Supporting Information

Coyne *et al.* 10.1073/pnas.0804220105

SI Materials and Methods

Assessment of *in Vitro* Growth Under Various Conditions. The bile salt mixture was an ~ 50:50 mix of sodium cholate and sodium deoxycholate. The conjugated bile salts were in the form of porcine bile extract. The porcine bile extract, hexadecyltrimethylammonium bromide, benzalkonium chloride, and deoxycholic acid were purchased from Sigma-Aldrich Co. Basal medium was prepared, autoclaved, and supplemented as usual, but at 2× concentration. Appropriate amounts of the additives to be tested were dissolved in this medium, and the final volume was adjusted with sterile water. Serial 2-fold dilutions were made with 1× supplemented basal medium to produce a range of concentrations for testing, and the aliquots were filter sterilized. Strains to be tested were grown overnight in basal medium, and these cultures were used to inoculate the anaerobic test medium. Primary growth in the test medium was assessed after incubation anaerobically overnight at 37°C.

Vigorously resuspended overnight cultures of the strains to be compared for growth rate were used to inoculate anaerobic basal medium in 16 × 100-mm screw-capped tubes. Growth was monitored by recording absorbance at 600 nm every hour using a Genesys 20 spectrophotometer (Thermo Electron Corp.). The readings for several replicate cultures were averaged and plotted to arrive at an accurate assessment of growth rate.

Sensitivity to acid pH was assessed by inoculating pH-adjusted basal medium or pH-adjusted normal saline with growth from overnight basal-medium cultures. Samples were removed from the normal saline cultures after exposure to acid conditions for 2, 4, 6, and 24 h and were used to inoculate the anaerobic basal medium. Growth of these postexposure cultures was assessed after overnight anaerobic incubation.

Gene Expression Analysis. Triplicate cultures of *B. fragilis* 9343 wild type and Δ*ungD1*Δ*ungD2*ΔPSH were grown overnight, diluted 1:100 into fresh supplemented basal medium, and allowed to reach mid-log phase (OD_{600} ~ 0.8). Total RNA was extracted from each culture using the RNeasy Mini Kit and RNAprotect reagent (Qiagen) according to the manufacturer's instructions. The RNA samples were subjected to on-column DNase digestion using the RNase-free DNase kit (Qiagen) as suggested. RNA quality and concentration was determined using an Agilent 2100 bioanalyzer at the Harvard Medical School-Partners Healthcare Center for Genetics and Genomics (Cambridge, MA).

For gene expression analysis, high-density oligonucleotide microarrays (14 24-mer primer pairs per gene, 3-fold technical redundancy per chip, using design 2005–08-12_B_fragili_24mer) were purchased from NimbleGen Systems. Array synthesis, Cy3 labeling, hybridization, and array scanning were performed by NimbleGen (Roche NimbleGen). Data extraction, background correction, quantile normalization (1), and RMA analysis (2) were performed using the NimbleScan software package, version 2.4.27.

The normalized, 3-per-gene replicate expression levels for each chip were averaged to reduce the data set to 1 expression level per gene per chip. These averaged data were log 2 transformed and analyzed by 1000 permutations of rank products (3), using a custom Perl script based on the Perl implementation of the rank products procedure as provided by the Bioinformatics Research Centre at the University of Glasgow. A fold-change score also was assigned to each gene by computing the ratio of the average of the 3 mutant expression values to the average of the 3 wild-type expression values. A gene was considered differentially expressed if its false discovery rate as calculated by rank products was ≤ 0.05 and if its fold-change ratio satisfied $0.5 \leq x \geq 2.0$.

1. Bolstad B, Irizarry R, Astrand M, Speed T. (2003) A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinformatics* 19(2):185–193.
2. Irizarry RA, *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249–264.
3. Breitling R, Armengaud P, Amtmann A, Herzyk P. (2004) Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 573:83–92.
4. Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. *Science* 278:631–637.

