidobacteria

ARTICLE *in* ANTONIE VAN LEEUWENHOEK · JUNE 2007

Impact Factor: 1.81 · DOI: 10.1007/s10482-006-9122-6 · Source: PubMed

5 AUTHORS, INCLUDING:

Marco Ventura
Università degli studi di Parma
129 PUBLICATIONS **4,376** CITATIONS
SEE PROFILE

Carlos A. Canchaya
University of Vigo
43 PUBLICATIONS **3,515** CITATIONS
SEE PROFILE

Radhey Gupta
McMaster University
315 PUBLICATIONS **10,948** CITATIONS
SEE PROFILE

Douwe Van Sinderen
University College Cork
315 PUBLICATIONS **11,187** CITATIONS
SEE PROFILE

Antonie van Leeuwenhoek (2007) 91:351–372
DOI 10.1007/s10482-006-9122-6

ORIGINAL PAPER

Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria

Marco Ventura · Carlos Canchaya · Gerald F. Fitzgerald · Radhey S. Gupta · Douwe van Sinderen

Received: 6 July 2006 / Accepted: 28 September 2006 / Published online: 28 October 2006

Abstract The field of microbiology has in recent years been transformed by the ever increasing number of publicly available whole-genome sequences. This sequence information has significantly enhanced our understanding of the physiology, genetics and evolutionary development of bacteria. Among the latter group of microorganisms, bifidobacteria represent important human commensals because of their perceived contribution to maintaining a balanced gastrointestinal tract microbiota. In recent years bifidobacteria have drawn much scientific attention because of their use as live bacteria in numerous food preparations with various health-related claims. For this reason, these bacteria constitute a growing area of interest with respect to genomics, molecular biology and genetics. Recent genome sequencing of a number of bifidobacterial species has allowed access to the complete genetic make-up of these bacteria. In this review we will discuss how genomic data has allowed us to understand bifidobacterial evolution, while also revealing genetic functions that explains their presence in the particular ecological environment of the gastrointestinal tract.

Keywords Genomics · Bifidobacteria · Prebiotic properties · Bifidobacterial taxonomy

M. Ventura (✉)
Department of Genetics, Anthropology and Evolution, University of Parma, Parco Area delle Scienze 11a, Parma 43100, Italy
e-mail: marco.ventura@unipr.it

C. Canchaya · G. F. Fitzgerald · D. van Sinderen
Alimentary Pharmabiotic Centre and Department of Microbiology, BioSciences Institute, National University of Ireland, Cork, Ireland

R. S. Gupta
Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Canada

Introduction

Whole-genome nucleotide sequencing has revolutionized the genetic, biochemical and molecular biology research on bacteria and, indeed, many higher organisms. Over the past decade, the sequences of more than 300 bacterial genomes have become available in the public domain. Considerable emphasis was initially placed on sequence determination of the genomes of pathogenic bacteria, including food-borne pathogens (NCBI source). However, in recent years, genome sequencing of gut commensals and symbionts, as well as food-grade bacteria, has become more established among sequencing projects, currently represented by genome sequences of more than 30 lactic acid bacteria (LAB) and bifidobacteria (Klaenhammer et al. 2005). Members of the

latter group encompass gastrointestinal tract (GIT) commensals with probiotic properties (for reviews see (Lievin et al. 2000; Ouwehand et al. 2002; Servin 2004; Ventura et al. 2004c). Bifidobacterial species are found in the GIT of mammals with the notable exception of species isolated from the human oral cavity (*Bifidobacterium dentium*), sewage (*Bifidobacterium minimum* and *Bifidobacterium subtile*), or the insect gut (*Bifidobacterium asteroides*, *Bifidobacterium indicum* and *Bifidobacterium coryneforme*). Bifidobacteria are representatives of the high G + C Gram positive group of bacteria, belonging to the *Actinobacteria* phylum, within which they form a distinct order, ''*Bifidobacteriales*''. This order is comprised of a single family *Bifidobacteriaceae*, which in turn consists of four genera, *Bifidobacterium*, *Gardnerella, Scardovia* and *Parascardovia* (Biavati and Mattarelli 2001; Garrity and Holt 2001; Stackebrandt and Schumann 2000). Except for the *Bifidobacterium* genus, which contains 29 species, the other genera each contain just a single species. In phylogenetic trees based on 16S rRNA as well as protein sequences (e.g. CTP synthetase), bifidobacteria form a deep branching lineage within the *Actinobacteria* phylum (Gao and Gupta 2005; Stackebrandt and Schumann 2000). The placement of bifidobacteria within *Actinobacteria* and their deep branching within this phylum is now strongly supported by comparative genomic studies (to be discussed later) (Gao and Gupta 2005, 2006). Bifidobacteria are predicted to constitute less than 10% of the total intestinal microbiota, although the true contribution of bifidobacteria to the intestinal microbiota may vary between individuals and appears to be dependent on age and diet (Favier et al. 2002; Lay et al. 2005; Vaughan et al. 2005; Zoetendal et al. 2006). Bifidobacteria can ferment a wide variety of oligosaccharides, some of which are not digested by their host and have thus found application in enhancing bifidobacterial numbers in situ (the prebiotic concept; Gibson 1999). Here we review how genomic information may be used to gain insights into bifidobacterial physiology and phylogeny, as well as to understand how these commensals have adapted to their specific ecological niche.

Genome evolution

The availability of complete genome sequences of closely related bacteria presents an exciting opportunity to reconstruct events of genome evolution. It has become clear that, in addition to nucleotide substitutions, other genetic events shape the genome. The principal forces that drive genome evolution include gene duplication, horizontal gene transfer, gene loss and chromosomal rearrangements. Gene duplication is considered a central mechanistic precursor of gene innovation that has promoted adaptation to new environments and exploiting new ecological niches. In the pre-genomics era it was thought that bacterial genomes had originated from a much smaller genome ancestor through numerous gene duplication events (Kunisawa 1995). However, an analysis based on the currently available bacterial genome data does not support this theory and shows that gene duplication is a phenomenon that contributes only in a modest way to genome evolution (Gevers et al. 2004). Besides gene duplication and ensuing functional divergence, bacteria may employ an alternative mechanism for genetic adaptation to their environment. The introduction of novel genes, in this context also called alien genes, by HGT allows for niche-specific adaptation, which might lead to bacterial diversification and speciation (Cohan 2001). The current concept of evolution is based on the interplay between vertically generated and horizontally acquired variation (Woese 2000). Taking this concept to its extreme, it has been suggested that two bacterial groups are more similar than a third one, not because they share a more recent ancestor but because they exchange genes more frequently (Gogarten et al. 2002). As claimed by Philippe and Douady (2003), three categories of genes have been distinguished based on their sensitivity to HGT: (i) a hard core of genes, which are recalcitrant to HGT; (ii) a soft core which includes genes that are not commonly susceptible to HGT; and (iii) the shell genes which are easily affected by HGT. The latter gene category encompasses genes found on plasmids, bacteriophages, genomic islands and transposable elements, such as insertion sequences (IS), transposons (Tn) and integrons. HGT events are

held responsible for the introduction of a variety of functions in bacterial genomes and have consequently enhanced the competitiveness of bacteria in their natural environments. Hence, genes encoding sugar transporters and carbohydrate hydrolases can represent a large portion of strain-specific genes that have been acquired by HGT. It has been suggested that selected genes involved in sugar metabolism as well as in the production of exopolysaccharides in *Bifidobacterium longum* biotype *longum* NCC2705, have been acquired via HGT, as part of the adaptation of this organism to a specific ecological niche. Moreover, it has been postulated that in this strain a region encoding for rhamnosyl transferases has been acquired through HGT from streptococci (Schell et al. 2002). Apart from these, two other regions containing genes encoding two different restriction modification systems appear to have been acquired by this strain through HGT.

The identification of HGT events is not an easy task, and is based on the divergence in the G + C content (GC skew), and/or dinucleotide differences and/or codon usage. Simple estimations of HGT frequency point out that it is low (below 10% of the total gene complement; Kunin and Ouzounis 2003; Snel et al. 2002). In the case of *B. longum* biotype *longum* NCC2705 less than 5% of its genome content seems to have been acquired by recent and therefore recognizable HGT (Schell et al. 2002).

Besides the events that influence variation in gene content, as described above, genome rearrangement is a force which is influencing genome organization. Gene order in prokaryotes is mantained to a much smaller degree than average protein sequence similarity (Wolf et al. 2001). These findings suggest that the relative gene position is not essential for gene function. Moreover, it has been found that the degree of genome reorganization constantly increases beside the time of divergence, and thus gene order conservation can be used as a phylogenetic measure to investigate bacterial relationships (Korbel et al. 2002).

Genomes displaying an elevated repeat density have higher rates of reorganizations leading to accelerated loss of gene order as a result of homologous recombination (Rocha 2003). In addition, chromosome development is influenced by large genome reshufflings, especially large inversions, which lead to the phenomenon of X-shaped patterns in the alignments of whole genomes. The most likely mechanism of generation of X-alignments involves large chromosomal inversions that reverse the genomic sequence symmetrically around the origin of replication. The finding of these X-alignments between many pairs of species suggests that chromosomal inversions around the origin are a common feature of bacterial genome evolution (Eisen et al. 2000). Remarkably, inversions that become fixed in natural populations tend to be symmetrical around the origin of replication or replication terminus, thereby reducing the disruption of the chromosomal structure relative to replication. Non-symmetrical large inversions are not often detected in natural strains, which indicate that they reduce the fitness of the bacterium in which they have occurred. A feature that could lead to the reduced impact of such chromosome rearrangements is the physical distance of the site of inversion from the origin of replication, as this determines the relative copy number of the gene in each cell of fast growing cultures of bacteria (Tillier and Collins 2000). Genes such as rRNA operons or insertion sequences (IS), which can be present in multiple identical copies, can easily produce homologous recombination. Inspection of the *B. longum* biotype *longum* NCC2705 genome revealed 16 intact IS elements, eight of which constitute a new member of five known families, including one rare IS607-type (Schell et al. 2002). Thus, such elements may be involved in chromosomal rearrangements in bifidobacteria in a manner similar to that previously described in *Lactococcus lactis*, the model LAB, where a major chromosomal inversion covering half of the chromosome was reported in two closely related strains that had probably arisen through recombination between two IS elements (Daveran-Mingot et al. 1998).

Extrachromosomal DNA elements

In contrast to the important advances obtained in bifidobacterial ecology, the lack of genetic tools has hampered molecular biological investigations

of members of this genus as compared to closely related bacteria such as *Streptomyces* or *Mycobacterium* species. The characterization of bifidobacterial plasmids might therefore be extremely useful in the development of a series of molecular tools to genetically manipulate these organisms. Generally, plasmids are rarely found in bifidobacterial species. So far, only small plasmids have been identified in a minority of tested bifidobacteria, such as *Bifidobacterium breve* (O'Riordan and Fitzgerald 1999; Park et al. 1997; Rossi et al. 1998; Sgorbati et al. 1982), *B. longum* biotype *longum* (Corneau et al. 2004), *B. asteroides* (Sgorbati et al. 1982), *B. indicum* (Sgorbati et al. 1982), *Bifidobacterium pseudocatenulatum* (Sgorbati et al. 1982) and *Bifidobacterium bifidum* (Yildirim et al. 1999). Most of these plasmids are cryptic, with the exception of a *B. bifidum* plasmid that was proposed to encode a bacteriocin, named bifidocin B (Yildirim et al. 1999). Plasmids pCIBb1 isolated from *B. breve* (O'Riordan and Fitzgerald 1999) and p4M from *B. pseudocatenulatum* (Sgorbati et al. 1982), seem to either possess genes, e.g. *rep*B, *tra*A and *mob* genes, whose protein products are indispensable for rolling-circle replication (RCR) and mobilization, or alternatively encode putative relaxases. Strikingly, the pCIBb1 and p4M plasmids encode a septal DNA translocator protein of the FtsK/SpoIIIE family, homologous to the *Streptomyces* Tra proteins (Park et al. 1997), whereas the other plasmids mentioned above all encode putative relaxases (Grohmann et al. 2003). This suggests that two different modes of plasmid segregation exist in bifidobacteria, one which employs a classical transfer system involving a relaxase and a single-stranded plasmid intermediate, while the other functions through a *Streptomyces*-type transfer system involving a septal DNA translocator protein (Grohmann et al. 2003).

