
# A Genomic View of the Human-*Bacteroides thetaiotaomicron* Symbiosis

Jian Xu *et al.*

*Science* **299**, 2074 (2003);
DOI: 10.1126/science.1080029

may enable phase variation via a slippage type mechanism (table S7).

An unprecedented amount of the *E. faecalis* V583 genome consists of intact or partial mobile elements. Many of these regions have complex mosaic structures comprised of different elements, suggesting they are "hotspots" or "graveyards" for mobile element insertion. This apparent propensity for the incorporation of mobile elements probably contributed to the rapid acquisition and dissemination of drug resistance in the enterococci and suggests that they act as a reservoir for the further dissemination of drug resistance traits such as vancomycin resistance via mobile elements and/or conjugative plasmids. The complete genome sequence of *E. faecalis* V583 has enabled the identification of numerous predicted virulence factors and surface-exposed proteins that may facilitate the development of therapeutic approaches to combat this important nosocomial pathogen.

**References and Notes**


1. I. Klare, G. Werner, W. Witte, *Contrib. Microbiol.* **8**, 108 (2001).
2. S. A. Tailor, E. M. Bailey, M. J. Rybak, *Ann. Pharmacother.* **27**, 1231 (1993).
3. M. M. Huycke, D. F. Sahm, M. S. Gilmore, *Emerg. Infect. Dis.* **4**, 239 (1998).
4. CDC, *Morb. Mortal. Wkly. Rep.* **51**, 902 (2002).
5. Materials and methods are available as supporting material on *Science* Online.
6. D. F. Sahm *et al.*, *Antimicrob. Agents Chemother.* **33**, 1588 (1989).
7. GenBank accession numbers of the chromosome and plasmids are as follows: AE016830 (chromosome), AE016833 (pTEF1), AE016831 (pTEF2), AE016832 (pTEF3).
8. F. Garnier, S. Taourit, P. Glaser, P. Courvalin, M. Galimand, *Microbiology* **146 (no. 6)**, 1481 (2000).
9. M. Arthur *et al.*, *Antimicrob. Agents Chemother.* **36**, 867 (1992).
10. R. Novak, B. Henriques, E. Charpentier, S. Normark, E. Tuomanen, *Nature* **399**, 590 (1999).
11. G. T. Robertson *et al.*, *J. Bacteriol.* **184**, 6987 (2002).
12. N. Shankar, A. S. Baghdayan, M. S. Gilmore, *Nature* **417**, 746 (2002).
13. B. A. Bensing, I. R. Siboo, P. M. Sullam, *Infect. Immun.* **69**, 6186 (2001).
14. G. M. Dunny, B. A. Leonard, *Annu. Rev. Microbiol.* **51**, 527 (1997).
15. M. V. Francia *et al.*, *Plasmid* **46**, 117 (2001).
16. D. B. Clewell, F. Y. An, S. E. Flannagan, M. Antiporta, G. M. Dunny, *Mol. Microbiol.* **35**, 246 (2000).
17. S. D. Sussmuth *et al.*, *Infect. Immun.* **68**, 4900 (2000).
18. M. M. Huycke, V. Abrams, D. R. Moore, *Carcinogenesis* **23**, 529 (2002).
19. C. L. Wells, R. P. Jechorek, S. L. Erlandsen, *J. Infect. Dis.* **162**, 82 (1990).
20. L. T. Pontius, D. B. Clewell, *Plasmid* **26**, 172 (1991).


21. This work was supported by the NIH, National Institute of Allergy and Infectious Disease Grant AI40963-02. We thank M. Heaney, S. Lo, M. Holmes, B. Lee, R. Karamchedu, and V. Sapiro for database and information technology support at TIGR, and we thank the TIGR faculty and sequencing core for expert advice and assistance. We thank M. Gilmore for generously providing *E. faecalis* V583. We thank D. Clewell for providing the complete sequence of pAD1 for comparative analysis.


