Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging

Francesca Turroni[a], Francesca Bottacini[a,b,c], Elena Foroni[a], Imke Mulder[d], Jae-Han Kim[e], Aldert Zomer[b], Borja Sánchez[f], Alessandro Bidossi[g], Alberto Ferrarini[h], Vanessa Giubellini[a], Massimo Delledonne[h], Bernard Henrissat[i], Pedro Coutinho[i], Marco Oggioni[g], Gerald F. Fitzgerald[b,c], David Mills[e], Abelardo Margolles[f], Denise Kelly[d], Douwe van Sinderen[b,c,1], and Marco Ventura[a,1,2]

[a]Laboratory of Probiogenomics, Department of Genetics, Biology of Microorganisms, Anthropology and Evolution, University of Parma, 43100 Parma, Italy; [b]Alimentary Pharmabiotic Centre and [c]Department of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland; [d]The Rowett Institute of Nutrition and Health, College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, AB21 9SB United Kingdom; [e]Department of Viticulture and Enology, The Robert Mondavi Institute for Wine and Food Sciences, University of California, Davis, CA 95616; [f]Departamento de Microbiologia y Bioquimica de Productos Lacteos, Instituto de Productos Lácteos de Asturias Consejo Superior de Investigaciones Científicas, Villaviciosa, 3300 Asturias, Spain; [g]Department of Molecular Biology, University of Siena, 53100 Siena, Italy; [h]Dipartimento Sientifico Tecnologico, Università degli Studi di Verona, 37134 Verona, Italy; and [i]Glycogenomics, Databases and Bioinformatics, Architecture et Fonction des Macromolécules Biologiques, Universités Aix-Marseille, 13288 Marseille, France

Edited* by Todd R. Klaenhammer, North Carolina State University, Raleigh, NC, and approved October 4, 2010 (received for review July 29, 2010)

The human intestine is densely populated by a microbial consortium whose metabolic activities are influenced by, among others, bifidobacteria. However, the genetic basis of adaptation of bifidobacteria to the human gut is poorly understood. Analysis of the 2,214,650-bp genome of *Bifidobacterium bifidum* PRL2010, a strain isolated from infant stool, revealed a nutrient-acquisition strategy that targets host-derived glycans, such as those present in mucin. Proteome and transcriptome profiling revealed a set of chromosomal loci responsible for mucin metabolism that appear to be under common transcriptional control and with predicted functions that allow degradation of various O-linked glycans in mucin. Conservation of the latter gene clusters in various *B. bifidum* strains supports the notion that host-derived glycan catabolism is an important colonization factor for *B. bifidum* with concomitant impact on intestinal microbiota ecology.

coevolution | genomics | host-glycans metabolism | human gut intestinal bacteria | mucin

Bifidobacteria are one of the dominant bacterial groups found in the intestinal microbiota of infants (1). The intestinal microbiota has been compared with a metabolic organ, extracting energy from dietary components that otherwise would escape metabolism (2). For this purpose the intestinal microbiota produces various glycolytic enzymes to ferment otherwise indigestible glycans into short-chain fatty acids that in turn are used by the host (3). The molecular strategies used by members of the intestinal microbiota to harvest and degrade complex glycans are important for understanding the genetic and associated metabolic properties that underpin ecological fitness in and adaptation to the human intestinal environment. Apart from dietary components, host-derived glycans are believed to constitute a nutrient resource for (certain) members of the intestinal microbiota (3) and thus may influence the composition and activities of this complex microbial consortium. Indeed, in the absence of dietary nutrients colonization of the intestinal microorganism *Bacteroides thetaiotaomicron* is reliant on host-derived glycans, which it metabolizes by means of polysaccharide utilization loci (4), showing that under such circumstances endogenous carbohydrates influence the composition of the intestinal microbiota (3). Recently, another constituent of human gut microbiota, *Akkermansia muciniphila*, was identified as an important mucin degrader (5, 6), but little is known regarding the genetic elements required for this property.

Breast-fed infants develop an enteric microbiota that typically contains high levels of bifidobacteria (1, 7) in which species like *Bifidobacterium bifidum* are abundant (8). The ability of *B. bifidum* to grow on mucin has been noted previously (9), and this organism also possesses a metabolic pathway for the degradation of lacto-*N*-biose and galacto-*N*-biose (10, 11), which constitute the building blocks of human milk oligosaccharides (HMO) and the core 1 structure of mucin-type O-glycan, respectively (12). However, information on the molecular mechanisms governing the complete metabolism of these glycans by bifidobacteria is limited, and *B. bifidum*, *Bifidobacterium longum* subsp. *infantis* (13), and *B. thetaiotaomicron* (14) thus may serve as models for studying the interaction between members of the intestinal microbiota and the host mucosa to uncover host features that determine intestinal colonization.

Genome sequences of various commensal bifidobacteria are publicly available (13, 15–18), and such sequences will be crucial in unraveling the intricate interactions that must exist between host and its resident (bifido)bacteria. Here, we report on the genome analysis of *B. bifidum* PRL2010, revealing putative metabolic traits that underpin this case of host–microbe coevolution.