A *B. longum* biotype *longum* plasmid, pMB1, was used to construct recombinant plasmids (Rossi et al. 1998) and the first generation of bifidobacterial cloning vectors (Missich et al. 1994; Rossi et al. 1998).

Recently, the complete sequence of two plasmids, pDOJH10L and pDOJH10S, isolated from *B. longum* biotype *longum* DJO10A has been described (Lee and O'Sullivan 2006). Interestingly, pDOJH10S does not contain any features consistent with RCR, and instead contain sequences homologous to replication functions of plasmids that replicate via the theta-type replication mode. Notably, based on the analysis of gene and G + C content of this plasmid, Lee and O'Sullivan (2006), speculated that pDOJH10S has been acquired through HGT from another *Actinobacteridae* member, possibly *Rhodococcus rhodochrous*. The pDOJH10S plasmid was used as the basis for a shuttle cloning vector and was transformed successfully into *Escherichia coli* and *B. longum* species, but could not be introduced into LAB such as *L. lactis* and *Lactobacillus delbrueckii* subsp. *bulgaricus*, which suggests that this plasmid displays a narrow host range, consistent with previous replication studies on cryptic bifidobacterial plasmids (Matsumura et al. 1997; Park et al. 1997; Rossi et al. 1998).

Prophage and putative bacteriophage in bifidobacteria as agents of HGT

Bifidobacteria have long been considered free from phage infection. Only very recently, the identification of three prophage-like elements that are present in the genomes of *B. breve* UCC2003, *B. longum* biotype *longum* NCC2705 and *B. longum* biotype *longum* DJO10A, designated Bbr-1, Bl-1 and Blj-1, respectively, has changed this perception (Ventura et al. 2005b). These prophage like-elements exhibit homology with genes of double-stranded DNA (dsDNA) bacteriophages that infect a broad phylogenetic range of bacteria. Surprisingly, using a novel method to investigate the evolutionary development of bacteriophages, i.e. the proteomic tree (Rohwer and Edwards 2002), it was shown that the Bbr-1, Bl-1 and Blj-1 prophage like-elements exhibit a close phylogenetic relationship with phages infecting low-G + C bacteria (e.g., lactococcal and staphylococcal phages; Ventura et al. 2005b). This suggests that these bacteria have shared the same ecological niche (e.g. the animal GIT) during their evolution and thus exchanged sequences between members of these groups. This may therefore represent evidence pointing to an ancient exchange of DNA sequences between

low and high G + C bacteria. The clustering of these prophage-like elements found in bifidobacteria may also be explained by assuming that the latter's originally infecting the ancestor of high G + C Gram positive bacteria, which is in agreement with the argument for an evolutionary origin of high G + C Gram positive bacteria from low G + C ancestors (Gupta 2001).

All three prophage-like elements are integrated in a $tRNA_{Met}$ gene, which had not previously been shown to act as an *att*B site in Gram-positive bacteria (Campbell 1992). Analysis of the distribution of this integration site in many bifidobacterial species revealed that the *att*B sites are well conserved. This analysis highlighted positions where sequence variation may be tolerated in sites that are active for recombination. The use of these conserved sequences and the *int* gene sequences of bifidophages might allow the construction of an efficient recombination module, analogous to the *Streptomyces* integrating plasmid pSE211 (Brown et al. 1990). Such a recombination module may represent the ideal source for the construction of vector integration systems that enable the food-grade introduction of foreign DNA sequences in single copy at a specific site within the host chromosome without disturbing any host function, similar to systems developed for lactic acid bacteria, such as *Lactobacillus* (Martin et al. 2000), and for high G + C content bacteria such as *Streptomyces* and *Mycobacterium* (Combes et al. 2002).

The Blj-1 prophage is 36.9-kb long and is excised from the chromosome when a *B. longum* biotype *longum* DJO10A culture is exposed to mitomycin C or hydrogen peroxide (Ventura et al. 2005b). Thus, Blj-1 appears to constitute the first reported inducible and molecularly characterized prophage within the *Bifidobacterium* genus, thereby presenting possibilities for further studies on the biology of bifidophages. In contrast, the Bbr-1 and Bl-1 prophage-like elements appear to represent two deficient bacteriophages. However, they may constitute functional satellite phages, whose mobility depends on helper phages in a manner similar to that described for the cryptic mycophages Rv1 and Rv2 (Hendrix et al. 1999).

The Bbr-1 prophage-like element appears to consist of a 28.5-kb chimeric DNA fragment composed of a composite mobile element inserted into prophage-like sequences, which do not appear to be widely distributed among *B. breve* strains. Finally, these prophage-like sequences contain a gene that is involved in sugar metabolism (Ventura et al. 2005b).

Bifidobacterial genomics: metabolism and biosynthetic abilities

The first complete genome of a bifidobacterial strain, *B. longum* biotype *longum* NCC2705, was made public in 2002 (Schell et al. 2002). Recently, the partly assembled genomes of two other bifidobacterial strains or species (viz. *B. longum* biotype *longum* DJO10A and *Bifidobacterium adolescentis* ATCC 15703) have become available in the NCBI database. Furthermore, the genomes of additional bifidobacterial species/strains [*B. breve* UCC2003, *B. breve* M-16V, *B. breve* Yacult, *Bifidobacterium animalis* subsp. *lactis*, *B. longum* biotype *longum*, *B. longum* biotype *infantis* and *B. dentium* Bd1 (Liu et al. 2005); our own unpublished results] are at various degrees of completion and detailed sequence information for some of these genomes is expected to become publicly available in the near future. These genomes range in size from 1.9 to 2.9 Mb and possess a large number of tRNA molecules (e.g. from 54 to 58), while displaying common architectural features of a typical bacterial chromosome (in cases where this could be determined), such as the co-orientation between gene transcription and DNA replication (McLean et al. 1998) and the asymmetric bias in the nucleotide composition of the leading DNA strand (Frank and Lobry 1999), which is generally enriched in guanines and depleted in cytosines. Several of these genomes are derived from different strains within the same species, i.e. *B. breve* and *B. longum*. Comparative studies on these genomes should prove extremely helpful in identifying strain-specific DNA sequences that may have been recently evolved or acquired, and may consequently be responsible for adaptations of these strains to their specific niches/lifestyle. Similar comparative studies using all available bifidobacterial genomic information is expected

to provide valuable information regarding genetic characteristics that are shared by bifidobacterial species and that distinguish them from other bacteria.

Genome sequencing of a probiotic strain also constitutes an essential step to generate primary information for downstream comparative and functional purposes, such as comparative genomics, transcriptomics and/or proteomics, which in turn can be utilized to address many fundamental and applied questions that need to be answered before the probiotic concept can be fully accepted. So far, the only available genome-based information relating to biosynthetic abilities of bifidobacteria is limited to *B. longum* biotype *longum* NCC2705 strain (Schell et al. 2002). This genome was reported to harbor genes for the synthesis of at least 19 amino acids from ammonia and a number of biosynthetic precursors (phosphoenolpyruvate, oxaloacetate, oxoglutarate, and fumarate), the latter provided by a partially present Krebs cycle, which lacks fumarase, oxoglutarate dehydrogenase and malate dehydrogenase (Schell et al. 2002). The ability for cysteine biosynthesis is not fully clear in bifidobacteria, as genes for the sulfate/sulfite assimilation pathway are not present in *B. longum* biotype *longum* NCC2705 (Schell et al. 2002). This means that this biosynthetic pathway is intact but that a reduced sulfur source is required (Grundy and Henkin 1998). This reduced sulfur compound may be provided in the form of hydrogen sulfide, which is produced by sulfur-reducing colonic microbiota, suggesting a synergistic relationship. Of note is that *B. longum* biotype *longum* NCC2705 appears to use the gatABC/asparaginyl-tRNA-dependent route to produce asparaginyl-tRNA using an aspartate-dependent route (Min et al. 2002). Furthermore, this strain possesses all of the enzymes needed for the biosynthesis of pyrimidine and purine nucleotides from glutamine as well as those needed for the synthesis of folic acid, thiamine, and nicotinate (Schell et al. 2002). In contrast, pathways needed for the *de novo* biosynthesis of riboflavin, biotin, cobalamin, pantothenate, lipoate and pyridoxine appear to be partially or completely absent (Schell et al. 2002).

The biosynthetic capabilities seem to indicate that *Bifidobacterium* species have adapted to an environment where they cannot rely on an extraneous source of amino acids, nucleotides and certain vitamins. Preliminary comparative genome studies complemented by the development of defined synthetic media for other bifidobacteria, such as *B. breve* UCC2003, have confirmed many of these prototrophic and auxotrophic characteristics (O'Connell-Motherway, Forde, Fitzgerald and van Sinderen, unpublished results). Similar to *B. longum* biotype *longum* NCC2705, *B. breve* UCC2003 has all the homologues needed for the synthesis of pyrimidines and purine nucleotides from glutamine.

With respect to the carbohydrate fermentation capabilities of *B. longum* biotype *longum* NCC2705, genome analysis has indicated that this strain possesses all the gene homologs to encode enzymes needed to feed fructose, galactose, N-Ac-glucosamine, N-Ac-galactosamine, arabinose, xylose, ribose, sucrose, lactose, cellobiose, melibiose into the fructose-6-phosphate shunt (Schell et al. 2002). It is notable that of all genes contained in the genome of *B. longum* biotype *longum* NCC2705, nearly 10 % encode for proteins that are predicted to be involved in carbohydrate metabolism and transport (Table 1). Such an extensive genetic adaptation to carbohydrate metabolism is shared by other intestinal commensal bacteria, i.e. *Bacteroides* (Table 1), and is likely to represent a specific adaptation to living in the GIT (see below).

Moreover, cell cycle control and cell division COGs are over-represented in bifidobacteria compared to other bacteria (Table 1). These include mainly proteins such as those involved in cell septation (Fic, FtsE, FtsK, FtsX, FtsZ) and in chromosome partioning (e.g. ParA and Smc). However, the significance of these findings is still not clear; although it is possible that it signifies a link to the ability of coping with certain types of environmental stresses (e.g. osmotic stress). Moreover, a large proportion of the analyzed bifidobacterial genomes is dedicated to transcriptional regulation. In the case of *B. longum* biotype *longum* NCC2705 almost 70% of its predicted transcriptional regulators appear to be repressors (Schell et al. 2002). On that basis it has been speculated that repressors allow a quicker and more stringent response to environmental

Table 1 COGs functional annotation of genomes from bifidobacteria and other human intestinal bacteria

COG description	% in Bifidobacterium*	% in Bacteroides*	% in Actinobacteria*	% in Bacteria*
Translation	6.61	3.35	4.13	3.65
RNA processing and modification	0.05	0.00	0.02	0.01
Transcription	6.81	3.31	7.06	4.94
Replication, recombination and repair	5.58	4.09	4.63	4.42
Chromatin structure and dynamics	0.00	0.00	0.01	0.01
Cell cycle control, mitosis and meiosis	1.32	0.72	0.83	0.69
Defense mechanisms	2.30	1.70	1.40	1.15
Signal transduction mechanisms	3.62	3.27	3.48	3.44
Cell wall/membrane biogenesis	3.62	6.35	3.13	4.10
Cell motility	0.78	0.28	0.43	1.37
Cytoskeleton	0.05	0.00	0.01	0.01
Extracellular structures	0.19	0.00	0.02	0.02
Intracellular trafficking and secretion	1.37	1.48	0.93	1.72
Posttranslational modification, protein turnover, chaperones	2.94	2.14	2.48	2.61
Energy production and conversion	2.64	3.99	4.53	4.37
Carbohydrate transport and metabolism	9.79	5.65	4.23	3.48
Amino acid transport and metabolism	9.79	4.39	7.14	5.90
Nucleotide transport and metabolism	3.28	1.63	1.89	1.63
Coenzyme transport and metabolism	2.59	3.04	3.13	2.65
Lipid transport and metabolism	2.01	1.69	4.07	2.66
Inorganic ion transport and metabolism	4.85	4.06	4.74	3.62
Secondary metabolites biosynthesis, transport and catabolism	0.54	0.89	3.64	1.79
General function prediction only	11.56	7.58	11.13	8.19
Function unknown	4.80	3.60	5.06	4.81
Not in COGs‡	12.88	38.61	20.41	32.68

* These data were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/)

‡ Not in COGs category includes mainly hypothetical proteins

change, which is consistent with the need to adapt to rapidly fluctuating substrate availability in the GIT (Schell et al. 2002).