**Supporting Online Material**

www.sciencemag.org/cgi/content/full/299/5615/2071/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 to S7

19 November 2002; accepted 12 February 2003

# A Genomic View of the Human–*Bacteroides thetaiotaomicron* Symbiosis

**Jian Xu, Magnus K. Bjursell, Jason Himrod, Su Deng, Lynn K. Carmichael, Herbert C. Chiang, Lora V. Hooper, Jeffrey I. Gordon***

**The human gut is colonized with a vast community of indigenous microorganisms that help shape our biology. Here, we present the complete genome sequence of the Gram-negative anaerobe *Bacteroides thetaiotaomicron*, a dominant member of our normal distal intestinal microbiota. Its 4779-member proteome includes an elaborate apparatus for acquiring and hydrolyzing otherwise indigestible dietary polysaccharides and an associated environment-sensing system consisting of a large repertoire of extracytoplasmic function sigma factors and one- and two-component signal transduction systems. These and other expanded paralogous groups shed light on the molecular mechanisms underlying symbiotic host-bacterial relationships in our intestine.**

A major theme of life on our planet is the complex and beneficial interactions that occur between eukaryotes and prokaryotes. Humans are no exception. As adults, we harbor diverse communities of microorganisms whose total number exceeds the sum of all of our somatic and germ cells (*1*). As yet, the ways in which these communities contribute to normal postnatal development and adult physiology are largely unexplored. The human gut contains the largest such collection of microbes [$10^{11}$ organisms per ml proximal colonic contents (*1*)]. An estimated 2 to 4 million genes are embedded in the aggregate genome (microbiome) of an intestinal community of ~500 to 1000 bacterial species (*2*). The products of these genes provide metabolic capacities not encoded in our own genome (*3*).

The gut microbiota is a key regulator of the human immune system; it acts to induce tolerance to microbial epitopes and thus to reduce responses to commonly encountered foodstuffs and other environmental antigens (*4*). Functional genomic studies of germfree mice colonized with components of the human intestinal microbiota are revealing other functions affected by indigenous bacteria, including fortification of the mucosal barrier and angiogenesis (*5–7*). These observations emphasize the need to understand more about the roles played by the microbiota in host biology, as well as the potential for control and modulation.

Here, we describe the complete 6.26-Mb genome sequence of the Gram-negative anaerobe, *Bacteroides thetaiotaomicron* (figs. S1 to S5 in supporting online material). This genetically manipulatable organism is a predominant member of the normal human (and murine) distal small intestinal and colonic microbiota (*8*) and has been used as a model for understanding the impact of constituents of the microbiota on gut gene expression (*5, 9*). The genome sequences of members of the Bacteroidetes phylum, which diverged early in the evolution of Bacteria (*10*), have not yet been reported.

The *B. thetaiotaomicron* type strain, VPI-5482 (ATCC 29148), was originally isolated from the feces of a healthy adult human. Of the 4779 predicted proteins in its proteome, 2782 (58%) were assigned putative functions on the basis of homology to other known proteins. Of the predicted proteins, 848 (18%) have homology to proteins with no known function, whereas 1149 (24%) have no appreciable homology to entries in public databases. The most markedly expanded paralogous groups are involved in polysaccharide uptake and degradation (glycosylhydrolases, cell-surface carbohydrate-binding proteins); capsular polysaccharide biosynthesis (e.g., glycosyltransferases); environmental sensing and signal transduction [one- and two-component systems; extracytoplasmic function (ECF)-type sigma factors]; and DNA mobilization (transposases, conjugative transposons) (table S1). These expansions reveal strategies used by *B. thetaiotaomicron* to

Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO 63110, USA.

*To whom correspondence should be addressed. E-mail: jgordon@molecool.wustl.edu

survive and to dominate in the densely populated intestinal ecosystem.