Results and Discussion

General Genome Features of *B. bifidum* PRL2010. The chromosome of *B. bifidum* PRL2010 consists of 2,214,650 bp, with a guanine and cytosine (G+C) content of 62.66%, higher than that of other bifidobacterial genomes (13, 15–18) although still within the typical range of *Actinobacteria* (19). Genome features are presented in Table S1; functional distribution of gene products assigned to clusters of orthologous groups (COGs) of proteins is typical of that of a bifidobacterial genome, and functional roles were assigned to 67.5% of the predicted ORFs. Homologs from

Author contributions: G.F.F., D.M., A.M., D.K., D.v.S., and M.V. designed research; F.T., F.B., E.F., and I.M. performed research; J.H.K., B.S., A.F., and M.D. contributed new reagents/analytic tools; A.Z., A.B., V.G., B.H., P.C., and M.O. analyzed data; and D.v.S. and M.V. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The sequence reported in this article has been deposited in the GenBank database (accession number CP001840). The array data have been deposited in the GEO database (accession no. GSE21976).

[1]D.v.S. and M.V. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: marco.ventura@unipr.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1011100107/-/DCSupplemental.

other bacterial species with unknown function were identified for an additional 23.4% of the *B. bifidum* PRL2010 ORFs; the remaining 9.1% appears to be unique to *B. bifidum,* considerably higher than in other bifidobacterial genomes (13, 15–18). Not surprisingly, the PRL2010 genome is very similar (96% at the nucleotide level) to the partially sequenced *B. bifidum* NCIMB 41171 (National Center for Biotechnology Information, NZ_ABQP00000000). The *B. bifidum* PRL2010 genome also is quite similar (89% at the nucleotide level) to that of *B. longum* subsp. *infantis* ATCC15697 (13). *B. bifidum* PRL2010 and *B. longum* subsp. *infantis* ATCC15697 belong to two distinct phylogenetic clusters, and their similarity may reflect their sharing a common ecological niche (19). A large proportion (36%) of genes common to both *B. bifidum* genomes but absent in other bifidobacteria can be assigned to the COG family of carbohydrate metabolism and transport, including genes predicted to be involved in mucin metabolism (as discussed below). To investigate the phylogenetic position of *B. bifidum* PRL2010 within the genus *Bifidobacterium* (19) and for comparison with other members of the human intestinal microbiota, a BLAST heatmap based on the nonredundant BLAST database of currently available genome sequences of bifidobacteria and human gut microbial genomes (http://ncbi.nlm.nih.gov) (20, 21) was produced. The organism distribution at species level was determined for each predicted ORF, and corresponding similarity values were grouped into predefined ranges (Fig. S1 *A* and *B*). Not surprisingly, the most intense heat flare, which corresponds to DNA regions of the genome against which PRL2010 was paired, was obtained from the comparison with *B. bifidum* NCIMB 41171. A second dominant flare was detected from the comparative analysis with the *B. longum* subsp. *infantis* genome (Fig. S1*A*). In contrast, fewer and less intense heat flares were observed in the heatmap produced using information from the available sequence information of human gut microbiome projects (Fig. S1*B*). Notably, the majority of the BLAST hits had E-values ranging between 10^{-10} and 10^{-50} (Fig. S1*B*), suggesting an ancient diversification and/or substantial gene loss/gain during evolution of the *B. bifidum* taxa. Apart from the genes that predict the ability of *B. bifidum* PRL2010 to use mucin (discussed below), there is genomic evidence for other interactions between *B. bifidum* PRL2010 and its human host, such as the putative adhesion factor BBPR_0612 possessing 98% identity to BopA from *B. bifidum* MIMBb75 (22) (Fig. S2*B*). Analysis of the *B. bifidum* PRL2010 genome identified four additional loci encoding proteins with domains related to adhesion and host colonization (Fig. S2); three of these loci encode pilus-like structures, which may have a role in bacterial adhesion to intestinal mucus (23).

Basic Metabolism. Homologs of all enzymes necessary for the fermentation of glucose and fructose to lactic acid and acetate by means of the characteristic fructose-6-phosphate shunt (bifid shunt), as well as a partial Embden–Meyerhoff pathway (24), were annotated in the genome of *B. bifidum* PRL2010, which appears to be typical of bifidobacteria. Carbon metabolism has been characterized quite extensively in bifidobacteria (25), but the provision and utilization of nitrogen by bifidobacteria has received much less scientific scrutiny. Inorganic nitrogen is likely to be the preferred nitrogen source for most bifidobacteria, because ammonium is predicted to be imported by a dedicated transporter (BBPR_1693), homologs of which are present in all sequenced bifidobacterial genomes. Moreover, *B. bifidum* PRL2010 may use deamination of *N*-acetylglucosamine and *N*-acetylgalactosamine present in mucin or other human- or diet-derived hexosamines (e.g., HMOs), or it may use peptide degradation by dedicated peptidases, for which encoding genes are present in abundance on its genome (2.06% of the total ORFs).

The genome of *B. bifidum* PRL2010 contains 20 complete ATP-binding cassette (ABC) transporters predicted to be involved in the transport of dietary carbohydrates on the basis of the transporter classification (TC) database (26) and four complete phosphoenolpyruvate-phosphotransferase systems, a smaller and larger number, respectively, than in most other characterized bifidobacterial genomes. A much larger proportion of the ABC transporters is dedicated to efflux than to uptake.

Similar to other bifidobacterial genomes, complete biosynthetic pathways for purines and pyrimidines from glutamine, as well as for riboflavin, thiamine, and folate, were identified; no homologs were present for complete pathways of other B vitamins (13, 15–18).

Analysis of the *B. bifidum* PRL2010 genome revealed the presence of conventional mobilome candidates that may have been acquired through horizontal gene transfer (details are given in *SI Text* and Fig. S1*C*) and may provide important ecological advantages and also influence chromosome structure (19).