The bifidobacterial genome content reflects adaptation to the intestinal niche

As mentioned above, genome sequencing provides the gold standard for examining the full genetic complement of an organism, and gives predictive clues regarding the genetic determinants that specify adaptive functions specific to the environment in which the organism lives. For example, the food-ingredients, e.g. carbohydrates, that are present in the distal part of the colon have shaped the genomes of the microbiota living in this GIT compartment.

Sugar availability in the colon of mammals

Mammals are well equipped to adsorb simple sugars (e.g. glucose, galactose) and to hydrolyze particular disaccharides (e.g. sucrose, lactose and maltose) and certain polysaccharides (e.g. starch). It contrast, they cannot digest other dietary fiber-derived polysaccharides (e.g. pectin, and xylan- and arabinose-containing polysaccharides). The indigestibility of these sugars reflects the paucity of host enzymes required for their degradation. The human genome encodes just a single predicted glycoside hydrolase represented in the nine families of enzymes known in nature with xylanase, arabinosidase, pectinase, or pectin lyase activities, whereas the mouse genome does not appear to specify any glycolytic enzymes at all (Sonnenburg et al. 2005).

Many of the sugars that escape digestion by the host's enzymes are candidate prebiotic compounds and include fructo-oligosaccharides (FOS), galacto-oligosaccharides (GOS), gluco-oligosaccharides, xylo-oligosaccharides, lactulose and raffinose (Guarner and Malagelada 2003). All the above mentioned dietary compounds would have been lost if the distal intestine of mammals had not have been colonized by diverse anaerobic microbial communities, represented by, among others, *Bacteroides*, *Bifidobacteria*, *Clostridia*, *Enterobacteria* (for a review see Vaughan et al. 2005), that possess the metabolic capacity to degrade such oligo- and polysaccharides. The energy needs of these bacteria are met by the fermentation of these substrates that have been left undigested in the small intestine (Cummings and Englyst 1987; Delzenne and Williams 2002; Sinha and Kumria 2001). In this mutualistic relationship, the hosts gain carbon and energy, through short-chain fatty acids, and the microbes are provided with a rich buffet of glycans and a protected anoxic environment.

The ability to degrade complex carbohydrates, with respect to level of polymerization and type of glycosidic bonds, contributes to the competitiveness of a given strain in the GIT, as exemplified for *E. coli*, where it was shown that mutations in carbohydrate metabolic pathways affect intestinal colonization (Chang et al. 2004). Thus, simple sugars are metabolized rapidly in the upper GIT, whereas complex carbohydrates remain abundant in the distal part of the GIT. These complex carbohydrates can be dietary compounds, e.g. starch, cellulose, hemicellulose, xylan and pectin, or host derived compounds, such as mucins, glycosphingolipids, hyaluronic acid and heparin (Hooper et al. 2002). The type of sugar available may influence the species composition and abundance of the microbiota along the GIT. In this context, bacteria such as *Lactobacillus* are particularly prevalent in the upper GIT (Vaughan et al. 2005), where they ferment relatively simple mono-, di- and trisaccharides. In contrast, bacteria active in the lower parts of the colon such as bifidobacteria probably derive their localized ecological success from their capacity to metabolize complex carbohydrates (Hooper et al. 2002).

It comes as no surprise that a significant proportion of the genome of a common inhabitant of a mammalian colon such as bifidobacteria is dedicated to sugar metabolism, thus reflecting a critical adaptation to this highly competitive ecological niche.

As discussed above the bifidobacterial genomes have a plethora of predicted proteins assigned to COGs in the carbohydrate transport-metabolism category, nearly 10% of the total predicted proteins. This is 30% more than other GIT inhabitant such as *E. coli*, *Enterococcus faecium* and *L. lactis* (Schell et al. 2002). However, bifidobacteria possess sugar fermentative capacity similar to another intestinal commensal bacterium such as *Bacteroides* (Xu et al. 2003).

Carbohydrate utilization

The gene composition of the *B. longum* biotype *longum* NCC2705 genome accurately reflects an intrinsic adaptation to the intestinal niche. Interestingly, in this bacterium many of the genes involved in sugar utilization appear to be clustered in a specific ''life style adaptation'' region in its chromosome, which seems to display a high degree of variability in different strains of this species (Klijn et al. 2005). This region includes α-mannosidases and endo-β-NAc-glucosaminidase, that might well contribute to the ecological fitness of NCC2705 strain in the GIT (Klijn et al. 2005) by increasing the metabolic capacity of this strain with regards to complex carbohydrates.

In the case of the NCC2705 strain an enormous potential for degradation and utilization of complex carbohydrates can be predicted from its genome content, including a variety of glycosyl hydrolases that are required for utilization of diverse plant-derived dietary fibre or complex carbohydrate structures produced by the host. According to the sequence-based classification of carbohydrate-active enzymes (CAZy; Coutinho and Henrissat 1999), about 10% of the annotated genes of *B. longum* biotype *longum* NCC2705 encode enzymes involved in the transport and metabolism of carbohydrates (Fig. 1). The *B. longum* biotype *longum* NCC2705 genome is predicted to contain more than 40 glycosyl

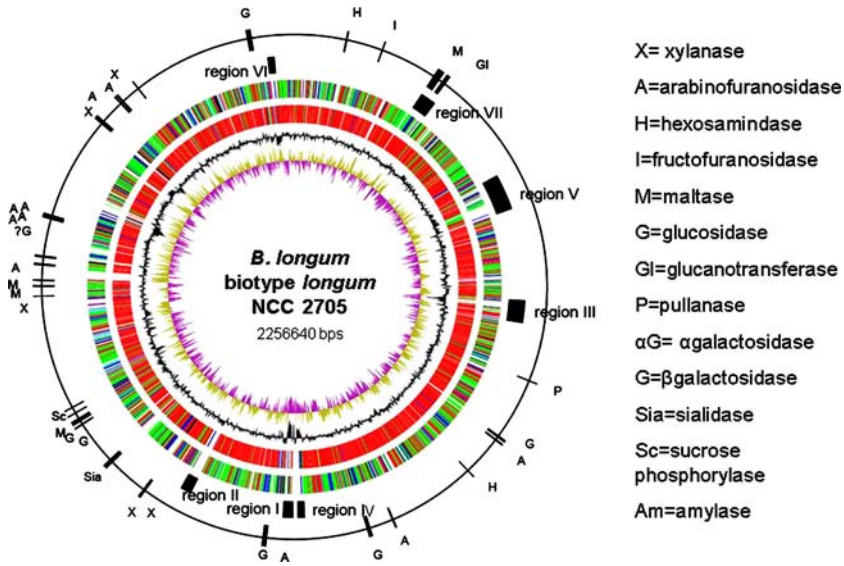


Fig. 1 Circular plot of genome diversity found in *Bifidobacterium*. From inside to outside: ring 1, GC deviation; ring 2, G + C content; ring 3, atlas of *B. longum* biotype *longum* NCC2705 strain, ring 4, comparison to the unfinished genome sequences of *B. adolescentis* ATCC 15703, red indicates homologies >95%, green indicates Homologies 90<X>95, blue indicates homologies <90%; ring 5, genes involved in carbohydrate metabolism present in the genome of *B. longum* biotype *longum* NCC2705 strain. Regions I to VII indicate variable genome sequences of *B. longum* biotype *longum* (Klijn et al. 2005)

hydrolases whose assumed substrates represent a wide range of di-, tri-, and higher order oligosaccharides (Fig. 1). Several of these glycosyl hydrolases are assumed to exert their activity outside the cytoplasm as secreted enzymes. Hence, they might not only be required for the nutritional demand of the enzyme-producing bacterium, but may also shape the metabolic scene of the colon to sustain a microbiota that indirectly provides the host with approximately 10–15% of its calories from the degradation of complex carbohydrates through short-chain fatty acids (Vaughan et al. 2005). Notably, in *B. longum* biotype *longum* NCC2705, as well as in other unpublished bifidobacterial genome sequences, many of the genes encoding predicted glycosyl hydrolases are associated with genetic units that encode MalEFG-type oligosaccharide transporters, and are organized in clusters that display a conserved modular organization. Each of these clusters specifies a LacI-type, sugar-responsive repressor, an ABC-type MalEFG oligosaccharide transporter and one to six genes encoding various classes of glycosyl hydrolases.

Previous work on several bifidobacterial species has identified many genes encoding carbohydrate-modifying enzymes that degrade, modify or create glycosidic bonds (Table 2).

Certain bifidobacteria can indeed utilize complex carbohydrates such as hog gastric mucin (Hoskins et al. 1985), pectin (Slovakova et al. 2002) and other complex plant oligosaccharides (Van Laere et al. 2000a, b). Hence, such bifidobacteria have acquired adaptations to process a variety of otherwise indigestible components of the human or animal diet. With respect to host-produced sugar-containing substances, bifidobacteria display the capacity to degrade a variety of host-derived glycoconjugates including mucin oligosaccharides and glycosphingolipids (Hoskins et al. 1985). In this way the host itself may provide substrates to bifidobacteria, thereby representing a remarkable synergistic relationship (Corfield et al. 1992).

Other nutrients which have partly or completely escaped digestion in the upper human GIT are plant-derived high molecular weight carbohydrates such as starch and pectin. Limited knowledge is available relating to starch degradation by bifidobacteria. An extracellular α-amylase produced by a human isolate of

Table 2 Sugar metabolic genes

Gene	Encoded enzymes	Substrate	Bifidobacterial species	References
*scr*P	Sucrose-phosphorylase	Sucrose	*B. animalis* subsp. lactis; *B. longum* bioype longum NCC2705	Trindade et al. (2003); Kullin et al. (2006)
β-gal	β-galactosidase	Lactose	*B. longum* biotype *infantis* HL96 *B. adolescentis* DSM20083	Van Laere et al. (2000; Zarate et al. (2000); Rossi et al. (2000)
*lnp*A	Lacto-*N*-biose phosphorylase	Galactose	*B. bifidum* JCM1254	Kitaoka et al. (2005)
fos operon	β-fructofuranosidase	FOS, sucrose	*B. breve* UCC2003	Ryan et al. (2005)
*bfr*A	β-fructofuranosidase	FOS, sucrose, inulin, raffinose	*B. animalis* subsp. *lactis* DSM10140	Ehrmann et al. (2003); Janer et al. (2004)
*gal*A	Endogalactanase	GOS, (arabino)galactans	*B. longum* biotype *longum* NCC2705	Hinz et al. (2005)
*apu*B	amylopullulanase	Amylopectin, pullulan	*B. breve* UCC2003	Ryan (2006)
*afc*A	1,2-α-L-fucosidase	Fucose-containing bioactive glycoconjugates	*B. bifidum* JCM1254	Katayama et al. (2004)

B. adolescentis was shown to hydrolyze starch to produce maltotriose and maltose. The latter compound was found to be a strong inhibitor of starch hydrolysis, suggesting the existence of end-product inhibitor (Lee et al. 1997). In a recent study, it was shown that starch, amylopectin and pullulan were utilized by 11 out of 42 different strains representing different species of bifidobacteria (Ryan 2006). Notably, all *B. breve* strains tested possessed amylopullulanase activity, which may suggest that the combined extracellular α-(1–4) and α-(1–6) glucosidase activities may be characteristic for this *B. breve* strain and may have some biological relevance for this organism in the gut (Ryan et al. 2006). *B. breve* constitutes one of the dominant bacteria in the infant microbiota (Ventura et al. 2004c; Scardovi et al. 1971) and so this enzyme may be important during weaning when non-milk foods are supplemented to the diet and infants are, for the first time, exposed to complex carbohydrates different from those present in mother's milk.

