

The *Bifidobacterium dentium* Bd1 Genome Sequence Reflects Its Genetic Adaptation to the Human Oral Cavity

Marco Ventura^{1*}, Francesca Turrone¹, Aldert Zomer², Elena Foroni¹, Vanessa Giubellini¹, Francesca Bottacini¹, Carlos Canchaya¹, Marcus J. Claesson², Fei He³, Maria Mantzourani⁴, Laura Mulas⁵, Alberto Ferrarini⁶, Beile Gao⁷, Massimo Delledonne⁶, Bernard Henrissat⁸, Pedro Coutinho⁸, Marco Oggioni⁵, Radhey S. Gupta⁷, Ziding Zhang³, David Beighton⁴, Gerald F. Fitzgerald², Paul W. O'Toole², Douwe van Sinderen^{2*}

1 Laboratory of Probiogenomics, Department of Genetics, Biology of Microorganisms, Anthropology, and Evolution, University of Parma, Parma, Italy, **2** Alimentary Pharmabiotic Centre and Department of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland, **3** College of Biological Sciences, China Agricultural University, Beijing, China, **4** Department of Microbiology, The Henry Wellcome Laboratories for Microbiology and Salivary Research, Kings College London Dental Institute, London, United Kingdom, **5** Department of Molecular Biology, University of Siena, Siena, Italy, **6** Dipartimento Scientifico Tecnologico, Università degli Studi di Verona, Verona, Italy, **7** Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada, **8** Glycogenomics, Databases, and Bioinformatics, Architecture et Fonction des Macromolécules Biologiques, Universités Aix-Marseille, Marseille, France

Abstract

Bifidobacteria, one of the relatively dominant components of the human intestinal microbiota, are considered one of the key groups of beneficial intestinal bacteria (probiotic bacteria). However, in addition to health-promoting taxa, the genus *Bifidobacterium* also includes *Bifidobacterium dentium*, an opportunistic cariogenic pathogen. The genetic basis for the ability of *B. dentium* to survive in the oral cavity and contribute to caries development is not understood. The genome of *B. dentium* Bd1, a strain isolated from dental caries, was sequenced to completion to uncover a single circular 2,636,368 base pair chromosome with 2,143 predicted open reading frames. Annotation of the genome sequence revealed multiple ways in which *B. dentium* has adapted to the oral environment through specialized nutrient acquisition, defences against antimicrobials, and gene products that increase fitness and competitiveness within the oral niche. *B. dentium* Bd1 was shown to metabolize a wide variety of carbohydrates, consistent with genome-based predictions, while colonization and persistence factors implicated in tissue adhesion, acid tolerance, and the metabolism of human saliva-derived compounds were also identified. Global transcriptome analysis demonstrated that many of the genes encoding these predicted traits are highly expressed under relevant physiological conditions. This is the first report to identify, through various genomic approaches, specific genetic adaptations of a *Bifidobacterium* taxon, *Bifidobacterium dentium* Bd1, to a lifestyle as a cariogenic microorganism in the oral cavity. *In silico* analysis and comparative genomic hybridization experiments clearly reveal a high level of genome conservation among various *B. dentium* strains. The data indicate that the genome of this opportunistic cariogen has evolved through a very limited number of horizontal gene acquisition events, highlighting the narrow boundaries that separate commensals from opportunistic pathogens.

Citation: Ventura M, Turrone F, Zomer A, Foroni E, Giubellini V, et al. (2009) The *Bifidobacterium dentium* Bd1 Genome Sequence Reflects Its Genetic Adaptation to the Human Oral Cavity. PLoS Genet 5(12): e1000785. doi:10.1371/journal.pgen.1000785

Editor: Josep Casadesús, Universidad de Sevilla, Spain

Received: September 21, 2009; **Accepted:** November 23, 2009; **Published:** December 24, 2009

Copyright: © 2009 Ventura et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was financially supported by the Italian Award for Outstanding Young Researcher scheme Incentivazione alla mobilità di studiosi stranieri e italiani residente all'estero 2005–2009 and a Marie Curie Reintegration Grant (MERG-CT-2005-03080) to MV; by the Science Foundation Ireland Alimentary Pharmabiotic Centre located at University College Cork to DvS, GFF, and PWOT; and by an EMBARK postdoctoral grant to AZ. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: marco.ventura@unipr.it (MV); d.vansinderen@ucc.ie (DvS)

Introduction

Bifidobacteria are relatively abundant inhabitants of the gastrointestinal tract (GIT) of humans and animals [1]. Many bifidobacterial species, in conjunction with other members of the intestinal microbiota are believed to contribute to host nutrition, while also impacting on intestinal pH, cell proliferation and differentiation, development and activity of the immune system, and innate and acquired responses to pathogens [2–8]. These perceived beneficial health effects have driven commercial exploitation of bifidobacteria as live components of many functional foods and therapeutic adjuncts. However, bifidobac-

teria have also been isolated from the human oral cavity, where their presence is linked to the progression of tooth decay: bifidobacteria have been detected in high numbers in infected dentine from carious lesions in children [9] and have been associated with childhood dental caries [10]. *B. dentium* can be found as part of the microbiota implicated in human dental caries [10–16]. In recent surveys of oral bifidobacteria associated with coronal caries in adults and children [17] and root caries in adults [18], *B. dentium* was the most frequently isolated *Bifidobacterium* species, representing approximately eight percent of the culturable bacteria isolated from active carious lesions. This species is capable of acidogenesis to produce a final pH in glucose-containing media

Author Summary

The accessibility of complete bacterial genome sequences has provided important changes to the field of microbiology by significantly enhancing our understanding of the physiology, genetics, and evolutionary development of bacteria. Bifidobacteria are among such microorganisms, being mammalian commensals of biotechnological significance due to their perceived role in maintaining a balanced gastrointestinal (GIT) microflora. Bifidobacteria are therefore often applied as health-promoting or probiotic components in functional food products and represent a growing area of scientific interest. However, within the genus *Bifidobacterium* not all species provide beneficial effects on the host's health. In fact, the *Bifidobacterium dentium* species is considered an opportunistic pathogen since it has been associated with the development of dental caries. In this manuscript, we describe the complete genetic make-up of the *B. dentium* Bd1 genome and discuss functions that explain how this microorganism has adapted to the oral human cavity and imparts a cariogenic phenotype. Moreover, we performed comparative genomic analyses of *B. dentium* genome with other bifidobacterial genomes in order to trace genetic differences/similarities between the opportunistic oral pathogen *B. dentium* Bd1 and closely related intestinal bifidobacteria.

below pH 4.2 [19], sufficient to cause extensive demineralisation of tooth tissues [20]. *B. dentium* may therefore significantly contribute to the pathogenesis of dental caries which is one of the most common chronic diseases, remaining untreated in many underdeveloped countries where dental pain is often alleviated only by the loss or extraction of the affected tooth [21].

The ecological plaque hypothesis was formulated to explain the composition and phenotypic properties of the microbiota associated with caries initiation and progression [22]. This hypothesis envisages that caries is the result of environmental changes, particularly as a result of reduced intra-oral pH as a consequence of bacterial fermentation of dietary carbohydrates. When this occurs in the oral cavity, it selects for a microbiota which is more aciduric and more acidogenic than that present in the absence of caries. The environmental change results in a significant alteration in the composition of the commensal microbiota, with taxa including bifidobacteria, lactobacilli, *Actinomyces* and streptococci proliferating [11].

Complete genome sequences of relatively few human intestinal commensal bifidobacteria have been determined, being largely motivated by their perceived health-promoting activity. These include *Bifidobacterium longum* subsp. *longum* NCC2705, *B. longum* subsp. *longum* DJO10A and *B. longum* subsp. *infantis* ATCC15697, *Bifidobacterium animalis* subsp. *lactis* DSM10140 and *B. animalis* subsp. *lactis* ADO11 [23–27]. Here, we describe the sequence analysis of the *B. dentium* Bd1 genome. This strain was originally isolated from human dental caries [28]. Analysis of the predicted proteome together with comparisons of the genome sequence to those of intestinal bifidobacteria revealed that this bacterium has undergone specific genetic adaptations for colonization and survival in the oral cavity.

Results

General genome features

The genome of *B. dentium* Bd1 is one of the largest bifidobacterial genomes reported to date, with a single circular

chromosome consisting of 2,636,368 base pairs (Figure 1). The average GC content of 58.54% is similar to that of other sequenced bifidobacterial genomes and is consistent with the range of G+C mol% values for the *Actinobacteria* [1]. For protein-encoding DNA regions, the G+C contents of codon positions 1, 2 and 3 were determined to be 61%, 43%, 74%, respectively, the latter value somewhat deviating from the expected value (70%), as based on a survey of 696 eubacterial and 56 actinobacterial genomes or bifidobacterial genomes (NCBI source) (Figure S1 and data not shown).

The genome of *B. dentium* Bd1 possesses 55 tRNAs and four rRNA operons, which are located in proximity of the *oriC*. While *B. dentium* contains tRNAs for every amino acid, the corresponding genes for aminoacyl-tRNA synthetases for asparagine and glutamine appear to be absent. An alternative route is a pathway described for *Fusobacterium nucleatum* [29] that utilizes Gln- and Asn-tRNA amidotransferases, which amidate misacylated Gln-tRNA or Asn-tRNA charged with Glu or Asp to produce Gln-tRNA Gln or Asn-tRNA-Asn, respectively. Homologous genes that specify subunits for the Gln- and Asn-tRNA amidotransferase are present on the Bd1 genome and are also present on other sequenced bifidobacterial genomes (data not shown).

Identification of protein-coding sequences revealed 2143 open reading frames (ORFs) with an average length of 1059 bases and constituting 89% of the genome, the remainder representing intergenic regions with an average length of 143 bp. This latter value is lower than those calculated for other known bifidobacterial genomes, whose combined average intergenic region length is 191 bp, indicating that *B. dentium* Bd1 has a more compact genome. Such results are not biased due to the methods used for *B. dentium* Bd1 genome annotation (see Text S1), since it employed an ORF identification protocol, with cut-off values that are similar to those used for the annotation of the so far published bifidobacterial genome sequences [23–27]. The ORFs are organised in a typical bacterial configuration, so that transcription is frequently in the same direction as DNA replication. A functional assignment was made for 78.5% of the predicted ORFs, while homologs with no known function from other bacterial species were identified for an additional 13% of the *B. dentium* Bd1 ORFs. The remaining 8.4% appears to be unique to *B. dentium*. The ATG start codon is preferred (78.9% of the time), while GTG and TTG are less frequently used start codons at 18.9% and 2%, respectively.

The presumed origin of replication (*oriC*) of the *B. dentium* Bd1 chromosome, including the adjacent and conserved *dnaA*, *dnaN* and *recA* gene configuration, was identified on the basis of common features to corresponding regions in other bacterial chromosomes [30–31]. The *oriC* was located proximal to the *dnaA* gene, in an AT-rich sequence containing characteristic DnaA boxes, while the position of the replication terminus (*terC*) was inferred using GC skew analysis (Figure 1).

The predicted *B. dentium* Bd1 proteins were functionally categorized and the proportions in each category were compared with those of other bifidobacterial genomes (Figure 2). It is notable that approximately 14% of the genes identified in the *B. dentium* Bd1 genome encode proteins that are predicted to be involved in carbohydrate metabolism and transport. Such an extensive genetic adaptation to carbohydrate metabolism is shared, to a similar degree, with enteric bifidobacteria, and likely represents a specific genetic adaptation of bacteria residing in the GIT, apparently both in the upper region (the oral cavity) as well as in the distal tract (the colon) of the GIT.

Furthermore, 3D-structure prediction of 1955 of the 2143 deduced proteins that constitute the predicted proteome of *B. dentium* Bd1 using the Fugue fold recognition method allowed a