Mucin Utilization by Bifidobacteria. The fermentation ability of *B. bifidum* PRL2010 seems to be limited to a relatively small number of carbohydrate sources but does include complex sugars such as HMO and mucin (Fig. S3*A*). Although a large selection of representative strains for many species of the genus *Bifidobacterium* was used to evaluate their ability to use mucin as a sole carbon source (19), only strains belonging to the *B. bifidum* species were capable of growing on mucin-based medium (Fig. S3*B*). No major differences in growth on mucin-based medium were evident among the *B. bifidum* isolates, although strains 85B and L22 exhibited the least and strain PRL2010 exhibited the best growth on mucin (Fig. S3*B*). Therefore strain PRL2010 was selected as a model to study mucin metabolism in *Bifidobacterium*.

Genomics of Host-Glycan Utilization. Mucus functions as a protective, semipermeable barrier located on the epithelial surfaces of the gastrointestinal tract, where it provides critical functions supporting the health status of the host. The mucus layer also constitutes an important site for adhesion and colonization of gut bacteria, especially in the outer, loose mucus layer (27), as well as a rich nutritional reservoir of carbohydrates for enteric bacteria and a barrier against the penetration of pathogens to gut epithelial surfaces (27). Mucin consists of glycoprotein components with a peptide structure that contains alternating O-linked glycosylated and N-linked glycosylated domains. N-glycans consist of oligosaccharides usually connected through an *N*-acetylglucosamine (GlcNAc) linkage to asparagine, whereas O-linked oligosaccharides can be variously attached to serine or threonine through fucose, glucose, mannose, xylose, fucose, arabinose, and other sugars, including the most commonly detected mucin-type O-linked *N*-acetylgalactosamine (GalNAc) (12). The principal constituent mucin monosaccharides, GlcNAc, GalNAc, and galactose, often are decorated with fucose, sialic acid, and sulfate groups (28). Given the diversity and complexity of mucin structures found within the gut (12, 29), specific strategies for deconstructing these molecules must be inherent features in the genomes of mucin-using bacteria. The *B. bifidum* PRL2010 genome encodes various glycosyl hydrolases putatively implicated in degradation of mucin-derived oligosaccharides, including a predicted cell wall-anchored endo-α-*N*-acetylgalactosaminidase (BBPR_0264), an enzyme that has been shown previously to catalyze the hydrolysis of the O-glycosidic α-linkage between GalNAc and serine/threonine residues of various mucin-type glycoproteins (30–32). Moreover, the genome of *B. bifidum* PRL2010 encodes a putative 1,2-α-L-fucosidase (BBPR_0193), as well as a predicted 1,3/4-α-L-fucosidase (BBPR_1360), which releases various α-linked L-fucoses from the oligosaccharide core of the mucin structure (33–35). Both fucosidases contain a signal peptide, but only BBPR_0193 contains an LPXTG motif, suggesting that this enzyme is secreted and anchored to the cell wall, whereas the presumed fucosidase encoded by BBPR_1360 contains two transmembrane domains, indicating

that it is bound to the cell membrane. Additional glycosyl hydrolases involved in the breakdown of mucin-derived oligosaccharides were detected in the predicted proteome of *B. bifidum* PRL2010, including four putative *N*-acetyl-β-hexosaminidases (BBPR_0026, BBPR_1018, BBPR_1514, and BBPR_1529), four predicted β-galactosidases (BBPR_150, BBPR_0482, BBPR_1355, and BBPR_1460), and two putative exo-α-sialidases (BBPR_1793 and BBPR_1794) scattered across the PRL2010 chromosome. Interestingly, homologs of two putative *N*-acetyl-β-hexosaminidases (BBPR_1018 and BBPR_1529), as well as a predicted β-galactosidase (BBPR_0482) of *B. bifidum* JCM1254, are involved in the degradation of lacto-*N*-neotetraose (Galβ1–4GlcNAcβ1–3Galβ1–4Glc) present in the core structures of HMOs as well as in core 2 mucin-type O-glycans (36). Notably, and similar to what was described for their homologs (36), the galactosidase (BBPR_0482) and both *N*-acetyl-β-hexosaminidases (BBPR_1018 and BBPR_1529) possess an N-terminal signal sequence and a C-terminal transmembrane region, suggesting that they act as extracellular membrane-bound enzymes. In contrast, the two other putative *N*-acetyl-β-hexosaminidases (BBPR_0026 and BBPR_1514) are predicted to represent intracellular proteins. Moreover, apart from this glycolytic potential, the molecular basis for transmembrane import of mucin-derived oligosaccharides is evident from various ABC transport systems whose corresponding genes are located in the vicinity of each of the above-mentioned genes. Only *B. bifidum* and *B. longum* subsp. *infantis* species are known to be capable of degrading galacto-*N*-