Carbohydrate transport appears to be facilitated predominantly by ABC transporters, permeases and proton symporters, rather than so-called phosphoenolpyruvate-phosphotransferase systems (PEP-PTS; Schell et al. 2002). Glucose transport in *B. breve* has previously been established to be mediated via PEP-PTS (Degnan and Macfarlane 1993). The PTS acts through the concomitant internalization and phosphorylation of carbohydrates. The transfer of phosphate from PEP to the incoming sugar is mediated via a phosphorylation chain, which involves Enzyme I (EI), histidine-containing protein (HPr) and Enzyme II (EII). Interestingly, *B. longum* biotype *longum* NCC2705 is predicted to contain just a single EII-encoding homolog, whereas *B. breve* UCC2003 contains four EII-encoding loci (Maze et al. 2006). The biological significance of this difference in PEP-PTS's is unclear, but may indicate that *B. breve* may more frequently encounter less complex sugars. It is possible that the different carbohydrate transport capability of *B. breve* and *B. longum* biotype *longum* is linked to the different ecological origin of these bacteria. In fact, *B. breve* strain are more frequently found in the infant intestine, whereas *B. longum* biotype *longum* isolates are more commonly associated with the adult GIT. Consequently the different sugar-containing diets to which infant and adult are subject may have shaped the genome of their intestinal commensals, such as different bifidobacterial species, also with respect their sugar metabolizing capacity.

Prebiotic properties of bifidobacteria

Prebiotics are indigestible food ingredients that beneficially affect the host by selectively stimulating growth of beneficial commensals such as bifidobacteria. For the latter bacteria, such growth-promoting effects have been termed as bifidogenic activity. Various oligosaccharides are known prebiotics, including FOS and GOS (Guarner and Malagelada 2003). FOS represents the most widely used commercial prebiotic, being composed of short to medium size chains (degree of polymerization, 4 to 60) of fructose moieties connected by β(2–1) linkages, which are in turn attached to a terminal glucose unit by a β(2–1) bond. Because of these β(2–1) linkages FOS are resistant to mammalian enzymes and thus are able to reach the colon (Gibson 1999). Inulin, a long chain FOS has received a significant amount of scientific attention. This is in part due to the fact that inulin is a naturally occurring polysaccharide found in many plants, representing a common component of the mediterranean diet. The inspection of bifidobacterial genomes for the presence of operons involved in prebiotic metabolism revealed a rich arsenal of such genes. In the genome of *B. breve* UCC2003, a *fos* operon, which encodes a putative permease, a conserved hypothetical protein, and a β-fructofuranosidase, was identified and shown to be involved in fructooligosaccharide breakdown (Ryan et al. 2005). Characterized bifidobacterial β-fructofuranosidases display different substrate kinetics and biochemical properties (Ehrmann et al. 2003; Imamura et al. 1994; Janer et al. 2004; Warchol et al. 2002).

Transcriptional analysis of the *B. breve* UCC2003 *fos* operon showed that induction of the operon occurred when the substrates sucrose and 'Actilight' were present, but only in the absence of glucose or fructose, suggesting that a regulatory mechanism is operational in *B. breve* UCC2003 for differential carbohydrate utilization that depends on substrate availability in the absence of a preferred sugar (Ryan et al. 2005; Titgemeyer and Hillen 2002).

Human milk is rich soluble oligosaccharides (e.g. GOS) that have been proposed to function as a prebiotic (Ward et al. 2006). These human milk oligosaccharides range in their degree of polymerization from three to over 32 moieties, and are comprised of glucose, galactose, *N*-acetylglucosamine, fucose and sialic acid (Kunz et al. 2000). Previous work on *B. longum* biotype *longum* NCC2705 has identified a gene, *gal*A, which codes for an extracellular enzyme that can degrade GOS (Hinz et al. 2005).

Besides oligosaccharides, human milk peptides represent another class of diet-derived compounds, eliciting bifidogenic effects (Liepke et al. 2002). These include small peptides, which are derived from the digestion of milk protein with the gastric protease pepsin, and which contain a pair of cysteine residues forming a disulfide bond and two small hydrophobic domains located C-terminally to the two cysteines (Liepke et al. 2002). The presence of such peptides was shown to be 100 times more effective on a molar basis in the stimulation of the growth of bifidobacteria as compared to certain oligosaccharides, and it has been speculated that the bifidogenic activity of breast milk through the presence of specific peptides for exceeds that of milk oligosaccharides (e.g. GOS).

Extracellular structures and bifidobacteria-GIT interaction

Gut commensals, such as bifidobacteria, may be expected to interact with the host through direct contact between bacteria and host epithelial cells. For several intestinal pathogens, including *Listeria monocytogenes* and *Salmonella* spp., these physical contacts have been studied in detail and their role in the virulence of these bacteria has been established (Lecuit et al. 2001; Mengaud et al. 1996). In the context of exploring possible interactions between bifidobacteria and human host cells at the molecular level, knowledge on exported proteins represents a key aspect. MacConaill et al. (2003) reported that the protein export machinery of *B. breve* is comparable to that of other Gram-positive bacteria. It was concluded from this study that proteins exported by *B. breve* UCC2003 contain either a signal peptide or a number of transmembrane regions (MacConaill et al. 2003). Notably, in both *B. longum* biotype *longum* NCC2705 and *B. longum*

biotype *longum* DJO10A genomes, genes are found that are predicted to encode glycoprotein-binding fimbriae-like structures, which could be involved in direct host interaction. In addition, this bacterium as well as other bifidobacterial strains (e.g. *B. breve* UCC2003 and *B. longum* biotype *longum* DJO10A) appear to encode a serpin-like protease inhibitor that has been demonstrated to contribute to host interaction in the GIT (Ivanov et al. 2006). In fact, bifidobacteria may encounter both pancreatic elastase and neutrophil elastase in their natural environment and protection against exogenous proteolysis may play an essential role in the interaction between these commensal bacteria and their host. It was shown that the *B. longum* biotype *longum* NCC2705 serpin is an efficient inhibitor of human neutrophil elastase and pancreatic elastase, whose release by activated neutrophils at the sites of intestinal inflammation represents an interesting mechanism of innate immunity (Ivanov et al. 2006). Such properties could promote the use of bifidobacteria in the pro and prebiotic therapy of ulcerative colitis (Suzuki et al. 2006). In a recent study, it was demonstrated that the supply of a bifidogenic compound to patients suffering from ulcerative colitis mediated competitive interactions and production of antimicrobial metabolites, while positively influencing the epithelium and effecting immunostimulation (Suzuki et al. 2006). The restoration and enforcement of the natural microbial community using prebiotic components thus represents a very attractive physiological and non-toxic way to prevent or treat ulcerative colitis.

Another possibility for interactions between bifidobacteria and its host is represented by the bifidobacterial ability to produce exopolysaccharides (EPS). The role of EPS produced by non-pathogenic colonic bacteria is poorly understood, but it is likely that these extracellular structures are important for bacteria to establish themselves within the host (Ruas-Madiedo et al. 2006). Notably, the commensal microorganism *Bacteroides tetaiotamicron* has evolved a genetic mechanism allowing it to change its surface make-up by producing at least eight distinct capsular polysaccharides, which may help the microorganism to escape recognition by the host (Krinos et al. 2001). There are only a few reports concerning these extracellular structures in bifidobacteria. However, Nagaoka et al. (1994) reported the production of rhamnose-rich EPS linked to the cell wall in both *B. breve* and *B. bifidum* strains, and a cell-associated polysaccharide consisting of glucose, ribose, galactose, rhamnose and chiroinositol in *B. bifidum* (You et al. 2004).

Genome analysis of *B. longum* biotype *longum* NCC2705 revealed the presence of two regions related to polysaccharide biosynthesis. Interestingly, both regions are flanked by IS elements and show a strong divergence in G + C content relative to the remainder of the genome, indicating that these regions have been acquired by horizontal gene transfer. This genome region appears to be unique to NCC2705 genome and bioinformatics analyses suggested that it has been acquired though HGT (Schell et al. 2002). Similarly, analyses of both *B. breve* UCC2003 and *B. adolescentis* ATCC15703 genomes revealed the presence of DNA regions encoding for EPS synthesis which appear to be strain specific or not widely distributed in all the strains analyzed.

Bifidobacteria and stress response

Another example of how the genome content of bifidobacteria reflects its adaptation to the GIT is represented by genes coding for proteins involved in stress responses (for a review see Ventura et al. 2006a). Genes encoding for the principal molecular chaperones as well proteases families, i.e. *gro*EL-*gro*ES, *dna*K, *grp*E, *dna*J, *clp*B, *clp*C, *clp*P, were isolated from almost all bifidobacterial taxa. However, in contrast to other *Actinobacteriae*, i.e. soil living *Actinobacteriae* or plant pathogenic *Actinobacteriae*, bifidobacteria, except for *dna*J and *clp*P, do not carry multiple copies (e.g. due to paralogy) of the above mentioned genes in their genome (see Ventura et al. 2006a). Interestingly, bifidobacteria possess one of the smallest set of genes coding for heat-shock-proteins (Hsp), that include molecular chaperones and proteases to deal with thermal damage, relative to other high G + C Gram positive bacteria (Ventura et al. 2006a). In fact, in their more or less isothermal niche, i.e. predominantly warm blooded

mammalian hosts, one would predict that an elaborate system responding to temperature fluctuations may become partly obsolete (Ventura et al. 2006a). In contrast, the intestinal ecological niche, where bifidobacteria normally live, is exposed to significant fluctuations in osmotic conditions, e.g. due to diet composition, and bifidobacteria seem to have evolved a more elaborate system of protection against osmotic or bile salt stress (Sanchez et al. 2005; Ventura et al. 2005a, 2006a).