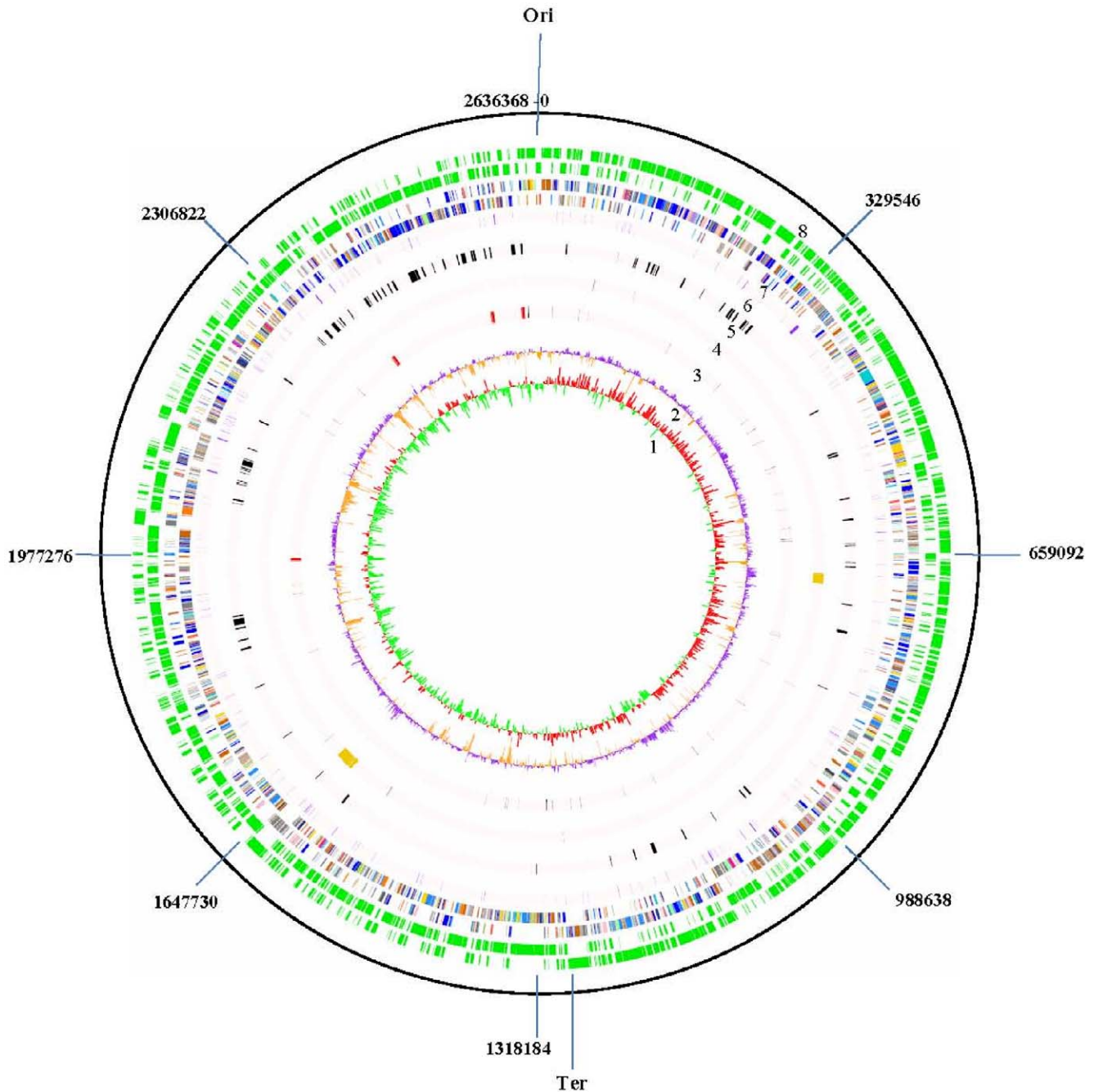


Figure 1. Circular genome map of *B. dentium* Bd1. From innermost circle, circle (1) illustrates GC skew ($G-C/G+C$), values >0 are in red and <0 in green. Circle (2) highlights G+C% deviation from the mean (58.54%). Circle (3) indicates rRNAs (depicted in red) and tRNAs (depicted in blue). Circle (4) denotes IS and prophages (depicted in orange). Circle (5) depicts genes involved in sugar metabolism according to the CAZY database. Circle (6) denotes conserved ORF distribution. Circle (7) shows coding regions by strand with color corresponding to the COG functional assignment (the color code used is the same indicated in Figure 2). Circle (8) displays the ORF distribution by strand. doi:10.1371/journal.pgen.1000785.g001

functional attribution of these predicted structures by means of the SCOP domain annotation into superfamilies (Figure 3). Such an analysis was also performed for the intestinal *B. longum* subsp. *longum* NCC2705 strain. Notably, both genomes possess a similar protein superfamily content distribution except for proteins assigned to the Toxin-defence group: the *B. dentium* Bd1 genome encodes nine times more proteins with this annotated function as compared to the *B. longum* subsp. *longum* NCC2705 genome (Figure 3). There are 18 genes encoding predicted sensor histidine

protein kinases (HPK), 15 of which are located adjacent to a putative response regulator-encoding gene (one of which is separated by just a single gene), distributed throughout the Bd1 chromosome, which is somewhat more than one would predict based on its genome size and the number present in other, similarly sized bifidobacterial genomes (ranging from 5 to 17) [32]. This suggests that the relative abundance of two-component systems (2CSs) in a micro-organism is an indicator of its ability to sense dynamic environmental cues and to modulate appropriate

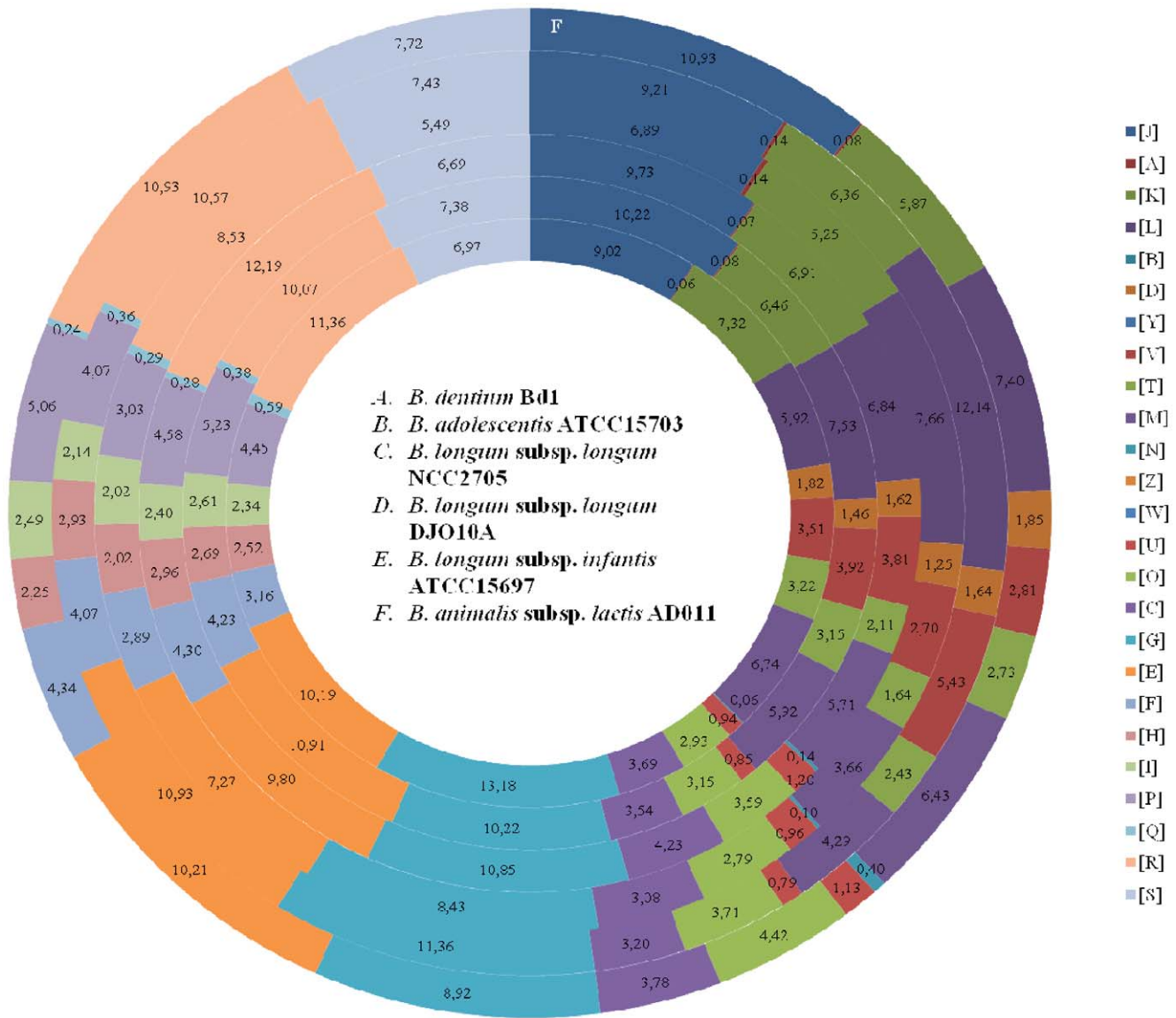


Figure 2. Comparison of COG functional categories between completely sequenced bifidobacterial genomes. Each coloured segment indicates the relative contribution of a functional category as a percentage of total COGs. Each ring indicates a different bifidobacterial genome as labelled. The color of each COG family is indicated in the Figure. The name of the bacterial genomes are indicated in the Figure. doi:10.1371/journal.pgen.1000785.g002

physiological responses, a notion also exemplified by the high number of 2CSs found in *Bacteroides thetaiotaomicron* [33].

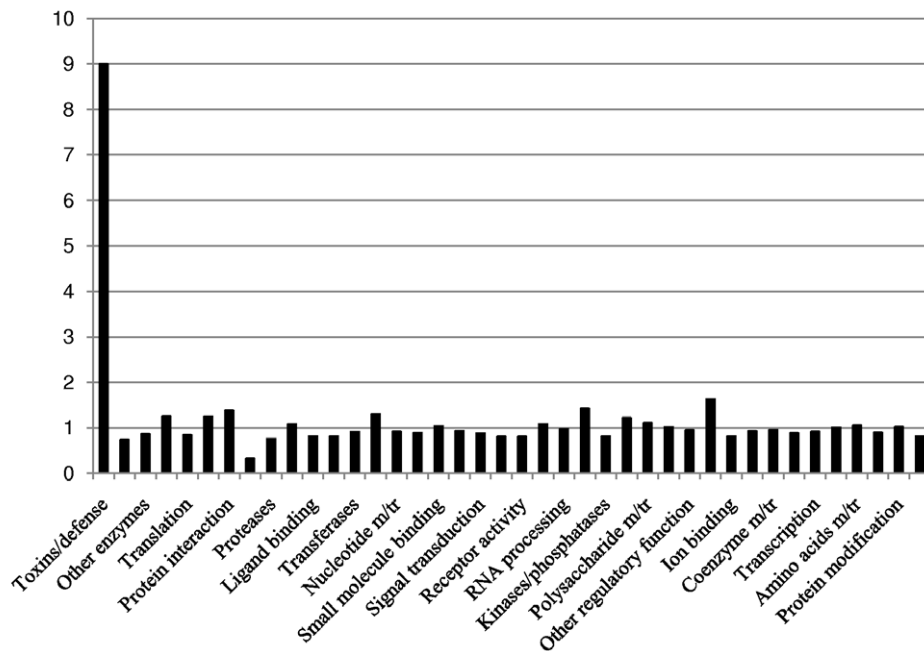
Genome differences in bifidobacteria revealed by genome sequence alignments

Comparative genomics of intestinal bifidobacteria may elucidate genomic regions involved in the maintenance of physiological homeostasis that is attributed to these bacteria. The genomic structure of *B. dentium* Bd1 is highly syntenic with that of the recently sequenced genome of *B. dentium* ATCC27678 (accession no. ABIX00000000; Figure 4) with an average nucleotide identity of 99% across these two genomes. This Bd1 versus ATCC27678 genome comparison was scrutinized for the identified ORFs with particular consideration of nucleotide changes occurring at particular positions for every codon in the Bd1 genome. The highest substitution rate occurred at the third codon position (36.3% vs. 29% at the second nucleotide and 34.7% at the first

nucleotide). Furthermore, a survey of DNA sequence similarity at intergenic regions between both strains revealed a lower level of DNA conservation (98%) compared to that identified between coding regions (99%). Thus, in the *B. dentium* taxa, as is generally the case for *Eubacteria*, the intergenic regions have experienced a higher rate of nucleotide substitution compared to that of the coding regions.

Furthermore, sequence identity varied between the ORFs shared by both genomes (2133); with the large majority displaying an identity of 100%, and just 19 ORFs showing a similarity of less than 95%. The two genomes were shown to contain an identical repertoire of prophage-like and IS elements, although a small number of nucleotide differences (e.g., deletions or substitutions) were noticed for two IS elements, *isblo10* and *isblo3-2*, and for the prophage-like elements *Bdent-1* and *Bdent-2*, suggesting that these strains are very closely related but genetically distinct.

A



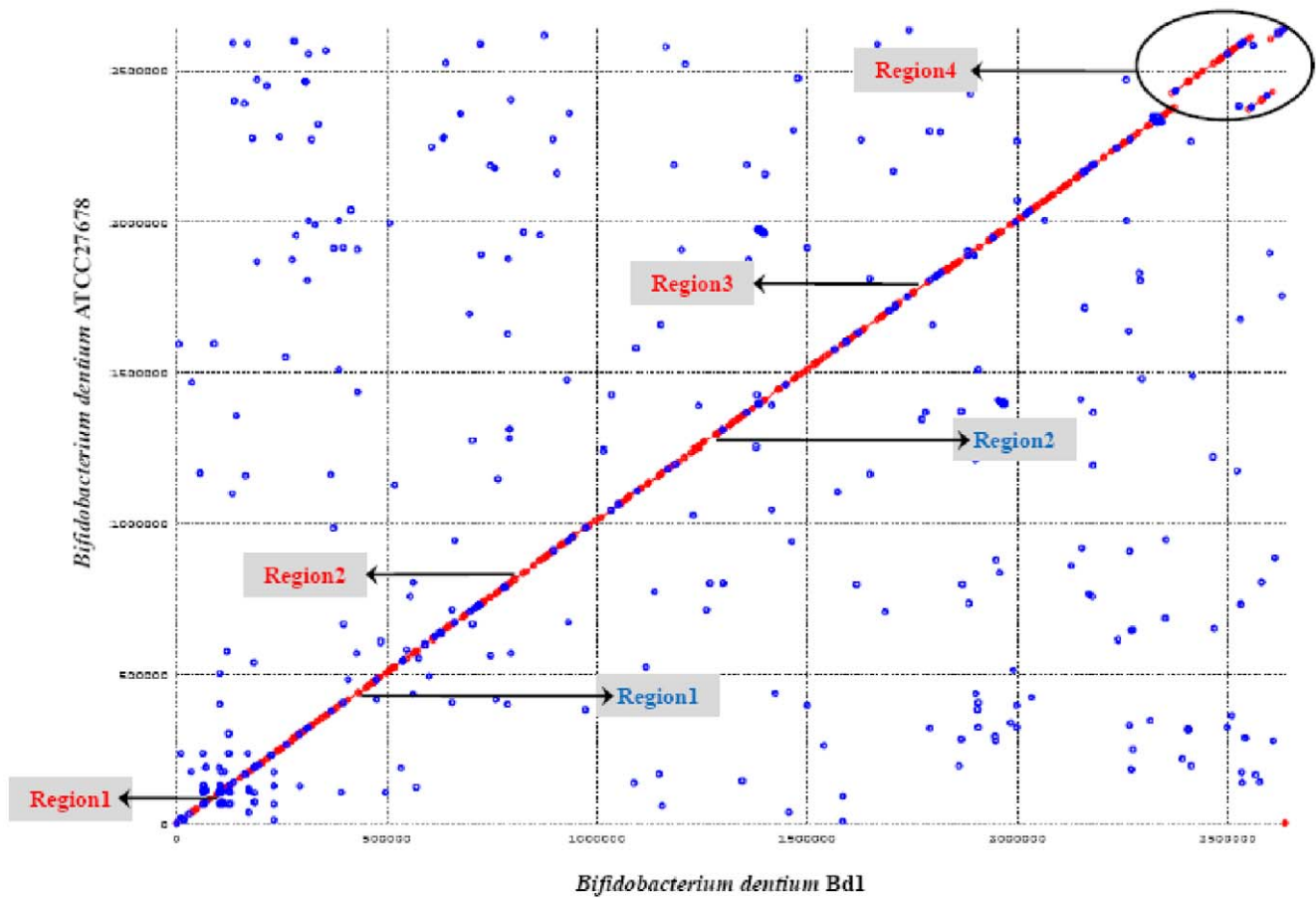
B

ORF	Fugue hit	Z-score	SCOP family
BDP_0320	hs1e68a	4.26	a.64.2.1
BDP_0674	hs1y29a	4.42	g.3.6.2
BDP_0842	hs1e68a	4.78	a.64.2.1
BDP_0935	hs1e68a	4.25	a.64.2.1
BDP_1282	hs1ugla	4.22	g.3.7.5
BDP_1285	hs1i25a	4.44	g.3.6.2
BDP_1638	hs1f3ka	4.47	g.3.6.1
BDP_1717	hs1d1ha	4.11	g.3.6.2
BDP_2014	hs1ixta	5.15	g.3.6.1
BDP_2150	hs1e4qa	4.38	g.9.1.1
BDP_2174	hs1i25a	4.07	g.3.6.2