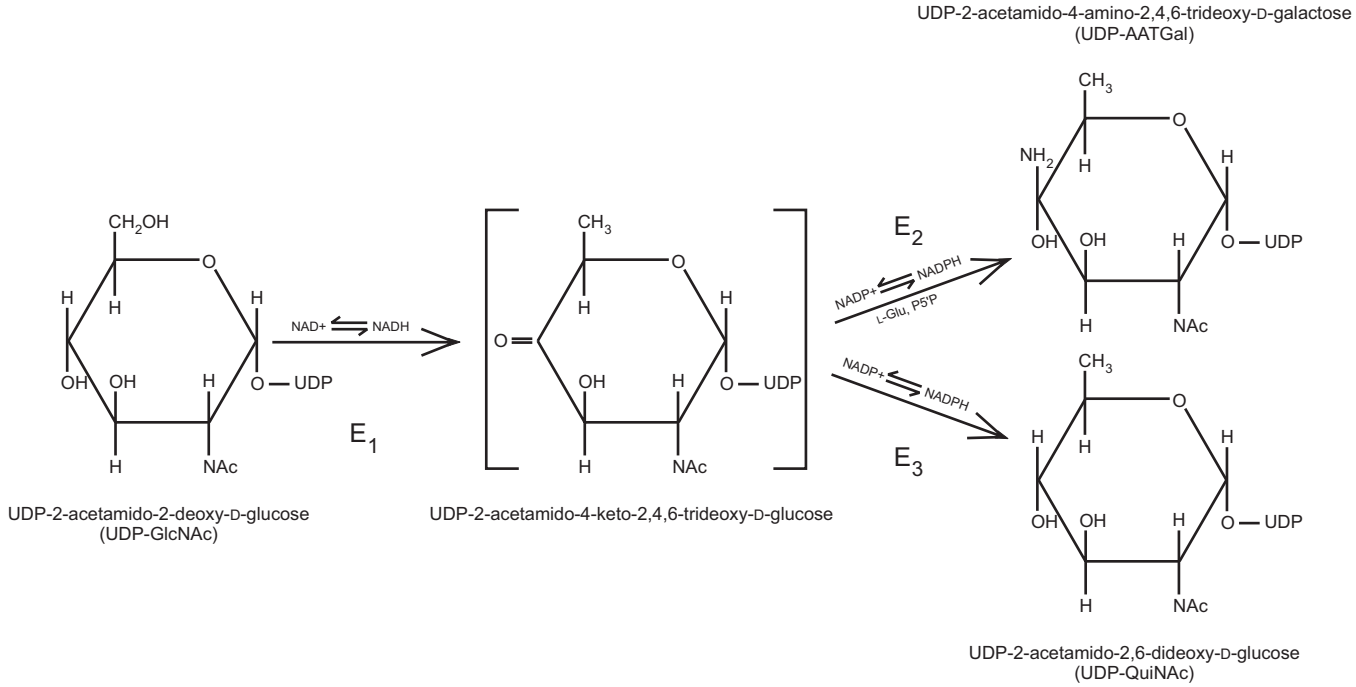

Fig. S1. Proposed involvement of UngD2 (E1, UDP-2-acetamido-2-deoxy-D-glucose 4,6-dehydratase) in the formation of a UDP-2-acetamido-4-keto-2,4,6-trideoxy-D-glucose intermediate, with subsequent formation of UDP-2-acetamido-4-amino-2,4,6-trideoxy- D-galactose or UDP-2-acetamido-2,6-dideoxy- D-glucose by the actions of a aminotransferase (E2) or a 4-reductase (E3), respectively. In the pyroxidal-5′-phosphate (P5′P)-dependant aminotransferase reaction, L-glutamate (L-Glu) is converted by oxidative deamination to α-ketoglutarate.