The impact of genomics on taxonomy

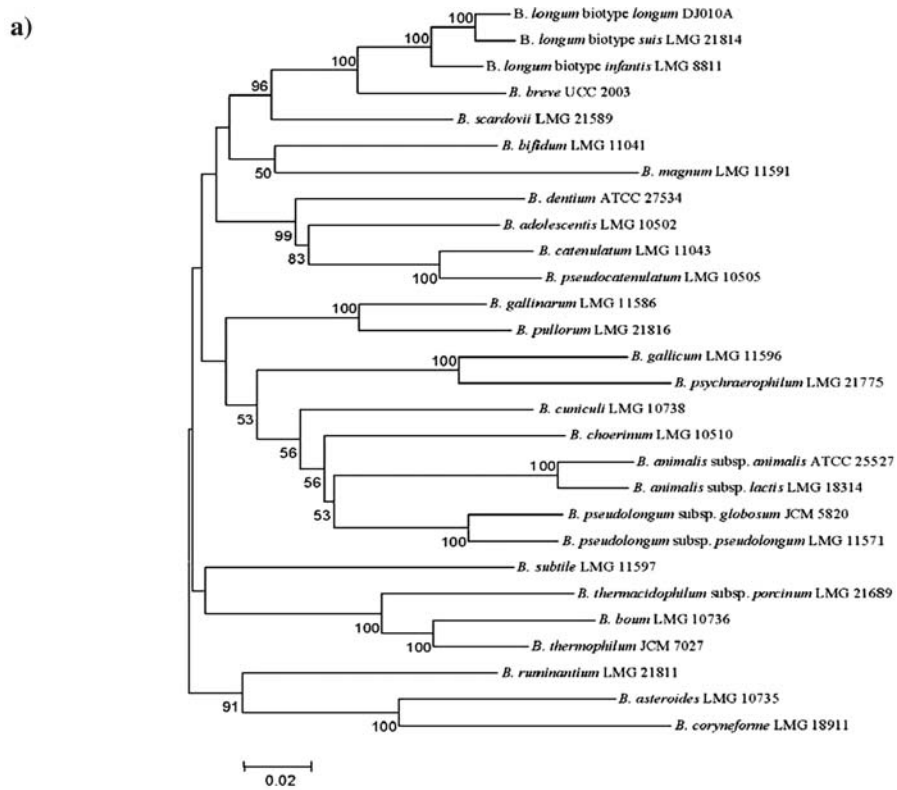
In the last decade prokaryotic taxonomy has benefited from the enormous advances in microbial population genetics, ecology and genomics, and in particular by the ease with which sequence data can be obtained. Prokaryotic species are currently characterized using a polyphasic approach that incorporates genotypic and phenotypic (including chemotaxonomic) properties (Stackebrandt et al. 1997; Vandamme et al. 1996). Since the 1970s, physicochemical DNA-DNA hybridization methods have been considered the cornerstone of bacterial taxonomy and at present a bacterial species is ''a category that circumscribes a genomically coherent group of individual isolates/strains sharing many unconnected features, comparatively tested under highly standardized conditions'' (Coenye et al. 2005). In this approach, the overall genetic similarity among strains is assessed by the degree with which their genomes hybridize under standardized conditions. Isolates that display a DNA–DNA hybridization value of more than 70%, while exhibiting less than 5% difference in their melting temperature are considered to belong to the same species. Conversely, isolates, sharing a level of genomic DNA hybridization that is less than a 50%, do not belong to the same species. Comparative sequence analysis of the 16S rRNA gene is commonly used to determine the phylogenetic position of novel isolates. Strains that show over 97% of 16S rRNA sequence identity are considered to belong to the same species, which is in most cases consistent with a minimal genomic hybridization level of 70% (Vandamme et al. 1996).

As claimed by the ad-hoc committee for the revaluation of the species definition in bacteriology (Stackebrandt et al. 1997), the introduction of novel taxonomic methods is expected to provide increased sensitivity, consistency and robustness to the systematics of prokaryotes. One of the particularly interesting developments includes the analysis of complete genome sequences. The classical genotypic methods, described above, which are used in bacterial taxonomy, will not be significantly influenced by either of the evolutionary forces that shape bacterial genomes.

In recent years, comparative sequence analysis of genes that specify more or less universally conserved macromolecules like the 16S rRNA gene (Matsuki et al. 1998, 1999; Ventura et al. 2001a, b, 2003b; Ventura and Zink 2002), or housekeeping genes such as *rec*A (Kullen et al. 1997; Ventura and Zink 2003c), *tuf*A (Ventura et al. 2003a), *atp*D (Ventura et al. 2004a), *gro*EL (Jian et al. 2001; Ventura et al. 2004b), *dna*K (Ventura et al. 2005d), *grp*E (Ventura et al. 2005d), *clp*P (Ventura et al. 2005c), and *hrc*A (Ventura et al. 2005a) have become standard practice in bifidobacterial taxonomy. Nevertheless, several studies raised concerns that single-gene trees may not adequately reflect phylogenetic relationships, because of the possibility of HGT, incongruous mutation rates and variable rates of recombination. Consequently, a phylogenetic tree based on a set of combined alignments of conserved orthologous proteins [so-called supertrees (Bininda-Emonds 2004)] generates the most reliable picture of evolutionary relationships between bacteria (Brown et al. 2001).

Recently, a phylogenetic analysis of the genus *Bifidobacterium* using a multigene concatenation approach, revealed the increase in discriminatory power and robustness of the derived phylogenetic tree (Ventura et al. 2006b). Moreover, this analysis showed that the projected ancestor of all recognized bifidobacteria groups was most closely related to the current *B. asteroides* species (Fig. 2a).

In addition to the conventional 16S rRNA-based methods for detection of bifidobacterial species, the availability of genome sequences has led to identification of numerous other molecular



b)

BL0004

```
B. longum NCC2705     ITDTENLLGSHAAEVTDAIAKTEKETGVHVHLLYLSSFNSQQKPGDWAATVMESMNPKPNTVMLAVASNDGNLVVVVSKNSDKWLL
B. longum DJ010A     ITDTENLLGSHAAEVTDAIAKTEKETGVHVHLLYLSSFNSQQKPGDWAATVMESMNPKPNTVMLAVASNDGNLVVVVSKNSDKWLL
B. longum infantis    ITDTENLLGSHAAEVTDAIAKTEKETGVHVHLLYLSSFNSQQKPGDWAATVMESMNPKPNTVMLAVASNDGNLVVVVSKNSDKWLL
B. breve              ITDTENLLGSHATEVSDAIAKTEKETGVHVHLLYLSSFNSEQTPQKWASTVLESTKPKANTVLLAVAANDGNLVVAVSSNSDEWLK
B. adolescentis       ITDTQNLLGSNVSEITDAINQTKKETGVTVRLVYLDSFNAKGKPDKWVSALLESLNPKPNTVMLAVASNDGNLVVAVSSNSEEWLK
B. animalis           ITDPQNLLGDNLTKVIDAANRTKQETGVTVRLLYLESFNTTEKPEQWASQLLESLNPPANTVMLAVASHDGNLVVVVSPNSDEWLR
                      ***.:****.: ::: **  :*::**** *:*:**.***:  .* .*.: ::** :* .***:****::******.** **::**

B. longum NCC2705     DNKTVDKLSEAAQQPLLENP----PSWSGAATAMMDQIVKSKKASTSSSTVIVGIIIMGVVLVALVIIIVVMVVIH-------RRK
B. longum DJ010A     DNKTVDKLSEAAQQPLLENP----PSWSGAATAMMDQIVKSKKASTSSSTVIVGIIIMGVVLVALVIIIVVMVVIH-------RRK
B. longum infantis    DXKTVDKLSEAAQQPLLENP----PSWSGAATAMMDQIVKSKKASTSSSTVIVGIIIMGVVLVALVIIIVVMVVTH-------RRK
B. breve              QQKTVDALSEAAQQQLMEST----PNWSGAATAMMDQIAKSKKTSTSSSTVTIGVITMGVALAALVIIIVVMVIVR-------RRR
D. adolescentis       KKSTVDALSDAAQKPLMESD----PDWAKSATDMMNQISVQKKTSTSSRTIWIGVGGFAGALLLLVVILVVFHILRKKDVIGHRRK
B. animalis           SQSTVDQLSEAASKPLLDQQAGQGPNWSQSATDMMDATASAKSTATTHAVSSVSITVMIIVFVVLIVVVLVMIFMR-------RKR
                      . .*** **:**.: *::.     *.*: :*  **: *   *.::*:  .  :.:  :  .:  *:::::*: .::       *::
```

BL1532

```
B. longum NCC2705     AAFGADLRSQIREFSWAWNQWTGLKQPYIGFFSFVPMDIYPTAHYMWPSDPTYLTDQHNVVLTVFYGAMVTFARHLTGSNDAGIVTLAAL
B. longum DJ010A      AAFGADLRSQIREFSWAWNQWTGLKQPYIGFFSFVPMDIYPTAHYMWPSDPTYLTDQHNVVLTVFYGAMVTFARHLTGSNDAGIVTLAAL
B. adolescentis       AAYGADIRSQIREFSWALNQWTGLKQPYIGFFSFVPMDIYPTAHYMWPANPTYLTDQHNIVLTVFYGAFAAVSRYFTDSNDAGIVALAAL
                      **:***:********** ********************************::*********:********:.:.:*::*.*******:****

B. longum NCC2705     QTLFAVFCCAAAANRFLNRPW--IGKTATDSAAPP----------------------QAGGLARFLILLFFMVCPLAVFSTISITKSP
B. longum DJ010A      QTLFAVFCCAAAANRFLNRPW--IGKTATDSAAPP----------------------QAGGLARFLILLFFMVCPLAVFSTISITKSP
B. adolescentis       QWLFAAFCCAATANRFFNLPWRRLGVGTFDFSHPERHDCWNMRDFTHPEAGVVARPSRLRAGAKTRFVILLFFMVCPLAVFATISLTKSP
                      * ***.*****:****:* **  :*  : * : *                              :**. :**:**************:***:****

B. longum NCC2705     LFAFSFVWWFSVWYELVQTWHPAGTRKHPQTPAIATPVHLPRHSFIAFILATSVMLISAKYAWYIIALQIVLALIADRKRWATYVVALLI
B. longum DJ010A      LFAFSFVWWFSVWYELVQTWHPAGTRKHPQTPAIATPVHLPRHSFIAFILATSVMLISAKYAWYIIALQIVLALIADRKRWATYVVALLI
B. adolescentis       LFAFAFVWWFGVWYELIMTHIKALPTIN------GKPMKLRKRSLAALFMSSCVMLISAKYAWYIILFAALLAIINDRKRWKTYVVALML
                      ****:*****.*****  *   * .  :      ..*::' ::*: *:::::.************** :  :**:* ***** *******::

B. longum NCC2705     PTVLIHGGISFAISSGAIIGGDPIESRGVQLQMIARVAQRNPDGISDEAKKNLAPVFNLDQMADAYSQQDADPVKSSGIQAKKVSYKWRT
B. longum DJ010A      PTVLIHGGISFAISSGAIIGGDPIESRGVQLQMIARVAQRNPDGISDEAKKNLAPVFNLDQMADAYSQQDADPVKSSGIQAKKVSYKWRT
B. adolescentis       PTVLIHGGLVYLVNSGAVIGGDPIESRGIQLQQIARVAKYNPQGIPEDAAKKLAPVFNLDQMAESYFQQDADPVKSSGIQSKKVSYKWRT
                      ********: : :.***:**********:*** *****: **:**.::* *:************::* **************:*********

B. longum NCC2705     VTPEDMTNFNKAWFEIVKDNPIIALDALLAKCFGYFNVNDQPYVSMDYYVTSDYVQKNSTWIKDYNHDWREHIAGFTRVWGGIPVLGWPT
B. longum DJ010A      VTPEDMTNFNKAWFEIVKDNPIIALDALLAKCFGYFNVNDQPYVSMDYYVTSDYVQKNSTWIKDYNHDWREHIAGFTRVWGSIPVLGWPA
B. adolescentis       VTKDDMKDFNDAWWQIVKANPQIALDALFAECFGYFNVTDLPYVSMDYYVNNDYVQSDNEWIHLYNHQWRNQVAGFAKKWGQIPVLGWVT
                      ** .** .** ***:*** ** ********.********* * *********  **** .  **. ***********.... ** ****** .
```

◀ **Fig. 2** Phylogenetic tree of the genus *Bifidobacterium* computed from the concatenation of *clp*C, *dna*B, *dna*G, *dna*J1, *pur*F, *rpo*C and *xfp* gene sequences by the neighbour-joining method and Kimura's two parameter model as the substitution model (panel a). In panel b, partial sequence alignments of two proteins BL004 (Accession no. NP_695235) and BL1532 (Accession no. NP_696688) showing regions of similarity. The asterisks (*) denote residues that are completely conserved, where (:) and (.) indicate positions that contain similar or related amino acids. A few conserved indels (i.e. inserts or deletions) that appear specific for certain bifidobacterial species are also present in these alignments. The homologs of these proteins are presently not found in any other organisms and hence primers based on conserved regions in these proteins can be used for identification or detection bifidobacterial species

markers that provide powerful and highly specific means for detection of these bacteria. In this regard, systematic BLAST searches have been carried out on each of the proteins or open reading frames (ORFs) from the genome of *B. longum* biotype *longum* NCC2705 against all available proteins in the NCBI database to identify proteins, which are either (1) unique to this strain, (2) limited to the two sequenced strains of *B. longum* biotype *longum* (NCC2705 and DJO10A), (3) present in all currently available bifidobacterial genomes, or (4) are only found in some or all *Actinobacteria*, including bifidobacteria (Gao and Gupta 2006). These studies have identified >30 proteins which are specific for the *B. longum* biotype *longum* NCC2705 strain, >90 proteins that at present are uniquely found in the two *B. longum* biotype *longum* strains, and additional >70 proteins that appear to be distinctive characteristics of various available bifidobacterial species (mainly *B. longum* biotype *longum* NCC2705, *B. longum* biotype *longum* DJO10A and *B. adolescentis*) (Gao and Gupta, R.S., unpublished results). When blastp searches are carried out using these proteins as queries, all significant hits obtained correspond to the indicated strains or species, and no significant hits are detected from any other organisms. The detailed results of these analyses indicating the observed E-values for all significant hits from blastp searches can be found at the Gupta lab bacterial phylogeny website (www.bacterialphylogeny.com) under actinobacteria signatures. Partial sequence alignments for two of these proteins for all available homologs are presented in Fig. 2b. Both of these proteins are highly conserved in bifidobacterial species and there are various long stretches where the amino acid sequence is completely conserved in all available species. The development of PCR primers based on these conserved regions should provide highly specific means for the identification of both known as well as new bifidobacterial species in different environments and for their in situ detection by methods such as FISH, with minimal possibility of a false-positive signal.