Figure 3. Functional annotation and assignment of the encoded proteins from the genomes of *B. dentium* Bd1 and *B. longum* subsp. *longum* NCC2705 in different superfamily categories according to the Fugue fold recognition method. In (A), each bar represents the odd ratio between *B. dentium* Bd1 and *B. longum* subsp. *longum* NCC2705. The bar corresponding to the toxins/defence superfamily is indicated. In the y-axis is indicated the value of the odd ratio between *B. dentium* Bd1 and *B. longum* subsp. *longum* NCC2705. (B) indicates the different *B. dentium* Bd1 ORFs classified in the toxins/defence superfamily. doi:10.1371/journal.pgen.1000785.g003

Points of disruption of the gene conservation between the two *B. dentium* genome sequences corresponds to the presence or absence of integrated elements (IS elements) and genes with a predicted function in sugar metabolism (Figure 4). For example, the *B. dentium* ATCC27678 genome contains a putative rhamnosyltransferase-encoding gene, which is located close to an IS element, and which is lacking in the corresponding position on the *B. dentium* Bd1 genome. This suggests that this genetic element in *B. dentium* ATCC27678 was acquired by HGT or it may be lost due to the

presence of this mobile element. The conserved gene order was not limited to *B. dentium* ATCC27678, but can be expanded to *B. adolescentis* ATCC15703 (Figure S2). The degree of alignment between the genomes of different bifidobacterial species varied depending on the phylogenetic distance of the genomes being compared. Thus, the alignments of *B. dentium* Bd1 with the genomes of *B. longum* subsp. *longum* NCC2705, *B. longum* subsp. *longum* DJO10A or *B. longum* subsp. *infantis* ATCC15697 display an clearly reduced colinearity, resulting in an X-shaped plot diagram



Region	Function
Region1	Hypothetical protein
Region2	IS150 putative transposase
Region1	beta-glucosidase
Region2	rhamnosyltransferase
Region3	solute-binding protein of ABC transporter
Region4	calcium-transporting ATPase

Figure 4. Dot plot of *B. dentium* Bd1 versus *B. dentium* ATCC27678. The visible areas of divergence are indicated. Blue type indicates sequences unique to Bd1 compared with ATCC27678, whereas red type shown sequences absent in Bd1. The deduced functions of the divergent regions are described in the table.

doi:10.1371/journal.pgen.1000785.g004

(Figure S2). This is indicative of multiple large rearrangements around the origin-terminus axis of the genome following divergence from a common ancestor [34]. When the genome of a *Bifidobacterium* strain outside the *B. longum* and *B. adolescentis*

phylogenetic groups (e.g., *B. animalis* subsp. *lactis* ADO011) was thus aligned, sequence identities were restricted to very small genome segments (Figure S2). The fact that long-range genome alignments of *Bifidobacterium* could not be produced at the DNA

level is a significant indication of profound intra-genus diversity, similar to that found for the lactobacilli [35], but contrasting for example with DNA-DNA interspecies alignments among other genera belonging to the *Actinobacteria* [1].

Genome alignments using PROmer allows the reconstruction of broad phylogenetic relationships between prokaryotic genomes [35–37]. A previous phylogenomics analysis based on 123 protein sequences representing the minimal core proteins of the *Actinobacteria* phylum highlighted the relatedness of *B. longum* subsp. *longum* NCC2705 to propionibacteria, *Leifsonia* and *Tropheryma* [1]. Here, we included in such an analysis the *B. dentium* Bd1 and other bifidobacterial genome sequences published to date. Interestingly, the resulting neighbour-joining tree revealed a clear evolutionary split of these bifidobacterial sequences with those derived from *Propionibacterium acnes*, *Leifsonia xyli* subsp. *xyli* and *Tropheryma whipplei* (Figure S3), and indicating that bifidobacteria are derived from a deep ancestor of the *Actinobacteria* phylum.

A comparative study was undertaken to determine putative orthology between the *B. dentium* Bd1 CDSs with those of five other completely sequenced bifidobacterial genomes, resulting in 908 putative orthologs that were shared between all these genomes (Figure S4). The most common functional classes represented by these core proteins were those involved in housekeeping functions including information processing, DNA replication, repair, cell division, transcription, translation and secondary metabolite biosynthesis, transport and catabolism. Proteins belonging to functional categories representing sugar and amino acid metabolism, and uptake by ABC transporters were the second largest commonly found group, emphasizing their apparent importance to bifidobacteria (Figure 3). When the genome sequence of *B. dentium* ATCC27678 was included in this analysis a total of 692 CDSs were found that have no matches in currently available bifidobacterial genomes, thus representing *B. dentium*-specific proteins. Over half of these are hypothetical proteins, whereas the remainder have their best matches in sequenced members of the *Actinobacteria* and/or *Firmicutes*, including bacteria of the oral microbiota such as *Actinomyces* spp., *S. mutans* and *Treponema denticola* [38–39]. Notably present among these *B. dentium*-specific genes are two adjacent ORFs (BDP_1871–BDP_1872) with homology to the *hip* operon found in *Enterobacteria*, which allows increased survival following various stress conditions [40]. Thus, the *B. dentium hip* operon may positively influence persistence in the oral environment upon exposure to stress conditions, e.g. the fluctuating acid environment that accompanies caries initiation. A subset, i.e. 181, of the 692 *B. dentium*-specific proteins are conserved in both *B. dentium* genomes but do not have significant matches in other currently available genome sequences, and may thus be responsible for certain unique adaptive properties. When these hypothetical proteins were scanned against a database of structural profiles using the FUGUE program, which can recognize distant homologues by sequence-structure comparisons [41], we identified a number of potential homologs with a significant Z-score subdivided in various clusters according to their predicted gene function (Figure S5). Interestingly, this analysis revealed that part (13.33%) of these *B. dentium*-specific proteins clustered in the toxin/defence family, suggesting that these proteins provide protection against host defensins, such as cationic and cysteine-rich peptides [42].

The mobilome of *B. dentium* genome

The coexistence of *B. dentium* within the oral biofilm which consists of over 900 taxa may facilitate exchange of genetic information that is mediated by mobile genetic elements. Such elements are present in virtually all bacterial genomes, and in some

organisms the associated genes may contribute to the metabolism or pathogenic potential of the organisms. Analysis of the *B. dentium* Bd1 genome revealed the presence of conventional mobilome candidates that may have been acquired through Horizontal Gene Transfer (HGT). Analysis of G+C content (G+C), amino acid usage [43], BLASTP best-match and codon preference of the *B. dentium* Bd1 chromosome indicated that considering its total DNA content just 93,300 bp of the *B. dentium* Bd1 genome display a significant deviation (>2-fold difference) from the average values of the parameters as indicated above and may have been recently acquired by HGT (Figure 5). This suggests that in contrast to other bifidobacterial genomes [44], HGT is not the main force driving genome evolution in *B. dentium* species.

Some representative mobile elements as well as DNA regions acquired by HGT will be discussed below. In *silico* analysis of the *B. dentium* Bd1 genome revealed the presence of two prophage-like elements designed Bdent-1 and Bdent-2, which exhibit a close phylogenetic relationship with phages infecting bacteria belonging to the *Firmicutes* phylum (Ventura et al, AEM in press, published now). The *B. dentium* Bd1 chromosome harbours seven insertion sequences (IS), belonging to three IS families, ISL3, ISL10, ISSdel and IS3 like (Table 1), a number which is much lower than those in other sequenced bifidobacterial genomes (data not shown).

Additional putative mobile elements identified in the *B. dentium* Bd1 genome are represented by two Clustered of Regularly Interspersed Short Palindromic Repeats (CRISPR) loci, named CRISPR1 and CRISPR2, with adjacent CRISPR-associated *cas* genes, CRISPR-Cas1 and CRISPR-Cas2, respectively. When CRISPR-Cas1 and CRISPR-Cas2 were compared to identified lactic acid bacteria CRISPR loci [45], they clustered into two different CRISPR families, Blon1 and Lle11, respectively (Figure S6), suggestive of two independent HGT events. CRISPRs represent the most widely distributed prokaryotic family of repeats [46,47], and act as defence systems against invasion of foreign genetic material, in particular phages [48,49].

Genome diversity in *B. dentium*

The genome variability among different strains of *B. dentium* was investigated by Comparative Genomic Hybridization (CGH) experiments using *B. dentium* Bd1-based microarrays. We determined which and how many ORFs from the sequenced *B. dentium* Bd1 strain did or did not hybridize with total genomic DNA extracted from ten *B. dentium* strains from different origins (dental caries from adult or child, from saliva and from fecal samples). Overall, DNA from the tested *B. dentium* strains failed to efficiently hybridize to between 1% and 12% of the probes from the reference *B. dentium* strain Bd1. These values are small compared to those described for other bacterial species [50–56], including bifidobacteria, such as *B. longum* subsp. *longum* [57]. Such findings suggest that the *B. dentium* genome is only slowly evolving compared to other bacteria, including bifidobacterial species residing in the distal tract of the human GIT. Nevertheless, CGH cannot identify regions present in the tested strains but absent from the *B. dentium* Bd1 strain, while it also does not analyze the synteny of the genome. Consequently, caution needs to be employed when applying the term “divergent” to CGH studies. When projected on the genome map of *B. dentium* Bd1, the CGH results highlight clustering of conserved and variable ORFs (Figure 6). The region between the origin of replication and the terminus of replication in the clockwise direction represents the largest genome segment of relative high gene conservation (denoted as I). In contrast, the region between the replication terminus and the origin of replication in the clockwise direction was shown to be a major area of genetic diversity (II in Figure 6).

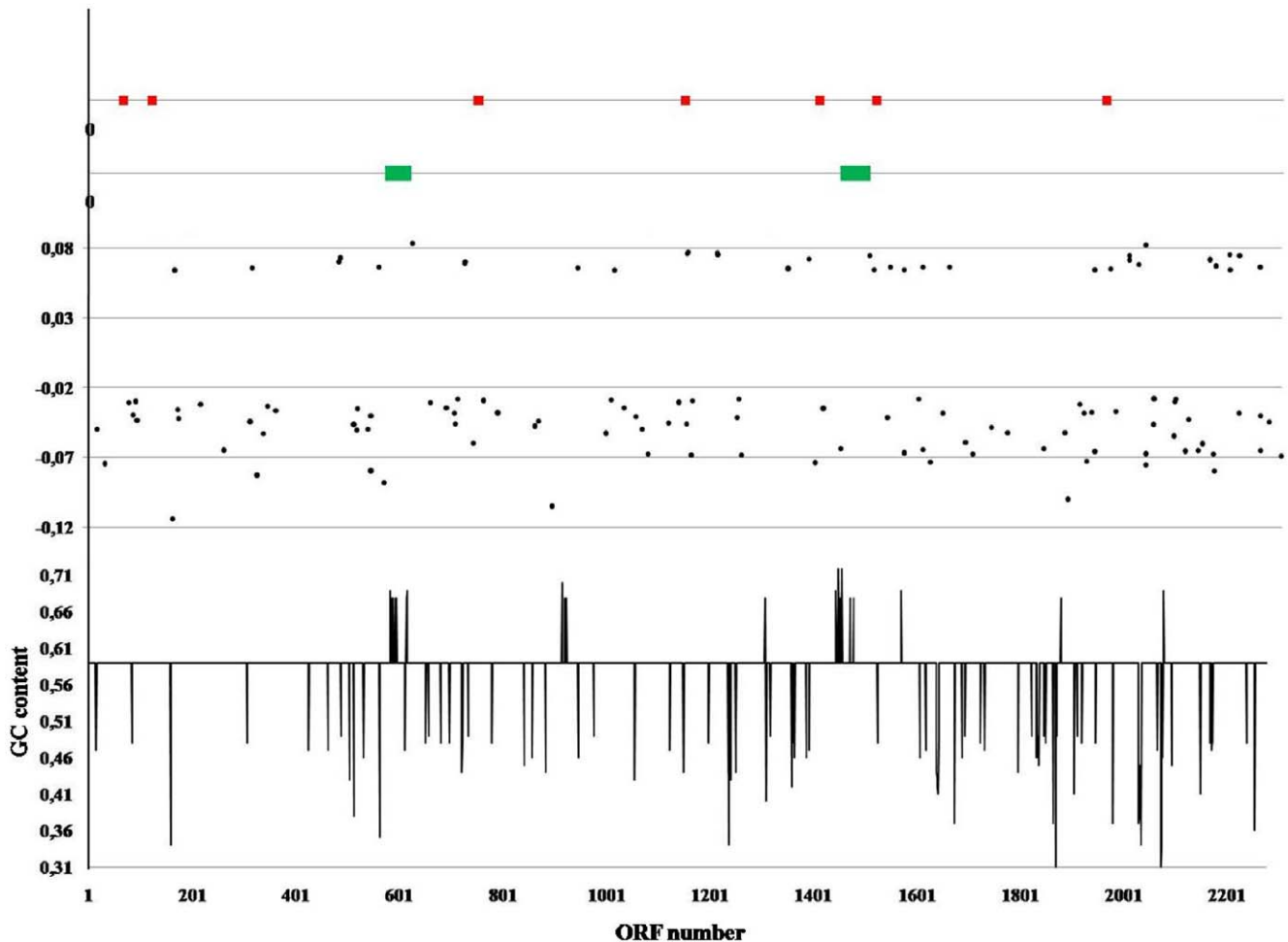


Figure 5. Mobile genetic elements of the *B. dentium* Bd1 genome. IS elements and predicted prophage-like elements are labelled in red and green, respectively. The first plot from the bottom indicates the deviation of the G+C content of each ORF of the *B. dentium* Bd1 genome from the mean average (58.54%). In the second plot each dot represents an ORF displaying a biased codon usage determined by factorial correspondence analysis of codon usage.
doi:10.1371/journal.pgen.1000785.g005