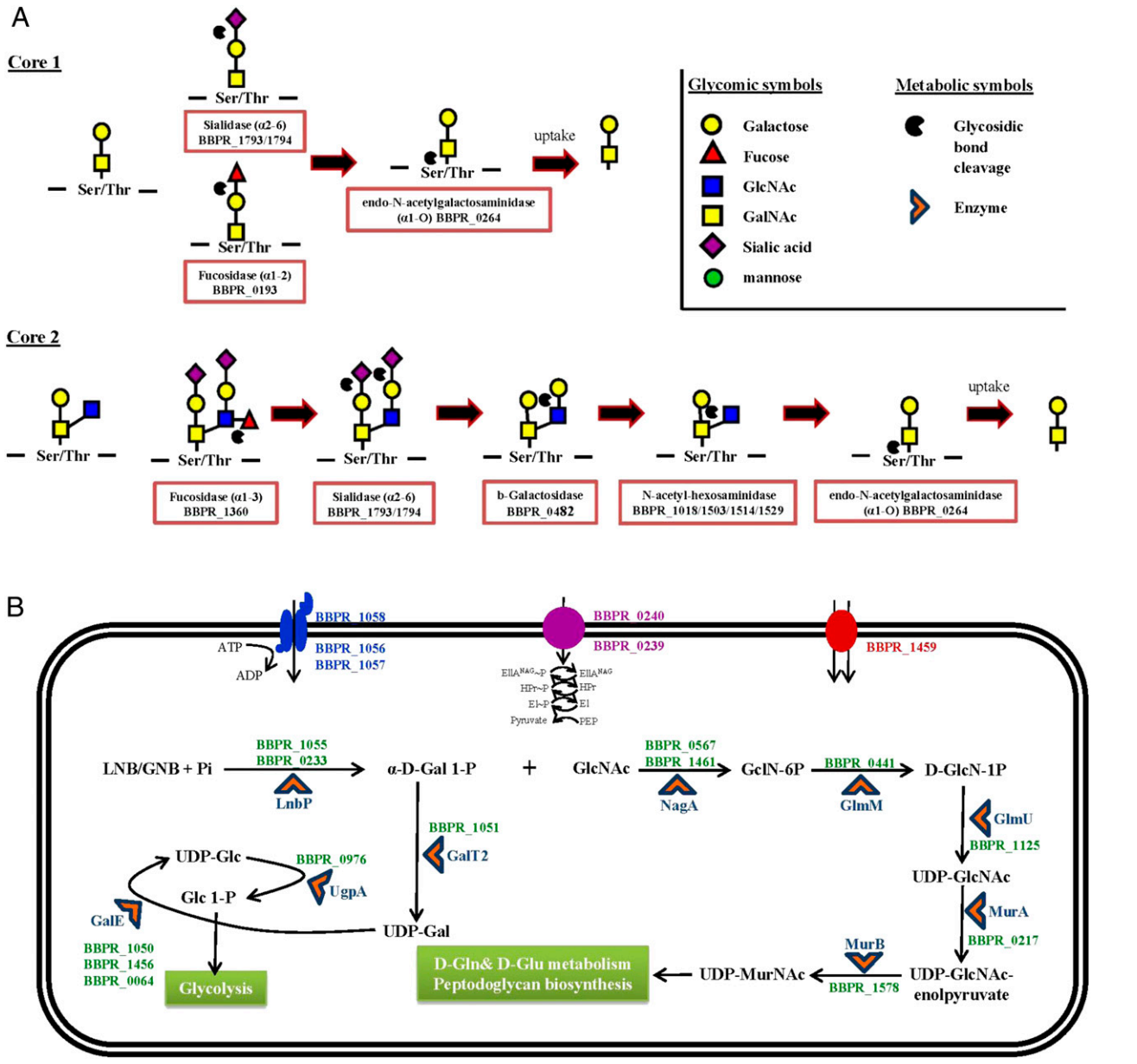

Fig. 1. *B. bifidum* PRL2010 metabolism of the major core types of mucin O-glycans (39). (*A*) Predicted target sites of various enzymes involved in the degradation of major core 1 and 2 O-glycans. Ser, serine; Thr, threonine. (*B*) Schematic representation of a cell and metabolic pathways for mucin-derived oligo- and monosaccharides. Carbohydrates derived from mucin are internalized by ABC transporters (shown in blue), PTS systems (shown in violet), and a symporter (depicted in red), after which glycosyl hydrolases process such oligosaccharides to monosaccharides. The different ORFs of *B. bifidum* PRL2010 encoding the presumed enzymes involved in the breakdown of O-glycans are indicated (see Table S2 for details about function). Nomenclature was taken from www.functionalglycomics.org. UDP-MurNAC, uridine diphosphate-N-acetylmuramic acid.

tetraose (Gal β1,3GalNAc β1,3Galβ1,4GlcNAc, GNT), which constitutes a common mucin oligosaccharide structure present in extended core 2 mucin-type O-glycans (29) and lacto-*N*-tetraose (Galβ1–3GlcNAcβ1–3Galβ1–4Glc), which forms the core component of HMO, via galacto-*N*-biose (Galβ1–3GalNAc, GNB) or lacto-*N*-biose (Galβ1–3GlcNAc, LNB) intermediates, respectively (11). A clear homolog of the operon that is responsible for GNB/LNB utilization in *B. longum* JCM1217 (10, 11, 37, 38) is present on the genome of *B. bifidum* PRL2010 (BBPR_1050–BBPR_1058) (Fig. 1, Table S2, and Fig. S4). This locus encodes a putative ABC-type oligosaccharide uptake system (BBPR_1056–BBPR_1058), whereas the adjacent gene (BBPR_1055) is a clear homolog of *lnp*A previously shown to be involved in the breakdown of disaccharides such as LNB and GNB (37). This observation suggests a specific LNB-utilization pathway for *B. bifidum* that is different from that found in *B. longum* subsp. *infantis* ATCC15697 (13), in which *B. bifidum* PRL2010 acquires mucin- or HMO-associated LNB/GNB by extracellular enzymes [such as 1,2-α-L-fucosidase (BBPR_0193) and α-1,3/4-fucosidase (BBPR_1360)] that catalyze defucosylation of LNB/GNB-containing oligosaccharides/O-glycans to facilitate further degradation by the action of a lacto-*N*-biosidase (BBPR_1438), endo-α-*N*-acetylgalactosaminidase (BBPR_0264), and perhaps other enzymes (11, 39). This finding confirms the idea of two converging metabolic strategies for the utilization of both mucin and HMOs by this bacterium, as was proposed previously (11).