For most of the proteins that were indicated to be either *B. longum*- or bifidobacteria-specific, sequence information is available from only a limited number of bifidobacterial species. Hence, it is important to obtain sequence information for these proteins from other members of this genus. It is likely that while many of these proteins will prove to be strain-, species- or genus-specific, some may prove to be specific for a subset of bifidobacterial strains or species revealing their evolutionary relationships and important differences amongst them. Virtually all of the bifidobacteria-specific proteins identified by these studies are of unknown function. These proteins likely carry out novel functions that make bifidobacteria biochemically and physiologically distinct from other bacteria. Some of these functions may also be responsible for the probiotic activities of bifidobacterial species in the GIT. Hence, a major challenge for the future is to understand the cellular functions of these bifidobacterial-specific proteins, which may also shed light on the probiotic activities of these bacteria.

Although bifidobacteria are placed in the phylum *Actinobacteria*, no molecularly defined characteristic was known that was uniquely shared by these bacteria and various other actinobacteria. However, a specific relationship of bifidobacteria to the *Actinobacteria* is now strongly supported by a number of conserved indels (i.e. inserts or deletions) in protein sequences (cytochrome c oxidase subunit 1, CTP synthetase and glutamyl-tRNA synthetase) and the 23S rRNA that are uniquely shared by bifidobacteria and other *Actinobacteria*, but which are not found in any other bacterial groups (Gao and Gupta 2005). Recent comparative analyses of actinobacterial genomes

have also identified many proteins that are uniquely found in all actinobacteria, and homologs of which are not present in any other prokaryotic or eukaryotic organism (Gao and Gupta 2006). These proteins provide novel and distinctive molecular signatures for the phylum *Actinobacteria* based on whose presence actinobacterial species can now be clearly distinguished from other bacteria. The bifidobacterial species are currently placed in a distinct order "*Bifidobacteriales*" within the phylum *Actinobacteria*, but their relationship to other orders or subgroups within this phylum is not fully understood. However, recent evidence from several lines of investigations (viz. branching pattern in phylogenetic trees based on concatenated sequences for many proteins, conserved indels and whole proteins that are uniquely shared) now strongly suggest that species belonging to the *Bifidobacteriales* order are related to species that are members of the *Micrococcineae* order (viz. *Arthrobacter, Tropheryma, Leifsonia, Kinenococcus*, etc. Gao and Gupta 2006). Both of these groups are shown to form deep branching lineages within the *Actinobacteria phylum* (Gao and Gupta 2006).

Novel techniques to monitor bifidobacteria in their natural environment

As outlined above, the application of new molecular markers or gene sets for bifidobacterial species identification and tracing represents an active field of genome-based bifidobacterial research. The natural environment of bifidobacteria is the GIT, so the enumeration of these bacteria is based on an initial cultivation step followed by selective plating (Hartemink and Rombouts 1999; Roy 2001) or PCR-mediated detection methods (for a review see Ventura et al. (2004c). However, it is recognized that not all microbiota of the human GIT can be cultivated (Favier et al. 2002; Zoetendal et al. 2004). In recent years, analysis methods using 16S rRNA gene sequences have been widely used in place of conventional culture methods for the structural analysis of intestinal microbiota (Amann et al. 1995). In complex mixed populations, 16S rRNA-targeted oligonucleotide probes have been applied to fluorescence in situ hybridization (FISH) as a culture-independent method (Franks et al. 1998; Langendijk et al. 1995). FISH involves whole cell hybridization with fluorescent oligonucleotide probes targeted against specific bifidobacterial species. This technique estimates that 4.4% of the adult fecal flora is represented by bifidobacteria (Lay et al. 2005). The FISH technique has also been adapted for use with flow cytometry (FCM), which allows for high-throughput analysis such as monitoring the metabolic activity of stressed and starved bacteria (Zoetendal et al. 2002). Moreover, FCM allows monitoring of bacterial heterogeneity at the single-cell level and provides a means to sort subpopulations of interest for further molecular analysis (Davey and Winson 2003).

Another semi-quantitative fingerprinting technique such as the denaturing gradient gel electrophoresis (DGGE) has been widely used to rapidly monitor the microbiota community shifts and compare the varying bifidobacterial communities that exist between different individuals and various intestinal locations, as well as variations that may occur as a result of a change in diet. The DGGE technique and its variations, such as temperature gradient gel electrophoresis (TGGE) or combinations of DGGE and TGGE, are based on 16S rRNA sequence-specific melting behaviour of PCR products of complex mixtures of bacteria (see review (Muyzer and Smalla 1998), DGGE was used to investigate the bifidobacterial content and composition in the human microbiota of infants versus adults. In this context the adult bifidobacterial microbiota was found to be dominated by *B. adolescentis*, *B. longum* biotype *longum* and *B. catenulatum*, whereas the bifidobacterial microbiota of infant is dominated by *B. breve* and *B. longum* biotype *infantis* (Satokari et al. 2001). These findings were also corroborated by a combined real-time PCR with *Bifidobacterium* genus- and species-specific primers (Bartosch et al. 2004; Gueimonde et al. 2004; Matsuki et al. 2004). The transaldolase-encoding gene found in all known *Bifidobacterium* species has been used as an alternative target for real-time PCR, and appears to be superior to the 16S rRNA gene in quantifying bifidobacterial populations in infants (Requena et al. 2002). Another

community fingerprinting technique, terminal restriction fragment length polymorphism (T-RFLP), has been adapted for the purpose of characterizing human fecal bifidobacteria (Sakamoto et al. 2003).

DNA microarray technology advances rapidly and offers a fast, highly throughput option for detection and estimation of microbes in a complex ecosystem (Bodrossy and Sessitsch 2004). These microarrays have been termed phylochips, microbial diagnostic microarrays or identification arrays. Typically, such microarrays contain hundreds of oligonucleotide probes, usually based on the 16S rRNA gene, specific for different strains or species or genera of bacteria that are detected in a single assay.

Concluding remarks

Many issues regarding sequencing of complete bifidobacterial genomes remain at present unresolved. A first important issue is that the small number of bifidobacterial strains that have been sequenced today are by no means a proper representation of the total bifidobacterial diversity, and, so far, have been biased towards organisms of biotechnological importance. It is probable that with the accumulation of more genome data, not merely representing bifidobacterial strains with industrial interest, we will observe fascinating exceptions to the ''rules'' that govern the identities and capabilities of bifidobacteria. We can also remain optimistic with regards to achieving a more quantitative and rigorous genome-based taxonomy as opposed to what is possible today with only a handful of molecular chronometers.

In addition to other bacteria, commensals such as bifidobacteria must co-exist with their host and must evade or survive the diversity of responses that the host has developed to eliminate pathogenic bacteria. Understanding how the human immune system can differentiate between beneficial and harmful bacteria is a serious challenge for the future.

Despite the advanced insights into microbial composition, bifidobacterial activity and bifidobacteria-host interactions in the GIT, these new findings also underline our limited understanding of the processes ongoing in this environmental niche. Recently, genes induced upon exposure to acid (Azcarate et al. 2005) and bile salts (Bron et al. 2004) were identified in LAB that might confer important roles in GIT survival. Follow-up in vivo studies evaluating the survival of acid- and bile-tolerance gene-deletion mutants in relation to wild-type strains, together with the corresponding host-cell responses, would provide a useful adjunct to understand the importance of acidic and/or bile tolerance on probiotic functionality in these bacteria. Another useful means to explore the genetic determinants involved in bacterial-host interaction is represented by whole-genome transcript profiling with DNA microarrays. Clinically relevant, global transcript profiles in response to ingestion of commensal strains are starting to become available for LAB such as *Lactobacillus*. Transcriptome profiles of duodenal mucosal cell biopsy samples collected from humans consuming *Lactobacillus rhamnosus* LGG revealed distinct sets of probiotic-regulated genes involved in immune modulation, cell growth and cell signaling (Di Caro et al. 2005). Similar studies for bifidobacteria will be extremely useful not only to understand how gut residence is achieved but also to assign functions to the many unknown proteins encoded by bifidobacterial genomes.

Acknowledgements This work was financially supported by the Italian Award for Outstanding Young Researcher scheme ''Incentivazione alla mobilita' di studiosi stranieri e italiani residenti all'estero'' and to the Marie Curie Reintegration Grant (MERG-CT-2005-03080) to MV, by an IRCSET Embark postdoctoral fellowship scheme 2005 to CC, by the Higher Education Authority Programme for Research in Third Level Institutions and by the Science Foundation Ireland Alimentary Pharmabiotic Centre located at University College Cork. The research work from RSG's lab was supported by a grant from the National Science and Engineering Research Council of Canada.