According to the Bd1 gene annotation, the types of genomic diversity thus identified can be assigned to two classes (i) mobile DNA that constitutes the *B. dentium* mobilome previously identified by *in silico* analyses; (ii) plasticity regions of *B. dentium* genome, which may underlie specific adaptations of the investigated strains, and which could represent laterally acquired DNA or remnants of ancestral DNA that have not (yet) been lost. Plasticity regions which are preferred sites for acquisition of strain-specific DNA are well recognized in the genomes of pathogens like *H. pylori* [58], where array-based CGH has similarly been used to highlight regions involved in adaptation to different pathological roles [59]. Five large DNA segments, which are conserved in *B. dentium* Bd1 and in the closely related strains ATCC27678 and LMG10585, clearly represent mobile DNA: two prophage-like elements, Bdent-1 and Bdent-1, the CRISPR elements and the cytosolic proteins (BDP_1391–1394) (Figure 6). In some strains, stretches of hybridizing prophage genes matched individual modules of the prophage (Figure 6, Bdent-2 prophage-like element), an observation which agrees with the hypothesis of modular phage evolution [60]. Within the variable regions of the CGH map, indicated as plasticity regions, genes associated with bacterium-environment interaction and metabolic abilities appear to be particularly

enriched. These include the *eps* clusters, a putative fimbrial-biosynthesis gene cluster and membrane-associated transporters. The *eps* clusters of the Bd1 strain are associated with the dTDP-rhamnose biosynthesis locus, and represent the largest genome segment with substantial inter-strain genetic variability.

When the CGH data are expressed on a \log_2 scale according to the mean ratios of the normalized results, a major peak is noticed at the same position for all the *B. dentium* strains tested, indicative of very similar DNA sequences (Figure 6A–6L). In addition to the inset of Figure 6, which globally quantifies the similarity of the test strains versus *B. dentium* Bd1, a clustering of the microarray data was performed in order to extract qualitative information about the presence of each gene. A phylogenetic tree based upon these CGH scores identified *B. dentium* ATCC27678 and *B. dentium* LMG 10585 as the closest relatives of *B. dentium* Bd1 (Figure S7). Moreover, CGH clustering produced four groups of *B. dentium* strains based on varying levels of genetic diversity, and largely corresponding to their ecological origin (Figure S7). As a complement to the CGH analysis, we performed multilocus sequence analyses for the same strains, using the genes for the Clp ATPase (*clpC* gene), two F6P-phosphoketolases (*xfp* gene), DnaJ chaperone (*dnaJ1*), and DNA-directed RNA polymerase B' subunit

Table 1. Genome features of *B. dentium* Bd1.

Trait	Number/ value
Size (Mb)	2,636,368
G+C content	58.54%
Number of identified ORF	2143
Assigned function	1684
- Amino acid transport and metabolism ^a	259
- Carbohydrate transport and metabolism ^a	286
- Transcription ^a	174
- Translation ^a	159
- Replication, recombination and repair ^a	123
- Defence mechanisms ^a	60
- Signal transduction ^a	100
- Cell wall/membrane biogenesis ^a	124
- Post-traslational modification, protein turnover, chaperones ^a	55
- Energy production and conversion ^a	70
- Nucleotide transport and metabolism ^a	64
- Coenzyme transport and metabolism ^a	55
- Lipid transport and metabolism ^a	51
- Inorganic ion transport and metabolism ^a	119
Phage regions	2
IS transposase families	3
- ISL3	4
- IS3-like	1
- ISL10	1
- ISSdel	1
CRISPR	2
Fimbrial systems	10
Transporters	771
- ABC systems	298
- PTS systems	2

^aAccording to the COG families.

doi:10.1371/journal.pgen.1000785.t001

(*rpoC* gene) as phylogenetic markers [61]. As expected, the phylogenetic tree produced from these concatenated sequences confirmed the clustering of GCH based data (Figure S7).

Metabolism and transport of *B. dentium*

Homologs of all the enzymes necessary for the fermentation of glucose and fructose to lactic acid and acetate through the characteristic “fructose-6-phosphate shunt” [62], as well as a partial Embden-Meyerhoff pathway were annotated in the *B. dentium* Bd1 genome. These metabolic pathways are important for generation of pyruvate and re-oxidation of NADH, as well as for synthesis of an additional ATP molecule per glucose during the conversion of pyruvate to acetate. The enzymes responsible for pyruvate metabolism identified in the *B. dentium* genome include xylulose 5-phosphate/fructose-6-phosphate phosphoketolase, pyruvate dehydrogenase, pyruvate formate-lyase, phosphotransacetylase, acetate kinase, lactate dehydrogenase. *B. dentium* Bd1 possesses an incomplete tricarboxylic acid (TCA) cycle, which lacks oxoglutarate dehydrogenase, fumarase and malate dehydrogenase. The primary role of these TCA enzymes is most likely the

production of precursors for amino acid and nucleotide biosynthesis. Since *B. dentium* Bd1 can be cultured anaerobically with urea, arginine and cysteine as the sole nitrogen sources (unpublished data), it was not surprising that genes required for the biosynthetic pathways of all amino acids were identified in the genome. The way in which cysteine is synthesized is unclear, as the genes involved in sulphate/sulphite assimilation are not present in the *B. dentium* Bd1 genome. It may synthesize cysteine in a manner similar to that suggested for *B. longum* subsp. *longum* NCC2705 using homologs of the genes for cysteine synthase/cystathione beta synthase, O-acetylhomoserine aminocarboxypropyltransferase and cystathionine γ -synthase, and utilizing a reduced sulphur-containing compound as a sulphur source [23]. Genes encoding complete biosynthetic pathways for purines and pyrimidines from glutamine, as well as for riboflavin, thiamine and folate were identified, while no homologues were present for pathways to produce biotin, pyridoxine, cobalamin, panthotenate and niacin/nicotinic acid, which are also variably distributed in the genomes of other sequenced bifidobacterial genomes (data not shown).

Comparative analysis against the Transport and Classification Database [63] predicts that the *B. dentium* Bd1 genome contains 771 genes encoding (components of) transport systems, accounting for almost 34% of the total number of ORFs (Figure S8A).

Transport in *B. dentium* Bd1 is largely carried out by transporters or carriers (e.g., uniporters, symporters and antiporters) and by P-P-bond hydrolysis-driven transporters. A large proportion of the identified transporters are ATP-dependent, as expected for a microorganism lacking an electron transport chain [64]. Annotated solute-transporting ATPases include P-type, F-type and ABC-type. The P-type ATPases are predicted to be involved in the transport of calcium and potassium, whereas the F-type ATPases (e.g., F₀F₁ATPases) use an electrochemical gradient of H⁺ or Na⁺ to synthesize ATP, or hydrolyze ATP to reverse the electrochemical gradient [65]. As described below, a single predicted H⁺-transporting ATP synthase-ATPase is encoded by the *B. dentium* Bd1 genome. We identified 298 predicted ABC-type ATPases, of which about 70% are categorized as importers, representing the most abundant transport category, and accounting for almost 13% of all *B. dentium* Bd1 gene products. The ABC transporters identified have a predicted specificity for a wide variety of substrates, including amino acids, carbohydrates, oligopeptides, osmoprotectants (e.g., proline/glycine, betaine, choline), inorganic ions (e.g., Fe³⁺, Co²⁺, Mn²⁺, phosphate, nitrate, sulphate, and molybdenum) and antimicrobial peptides.

The vast majority of carbohydrate-modifying enzymes encoded by *B. dentium* Bd1 are predicted to be intracellular and so the uptake of sugars with a low degree of polymerization is a key component of *B. dentium* carbohydrate metabolism. The genome of *B. dentium* Bd1 encodes at least 167 ABC transport systems for dietary carbohydrates (Table S1). The *B. dentium* Bd1 genome also specifies two phosphoenolpyruvate-phosphotransferase systems (PEP-PTS), consisting of the two general energy-coupling components, enzyme I (EI) and a heat-stable protein (Hpr), and two different sugar-specific multiprotein permeases known as enzyme II (EII).

Adaptation to the oral cavity

The human oral cavity is a complex microbial ecosystem, the composition of which may vary depending on the frequency and nature of food ingestion with consequent fluctuations in biofilm pH. Compared to the distal bowel, where organisms are presented with a relatively consistent stream of molecules that cannot be metabolized or degraded by the more proximal microbiota, the oral cavity microbiota is exposed to the full contents of the ingested

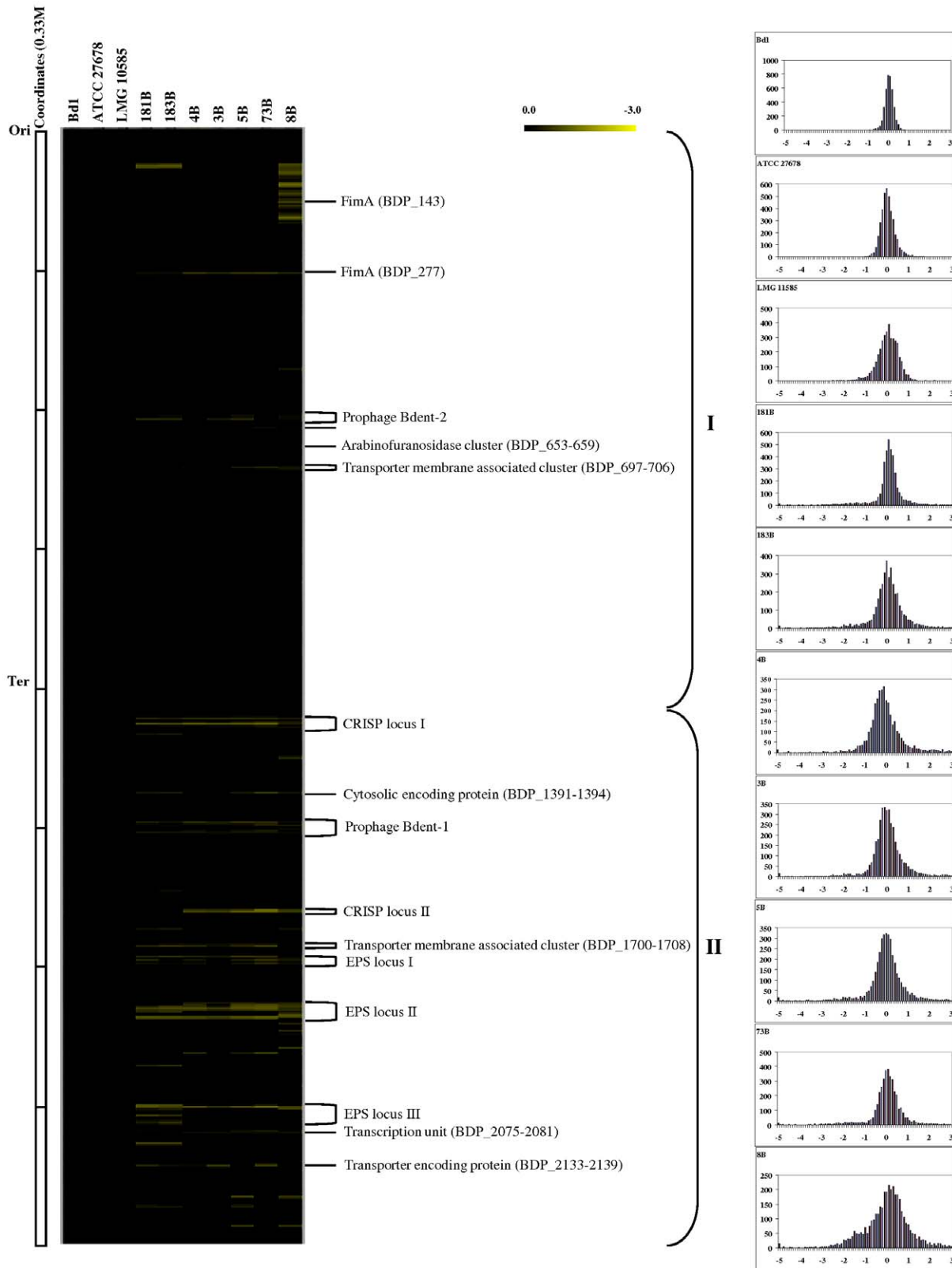


Figure 6. Genomic diversity in the *B. dentium* species with reference to the *B. dentium* strain Bd1 genome (Left) CGH data. Each horizontal row corresponds to a probe on the array, and genes are ordered vertically according to their position on the Bd1 genome. The columns represent the analysed strains, and strains are identified by their code numbers. The colour code corresponding to the presence/absence is given at the top right of the figure: the gradient goes from black to green to indicate the presence, divergence or absence of a gene sequence. The predicted function of some relevant genes are shown on the right-hand margin, ori: origin of replication; ter, terminus of replication (left-hand inset). Right-hand inset displays Signal ratio distribution of the CGH data. The reference is *B. dentium* strain Bd1. Ratios are expressed in a \log_2 scale. See text for details.
doi:10.1371/journal.pgen.1000785.g006