To determine the phylogenetic distribution of the presumed genetic determinants for mucin degradation, genome variability was investigated by comparative genome hybridization (CGH) using *B. bifidum* PRL2010-based microarrays (Fig. 2, and *SI Text*). Interestingly, the presumed mucin-utilization genes were shown to be conserved within the *B. bifidum* taxon, thus generalizing the current understanding of mucin metabolism in this human-associated species (Fig. 2). This notion also is supported by the mucin-dependent growth profiles shown in Fig. S3*B*. Interestingly, the absence of two genes encoding a putative carbohydrate transporter (BBPR_1056) and a predicted fucosidase (BBPR_0193) seems to be linked to strains that exhibit reduced growth on mucin-containing media (Fig. 2 and Fig. S3).

Proteome and Transcriptome Analyses of *B. bifidum* PRL2010 and Adaptation to Mucin Utilization. To substantiate the notion that *B. bifidum* PRL2010 contains specific genes dedicated to the utilization of mucin, we first investigated the proteome of *B. bifidum* PRL2010 grown in the presence of mucin or HMO as its sole carbon source and compared these results with data obtained when the strain was grown on lactose. Notably, five categories of enzymes were expressed specifically in cells grown on mucin relative to lactose. These proteins are assumed to possess amino sugar catabolic functions, and among such mucin-induced proteins we identified a variety of glycosyl hydrolases (Fig. S5 and Table S3). As expected, the expression levels of the proteins constituting the LNB pathway of *B. bifidum* PRL2010 were higher when cells were cultivated on mucin and HMOs than when grown on lactose (Table S3). The expression of various proteins implicated in ABC-type sugar transport and phosphotransferase (PTS) systems with predicted specificity for *N*-acetylglucosamine (BBPR_0240), glucose (BBPR_0239), and a possible PTS system for lactose/cellobiose uptake (BBPR_0030) was clearly induced when *B. bifidum* PRL2010 was cultivated in mucin-based medium, perhaps because the substrates for these PTS systems become available during mucin degradation and act as inducers. Furthermore, various predicted aminopeptidases, which may be involved in the proteolysis of the proteinaceous component of mucin, also exhibited increased expression (Table S3).

To expand our understanding of the genetic adaptation of *B. bifidum* PRL2010 to mucin utilization, we compared the transcriptome of *B. bifidum* PRL2010 grown on mucin with that of *B. bifidum* PRL2010 grown on lactose. (The complete list of genes whose expression pattern was altered following cultivation of *B. bifidum* PRL2010 on mucin is presented in Fig. S5 and Table S4.) In total, 56 genes displayed increased (more than fivefold) expression when PRL2010 was cultured in the presence

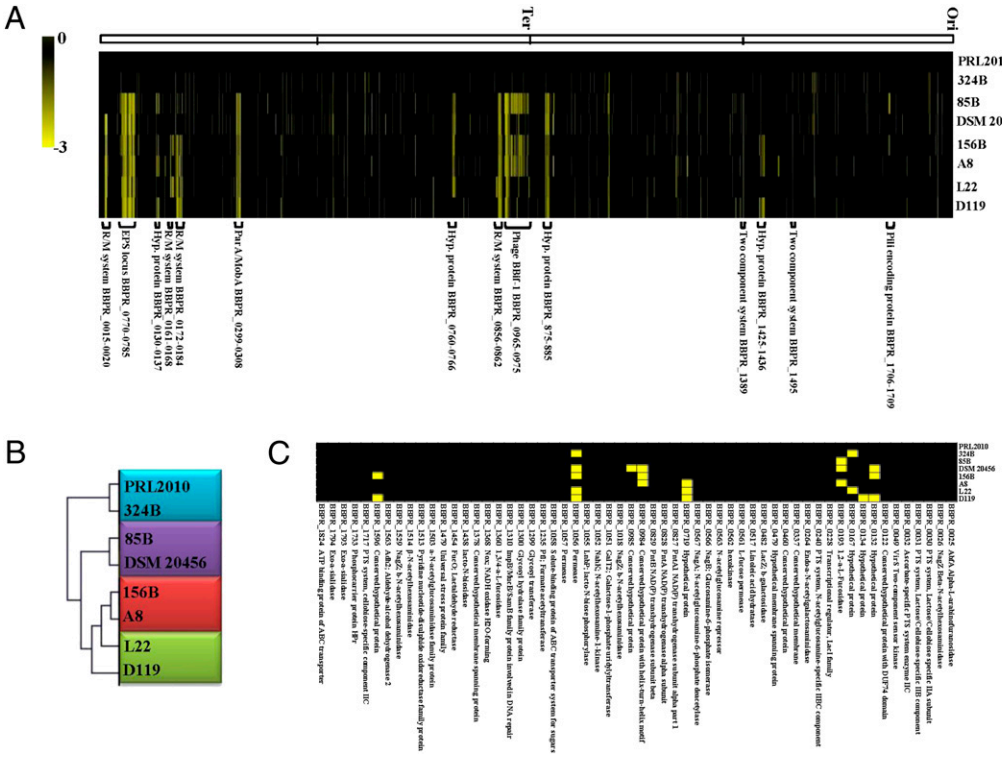

Fig. 2. Genomic diversity in the *B. bifidum* species with reference to the *B. bifidum* strain PRL2010 genome. (*A*) CGH data. Each horizontal row corresponds to a probe on the array, and genes are ordered vertically according to their position on the PRL2010 genome. Columns represent analyzed strains, which are indentified by their code numbers. The color code, which goes from black to yellow to indicate the presence, divergence, or absence of a gene sequence, is given at the left of the figure. The predicted functions of some relevant genes are shown in the right-hand margin. Ori, origin of replication; ter, terminus of replication. (*B*) CGH-based clustering data. A CGH-based clustering analysis was performed for the eight *B. bifidum* strains analyzed. (*C*) Presence (black) or absence (yellow) of genes predicted to be involved in mucin catabolism.