References

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169

Azcarate-Peril MA, Altermann E, Lick S, Russell WM, Klaehammer TR (2005) Microarray analysis of a two-component regulatory system involved in acid resistance and proteolytic activity in *Lactobacillus acidophilus*. Appl Environ Microbiol 71:5794–5804

Bartosch S, Fite A, Macfarlane GT, McMurdo ME (2004) Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. Appl Environ Microbiol 70:3575–3581

Biavati B, Mattarelli P (2001) The family *Bifidobacteriaceae*. In: Dworkin M, et al (eds) The prokaryotes: an evolving electronic resource for the microbiological community

Bininda-Emonds ORP (2004) The evolution of supertrees. Trends Ecol Evol 19:315–322

Bodrossy L, Sessitsch A (2004) Oligonucleotide microarrays in microbial diagnostics. Curr Opin Microbiol 7:245–254

Bron PA, Marco M, Hoffer SM, van Mullekom E, de Vos WM, Kleerebezem M (2004) Genetic characterization of the bile salt response in *Lactobacillus plantarum* and analysis of responsive promoters in vitro and in situ in the gastrointestinal tract. J Bacteriol 186:7829–7835

Brown DP, Idler KB, Katz L (1990) Characterization of the genetic elements required for site-specific integration of plasmid pSE211 in Saccharopolyspora erythraea. J Bacteriol 172:1877–1888

Brown JR, Douady CJ, Italia MJ, Marshall WE, Stanhope MJ (2001) Universal trees based on large combined protein sequence data sets. Nat Genet 28:281–285

Campbell AM (1992) Chromosomal insertion sites for phages and plasmids. J Bacteriol 174:7495–7499

Chang DE, Smalley DJ, Tucker DL, other authors (2004) Carbon nutrition of *Escherichia coli* in the mouse intestine. Proc Natl Acad Sci USA 101:7427–7432

Coenye T, Gevers D, Van de Peer Y, Vandamme P, Swings J (2005) Towards a prokaryotic genomic taxonomy. FEMS Microbiol Rev 29:147–167

Cohan FM (2001) Bacterial species and speciation. Syst Biol 50:513–524

Combes P, Till R, Bee S, Smith MC (2002) The streptomyces genome contains multiple pseudo-*att*B sites for the (phi)C31-encoded site-specific recombination system. J Bacteriol 184:5746–5752

Corfield MS, Wagner SA, Clamp JR, Kriaris MS, Hoskins LC (1992) Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. Infect Immun 60:3971–3978

Corneau N, Emond E, LaPointe G (2004) Molecular characterization of three plasmids from *Bifidobacterium longum*. Plasmid 51:87–100

Coutinho PM, Henrissat B (1999) Life with no sugars? J Mol Microbiol Biotechnol 1:307–308

Cummings JH, Englyst HN (1987) Fermentation in the human large intestine and the available substrates. Am J Clin Nutr 45:1243–1255

Daveran-Mingot ML, Campo N, Ritzenthaler P, Le Bourgeois P (1998) A natural large chromosomal inversion in *Lactococcus lactis* is mediated by homologous recombination between two insertion sequences. J Bacteriol 180:4834–4842

Davey HM, Winson MK (2003) Using flow cytometry to quantify microbial heterogeneity. Curr Issues Mol Biol 5:9–15

Degnan BA, Macfarlane GT (1993) Transport and metabolism of glucose and arabinose in *Bifidobacterium breve*. Arch Microbiol 160:144–151

Delzenne NM, Williams CM (2002) Prebiotics and lipid metabolism. Curr Opin Lipidol 13:61–67

Di Caro S, Tao H, Grillo A, Elia C, Gasbarrini G, Sepulveda AR, Gasbarrini A (2005) Effects of *Lactobacillus* GG on genes expression pattern in small bowel mucosa. Dig Liver Dis 37:320–329

Ehrmann MA, Korakli M, Vogel RF (2003) Identification of the gene for beta-fructofuranosidase of *Bifidobacterium lactis* DSM10140(T) and characterization of the enzyme expressed in *Escherichia coli*. Curr Microbiol 46:391–397

Eisen JA, Heidelberg JF, White O, Salzberg SL (2000) Evidence for symmetric chromosomal inversions around the replication origin of bacteria. BCM Biol 1:1–11

Favier CF, Vaughan EE, De Vos WM, Akkermans AD (2002) Molecular monitoring of succession of bacterial communities in human neonates. Appl Environ Microbiol 68:219–226

Frank AC, Lobry JR (1999) Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms. Gene 238:65–77

Franks AH, Harmsen HJ, Raangs GC, Jansen GJ, Schut F, Welling GW (1998) Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 64:3336–3345

Gao B, Gupta RS (2005) Conserved indels in protein sequences that are characteristic of the phylum Actinobacteria. Int J Syst Evol Microbiol 55:2401–2412

Gao B, Gupta RS (2006) Signature proteins that are distinctive characteristics of Actinobacteria and their subgrups. Antonie van Leeuwenhoek 90:69–91

Garrity GM, Holt JG (2001) The road map to the manual. In: Boone DR, Castenholz RW (eds) Bergery's manual of systematic bacteriology. Springer-Verlag, Berlin

Gevers D, Vandepoele K, Simillon C, Van de Peer Y (2004) Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Microbiol 12:148–154

Gibson GR (1999) Dietary modulation of the human gut microflora using the prebiotics oligofructose and inulin. J Nutr 129:1438S–1441S

Gogarten JP, Doolittle WF, Lawrence JG (2002) Prokaryotic evolution in light of gene transfer. Mol Biol Evol 19:2226–2238

Grohmann E, Muth G, Espinosa M (2003) Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev 67:277–301, table of contents

Grundy FJ, Henkin TM (1998) The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. Mol Microbiol 30:737–749

Guarner F, Malagelada JR (2003) Gut flora in health and disease. Lancet 361:512–519

Gueimonde M, Tolkko S, Korpimaki T, Salminen S (2004) New real-time quantitative PCR procedure for quantification of bifidobacteria in human fecal samples. Appl Environ Microbiol 70:4165–4169

Gupta RS (2001) The branching order and phylogenetic placement of species from completed bacterial genomes, based on conserved indels found in various proteins. Int Microbiol 4:187–202

Hartemink R, Rombouts FM (1999) Comparison of media for the detection of bifidobacteria, lactobacilli and total anaerobes from faecal samples. J Microbiol Methods 36:181–192

Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci USA 96:2192–2197

Hinz SW, Pastink MI, van den Broek LA, Vincken JP, Voragen AG (2005) *Bifidobacterium longum* endogalactanase liberates galactotriose from type I galactans. Appl Environ Microbiol 71:5501–5510

Hooper LV, Midtvedt T, Gordon JI (2002) How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu Rev Nutr 22:283–307

Hoskins LC, Agustines M, McKee WB, Boulding ET, Kriaris M, Niedermeyer G (1985) Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. J Clin Invest 75:944–953

Imamura L, Hisamitsu K, Kobashi K (1994) Purification and characterization of beta-fructofuranosidase from Bifidobacterium infantis. Biol Pharm Bull 17:596–602

Ivanov D, Emonet C, Foata F, Affolter M, Delley M, Fisseha M, Blum-Sperisen S, Kochhar S, Arigoni F (2006) A serpin from the gut bacterium *Bifidobacterium longum* inhibits eukaryotic elastase-like serine proteases. J Biol Chem 281:17246–17252

Janer C, Rohr LM, Pelaez C, Laloi M, Cleusix V, Requena T, Meile L (2004) Hydrolysis of oligofructoses by the recombinant beta-fructofuranosidase from *Bifidobacterium lactis*. Syst Appl Microbiol 27:279–285

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

Katayama T, Sakuma A, Kimura T, Makimura Y, Hiratake J, Sakata K, Yamanoi T, Kumagai H, Yamamoto K (2004) Molecular cloning and characterization of Bifidobacterium bifidum 1,2-alpha-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95). J Bacteriol 186:4885–4893

Kitaoka M, Tian J, Nishimoto M (2005) Novel putative galactose operon involving lacto-*N*-biose phosphorylase in *Bifidobacterium longum*. Appl Environ Microbiol 71:3158–3162

Klaenhammer TR, Barrangou R, Buck BL, Azcarate-Peril MA, Altermann E (2005) Genomic features of lactic acid bacteria effecting bioprocessing and health. FEMS Microbiol Rev 29:393–409

Klijn A, Mercenier A, Arigoni F (2005) Lessons from the genomes of bifidobacteria. FEMS Microbiol Rev 29:491–509

Korbel JO, Snel B, Huynen MA, Bork P (2002) SHOT: a web server for the construction of genome phylogenies. Trends Genet 18:158–162

Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE (2001) Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature 414:555–558

Kullen MJ, Brady LJ, O'Sullivan DJ (1997) Evaluation of using a short region of the *rec*A gene for rapid and sensitive speciation of dominant bifidobacteria in the human large intestine. FEMS Microbiol Lett 154:377–383

Kullin B, Abratt VR, Reid SJ (2006) A functional analysis of the *Bifidobacterium longum csc*A and *scr*P genes in sucrose utilization. Appl Microbiol Biotechnol

Kunin V, Ouzounis CA (2003) The balance of driving forces during genome evolution in prokaryotes. Genome Res 13:1589–1594

Kunisawa T (1995) Identification and chromosomal distribution of DNA sequence segments conserved since divergence of *Escherichia coli* and *Bacillus subtilis*. J Mol Evol 40:585–593

Kunz C, Rudloff S, Baier W, Klein N, Strobel S (2000) Oligosaccharides in human milk: structural, functional, and metabolic aspects. Annu Rev Nutr 20:699–722

Langendijk PS, Schut F, Jansen GJ, Raangs GC, Kamphuis GR, Wilkinson MH, Welling GW (1995) Quantitative fluorescence in situ hybridization of *Bifidobacterium* spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. Appl Environ Microbiol 61:3069–3075

Lay C, Sutren M, Rochet V, Saunier K, Dore J, Rigottier-Gois L (2005) Design and validation of 16S rRNA probes to enumerate members of the *Clostridium leptum* subgroup in human faecal microbiota. Environ Microbiol 7:933–946

Lecuit M, Vandormael-Pournin S, Lefort J, Huerre M, Gounon P, Dupuy C, Babinet C, Cossart P (2001) A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. Science 292:1722–1725

Lee JH, O'Sullivan DJ (2006) Sequence analysis of two cryptic plasmids from *Bifidobacterium longum* DJO10A and construction of a shuttle cloning vector. Appl Environ Microbiol 72:527–535

Lee SK, Kim YB, Ji GE (1997) Note: purification of amylase secreted from *Bifidobacterium adolescentis*. J Appl Microbiol 83:267–272

Liepke C, Adermann K, Raida M, Magert HJ, Forssmann WG, Zucht HD (2002) Human milk provides peptides