*Bacteroides* spp. are known to break down a wide variety of otherwise indigestible dietary plant polysaccharides (e.g., amylose, amylopectin, and pullulan) (*3, 10*). The representation of predicted glycosylhydrolases (α-galactosidases, β-galactosidases, α-glucosidases, β-glucosidases, β-glucuronidases, β-fructofuranosidases, α-mannosidases, amylases, and endo-1,2-β-xylanases, plus 14 other activities) in the *B. thetaiotaomicron* proteome exceeds that in any other sequenced Bacteria, including other human gut commensals and symbionts [*Clostridium perfringens*, *Bifidobacterium longum*, and *Escherichia coli* (table S1)]. *B. thetaiotaomicron* has also evolved the capacity to use a variety of host-derived glycans, including chondroitin sulfate, mucin, hyaluronate, and heparin (*3*) (table S2). Sixty-one percent of its glycosylhydrolases are predicted to be in the periplasm or outer membrane or extracellular. This suggests that these enzymes are not only important for fulfilling the needs of *B. thetaiotaomicron* but may also help shape the metabolic milieu of the intestinal ecosystem in ways conducive to maintaining a microbiota that supplies us with 10 to 15% of our daily calories as fermentation products of dietary polysaccharides (*11*).

Seven capsular polysaccharide synthesis (CPS) loci were identified. Each locus contains one or two genes encoding conserved regulatory proteins (UpcY and UpcZ homologs) positioned upstream of open reading frames (ORFs) specifying carbohydrate biosynthetic enzymes, including a variety of putative glycosyltransferases (table S3). Regulation of the eight known CPS loci in *Bacteroides fragilis* occurs through promoter inversion (*12*). This mechanism may allow the organism to evade detection by the host immune system, but the machinery controlling inversion remains to be defined (*12*). Interestingly, the presence of a flipped promoter in two of the seven *B. thetaiotaomicron* CPS loci correlates with the presence of an integrase gene immediately upstream of UpcY/UpcZ (table S3). These integrases have weak homology to the phage integrase family (*13*) and appear to be highly specific to Bacteroides.

The genome encodes many outer membrane proteins (OMPs) that are likely to be involved in acquisition of oligo- and polysaccharides. The largest paralogous group in the genome contains 106 members with homology to the OMP SusC. Another 57-member group of paralogs has homology to SusD. SusC and SusD belong to a previously characterized eight-component *B. thetaiotaomicron* starch utilization system (Sus) (*14–16*) and mediate binding of starches to the bacterial cell surface so that they can be subsequently broken down by outer membrane and periplasmic α-amylases (*16*). In 56 cases, the SusC and SusD homologs are paired together as members of a multigene cluster. Twenty of these clusters consist of the SusC/SusD pair with an upstream gene encoding an ECF-type sigma factor. Twelve of the 20 clusters also contain downstream ORFs encoding glycosylhydrolases together with enzymes involved in sugar metabolism (see fig. S1 for the distribution of these 12 clusters in the genome, table S4 for a list of genes in all 12 clusters, and table S5 for genes immediately downstream of all 106 SusC homologs).

The presence of an ECF-type sigma factor in these clusters suggests that they are regulated in response to environmental cues. Bacterial sigma factor components of RNA polymerase complexes play key roles in coordinating transcriptional responses to various physiological stimuli (*17*). *B. thetaiotaomicron* has a remarkably expanded population of ECF-type sigma factors (50) (table S1). These genes are typically cotranscribed with one or more negative regulators, often a transmembrane protein that binds to and inhibits the cognate sigma factor. When a stimulus is received from the environment, the ECF-type sigma factor is released so that it can bind to RNA polymerase to stimulate transcription (*18*). Sixteen of 20 SusC- and SusD-containing clusters with an ECF-type sigma factor ORF have a gene encoding a predicted transmembrane protein interposed between the sigma factor and SusC (table S4; see table S6 for a listing of all ECF-type sigma factors and their immediate downstream genes). Regulation of nutrient processing by ECF-type sigma factors has not been reported for this or other Bacteria. However, given the environmental sensing functions of these factors, their deployment by *B. thetaiotaomicron* to regulate expression of its elaborate polysaccharide utilization apparatus is one feature that may confer an advantage over less well endowed members of the microbiota.