of mucin as the sole carbon source, and a portion of these mucin-induced genes also had been identified from the proteomics analysis (Fig. S5). Transcriptional profiling allowed identification of several mucin-induced genes encoding secreted or cell envelope-associated enzymes that had not been identified from the mucin-dependent *B. bifidum* PRL2010 proteome. In this way, PRL2010 transcriptome profiling identified the following mucin-induced genes that encode extracellular proteins: BBPR_1793 and BBPR_1794, both encoding putative exo-α-sialidases, and BBPR_0193 and BBPR_1360, which specify the 1,2-α-L-fucosidase and 1,3/4-α-L-fucosidase (Table S4). Other genes whose transcription was induced on mucin included sugar transport-encoding genes such as PTS systems (e.g., the locus spanning BBRP_0030–BBRP_0032) and ABC-type carriers (e.g., BBRP_1058) and specific permeases (e.g., the fucose permease encoded by BBRP_0561) for the intake of mucin-derived carbohydrates (Table S4). In addition, the transcription levels of two genes, BBRP_0025 and BBRP_1300, which encode predicted extracellular enzymes α-L-arabinofuranosidase and α-1,3-galactosidase, respectively, were shown to be more than 50-fold higher when grown in mucin than when grown in lactose (Table S4). Notably, arabinose is part of a specific type of O-linked glycan in which the monosaccharide is linked to the hydroxyl of the amino acid residue (40); thus the α-L-arabinofuranosidase may be involved in the liberation of arabinose from such an O-linked glycan, and the α-1,3-galactosidase may be active in the hydrolysis of terminal α-galactosyl moieties from glycoproteins (41). Furthermore, the transcription of three genes encoding putative transcriptional regulators (BBRP_0228, BBRP_0563, and BBRP_0984) (Table S4) was enhanced when *B. bifidum* PRL2010 was grown on mucin, suggesting roles for these proteins in the regulation of mucin utilization. Interestingly, bioinformatic analysis of the putative promoter regions located upstream of the genes of *B. bifidum* PRL2010, whose level of transcription was significantly higher when grown on mucin, revealed the presence of a conserved 24-bp inverted repeat in many such promoter regions (Fig. S6 and Table S4). This inverted repeat may represent a regulatory element involved in global control of mucin-dependent transcription in PRL2010.

To investigate if the porcine mucin used in the above in vitro experiments elicits a response similar to that of human mucin, we evaluated host gene expression using two human intestinal cell lines (Caco-2 and HT-29) exposed to *B. bifidum* PRL2010 (Fig. S5 *E* and *F* and *SI Text*). When we looked at genes involved in regulating host glycan synthesis, we observed a fourfold ($P < 0.014$) increase in the expression of UDP-GlcNAc:βGal-β-1,3-*N*-acetylglucosainyltransferase5 (*B3GNT5*) (Fig. S5*F*). The product of this gene plays a key role in the synthesis of lacto- or neolacto-series carbohydrate chains on glycoconjugates, notably participating in the biosynthesis of Lewis X carbohydrate structures (42). *B3GNT5* catalyses the formation of the Lc3 structure, which is the core of the lacto-series and promotes GlcNAc transfer to glycoconjugate substrates. These glycan structures can provide adhesion sites for bacteria and hence may be important for attachment and signaling responses to colonizing PRL2010 bacteria.

Conclusions

Several ecological studies have shown that bifidobacteria are a dominant bacterial group of the infant microbiota as well as a key component of the adult-type intestinal microbiota (8, 43). It is believed that bifidobacteria are well adapted to maximize metabolic access to a wide variety of diet-derived and/or host-derived sugars (15–17). The host-derived glycans act not only as an energy source but also as attachment sites for adhesion proteins produced by commensal and pathogenic bacteria (23). The binding specificities of such lectins may allow host-dependent selection of commensal gut microorganisms, although the mechanisms and effects of microbiota association with host glycans are not well understood. Bacteria belonging to various genera, including bifidobacteria, have been shown to degrade mucin (5, 44), although the metabolic and regulatory machinery responsible for mucin utilization by bifidobacteria has remained elusive so far (9). A thorough understanding of how mucin foraging contributes to bifidobacterial colonization and persistence is contingent on the ability to deconstruct the complex gut ecosystem and to decipher the contributions and activities of its component parts. Thus, the investigation of host glycan metabolism by *B. bifidum* PRL2010 provides a salient property to test their impact on colonization and succession. Although our findings are based to a considerable degree on *in silico* analyses rather than on the biochemical characterization of every enzyme involved in mucin utilization, we nevertheless have gathered sufficient evidence that proves the existence of specific *B. bifidum* strategies for the utilization of host glycans (e.g., using enzymes that remove sialic acid and fucose moieties from GNB and its extended derivatives present in various mucin O-glycans) (10, 29). The O-glycan structures present in mucin are diverse and complex and consist predominantly of core 1–4 mucin-type O-glycans. The enzymatic activities identified in *B. bifidum* PRL2010 are expected to allow degradation of many core 1 and 2 O-glycans and possibly core 3 and 4 O-glycans (Fig. 1*A*). Such O-glycans, especially if they are elongated and/or branched, may require the initial removal of terminal fucose and sialic acid residues by 1,2-α-L-fucosidase and 1,3/4-α-L-fucosidase and exo-α-sialidases and subsequent processing by the extracellular β-galactosidase and one or more of the extracellular β-*N*-acetylhexosaminidases before the endo-α-*N*-acetylgalactosaminidase can liberate GNB from its connected mucin glycoprotein (Fig. 1) (30). Once released, GNB and other degradation products are translocated across the cell membrane for further hydrolysis, phosphorylation, isomerization, and/or deacetylation and deamination. The resulting monosaccharides then feed into the central fructose-6-phosphate phosphoketolase pathway to generate ATP (37). Genome analyses indicate that *B. bifidum* PRL2010 encodes the enzymes necessary for the breakdown of the four main core structures of mucin O-glycans (39) and may be capable of degrading the highly complex extended derivatives of these core structures as identified in human colonic mucin (29). Notably, the carbohydrate core structure of mucin is similar to that of certain other human-derived glycans such as HMOs (45); thus common enzymatic pathways may be required for their degradation, a notion that is consistent with the results obtained from our proteomic analyses.