foods. Thus, possessing extensive catabolic abilities for carbohydrates is a potent energy-harvesting mechanism for *B. dentium* Bd1. Genomic data combined with our own data suggests that *B. dentium* Bd1 has a significantly larger arsenal of genes allowing for breakdown of sugars, also called glycobiome [44], as compared to other bifidobacterial species [23–25] or other characterized members of the oral microbiota (Figure 7C and 7D). Classification according to the Carbohydrate Active Enzymes (CAZy) system of Coutinho & Henrissat (1999) showed that the Bd1 genome specifies 117 carbohydrate-active genes including glycoside-hydrolases (GH), glycosyl-transferases (GT) and glycosyl-esterases (CE), which are distributed in 27 GH families, seven GT and three CE families (Figure 7A and 7B). The majority of the identified GH enzymes from *B. dentium* Bd1 are predicted to be intracellular, with a putative cellulase (BDP_2148) and a xylosidase (BDP_0236) predicted to be the only extracellular GH enzymes. Members of GH families that had previously not been detected in bifidobacterial genomes are GH78 (BDP_2152) and GH94 (BDP_2127), which are predicted to be involved in the metabolism of fucose (GH78) and cellobiose (GH94). Furthermore, the Bd1 genome

encodes a wide variety of enzymes to ferment different pentose sugars (e.g., xylose, ribose and arabinose).

The fermentation abilities of this oral strain are clearly broader than those of the phylogenetically related enteric *B. adolescentis* ATCC15703 (Figure S9), probably reflecting the ecological niche it occupies, which apparently contains a higher variety of available sugars as compared to the distal regions of the gastrointestinal tract.

In addition to the transient food components, the human oral cavity is coated with large amounts of viscous secretion produced by the acinar cells of the salivary glands. This secretion consists predominantly of a heterogeneous population of glycoproteins, commonly referred to as salivary mucins [66]. These large, heavily glycosylated glycoproteins play a major role in the maintenance of viscoelastic properties of saliva, participate in the formation of the protective oral mucosal mucus coat and tooth enamel pellicle [67]. Salivary mucins are comprised of 20–22% protein, 0.2% covalently bound fatty acids, and 68–72% carbohydrate [66]. Notably, the carbohydrate component consists mainly of fucose, mannose, galactose (Gal), N-acetylglucosamine (GlcNAc) and N-

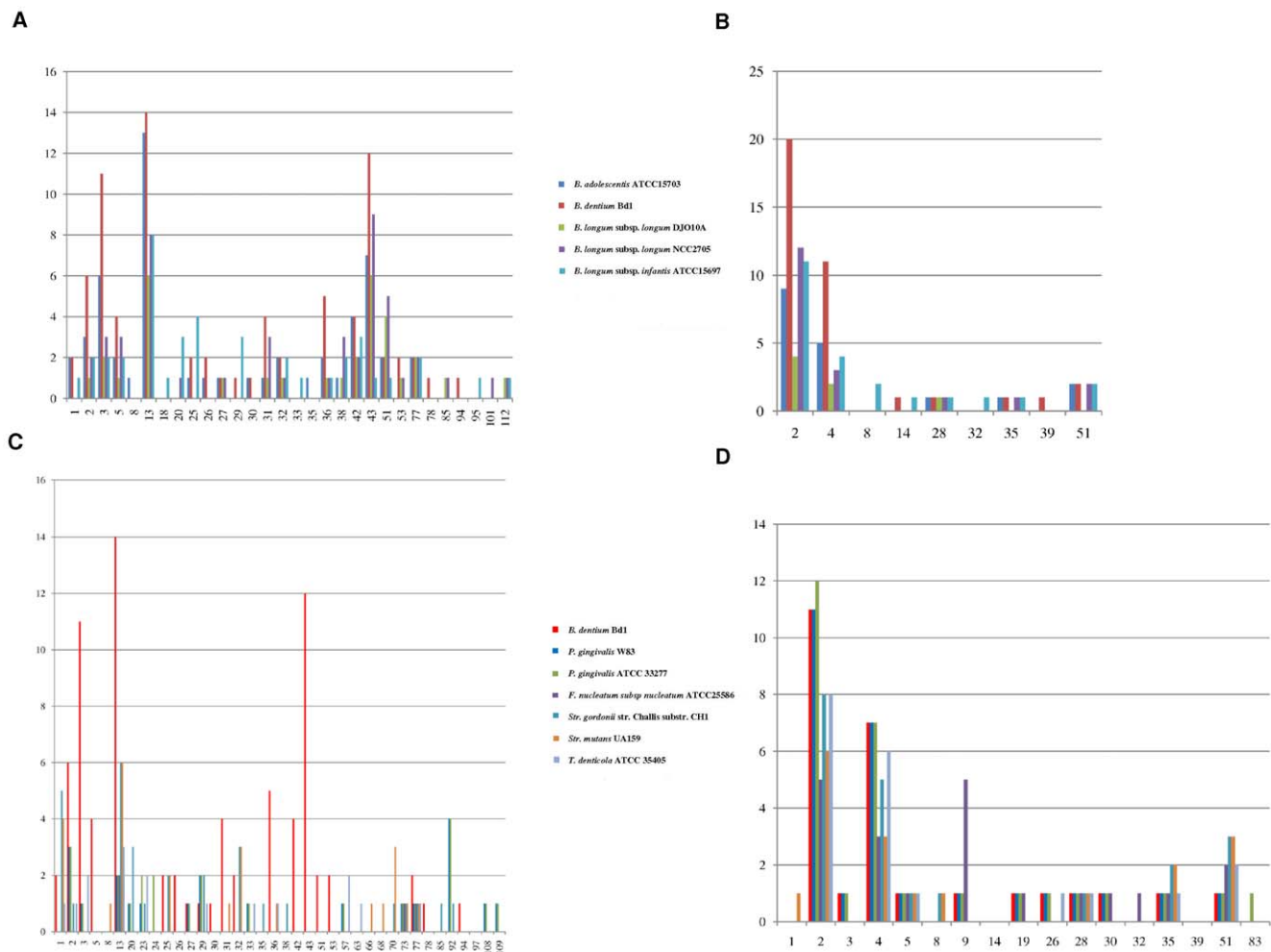


Figure 7. Analysis of the glycobiome of the *B. dentium* Bd1 genome by reference to CAZy database. (A) The glycoside-hydrolase (GH) families identified in the genome of *B. dentium* Bd1 and in enteric bifidobacterial genomes. (B) The glycosyl-transferase families (GT) encoded by the genome of *B. dentium* Bd1 and by enteric bifidobacterial genomes. (C) The GH families identified in the genome of *B. dentium* Bd1 and in other oral pathogens. (D) The GT families identified in the genome of *B. dentium* Bd1 and in other oral pathogens. In each panel, the X-axis represents the different GH families or GT families according to the CAZy database (Henrissat 1999), whereas the Y-axis indicates the abundance of each GH family expressed in percentage.

doi:10.1371/journal.pgen.1000785.g007

acetylgalactosamine (GalNAc). The *B. dentium* Bd1 genome contains an extensive gene repertoire that appears to be dedicated to the metabolism of the backbone of mucin-containing carbohydrate structures, such as Gal β -1,3-GalNAc or Gal β -1,4-GlcNAc disaccharides. This repertoire includes genes for predicted enzymes such as a glucosaminidase and β -galactosidase that could be involved in the removal of monomeric carbohydrates from mucins. Moreover, the presence of a gene encoding a putative fucosidase enzyme (BDP_2152) indicates that *B. dentium* Bd1 can probably degrade fucose-containing glycans, such as those present in salivary mucins [68].

Salivary glycoproteins also contain considerably quantities (3.8–4%) of sialic acid and sulphate which decorate the surface of the mucin sugar backbone [66]. Interestingly, a predicted O-sialoglycoprotein endopeptidase (BDP_1212) and a sialic-acid specific acetyltransferase (BDP_0122) were annotated in the genome of *B. dentium* Bd1. However, genes encoding a sialidase and additional enzymes to degrade sialic acid, which have been identified in other bifidobacterial genomes such as *B. longum* subsp. *infantis* ATCC15697 [25], do not appear to be present in the genome of *B. dentium* Bd1.

Dental caries is initiated by demineralization of the tooth surface due to the action of organic acid formed by dental plaque bacteria, arising from their fermentation of dietary carbohydrates. After fermentable carbohydrate intake, the plaque pH may decrease below the critical pH of 5.5, at which point human enamel undergoes demineralization, within minutes, and may remain acidified for several minutes up to several hours [69,70]. This rapid acidification may not only cause demineralization of tooth surface but also temporarily inhibit bacterial growth in the oral biofilm. Thus, a high level of inherent acid tolerance appears to be crucial for the cariogenicity of oral microorganisms [71]. When the intracellular pH maintained by *B. dentium* under varying external pH conditions was experimentally compared to those of other caries-associated oral bacteria, such as *Str. mutans* and *Lactobacillus paracasei*, *B. dentium* Bd1 displayed a superior ability to keep a more neutral internal pH compared to these two other bacteria (Figure S10A and Nakajo, Takahashi and Beighton, personal communication). Moreover, when *B. dentium* Bd1 was cultivated in a synthetic medium at different pH values the growth of Bd1 was not significantly reduced by the highest level of acidity tested (pH 4), a value which can be reached in the oral cavity after food ingestion (Figure S10B). Notably, other closely related bifidobacteria which occupy a different ecological niche (intestinal vs. oral) do not exhibit this aciduric property (Figure S10B and data not shown). Higher levels of inherent tolerance of oral bacteria to acidification have been related to the presence of a membrane-bound, acid-stable, proton-translocating F₁F₀ ATPase system whose activity has been considered crucial in maintaining the intracellular pH at 7.5 [71]. In the *B. dentium* Bd1 genome, the F₁F₀-ATPase is encoded by the *atp* operon, but this system is also encoded by other bifidobacteria [72]. However, the genome of *B. dentium* Bd1 contains two adjacently located genes (BDP_1749 and BDP_1750) encoding a glutamate decarboxylase (GadB) and a glutamate/ γ -aminobutyrate anti-porter (GadC), not present in other bifidobacterial genomes so far published [23–25], and known in other bacteria to form a glutamate-dependent acid resistance system 2 (AR2) [73].

Transcriptional profiling analysis of *B. dentium* Bd1 and adaptation to the oral lifestyle

Characteristics contributing to the ecological fitness in the oral cavity, such as utilization of different diet-derived carbohydrates, and stress tolerance to antimicrobial compounds and acidic

environments, should be discernible in *B. dentium* Bd1. To determine if *B. dentium* Bd1 functionally responds to stressful stimuli, we performed transcriptional profiling studies using Agilent arrays (Agilent, Palo Alto, Ca., USA) that contain oligonucleotides representing 2114 of the 2143 predicted *B. dentium* Bd1 protein-encoding genes.

(i) Identification of genes differentially expressed upon exposure to acidic environments. To study changes in gene expression occurring in *B. dentium* Bd1 in response to acidic conditions, cultures were exposed to acidic conditions simulating those occurring in active carious lesions (pH 4). In a time course experiment the global gene expression of *B. dentium* Bd1 cells exposed to a pH change from 7.0 to 4.0 was compared to gene expression of untreated cells. Samples were collected 30 minutes and 2 hours after the onset of acid stress, and submitted to transcriptome analysis. The complete list of the genes whose expression pattern was altered following acidic exposure is presented in Figure 8 and Table S2. In total, 46 genes displayed a profound change in mRNA expression levels (greater than tenfold) in response to acid stress. Genes BDP_1749 and BDP_1750, encoding a glutamate carboxylate and a glutamate: γ -aminobutyrate antiporter were upregulated 90 and 51 fold, respectively, consistent with their role in acid stress (see above). Also upregulated were genes encoding aspartate ammonium lyase (BDP_0309), formyl-coenzyme A transferase (BDP_1963), oxalyl-CoA decarboxylase (BDP_1966) and malate dehydrogenase (BDP_1119), all known to provide mechanisms to cope with acid stress [74]. Additionally, genes involved in arginine (i.e. ArgD) or glutamate metabolism were upregulated, perhaps linking this to the deamination of branched chain amino acids (BCAA) as a mechanism of maintaining the internal pH of the cells [75]. Furthermore, acetylornithine aminotransferase (ArgD) catalyzes a reaction that could serve as a way of recycling the 2-oxoglutarate formed by the BCAA aminotransferase as well as feeding glutamine synthesis.