Another manifestation of this symbiont's highly evolved capacity to sense and respond to environmental cues is the rich representation of one- and two-component signal transduction systems (table S1). A one-component system consists of a single protein that combines all the features of a two-component system necessary for coupling receipt of an environmental stimulus to regulation of gene expression. Twenty-two of the 32 one-component systems are adjacent to nutrient utilization genes (19 with oligo-polysaccharide hydrolases; three with sulfatases).

*B. thetaiotaomicron* has several types of mobile genetic elements: a 33-kb plasmid (fig. S1), 63 transposases (table S1), plus four homologs of the self-transmitting conjugative transposon CTnDOT (table S7). CTnDOT mediates the spread of tetracycline and erythromycin resistance among *Bacteroides* spp., and between *B. thetaiotaomicron* and other members of the normal gut microbiota (*19,20*). Although the VPI-5482 type strain does not harbor antibiotic resistance genes in its four conjugative transposons (CTns), the presence of these CTns, together with the broad host range of CTnDOT (*20*), suggests that they may contribute to horizontal transfer of DNA between *B. thetaiotaomicron* and other bacterial constituents of the distal gut, thereby promoting their microevolution.

Bacteroides is among the dominant groups of bacteria that coexist with adult humans. The genomewide view of *B. thetaiotaomicron* illustrates how symbiotic relationships between humans and bacteria can be forged on the basis of metabolic capabilities that allow an otherwise poorly accessible source of nutrients to be utilized. The microbe's ability to survive and prosper in our intestinal ecosystem appears to reflect highly evolved strategies for (i) sensing its luminal environment, (ii) acquiring dietary polysaccharides, and (iii) manipulating host gene expression in ways that establish and maintain a mutually advantageous partnership.

A large portion of the *B. thetaiotaomicron* proteome is dedicated to harvesting dietary polysaccharides and metabolizing their liberated sugars [e.g., 172 glycosylhydrolases, 163 homologs of SusC and SusD outer-membrane polysaccharide-binding proteins; 20 sugar-specific transporters plus 21 permease subunits of ATP-binding cassette (ABC) transporters]. The frequent colocalization of genes encoding polysaccharide utilization enzymes with genes specifying ECF-type sigma factors and one- and two-component systems provides a regulatory mechanism that presumably enables *B. thetaiotaomicron* to coordinate gene expression with nutrient availability.

Previous studies in germfree mice revealed that *B. thetaiotaomicron* stimulates angiogenesis during postnatal intestine development (*7*), thereby increasing the host's capacity for absorbing nutrients. *B. thetaiotaomicron* also regulates synthesis of various gut epithelial glycans, including those with terminal α-linked fucose (*9*), that can be harvested by its α-fucosidases (table S1). Control of epithelial fucosylated glycan production occurs through a bacterial regulatory system that senses fucose availability in the gut lumen and induces expression of host α1,2-fucosyltransferases and fucosylated glycans only when this pentose sugar is scarce (*9*). Regulation of epithelial glycan synthesis represents one strategy that *B. thetaiotaomicron* can deploy to create a habitable niche for itself that other organisms might exploit (*2*). An intriguing question is whether *B. thetaiotaomicron* is able to link carbohydrate availability in its niche with the types of capsular polysaccharide structures it adopts and the types of host epithelial glycans it helps create,

so as serve its own nutrient needs while at the same time camouflaging itself to avoid eliciting an adaptive host immune response.

The completed sequence of *B. thetaiotaomicron* should permit characterization of the bacterial messengers that influence host processes. Comparative genomic analysis with other major gut symbionts should help clarify their relative roles and contributions to the gut community and to the whole symbiosis. The results could reveal previously unknown entities important in human health and disease.