The present study provides a firm basis from which we can define further the factors that shape microbial ecology in health and disease and that influence interactions between microbiota and host. Arguably, mucin consumption by a specialized intestinal subpopulation of the microbiota may represent a means to enhance mucin production by the infant host (46). Thus, the relationship of *B. bifidum* and host-produced mucin constitutes an intriguing example of coevolution which should be taken into consideration in future studies of the infant intestinal microbiome.

Materials and Methods

Bacteria. *B. bifidum* PRL2010 was isolated from a fecal sample from a 3-mo-old, breast-fed, healthy infant. Further details can be found in *SI Text*.

Carbohydrate Growth Assays. Cell growth on semisynthetic de Man–Rogosa–Sharpe (MRS) medium supplemented with 1% (wt/vol) of a particular sugar was monitored at OD_{600} using a plate reader (Biotek). Further details are given in *SI Text*.

Genome Sequencing and Bioinformatics Analyses. The genome sequence of *B. bifidum* PRL2010 was determined by shotgun Sanger sequencing combined with pyrosequencing on a 454 FLX instrument (Roche) followed by assembly and gap closure (performed by Agencourt Genomic Services). The PRL2010 genome was sequenced to ≈25-fold coverage and assembled with Phred and Phrap incorporated in the Staden package (http://staden.sourceforge.net/). Automated gene modeling was achieved using multiple data-

bases and modeling packages as described previously (47). Additional information on sequencing and bioinformatics is provided in *SI Text*.

CHG. CGH was performed using a *B. bifidum* PRL2010 array that was hybridized using 2 μg of labeled DNA. Data analysis was performed according to the quantile normalization and log2 ratios of the reference sample and the samples analyzed. Further details can be found in *SI Text*.

Proteomic and Transcriptomics Experiments. Protein and RNA samples were prepared from *B. bifidum* PRL2010 cultivated at the exponential phase of growth. Proteins were extracted following mechanical lysis of the bacterial cell and tryptic digestion. The digested proteins were analyzed by LC-MS/MS. After database searches, spectral counting (48) was used to determine the relative expression of proteins. Further details can be found in *SI Text*. For transcriptome analysis, cDNA samples were prepared using the cDNA synthesis and labeling kit (Kreatech). Labeled cDNA was hybridized to a *B. bifidum* PRL2010 array using the protocols described in the manual for Agilent two-color microarray-based gene expression analysis. Data acquisition and analysis are described in *SI Text*.

ACKNOWLEDGMENTS. We thank Dr. David Sela (Stanford University, Stanford, CA) for his critical reading of the manuscript, and our students and coworkers for contributing data and for their enthusiasm. This work was supported by National Institutes of Health–National Institute of Child Health and Human Development Grant 1R01HD061923 (to D.M.). D.v.S., A.Z., and F.B. are members of the Alimentary Pharmabiotic Centre, a Center for Science and Technology funded by Science Foundation Ireland, through the Irish Government's National Development Plan Grant 07/CE/B1368.