The bile salt hydrolase-encoding gene, BDP_1106, was shown to be upregulated 97 fold after prolonged acid stress. It is tempting to speculate that *B. dentium* increases its bile salt hydrolase activity in response to imminent exposure to high levels of bile salt when it senses a low pH in its environment (e.g. when exposed to gastric juice during the passage from the oral cavity to the intestine). Other upregulated genes, especially after 2 hrs of acid stress, had in common that they are involved in proteolytic degradation or amino acid uptake or amino acid catabolism. Surprisingly the F₁F₀-ATPase, encoded by BDP_1952–1959, was not upregulated, in contrast to what has been observed for *B. animalis* subsp. *lactis* DSM10140 [72]. These data suggest that *B. dentium* manages its acid tolerance by amino acid degradation.

ii) Induction of genes related to the utilization of carbohydrates. Genome analysis of *B. dentium* Bd1, as supported by growth tests, revealed extensive genetic capabilities of this strain to ferment carbohydrates (see above and Figure S9). To validate the genetic basis for these observations, we investigated the transcriptome of *B. dentium* Bd1 grown on the most efficiently utilized carbohydrates (glucose, ribose, maltose, mannitol and amygdalin). As expected the transcriptional profiling data revealed induction of genes involved in sugar metabolism (Figure 8 and Table S3). Genes specifically induced when Bd1 was grown on ribose included *rbsR1*, *rbsR2*, *rbsA*, *rbsB*, *rbsC* and *rbsD* encoding transcriptional regulators, ABC transporters and ribose pyranase (Table S3), which are organized in a genetic cluster (BDP_170-BDP_175). In addition, the expression levels of another locus spanning BDP_2024 to BDP_2026, encoding a ribokinase, an inosine-

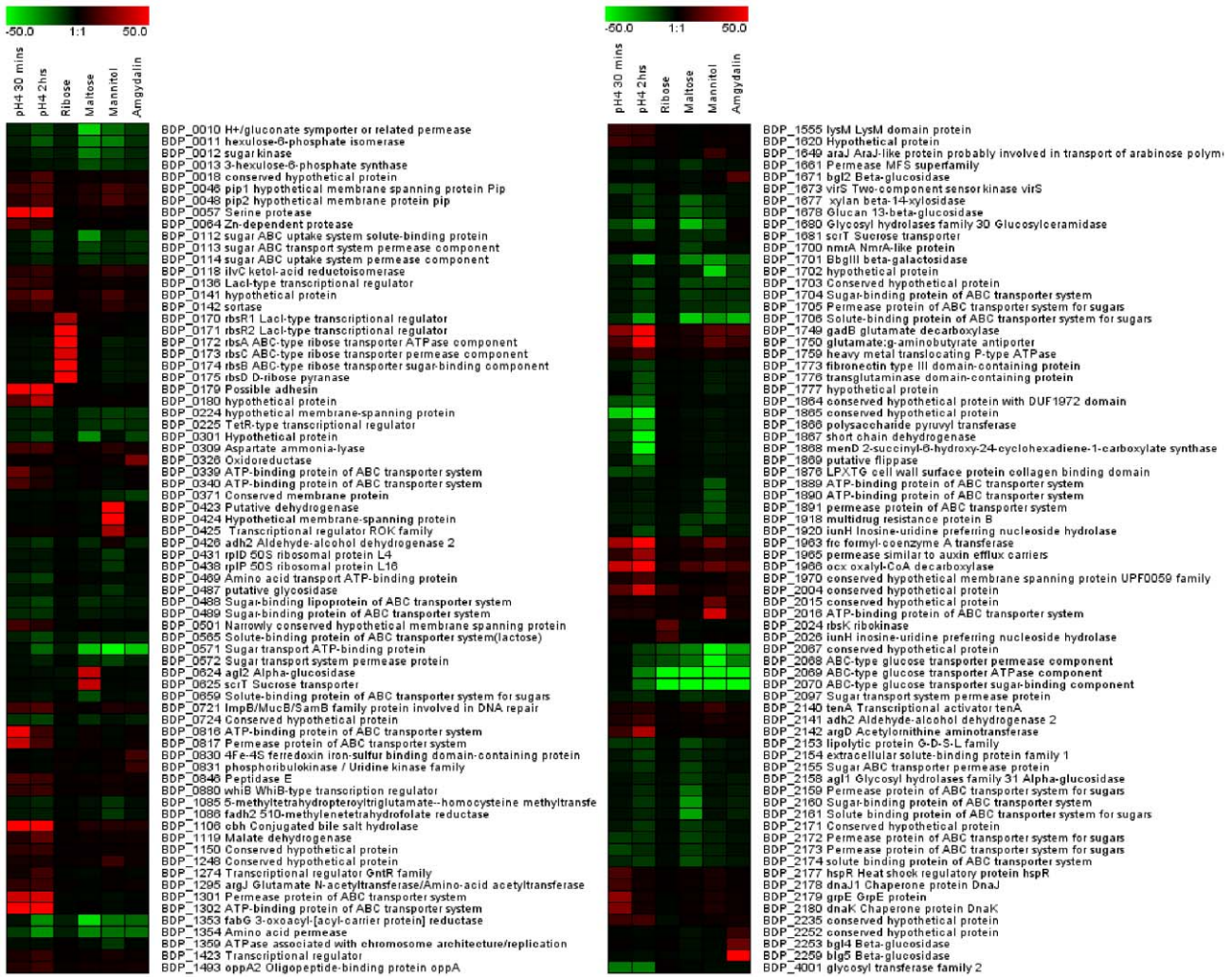


Figure 8. Identification of *B. dentium* Bd1 transcribed genes by DNA-micro array analysis. The heat-map indicates the change in the expression upon cultivation of Bd1 cells at low pH as well as represents selected genes that were up- or down-regulated when grown in media containing various carbohydrates as the sole carbon sources as compared to growth on glucose. Each row represents a separate transcript and each column represents a separate sample. Colour legend is on the top of each micro array plot, red indicates increased transcription levels, whereas green indicates decreased transcription level as compared to the reference samples (glucose-grown samples or growth at pH 6). doi:10.1371/journal.pgen.1000785.g008

uridine nucleoside hydrolase and a transporter, respectively, were shown to be more than 9-fold higher than their levels in the absence of ribose (Table S3). Maltose afforded excellent growth of *B. dentium* Bd1 and based on transcriptomics data, the genes related to the utilization of this sugar included two adjacent genes (BDP_0624-BDP_0625), which were highly up-regulated, and which encode a putative α -glucosidase and a sugar transporter, respectively (Table S3 and Figure S9).

The transcription of three genes (BDP_2252, BDP_2253 and BDP_2259) were shown to be enhanced when the Bd1 was grown on amygdalin, suggesting their role in the utilization of this complex sugar. Three adjacent genes (BDP_0423-BDP_0425), encoding a mannitol dehydrogenase, a mannitol permease and a transcriptional regulator, were shown to be up-regulated between 32- and 200- fold when Bd1 was grown in the presence of mannitol (Table S3).

iii) Evaluation of genes involved in biocide resistance. Oral bacteria are exposed to a large variety of antimicrobials/biociodes that are ingested with food (e.g.

bacteriocins and food preservatives) and/or are used in normal hygiene practices (e.g., mouth-washes). Transcriptional profiling was performed of the *B. dentium* Bd1 strain grown in the presence of diluted mouth-washes or of the commonly used mouth biocides such as chlorhexidine was performed. In total 112 genes were upregulated (fold >2.5) when cells were exposed to the mouth-wash for 2h or 8h (data not shown). These include genes encoding components of predicted major facilitator systems and extracellular solute binding proteins, which may chelate and extrude the cytotoxic compounds of the mouth-washes. As expected, other stress-related genes were induced upon exposure of Bd1 cells to mouthwash, such as those coding molecular chaperones, aminotransferase and glycosyltransferases. In contrast, only 2 genes were upregulated when the cells were exposed to chlorhexidine, while >1000 genes were down-regulated. Presumably exposure to chlorhexidine is lethal, thereby resulting in a drastic decrease of all mRNA. The upregulated genes include a putative multidrug transporter *mdrB* (BDP_1665) (data not shown).

Putative virulence factors

Although *B. dentium* is not an invasive, life-threatening pathogen it plays a role in tooth tissue destruction and infects tooth dentine, and there are a number of ORFs that code for potential colonization or virulence factors, such as adhesins, exoenzymes, protease- and cytokine-modulating molecules, as well as putative hemolysins (see Table 2 for an overview of putative virulence factors of *B. dentium* Bd1, some of which will be discussed below). The latter are similar to hemolysins from oral pathogens, including Hemolysin A from *S. mutans*, a coiled-coil myosin-like protein. Among the putative *B. dentium* Bd1 surface antigen proteins, the BDP_0164 protein displays 51% similarity to the *T. denticola* pathogen-specific surface antigen, and is flanked on one side by an ORF (BDP_0163) involved in iron metabolism (high-affinity iron permease) and on the other side by a gene (BDP_0165) encoding an integral membrane protein with high similarity to a protein encoded by the oral pathogen *Fusobacterium nucleatum*. This suggests that the gene cluster (BDP_0163-BDP_0165) is involved in iron acquisition and adhesion.

Oral bacteria can adhere to salivary agglutinin, other plaque bacteria, extracellular matrix and epithelial cell-surface receptors [76]. In the most intensely studied oral pathogen, *S. mutans*, two major adhesins mediate this attachment: cell-surface or adhesion proteins, such as SpaA adhesins [77], and sucrose-derived glucans (e.g., *gfbB*). A homolog (BDP_2059; 32% identity) of the gene encoding a major adhesin of viridans streptococci, SpaA, was identified in the genome of *B. dentium* Bd1 (Table 2). SpaA binds to human salivary agglutinin, collagen and cells of certain oral pathogens, such as *Actinomyces naeslundii* [78,79]. Notably, the BDP_2059 putative adhesin also contains a domain that is similar to a domain of a *Streptococcus gordonii* protein which mediates strong lactose-inhibitable coaggregation [80]. Furthermore, part of an operon that is required for the synthesis of cell wall polysaccharides in *S. mutans* UA159, i.e. *rgpA*, *rgpB*, *rgpD* and *rgpC* [81,82], exhibits clear homology with *B. dentium* Bd1 ORFs BDP_1864, BDP_1864a, BDP_2047 and BDP_2048, respectively. In *S. mutans* these genes play not only a crucial role in binding to human oral tissues [83,84] but they also participate in serotype determination [85]. These genes also show a strong divergence in G+C content relative to the remainder of the genome, indicating that this region has been acquired by horizontal gene transfer.

The *B. dentium* Bd1 genome also specifies surface proteins with domains that resemble those (Pfam number 31902) responsible for

inter-bacterial aggregation by a choline-binding domain. Such choline-binding motifs are present in the ligands of the most important pneumococcal virulence proteins, encoded by *pspC* and *pspA*, which presumably act as adhesins that bind to host factors such as IgA and factor H [86]. Notably, and in contrast to the available chromosomes of enteric bifidobacteria, the *B. dentium* Bd1 genome harbors five adjacent genes that encode proteins containing such predicted choline-binding domains (BDP_2045, BDP_2054, BDP_2056, BDP_2059 and BDP_2061).

Furthermore, a large number of predicted surface and extracellular proteins that may be involved in host attachment and interaction were identified in a similar fashion as described for other oral pathogens [87,88] (Table 2).

Other potential adhesion and virulence factors include glyco-protein-binding fimbriae that, in the oral cavity, may mediate the recognition of and adhesion to salivary proline-rich proteins that bind to tooth and mucosal epithelial cell surfaces. They may also bind to cell wall polysaccharides of certain oral bacteria [89–94]. So far bifidobacteria have not been shown to possess any fimbria-like structures on their cell surface, although homologs of fimbrial subunits have been identified in intestinal bifidobacteria [23,44]. Notably, four loci encoding homologs of known fimbrial subunits FimA, FszB and FszD were identified in the *B. dentium* Bd1 genome (Figure 9). FimA (BDP_0535 and BDP_1224) displays high identity to FimA homologs found in oral commensals such as *A. naeslundii* and *A. odontolyticus* [91,95].

Finally, *B. dentium* Bd1 encodes a number of putative proteases that may contribute to virulence by their ability to degrade host proteins for bacterial nutrition [96,97] (Table 2).

Discussion

We report here the *B. dentium* Bd1 genome sequence that constitutes the first genome based analysis of a bifidobacterial taxon recognized as an opportunistic pathogen. Although a large number of bacteria coexist in the oral cavity and upper respiratory tract in humans they have evolved to form a microbial community with complex physical and biochemical interactions. The analysis of the *B. dentium* Bd1 genome and comparisons with other bifidobacteria residing in the human intestine has revealed insights into the particular evolution and adaptive responses of the opportunistic pathogen *B. dentium* to the oral cavity. *B. dentium*, unlike its intestinal relatives that often are claimed to promote the health-status of their host, contributes to the destruction of the

Table 2. Putative virulence factors identified in the genome of *B. dentium* Bd1.