**References and Notes**

1. D. C. Savage, *Annu. Rev. Microbiol.* **31**, 107 (1977).
2. L. V. Hooper, J. I. Gordon, *Science* **292**, 1115 (2001).
3. L. V. Hooper, T. Midtvedt, J. I. Gordon, *Annu. Rev. Nutr.* **22**, 282 (2002).
4. C. Braun-Fahrlander *et al.*, *N. Engl. J. Med.* **347**, 869 (2002).
5. L. V. Hooper *et al.*, *Science* **291**, 881 (2001).
6. L. V. Hooper, T. S. Stappenbeck, C. V. Hong, J. I. Gordon, *Nature Immunol.* **4**, 269 (2003).
7. T. S. Stappenbeck, L. V. Hooper, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15451 (2002).
8. W. E. Moore, L. V. Holdeman, *Appl. Microbiol.* **27**, 961 (1974).
9. L. V. Hooper, J. Xu, P. G. Falk, T. Midtvedt, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9833 (1999).
10. A. A. Salyers, P. Valentine, V. Hwa, in *Genetics and Molecular Biology of Anaerobic Bacteria*, M. Sebald, Ed. (Springer-Verlag, New York, 1993), pp. 505–516.
11. E. N. Bergman, *Physiol. Rev.* **70**, 567 (1990).
12. C. M. Krinos *et al.*, *Nature* **414**, 555 (2001).
13. A. Bateman *et al.*, *Nucleic Acids Res.* **30**, 276 (2002).
14. J. N. D'Elia, A. A. Salyers, *J. Bacteriol.* **178**, 7180 (1997).
15. A. R. Reeves, G. R. Wang, A. A. Salyers, *J. Bacteriol.* **179**, 643 (1997).
16. K. H. Cho, A. A. Salyers, *J. Bacteriol.* **183**, 7224 (2001).
17. M. M. Wosten, *FEMS Microbiol. Rev.* **22**, 127 (1998).
18. J. D. Helmann, *Adv. Microb. Physiol.* **46**, 47 (2002).
19. G. Whittle, N. B. Shoemaker, A. A. Salyers, *J. Bacteriol* **184**, 3839 (2002).
20. N. B. Shoemaker, H. Vlamakis, K. Hayes, A. A. Salyers, *Appl. Environ. Microbiol.* **67**, 561 (2001).


21. We thank C. Baublite, C. Foushee, S. LaBrie, and D. Smoller, for assistance with BAC library construction and fingerprinting; A. Salyers and N. Shoemaker for providing p5482; P. Minx, H. Du, S.-p. Yang, J. Spieth, and J. Buhler for helpful discussions; C. Rexer for assistance with PCR; and P. Falk for continued support. This work was funded by AstraZeneca and the NIH (DK30292).


**Supporting Online Material**

www.sciencemag.org/cgi/content/full/299/5615/2074/DC1
Supporting Online Text
Figs. S1 to S5
Tables S1 to S8

1 November 2002; accepted 13 January 2003

# Pyogenic Bacterial Infections in Humans with IRAK-4 Deficiency


**Capucine Picard,[1] Anne Puel,[1] Marion Bonnet,[1] Cheng-Lung Ku,[1] Jacinta Bustamante,[1] Kun Yang,[1] Claire Soudais,[1] Stéphanie Dupuis,[1] Jacqueline Feinberg,[1] Claire Fieschi,[1] Carole Elbim,[2] Remi Hitchcock,[3] David Lammas,[4] Graham Davies,[5] Abdulaziz Al-Ghonaium,[6] Hassan Al-Rayes,[6] Sulaiman Al-Jumaah,[6] Sami Al-Hajjar,[6] Ibrahim Zaid Al-Mohsen,[6] Husn H. Frayha,[6] Rajivi Rucker,[3] Thomas R. Hawn,[7] Alan Aderem,[7] Haysam Tufenkeji,[6] Soichi Haraguchi,[3] Noorbibi K. Day,[3] Robert A. Good,[3] Marie-Anne Gougerot-Pocidalo,[2] Adrian Ozinsky,[7] Jean-Laurent Casanova[1,8]***