1. Favier CF, de Vos WM, Akkermans AD (2003) Development of bacterial and bifidobacterial communities in feces of newborn babies. *Anaerobe* 9:219–229.
2. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA (2008) Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. *Nat Rev Microbiol* 6: 121–131.
3. Martens EC, Chiang HC, Gordon JI (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* 4:447–457.
4. Bjursell MK, Martens EC, Gordon JI (2006) Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J Biol Chem* 281:36269–36279.
5. Derrien M, Collado MC, Ben-Amor K, Salminen S, de Vos WM (2008) The Mucin degrader *Akkermansia muciniphila* is an abundant resident of the human intestinal tract. *Appl Environ Microbiol* 74:1646–1648.
6. Collado MC, Derrien M, Isolauri E, de Vos WM, Salminen S (2007) Intestinal integrity and *Akkermansia muciniphila*, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. *Appl Environ Microbiol* 73: 7767–7770.
7. Kleessen B, Bunke H, Tovar K, Noack J, Sawatzki G (1995) Influence of two infant formulas and human milk on the development of the faecal flora in newborn infants. *Acta Paediatr* 84:1347–1356.
8. Turroni F, et al. (2009) Exploring the diversity of the bifidobacterial population in the human intestinal tract. *Appl Environ Microbiol* 75:1534–1545.
9. Ruas-Madiedo P, Gueimonde M, Fernández-García M, de los Reyes-Gavilán CG, Margolles A (2008) Mucin degradation by Bifidobacterium strains isolated from the human intestinal microbiota. *Appl Environ Microbiol* 74:1936–1940.
10. 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.
11. Wada J, et al. (2008) *Bifidobacterium bifidum* lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure. *Appl Environ Microbiol* 74:3996–4004.
12. Jensen PH, Kolarich D, Packer NH (2010) Mucin-type O-glycosylation—putting the pieces together. *FEBS J* 277:81–94.
13. Sela DA, et al. (2008) The genome sequence of *Bifidobacterium longum* subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci USA* 105:18964–18969.
14. Hooper LV, Gordon JI (2001) Commensal host-bacterial relationships in the gut. *Science* 292:1115–1118.
15. Schell MA, et al. (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci USA* 99: 14422–14427.
16. Lee JH, et al. (2008) Comparative genomic analysis of the gut bacterium *Bifidobacterium longum* reveals loci susceptible to deletion during pure culture growth. *BMC Genomics* 9:247.
17. Barrangou R, et al. (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. lactis DSM 10140 and Bl-04. *J Bacteriol* 191: 4144–4151.
18. Ventura M, et al. (2009) The Bifidobacterium dentium Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. *PLoS Genet* 5:e1000785.
19. Ventura M, et al. (2007) Genomics of Actinobacteria: Tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* 71:495–548.
20. Gill SR, et al. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355–1359.
21. Eckburg PB, et al. (2005) Diversity of the human intestinal microbial flora. *Science* 308: 1635–1638.
22. Guglielmetti S, et al. (2008) Implication of an outer surface lipoprotein in adhesion of *Bifidobacterium bifidum* to Caco-2 cells. *Appl Environ Microbiol* 74:4695–4702.
23. Kankainen M, et al. (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. *Proc Natl Acad Sci USA* 106:17193–17198.
24. de Vries W, Stouthamer AH (1969) Factors determining the degree of anaerobiosis of Bifidobacterium strains. *Arch Mikrobiol* 65:275–287.
25. van den Broek LA, Hinz SW, Beldman G, Vincken JP, Voragen AG (2008) Bifidobacterium carbohydrases—their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res* 52:146–163.
26. Saier MH, Jr, Yen MR, Noto K, Tamang DG, Elkan C (2009) The Transporter Classification Database: Recent advances. *Nucleic Acids Res* 37(Database issue): D274–D278.
27. Johansson ME, Holmen Larsson JM, Hansson GC (2010) Microbes and Health Sackler Colloquium: The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci USA*, 10.1073/pnas.1006451107.
28. Forstner JFOM, Sylvester FA (1995) *Production, Structure and Biologic Relevance of Gastrointestinal Mucins* (Raven, New York), pp 71–88.
29. Podolsky DK (1985) Oligosaccharide structures of isolated human colonic mucin species. *J Biol Chem* 260:15510–15515.
30. Fujita K, et al. (2005) Identification and molecular cloning of a novel glycoside hydrolase family of core 1 type O-glycan-specific endo-alpha-N-acetylgalactosaminidase from *Bifidobacterium longum*. *J Biol Chem* 280:37415–37422.
31. Katayama T, Fujita K, Yamamoto K (2005) Novel bifidobacterial glycosidases acting on sugar chains of mucin glycoproteins. *J Biosci Bioeng* 99:457–465.
32. Ashida H, et al. (2008) Characterization of two different endo-alpha-N-acetylgalactosaminidases from probiotic and pathogenic enterobacteria, *Bifidobacterium longum* and *Clostridium perfringens*. *Glycobiology* 18:727–734.
33. Nagae M, et al. (2007) Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from *Bifidobacterium bifidum*. *J Biol Chem* 282:18497–18509.
34. Ashida H, et al. (2009) Two distinct alpha-L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology* 19:1010–1017.
35. Katayama T, et al. (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.
36. Miwa M, et al. (2010) Cooperation of {beta}-galactosidase and {beta}-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure. *Glycobiology* 20:1402–1409.
37. Nishimoto M, Kitaoka M (2007) Identification of N-acetylhexosamine 1-kinase in the complete lacto-N-biose I/galacto-N-biose metabolic pathway in *Bifidobacterium longum*. *Appl Environ Microbiol* 73:6444–6449.
38. Nishimoto M, Kitaoka M (2007) Practical preparation of lacto-N-biose I, a candidate for the bifidus factor in human milk. *Biosci Biotechnol Biochem* 71:2101–2104.
39. Rose MC, Voynow JA (2006) Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* 86:245–278.
40. Mathews KC, van Holde KE (1990) *Biochemistry* (Benjamin/Cummings, Redwood City, CA).
41. Dean KJ, Sweeley CC (1979) Studies on human liver alpha-galactosidases. II. Purification and enzymatic properties of alpha-galactosidase B (alpha-N-acetylgalactosaminidase). *J Biol Chem* 254:10001–10005.
42. Togayachi A, et al. (2010) Lack of lacto/neolacto-glycolipids enhances the formation of glycolipid-enriched microdomains, facilitating B cell activation. *Proc Natl Acad Sci USA* 107:11900–11905.
43. Turroni F, et al. (2009) Microbiomic analysis of the bifidobacterial population in the human distal gut. *ISME J* 3:745–751.
44. Salyers AA, West SE, Vercellotti JR, Wilkins TD (1977) Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* 34:529–533.
45. 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.
46. Comstock LE (2009) Importance of glycans to the host-bacteroides mutualism in the mammalian intestine. *Cell Host Microbe* 5:522–526.
47. Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* 132:115–130.
48. Zybailov B, Coleman MK, Florens L, Washburn MP (2005) Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal Chem* 77: 6218–6224.