General function	Gene	Predicted function
Exotoxins	BD_0019, BDP_0878, BDP_1280	Hemolysin
Adhesins	BDP_0517	Myosin-cross reactive antigen
	BDP_523, BDP_643, BDP_1896_	Surface antigen
	BDP_2059	SpaA adhesin
	BDP_1864, BDP_1864a, BDP_2047, BDP_2048	Cell wall polysaccharides biosynthesis
	BDP_2045, BDP_2054, BDP_2056, BDP_2059, BDP_2061	Inter-bacterial aggregation protein
	BDP_0279, BDP_2189	Collagen adhesion proteins
	BDP_1773	Fibronectin/fibrinogen-binding protein
Exoenzymes/proteases	BDP_535, BDP_1224	FimA
	BDP_0549	Zn-dependent protease
	BDP_0375	Collagenase

doi:10.1371/journal.pgen.1000785.t002

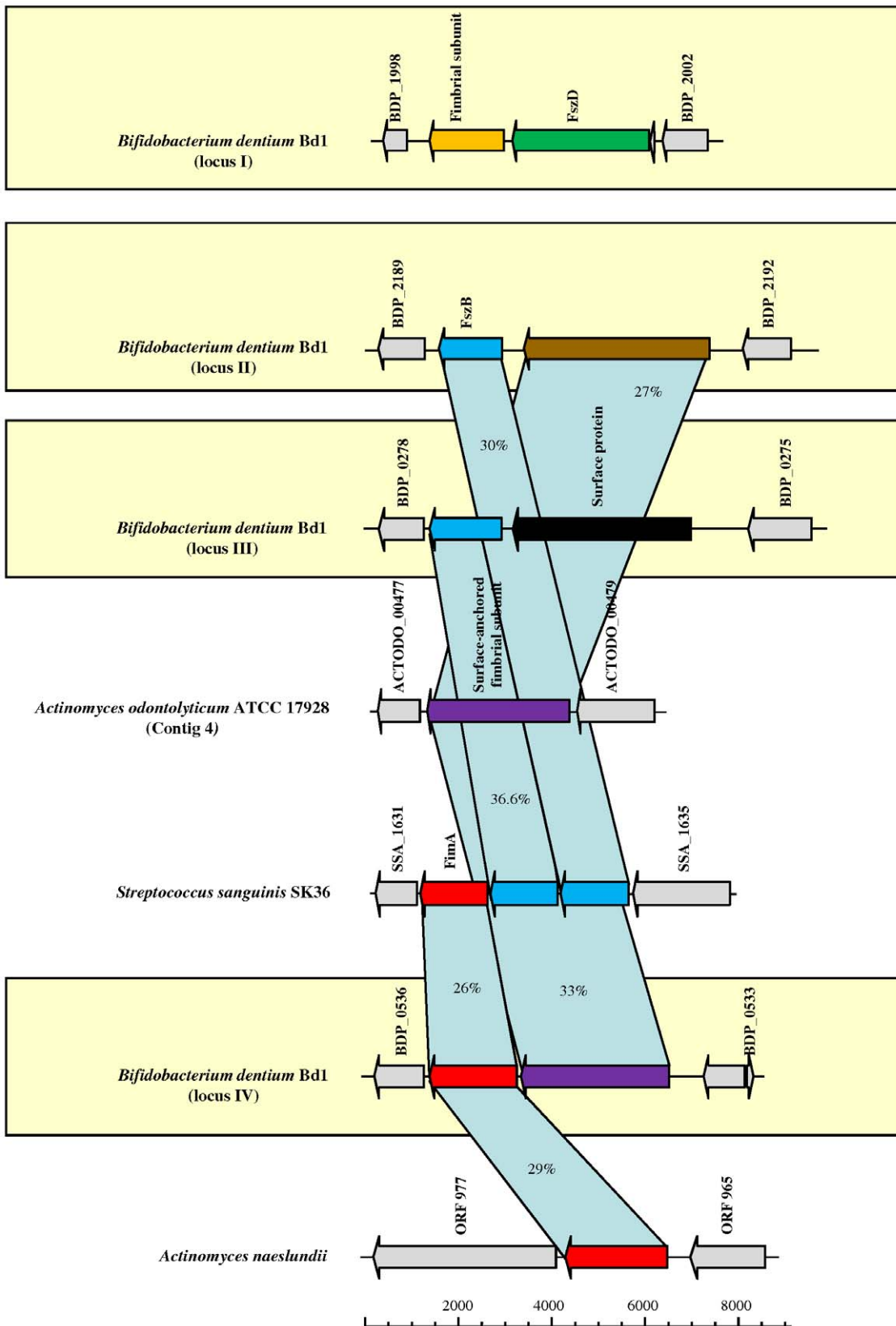


Figure 9. Comparison of presumptive fimbrial loci in *B. dentium* Bd1 with the corresponding loci in different oral bacteria. Each arrow indicates an ORF. The length of the arrow is proportional to the length of the predicted open reading frame. Corresponding genes are marked with the same colour. The putative function of the protein is indicated above each arrow.
doi:10.1371/journal.pgen.1000785.g009

dentition. The elucidation of a putative “cariogenic gene suite” would provide salient targets to test the relative contribution of *B. dentium* to a core cariogenic microbiome, its impact on microbial colonization and succession, and its phylogenetic distribution. It remains unknown if these genomic features constitute a unique competitive strategy evolved in *B. dentium*. As such, it differs from other known bifidobacteria in several aspects of its basic physiology and its adaptation to an ecological niche. As our genome analysis shows, *B. dentium* can metabolize a much larger variety of carbohydrates than other *Bifidobacterium* species sequenced so far and greater than the range of oral streptococcal including *S. mutans*. Genes encoding putative virulence factors associated with adhesins, acid tolerance, defense against toxic substances and capacity in utilizing saliva-derived components, represents genetic evidence of the capacity of *B. dentium* to colonize the oral cavity and to proliferate within active carious lesions. The genome sequence, when explored using functional genomics approaches, will permit the analysis of genes involved in colonization, survival, growth and pathobiology of *B. dentium* in this unique polymicrobial environment.

Materials and Methods

The strain used in this study *B. dentium* Bd1 is equivalent to the type strain of *B. dentium* species (ATCC27534 or LMG11045 or DSM20436 or JCM 1195).

The genome sequence of *B. dentium* Bd1 was determined by shotgun sequencing and subsequent gap closure (Agencourt Genomic Services, MA, USA). The Bd1 genome was sequenced to approximately 10-fold coverage and assembled with Phred [98], Phrap and the Staden package [99]. Automated gene modelling was achieved using multiple databases and modelling packages as described previously [100]. Additional information on sequencing, bioinformatic, and functional genomics analyses are provided in Text S1.

Nucleotide sequence accession numbers

The sequence reported in this article has been deposited in the GenBank database (accession number CP001750).

Supporting Information

Figure S1 Schematic representation of the GC content bias in 696 genomes. The *B. dentium* Bd1 GC content is circularized.
Found at: doi:10.1371/journal.pgen.1000785.s001 (0.34 MB TIF)

Figure S2 Genome colinearity of *B. dentium* Bd1 with *B. adolescentis* ATCC15703 (A), *B. longum* subsp. *infantis* ATCC15678 (B), *B. longum* subsp. *longum* DJO10A (C), *B. longum* subsp. *longum* NCC2705 (D), and *B. animalis* subsp. *lactis* ADO11 (E). Each dot matrix was calculated using MUMmer. The comparison window was 50 bp and the stringency was 30 bp.
Found at: doi:10.1371/journal.pgen.1000785.s002 (1.80 MB TIF)

Figure S3 Phylogenetic supertree based on the sequences of *Actinobacteria* core proteins, using SplitsTree.
Found at: doi:10.1371/journal.pgen.1000785.s003 (0.29 MB TIF)

Figure S4 Venn diagram of homologs shared between sequenced bifidobacterial genomes. Circle sizes are proportional to members contained in each set.
Found at: doi:10.1371/journal.pgen.1000785.s004 (0.31 MB TIF)

Figure S5 Analysis of the proteins encoded by the genome of *B. dentium* Bd1 according to the distribution of functions in terms of SCOP Domain Superfamilies. (A) shows a pie chart displaying the proportion of proteins encoded by the genome of *B. dentium* Bd1

classified according to general functional categories, while (B) shows a pie chart distribution of the proteins encoded by the genome of *B. dentium* Bd1 when classified according to more detailed functional categories.

Found at: doi:10.1371/journal.pgen.1000785.s005 (0.47 MB TIF)

Figure S6 Comparative analysis of Cas1 protein sequences. CRISPR repeat families are indicated within the shaded boxes. Bootstrap values are indicated at the nodes for a total of 1,000 replicates. The arrows pointed the Cas protein in the two CRISPR loci of *B. dentium* Bd1. Bootstrap values above 40 are shown.

Found at: doi:10.1371/journal.pgen.1000785.s006 (0.39 MB TIF)

Figure S7 Polyphasic analysis of the genetic diversity in the *B. dentium* species using the CGH clustering data and a phylogenetic tree of the ten *B. dentium* strains computed from the concatenation of *clpC*, *dnaJ1*, *rpoC*, and *xfp* gene sequences by the neighbour-joining method and Kimura’s two parameter model as the substitution model. In each tree, the strain is indicated at the right end of the branch, the colour typing indicates the different ecological origin, i.e., red from dental caries, green from saliva of patients with caries, blue from saliva of healthy patients and yellow from fecal samples. The numbers at the nodes relate to the bootstrap probabilities. The different clusters are boxed.

Found at: doi:10.1371/journal.pgen.1000785.s007 (0.34 MB TIF)

Figure S8 Predicted transport capabilities of *B. dentium* Bd1 compared to other bifidobacteria (A). Predicted compounds transported by the sequenced bifidobacteria (B).

Found at: doi:10.1371/journal.pgen.1000785.s008 (0.43 MB TIF)

Figure S9 Carbohydrate metabolizing capabilities of *B. dentium* Bd1. (A) shows the sugar fermentation profiles of *B. dentium* Bd1 and *B. adolescentis* ATCC15703 strains, respectively. Carbohydrates used are indicated. + indicates acid production; – indicates absence of acid production. (B) displays the growth curves of *B. dentium* Bd1 on different carbohydrates as their sole carbon source. The carbohydrates used are indicated next to each curve.

Found at: doi:10.1371/journal.pgen.1000785.s009 (6.37 MB TIF)

Figure S10 Ecological adaptability of *B. dentium* Bd1 to acidic environments. (A) indicates the intracellular pH of *B. dentium* Bd1, *S. mutans* UA159, and *Lb. paracasei* subsp. *paracasei* ATCC11974 at various extracellular pH values. The data obtained from three independent experiments were plotted. In (B), growth of *B. dentium* Bd1 and *B. longum* subsp. *longum* ATCC 15707 cultures maintained at different acidic conditions were monitored over 24 hours. The colour of the line indicates the pH value of the medium used: red, pH 5; blue, pH4; green, pH 3.

Found at: doi:10.1371/journal.pgen.1000785.s010 (0.90 MB TIF)

Table S1 Number of ABC exporter and importer system according to substrate type present in the bifidobacterial genome sequenced so date.

Found at: doi:10.1371/journal.pgen.1000785.s011 (0.05 MB DOC)

Table S2 Selected genes upregulated/downregulated upon acidic stress.

Found at: doi:10.1371/journal.pgen.1000785.s012 (0.07 MB DOC)

Table S3 Selected genes differentially expressed upon *B. dentium* Bd1 growth in different sugar-based media relative to growth in glucose.

Found at: doi:10.1371/journal.pgen.1000785.s013 (0.09 MB DOC)

Text S1 Predicted interactions between *B. dentium* Bd1 and host, genetic regions of oral pathogens present in the *B. dentium* Bd1 genome and detailed description of the experimental procedures. Found at: doi:10.1371/journal.pgen.1000785.s014 (0.13 MB DOC)

Acknowledgments

We thank Prof. Angelo Pavesi for his very valuable support on statistical analyses. We also thank Prof. Page Caufield for critical reading of the

manuscript. We wish to thank all students and coworkers who contributed data and their enthusiasm.

Author Contributions

Conceived and designed the experiments: MV. Performed the experiments: MV FT AZ EF MM LM AF MD. Analyzed the data: MV VG FB CC MJC FH BG BH PC MO RSG ZZ DB PWO. Contributed reagents/materials/analysis tools: MV GFF. Wrote the paper: MV DvS.