**Members of the Toll-like receptor (TLR) and interleukin-1 receptor (IL-1R) superfamily share an intracytoplasmic Toll–IL-1 receptor (TIR) domain, which mediates recruitment of the interleukin-1 receptor–associated kinase (IRAK) complex via TIR-containing adapter molecules. We describe three unrelated children with inherited IRAK-4 deficiency. Their blood and fibroblast cells did not activate nuclear factor κB and mitogen-activated protein kinase (MAPK) and failed to induce downstream cytokines in response to any of the known ligands of TIR-bearing receptors. The otherwise healthy children developed infections caused by pyogenic bacteria. These findings suggest that, in humans, the TIR-IRAK signaling pathway is crucial for protective immunity against specific bacteria but is redundant against most other microorganisms.**


The members of the mammalian Toll-like/interleukin-1 receptor superfamily characteristically have a TIR domain (*1*). This superfamily contains two classes of membrane receptors: TLRs, seven of which recognize known ligands derived from microorganisms (*2*, *3*) and interleukin-1 receptor and related receptors (IL-1Rs), two of which recognize known host cytokines, IL-1 (*4*) and IL-18 (*5*). Upon ligand binding, the receptor complex recruits, via its intracytoplasmic TIR domain, the TIR-containing cytosolic adapter proteins MyD88 (*6*) and TIRAP/Mal (*7–9*). These adapters in turn recruit the IRAK complex. Four IRAK molecules have been identified: IRAK-1 (*10*), IRAK-2 (*6*), IRAK-M (*11*, *12*), and IRAK-4 (*13–15*). IRAK-1 and IRAK-4 are active kinases, dissociating from the receptor-adapter complex upon phosphorylation and activating tumor necrosis factor receptor-associated factor-6 (TRAF-6) (*16*). TRAF-6 then activates at least two pathways, leading to the activation of NF-κB and MAPK (*16*). In the mouse, the TIR-IRAK signaling pathway plays an extensive role in immunity to infections (*1–5*).

We investigated three unrelated children (P1, P2, and P3) with recurrent infections and poor inflammatory response (for case reports, see supporting online text). Extracellular, pyogenic bacteria were the only microorganisms responsible for infection. Gram-positive *Streptococcus pneumoniae* and *Staphylococcus aureus* were the most frequently found and were the only pathogens identified in two patients. The infections began early in life but became less frequent with age, and the patients (now aged 6, 11, and 7 years) are well with no treatment. All known primary immunodeficiencies were excluded. In particular, the patients had normal serum antibody titers against protein and polysaccharide antigens, including those from *S. pneumoniae*. However, one of our three patients (P3) had previously been shown not to respond to lipopolysaccharide (LPS) and *Staphylococcus aureus* (*17*). The phenotype of the three patients was similar to that of another child described elsewhere, with impaired responses to lipopolysaccharide (LPS) and IL-1β, but not to tumor necrosis factor–α (TNFα) (*18*). This suggested that our three patients might be suffering from impaired TIR pathway signaling.

We first tested the response of the patients' monocytes to LPS, which is predominantly detected via TLR4 (*19*, *20*). As P3 did, neither P1


[1]Laboratoire de Génétique Humaine des Maladies Infectieuses, Université René Descartes-INSERM U550, Faculté Necker, 156 rue de Vaugirard, 75015 Paris, France. [2]Service d'Immunologie Cellulaire, Laboratoire d'Hématologie et d'Immunologie, INSERM U294, Faculté Xavier Bichat, 75018 Paris, France. [3]Department of Pediatrics, Division of Allergy and Immunology, University of South Florida and All Children's Hospital, St. Petersburg, FL 33701, USA. [4]MRC Center for Immune Regulation, The Medical School, University of Birmingham, Birmingham B15 2TT, UK. [5]Infectious Diseases Unit, Great Ormond Street Hospital for Children, London WC1N 3JH, UK. [6]Department of Pediatrics, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Kingdom of Saudi Arabia. [7]Institute for Systems Biology, 1441 North 34th Street, Seattle, WA 98103, USA. [8]Unité d'Immunologie et d'Hématologie Pédiatriques, Hôpital Necker-Enfants Malades, 149 rue de Sèvres, 75015 Paris, France.

*To whom correspondence should be addressed. E-mail: casanova@necker.fr