highly stimulating the growth of bifidobacteria. Eur J Biochem 269:712–718
Lievin V, Peiffer I, Hudault S, Rochat F, Brassart D, Neeser JR, Servin AL (2000) *Bifidobacterium* strains from resident infant human gastrointestinal microflora exert antimicrobial activity. Gut 47:646–652
Liu M, van Enckevort FH, Siezen RJ (2005) Genome update: lactic acid bacteria genome sequencing is booming. Microbiology 151:3811–3814
MacConaill LE, Fitzgerald GF, Van Sinderen D (2003) Investigation of protein export in *Bifidobacterium breve* UCC2003. Appl Environ Microbiol 69:6994–7001
Martin MC, Alonso JC, Suarez JE, Alvarez MA (2000) Generation of food-grade recombinant lactic acid bacterium strains by site-specific recombination. Appl Environ Microbiol 66:2599–2604
Matsuki T, Watanabe K, Tanaka R, Oyaizu H (1998) Rapid identification of human intestinal bifidobacteria by 16S rRNA-targeted species- and group-specific primers. FEMS Microbiol Lett 167:113–121
Matsuki T, Watanabe K, Tanaka R, Fukuda M, Oyaizu H (1999) Distribution of bifidobacterial species in human intestinal microflora examined with 16S rRNA-gene-targeted species-specific primers. Appl Environ Microbiol 65:4506–4512
Matsuki T, Watanabe K, Fujimoto J, Kado Y, Takada T, Matsumoto K, Tanaka R (2004) Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. Appl Environ Microbiol 70:167–173
Matsumura H, Takeuchi A, Kano Y (1997) Construction of Escherichia coli-*Bifidobacterium longum* shuttle vector transforming *B. longum* 105-A and 108-A. Biosci Biotechnol Biochem 61:1211–1212
Maze A, O'Connel-Motherway M, Fitzgerald GF, van Sinderen D (2006) Identification and characterization of a putative fructose PTS of *Bifidobacterium breve* UCC2003. FEMS Microbiol Lett (in press)
McLean MJ, Wolfe KH, Devine KM (1998) Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. J Mol Evol 47:691–696
Mengaud J, Ohayon H, Gounon P, Mege RM, Cossart P (1996) E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. Cell 84:923–932
Min B, Pelaschier JT, Graham DE, Tumbula-Hansen D, Soll D (2002) Transfer RNA-dependent amino acid biosynthesis: an essential route to asparagine formation. Proc Natl Acad Sci USA 99:2678–2683
Missich R, Sgorbati B, LeBlanc DJ (1994) Transformation of *Bifidobacterium longum* with pRM2, a constructed *Escherichia coli–B. longum shuttle* vector. Plasmid 32:208–211
Muyzer G, Smalla K (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie Van Leeuwenhoek 73:127–141
Nagaoka M, Hashimoto S, Watanabe T, Yokokura T, Mori Y (1994) Anti-ulcer effects of lactic acid bacteria and their cell wall polysaccharides. Biol Pharm Bull 17:1012–1017
O'Riordan K, Fitzgerald GF (1999) Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. FEMS Microbiol Lett 174:285–294
Ouwehand A, Isolauri E, Salminen S (2002) The role of the intestinal microflora for the development of the immune system in early childhood. Eur J Nutr 41(1), I32–37
Park MS, Lee KH, Ji GE (1997) Isolation and characterization of two plasmids from *Bifidobacterium longum*. Lett Appl Microbiol 25:5–7
Philippe H, Douady CJ (2003) Horizontal gene transfer and phylogenetics. Curr Opin Microbiol 6:498–505
Requena T, Burton J, Matsuki T, Munro K, Simon MA, Tanaka R, Watanabe K, Tannock GW (2002) Identification, detection, and enumeration of human bifidobacterium species by PCR targeting the transaldolase gene. Appl Environ Microbiol 68:2420–2427
Rocha EP (2003) DNA repeats lead to the accelerated loss of gene order in bacteria. Trends Genet 19:600–603
Rossi M, Brigidi P, Matteuzzi D (1998) Improved cloning vectors for Bifidobacterium spp. Lett Appl Microbiol 26:101–104
Rossi M, Altomare L, Gonzalez Vara y Rodriguez A, Brigidi P, Matteuzzi D (2000) Nucleotide sequence, expression and transcriptional analysis of the *Bifidobacterium longum* MB 219 *lac*Z gene. Arch Microbiol 174:74–80
Roy D (2001) Media for the isolation and enumeration of bifidobacteria in dairy products. Int J Food Microbiol 69:167–182
Rohwer F, Edwards R (2002) The phage proteomic tree: a genome based taxonomy for phage. J Bacteriol 184:4529–4535
Ruas-Madiedo P, Gueimonde M, de los Reyes-Gavilan CG, Salminen S (2006) Short communication: effect of exopolysaccharide isolated from ''viili'' on the adhesion of probiotics and pathogens to intestinal mucus. J Dairy Sci 89:2355–2358
Ryan SM, Fitzgerald GF, van Sinderen D (2005) Transcriptional regulation and characterization of a novel beta-fructofuranosidase-encoding gene from *Bifidobacterium breve* UCC2003. Appl Environ Microbiol 71:3475–3482
Ryan SM, Fitzgerald GF, van Sinderen D (2006) Screening and identification of starch, amylopectin and pullulan-degrading activities in bifidobacterial strains. Appl Environ Microbiol 72:5289–5296
Sakamoto M, Hayashi H, Benno Y (2003) Terminal restriction fragment length polymorphism analysis for human fecal microbiota and its application for analysis of complex bifidobacterial communities. Microbiol Immunol 47:133–142
Sanchez B, Champomier-Verges MC, Anglade P, Baraige F, de Los Reyes-Gavilan CG, Margolles A, Zagorec M (2005) Proteomic analysis of global changes in protein expression during bile salt exposure of

Bifidobacterium longum NCIMB 8809. J Bacteriol 187:5799–5808

Satokari RM, Vaughan EE, Akkermans AD, Saarela M, de Vos WM (2001) Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. Appl Environ Microbiol 67:504–513

Scardovi V, Trovatelli LD, Zani G, Crociani F, Matteuzzi D (1971) Deoxyribonucleic acid homology relationships among species of the genus *Bifidobacterium*. Int J Syst Bacteriol 21:276–294

Schell MA, Karmirantzou M, Snel B other authors (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. Proc Natl Acad Sci USA 99:14422–14427

Servin AL (2004) Antagonistic activities of lactobacilli and bifidobacteria against microbial pathogens. FEMS Microbiol Rev 28:405–440

Sgorbati B, Scardovi V, Leblanc DJ (1982) Plasmids in the genus *Bifidobacterium*. J Gen Microbiol 128:2121–2131

Sinha VR, Kumria R (2001) Polysaccharides in colon-specific drug delivery. Int J Pharm 224:19–38

Slovakova L, Duskova D, Marounek M (2002) Fermentation of pectin and glucose, and activity of pectin-degrading enzymes in the rabbit caecal bacterium *Bifidobacterium pseudolongum*. Lett Appl Microbiol 35:126–130

Snel B, Bork P, Huynen MA (2002) Genomes in flux: the evolution of archaeal and proteobacterial gene content. Genome Res 12:17–25

Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J, Buhler JD, Gordon JI (2005) Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science 307:1955–1959

Stackebrandt E, Schumann P (2000) Introduction to the taxonomy of Actinobacteria. In: Dworkin M, et al. (eds) The prokaryotes: an evolving electronic resource for the microbiological community

Stackebrandt E, Rainey FA, WardRainey NL (1997) Proposal for a new hierarchic classification system, Actinobacteria classis nov. Int J Syst Bacteriol 47:479–491

Suzuki A, Mitsuyama K, Koga H, Tomiyasu N, Masuda J, Takaki K, Tsuruta O, Toyonaga A, Sata M (2006) Bifidogenic growth stimulator for the treatment of ulcerative colitis: a pilot study. Nutrition 22:76–81

Tillier ER, Collins RA (2000) Genome rearrangement by replication-directed translocation. Nat Genet 26:195–197

Titgemeyer F, Hillen W (2002) Global control of sugar metabolism: a gram-positive solution. Antonie Van Leeuwenhoek 82:59–71

Trindade MI, Abratt VR, Reid SJ (2003) Induction of sucrose utilization genes from Bifidobacterium lactis by sucrose and raffinose. Appl Environ Microbiol 69:24–32

Van Laere KM, Abee T, Schols HA, Beldman G, Voragen AG (2000) Characterization of a novel beta-galactosidase from *Bifidobacterium adolescentis* DSM 20083 active towards transgalactooligosaccharides. Appl Environ Microbiol 66:1379–1384

Van Laere KM, Hartemink R, Bosveld M, Schols HA, Voragen AG (2000) Fermentation of plant cell wall derived polysaccharides and their corresponding oligosaccharides by intestinal bacteria. J Agric Food Chem 48:1644–1652

Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 60:407–438

Vaughan EE, Heilig HG, Ben-Amor K, de Vos WM (2005) Diversity, vitality and activities of intestinal lactic acid bacteria and bifidobacteria assessed by molecular approaches. FEMS Microbiol Rev 29:477–490

Ventura M, Elli M, Reniero R, Zink R (2001a) Molecular microbial analysis of *Bifidobacterium* isolates from different environments by the species-specific amplified ribosomal DNA restriction analysis (ARDRA). FEMS Microbiol Ecol 36:113–121

Ventura M, Reniero R, Zink R (2001b) Specific identification and targeted characterization of *Bifidobacterium lactis* from different environmental isolates by a combined multiplex-PCR approach. Appl Environ Microbiol 67:2760–2765

Ventura M, Zink R (2002) Rapid identification, differentiation, and proposed new taxonomic classification of *Bifidobacterium lactis*. Appl Environ Microbiol 68:6429–6434

Ventura M, Canchaya C, Meylan V, Klaenhammer TR, Zink R (2003a) Analysis, characterization, and loci of the *tuf* genes in lactobacillus and bifidobacterium species and their direct application for species identification. Appl Environ Microbiol 69:6908–6922

Ventura M, Meylan V, Zink R (2003b) Identification and tracing of Bifidobacterium species by use of enterobacterial repetitive intergenic consensus sequences. Appl Environ Microbiol 69:4296–4301

Ventura M, Zink R (2003c) Comparative sequence analysis of the *tuf* and *rec*A genes and restriction fragment length polymorphism of the internal transcribed spacer region sequences supply additional tools for discriminating *Bifidobacterium lactis* from *Bifidobacterium animalis*. Appl Environ Microbiol 69:7517–7522

Ventura M, Canchaya C, van Sinderen D, Fitzgerald GF, Zink R (2004a) *Bifidobacterium lactis* DSM 10140: identification of the *atp* (atpBEFHAGDC) operon and analysis of its genetic structure, characteristics, and phylogeny. Appl Environ Microbiol 70:3110–3121

Ventura M, Canchaya C, Zink R, Fitzgerald GF, van Sinderen D (2004b) Characterization of the *gro*EL and *gro*ES loci in *Bifidobacterium breve* UCC 2003: genetic, transcriptional, and phylogenetic analyses. Appl Environ Microbiol 70:6197–6209

Ventura M, van Sinderen D, Fitzgerald GF, Zink R (2004c) Insights into the taxonomy, genetics and physiology of bifidobacteria. Antonie Van Leeuwenhoek 86:205–223

Ventura M, Canchaya C, Bernini V, Del Casale A, Dellaglio F, Neviani E, Fitzgerald GF, van Sinderen D (2005a) Genetic characterization of the *Bifidobacterium breve* UCC 2003 *hrc*A locus. Appl Environ Microbiol 71:8998–9007

Ventura M, Lee JH, Canchaya C et al (2005b) Prophage-like elements in bifidobacteria: insights from genomics, transcription, integration, distribution, and phylogenetic analysis. Appl Environ Microbiol 71:8692–8705

Ventura M, Zhang Z, Cronin M, Canchaya C, Kenny JG, Fitzgerald GF, van Sinderen D (2005c) The ClgR protein regulates transcription of the *clp*P operon in *Bifidobacterium breve* UCC 2003. J Bacteriol 187:8411–8426

Ventura M, Zink R, Fitzgerald GF, van Sinderen D (2005d) Gene structure and transcriptional organization of the *dna*K operon of *Bifidobacterium breve* UCC 2003 and application of the operon in bifidobacterial tracing. Appl Environ Microbiol 71:487–500

Ventura M, Canchaya C, Zhang Z, Fitzgerald GF, van Sinderen D (2006a) How high G + C Gram-positive bacteria and in particular bifidobacteria cope with heat stress: protein players and regulators. FEMS Microbiol Rev 30:734–759

Ventura M, Canchaya C, Del Casale A, Dellaglio F, Neviani E, Fitzgerald GF, van Sinderen D (2006b) Analysis of bifidobacterial evolution using a multilocus approch. Int J Syst Evol Microbiol (in press)

Warchol M, Perrin S, Grill JP, Schneider F (2002) Characterization of a purified beta-fructofuranosidase from *Bifidobacterium infantis* ATCC 15697. Lett Appl Microbiol 35:462–467

Ward RE, Ninonuevo M, Mills DA, Lebrilla CB, German JB (2006) In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri. Appl Environ Microbiol 72:4497–4499

Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA 97:8392–8396

Wolf YI, Rogozin IB, Kondrashov AS, Koonin EV (2001) Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res 11:356–372

Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV, Gordon JI (2003) A genomic view of the human *Bacteroides thetaiotamicron* symbiosis. Science 299:2074–2076

Yildirim Z, Winters DK, Johnson MG (1999) Purification, amino acid sequence and mode of action of bifidocin B produced by *Bifidobacterium bifidum* NCFB 1454. J Appl Microbiol 86:45–54

You HJ, Oh DK, Ji GE (2004) Anticancerogenic effect of a novel chiroinositol-containing polysaccharide from *Bifidobacterium bifidum* BGN4. FEMS Microbiol Lett 240:131–136

Zarate G, Chaia AP, Gonzales S, Oliver G (2000) Viability and beta-galactosidase activity of dairy propionibacteria subjected to digestion by artificial gastric and intestinal fluids. J Food Protect 63:1214–1221

Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans AD, de Vos WM (2002) Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl Environ Microbiol 68:3401–3407

Zoetendal EG, Vaughan EE, de Vos WM (2006) A microbial worlds within us. Mol Microbiol 59:1639–1650

Zoetendal EG, Collier CT, Koike S, Mackie RI, Gaskins HR (2004) Molecular ecological analysis of the gastrointestinal microbiota: a review. J Nutr 134:465–472