References

- Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, et al. (2007) Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* 71: 495–548.
- Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. *Science* 307: 1915–1920.
- Guarner F, Malagelada JR (2003) Gut flora in health and disease. *Lancet* 361: 512–519.
- Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, et al. (2001) Molecular analysis of commensal host-microbial relationships in the intestine. *Science* 291: 881–884.
- Marco ML, Pavan S, Kleerebezem M (2006) Towards understanding molecular modes of probiotic action. *Curr Opin Biotechnol* 17: 204–210.
- O'Hara AM, Shanahan F (2007) Mechanisms of action of probiotics in intestinal diseases. *ScientificWorldJournal* 7: 31–46.
- Ouweland AC, Salminen S, Isolauri E (2002) Probiotics: an overview of beneficial effects. *Antonie Van Leeuwenhoek* 82: 279–289.
- Leahy SC, Higgins DG, Fitzgerald GF, van Sinderen D (2005) Getting better with bifidobacteria. *J Appl Microbiol* 98: 1303–1315.
- Becker MR, Paster BJ, Leys EJ, Moeschberger ML, Kenyon SG, et al. (2002) Molecular analysis of bacterial species associated with childhood caries. *J Clin Microbiol* 40: 1001–1009.
- Aas JA, Griffen AL, Dardis SR, Lee AM, Olsen I, et al. (2008) Bacteria of dental caries in primary and permanent teeth in children and young adults. *J Clin Microbiol* 46: 1407–1417.
- Beighton D (2005) The complex oral microflora of high-risk individuals and groups and its role in the caries process. *Community Dent Oral Epidemiol* 33: 248–255.
- Modesto M, Biavati B, Mattarelli P (2006) Occurrence of the family bifidobacteriaceae in human dental caries and plaque. *Caries Res* 40: 271–276.
- Chavez de Paz LE, Molander A, Dahlen G (2004) Gram-positive rods prevailing in teeth with apical periodontitis undergoing root canal treatment. *Int Endod J* 37: 579–587.
- Hojo K, Mizoguchi C, Taketomo N, Ohshima T, Gomi K, et al. (2007) Distribution of salivary *Lactobacillus* and *Bifidobacterium* species in periodontal health and disease. *Biosci Biotechnol Biochem* 71: 152–157.
- Love RM, Jenkinson HF (2002) Invasion of dentinal tubules by oral bacteria. *Crit Rev Oral Biol Med* 13: 171–183.
- Filоче SK, Soma KJ, Sissons CH (2007) Caries-related plaque microcosm biofilms developed in microplates. *Oral Microbiol Immunol* 22: 73–79.
- Mantzourani M, Gilbert SC, Sulong HN, Sheehy EC, Tank S, et al. (2009) The isolation of bifidobacteria from occlusal carious lesions in children and adults. *Caries Res* 43: 308–313.
- Mantzourani M, Fenlon M, Beighton D (2009) Association between Bifidobacteriaceae and the clinical severity of root caries lesions. *Oral Microbiol Immunol* 24: 32–37.
- van Houte J, Lopman J, Kent R (1996) The final pH of bacteria comprising the predominant flora on sound and carious human root and enamel surfaces. *J Dent Res* 75: 1008–1014.
- Hoshino E (1985) Predominant obligate anaerobes in human carious dentin. *J Dent Res* 64: 1195–1198.
- Anusavice KJ (2002) Dental caries: risk assessment and treatment solutions for an elderly population. *Compend Contin Educ Dent* 23: 12–20.
- Marsh PD (1994) Microbial ecology of dental plaque and its significance in health and disease. *Adv Dent Res* 8: 263–271.
- Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, et al. (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U S A* 99: 14422–14427.
- Lee JH, Karamychev VN, Kozyavkin SA, Mills D, Pavlov AR, 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.
- Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, et al. (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci U S A* 105: 18964–18969.
- Kim JF, Jeong H, Yu DS, Choi SH, Hur CG, et al. (2009) Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. *J Bacteriol* 191: 678–679.
- Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, et al. (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and BI-04. *J Bacteriol* 191: 4144–4151.
- Scardovi V, Crociani F (1974) *Bifidobacterium catenulatum*, *Bifidobacterium dentium*, and *Bifidobacterium angulatum*: three new species and their deoxyribonucleic acid homology relationships. *Int J Syst Bacteriol* 24: 6–20.
- Kapatral V, Anderson I, Ivanova N, Reznik G, Los T, et al. (2002) Genome sequence and analysis of the oral bacterium *Fusobacterium nucleatum* strain ATCC 25586. *J Bacteriol* 184: 2005–2018.
- Qin MH, Madiraju MV, Rajagopalan M (1999) Characterization of the functional replication origin of *Mycobacterium tuberculosis*. *Gene* 233: 121–130.
- Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, et al. (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390: 249–256.
- Ulrich DL, Kojetin D, Bassler BL, Cavanagh J, Loria JP (2005) Solution structure and dynamics of LuxU from *Vibrio harveyi*, a phosphotransferase protein involved in bacterial quorum sensing. *J Mol Biol* 347: 297–307.
- Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, et al. (2003) A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* 299: 2074–2076.
- Eisen JA, Heidelberg JF, White O, Salzberg SL (2000) Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol* 1: RESEARCH0011.
- Canchaya C, Claesson MJ, Fitzgerald GF, van Sinderen D, O'Toole PW (2006) Diversity of the genus *Lactobacillus* revealed by comparative genomics of five species. *Microbiology* 152: 3185–3196.
- Chan CX, Beiko RG, Ragan MA (2006) Detecting recombination in evolving nucleotide sequences. *BMC Bioinformatics* 7: 412.
- Henz SR, Huson DH, Auch AF, Nieselt-Struwe K, Schuster SC (2005) Whole-genome prokaryotic phylogeny. *Bioinformatics* 21: 2329–2335.
- Sakamoto M, Umeda M, Benno Y (2005) Molecular analysis of human oral microbiota. *J Periodontol Res* 40: 277–285.
- Tanner AC, Milgrom PM, Kent R, Jr., Mokeem SA, Page RC, et al. (2002) The microbiota of young children from tooth and tongue samples. *J Dent Res* 81: 53–57.
- Black PN (1991) Primary sequence of the *Escherichia coli* *fadL* gene encoding an outer membrane protein required for long-chain fatty acid transport. *J Bacteriol* 173: 435–442.
- Shi J, Blundell TL, Mizoguchi K (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* 310: 243–257.
- De Smet K, Contreras R (2005) Human antimicrobial peptides: defensins, cathelicidins and histatins. *Biotechnol Lett* 27: 1337–1347.
- Karlin A, Akabas MH (1998) Substituted-cysteine accessibility method. *Methods Enzymol* 293: 123–145.
- Ventura M, Canchaya C, Fitzgerald GF, Gupta RS, van Sinderen D (2007) Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria. *Antonie Van Leeuwenhoek* 91: 351–372.
- Horvath P, Coute-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, et al. (2008) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int J Food Microbiol*.
- Jansen R, Embden JD, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43: 1565–1575.
- Mojica FJ, Diez-Villasenor C, Soria E, Juez G (2000) Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* 36: 244–246.
- Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 1: 7.
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315: 1709–1712.
- Berger B, Pridmore RD, Barretto C, Delmas-Julien F, Schreiber K, et al. (2007) Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. *J Bacteriol* 189: 1311–1321.

51. Dobrindt U, Agerer F, Michaelis K, Janka A, Buchrieser C, et al. (2003) Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J Bacteriol* 185: 1831–1840.
52. Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM (2001) Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc Natl Acad Sci U S A* 98: 8821–8826.
53. Le Gall T, Darlu P, Escobar-Paramo P, Picard B, Denamur E (2005) Selection-driven transcriptome polymorphism in *Escherichia coli*/*Shigella* species. *Genome Res* 15: 260–268.
54. Leonard EE, 2nd, Tompkins LS, Falkow S, Nachamkin I (2004) Comparison of *Campylobacter jejuni* isolates implicated in Guillain-Barre syndrome and strains that cause enteritis by a DNA microarray. *Infect Immun* 72: 1199–1203.
55. Porwollik S, Wong RM, McClelland M (2002) Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis. *Proc Natl Acad Sci U S A* 99: 8956–8961.
56. Salama N, Guillemin K, McDaniel TK, Sherlock G, Tompkins L, et al. (2000) A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc Natl Acad Sci U S A* 97: 14668–14673.
57. Rezzonico E, VM, Cuanoud G, Pessi G, Gilberti G, Arigoni F (2003) DNA-array based analysis of genome variation and gene expression in bifidobacteria.; 2003; Baveno, Italy.
58. Alm RA, Ling LS, Moir DT, King BL, Brown ED, et al. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397: 176–180.
59. Romo-Gonzalez C, Salama NR, Burgeno-Ferreira J, Ponce-Castaneda V, Lazcano-Ponce E, et al. (2009) Differences in genome content among *Helicobacter pylori* isolates from patients with gastritis, duodenal ulcer, or gastric cancer reveal novel disease-associated genes. *Infect Immun* 77: 2201–2211.
60. Brussow H, Desiere F (2001) Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. *Mol Microbiol* 39: 213–222.
61. Ventura M, Canchaya C, Del Casale A, Dellaglio F, Neviani E, et al. (2006) Analysis of bifidobacterial evolution using a multilocus approach. *Int J Syst Evol Microbiol* 56: 2783–2792.
62. de Vries W, Stouthamer AH (1969) Factors determining the degree of anaerobiosis of *Bifidobacterium* strains. *Arch Mikrobiol* 65: 275–287.
63. Busch W, Saier MH, Jr. (2004) The IUBMB-endorsed transporter classification system. *Mol Biotechnol* 27: 253–262.
64. Driessen AJ, Nouwen N (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* 77: 643–667.
65. Saier MH, Jr., Yen MR, Noto K, Tamang DG, Elkan C (2009) The Transporter Classification Database: recent advances. *Nucleic Acids Res* 37: D274–278.
66. Slomiany BL, Murty VL, Piotrowski J, Slomiany A (1996) Salivary mucins in oral mucosal defense. *Gen Pharmacol* 27: 761–771.
67. Slomiany BL, Murty VL, Zdebska E, Slomiany A, Gwozdziński K, et al. (1986) Tooth surface-pellicle lipids and their role in the protection of dental enamel against lactic-acid diffusion in man. *Arch Oral Biol* 31: 187–191.
68. Klein A, Carnoy C, Wieruszski JM, Strecker G, Strang AM, et al. (1992) The broad diversity of neutral and sialylated oligosaccharides derived from human salivary mucins. *Biochemistry* 31: 6152–6165.
69. Igarashi K, Kamiyama K, Yamada T (1981) Measurement of pH in human dental plaque in vivo with an ion-sensitive transistor electrode. *Arch Oral Biol* 26: 203–207.
70. Takahashi N, Yamada T (1999) Acid-induced acid tolerance and acidogenicity of non-mutans streptococci. *Oral Microbiol Immunol* 14: 43–48.
71. Quivey RG, Kuhnert WL, Hahn K (2001) Genetics of acid adaptation in oral streptococci. *Crit Rev Oral Biol Med* 12: 301–314.
72. Ventura M, Canchaya C, van Sinderen D, Fitzgerald GF, Zink R (2004) *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.
73. Richard H, Foster JW (2004) *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J Bacteriol* 186: 6032–6041.
74. Teixeira-Gomes AP, Cloeckaert A, Zygmunt MS (2000) Characterization of heat, oxidative, and acid stress responses in *Brucella melitensis*. *Infect Immun* 68: 2954–2961.
75. Len AC, Hartly DW, Jacques NA (2004) Proteome analysis of *Streptococcus mutans* metabolic phenotype during acid tolerance. *Microbiology* 150: 1353–1366.
76. Ellen RP, Lepine G, Nghiem PM (1997) In vitro models that support adhesion specificity in biofilms of oral bacteria. *Adv Dent Res* 11: 33–42.
77. Jenkinson HF, Demuth DR (1997) Structure, function and immunogenicity of streptococcal antigen I/II polypeptides. *Mol Microbiol* 23: 183–190.
78. Jenkinson HF, Terry SD, McNab R, Tannock GW (1993) Inactivation of the gene encoding surface protein SspA in *Streptococcus gordonii* DL1 affects cell interactions with human salivary agglutinin and oral actinomyces. *Infect Immun* 61: 3199–3208.
79. Demuth DR, Duan Y, Brooks W, Holmes AR, McNab R, et al. (1996) Tandem genes encode cell-surface polypeptides SspA and SspB which mediate adhesion of the oral bacterium *Streptococcus gordonii* to human and bacterial receptors. *Mol Microbiol* 20: 403–413.
80. Eglund PG, Du LD, Kolenbrander PE (2001) Identification of independent *Streptococcus gordonii* SspA and SspB functions in coaggregation with *Actinomyces naeslundii*. *Infect Immun* 69: 7512–7516.
81. Ajdic D, McShan WM, McLaughlin RE, Savić G, Chang J, et al. (2002) Genome sequence of *Streptococcus mutans* UAI59, a cariogenic dental pathogen. *Proc Natl Acad Sci U S A* 99: 14434–14439.
82. Yamashita Y, Shibata Y, Nakano Y, Tsuda H, Kido N, et al. (1999) A novel gene required for rhamnose-glucose polysaccharide synthesis in *Streptococcus mutans*. *J Bacteriol* 181: 6556–6559.
83. Chia JS, Lin YL, Lien HT, Chen JY (2004) Platelet aggregation induced by serotype polysaccharides from *Streptococcus mutans*. *Infect Immun* 72: 2605–2617.
84. Tsuda H, Yamashita Y, Toyoshima K, Yamaguchi N, Oho T, et al. (2000) Role of serotype-specific polysaccharide in the resistance of *Streptococcus mutans* to phagocytosis by human polymorphonuclear leukocytes. *Infect Immun* 68: 644–650.
85. Shibata Y, Ozaki K, Seki M, Kawato T, Tanaka H, et al. (2003) Analysis of loci required for determination of serotype antigenicity in *Streptococcus mutans* and its clinical utilization. *J Clin Microbiol* 41: 4107–4112.
86. Hammerschmidt S (2006) Adherence molecules of pathogenic pneumococci. *Curr Opin Microbiol* 9: 12–20.
87. Nomura R, Nakano K, Taniguchi N, Lapirattanakul J, Nemoto H, et al. (2009) Molecular and clinical analyses of the gene encoding the collagen-binding adhesin of *Streptococcus mutans*. *J Med Microbiol* 58: 469–475.
88. Xu P, Alves JM, Kitten T, Brown A, Chen Z, et al. (2007) Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* 189: 3166–3175.
89. Clark WB, Wheeler TT, Cisar JO (1984) Specific inhibition of adsorption of *Actinomyces viscosus* T14V to saliva-treated hydroxyapatite by antibody against type 1 fimbriae. *Infect Immun* 43: 497–501.
90. Gibbons RJ, Hay DI (1988) Human salivary acidic proline-rich proteins and statherin promote the attachment of *Actinomyces viscosus* LY7 to apatitic surfaces. *Infect Immun* 56: 439–445.
91. Hofflack L, Yeung MK (2001) *Actinomyces naeslundii* fimbrial protein Orf977 shows similarity to a streptococcal adhesin. *Oral Microbiol Immunol* 16: 319–320.
92. Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, et al. (2006) A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* 103: 2857–2862.
93. McNab R, Jenkinson HF (1998) Altered adherence properties of a *Streptococcus gordonii* hppA (oligopeptide permease) mutant result from transcriptional effects on *csaA* adhesin gene expression. *Microbiology* 144(Pt 1): 127–136.
94. Stephenson AE, Wu H, Novak J, Tomana M, Mintz K, et al. (2002) The Fap1 fimbrial adhesin is a glycoprotein: antibodies specific for the glycan moiety block the adhesion of *Streptococcus parasanguis* in an in vitro tooth model. *Mol Microbiol* 43: 147–157.
95. Drobní M, Hallberg K, Ohman U, Birve A, Persson K, et al. (2006) Sequence analyses of fimbriae subunit FimA proteins on *Actinomyces naeslundii* genospecies 1 and 2 and *Actinomyces odontolyticus* with variant carbohydrate binding specificities. *BMC Microbiol* 6: 43.
96. Rosengren L, Winblad B (1976) Proteolytic activity of *Streptococcus mutans* (GS-5). *Oral Surg Oral Med Oral Pathol* 42: 801–809.
97. Cowman RA, Schaefer SJ, Fitzgerald RJ, Rosner D, Shklair IL, et al. (1979) Differential utilization of proteins in saliva from caries-active and caries-free subjects as growth substrates by plaque-forming streptococci. *J Dent Res* 58: 2019–2027.
98. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8: 186–194.
99. Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* 132: 115–130.
100. Chain PS, Denef VJ, Konstantinidis KT, Vergez LM, Agullo L, et al. (2006) *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc Natl Acad Sci U S A* 103: 15280–15287.