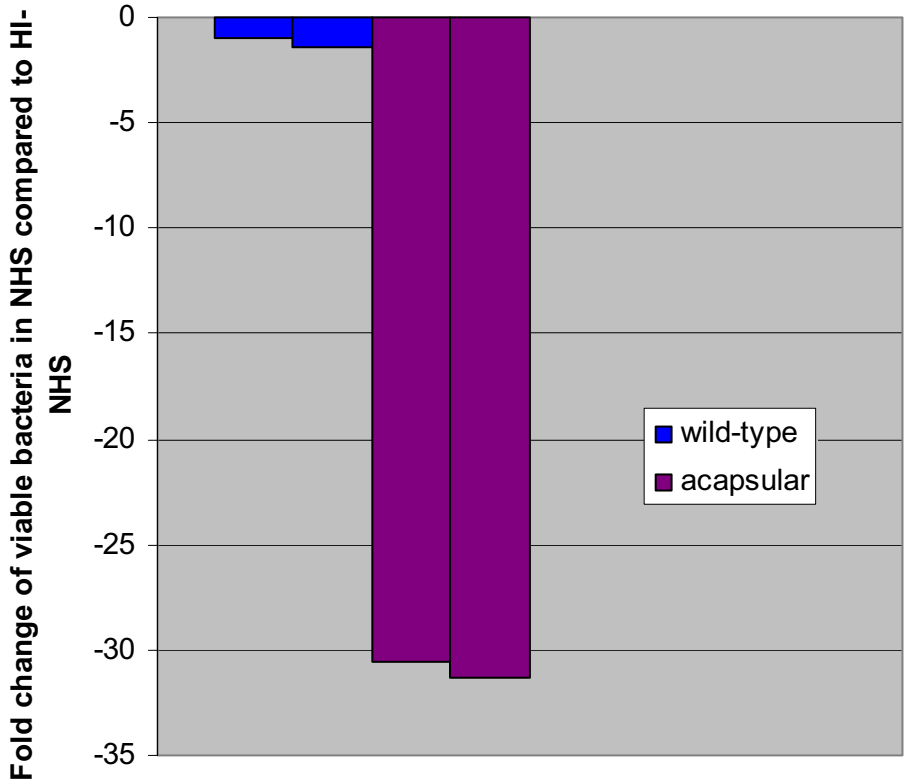

Fig. S2. Complement-mediated killing of wild-type *B. fragilis* and the Δ*ungD1*Δ*ungD2*ΔPSH acapsular mutant. The viability of bacteria incubated for 1 h in 10% normal human serum (NHS) was compared with bacteria incubated for 1 h in heat-inactivated normal human serum (HI-NHS), and the results are reported as fold change of viable bacteria. The results of 2 separate experiments are shown for each strain.

Table S1. Distribution among cluster of orthologous group (COG) categories of *B. fragilis* Δ*ungD1*Δ*ungD2*ΔPSH genes with altered expression levels relative to *B. fragilis* 9343/

COG	Function	9343 genome (4231 genes)	Differentially expressed* Up (62 genes)	Differentially expressed* Down (29 genes)
J	Translation	171	1	0
A	RNA processing and modification	0	0	0
K	Transcription	239	2	0
L	Replication, recombination and repair	204	2	0
B	Chromatin structure and dynamics	1	0	0
D	Cell cycle control, mitosis and meiosis	28	0	0
Y	Nuclear structure	0	0	0
V	Defense mechanisms	116	0	0
T	Signal transduction mechanisms	197	0	0
M	Cell wall/membrane biogenesis	338	6	7
N	Cell motility	2	0	0
Z	Cytoskeleton	0	0	0
W	Extracellular structures	0	0	0
U	Intracellular trafficking and secretion	72	3	0
O	Posttranslational modification, protein turnover, chaperones	104	1	0
C	Energy production and conversion	178	3	0
G	Carbohydrate transport and metabolism	262	1	2
E	Amino acid transport and metabolism	243	5	0
F	Nucleotide transport and metabolism	74	0	0
H	Coenzyme transport and metabolism	139	1	0
I	Lipid transport and metabolism	74	1	0
P	Inorganic ion transport and metabolism	281	0	0
Q	Secondary metabolites biosynthesis, transport and catabolism	35	1	0
R	General function prediction only	426	6	1
S	Function unknown	195	2	0
-	Not in COGs	1463	32	21

*Genes whose expression level in Δ*ungD1*Δ*ungD2*ΔPSH was greater (up column) or less than (down column) that in wild type with a false discovery rate ≤ 0.05 and an average fold-change of ≤ 0.5 or ≥ 2.0.

Table S2. List of genes with differential expression levels in Δ*ungD1*Δ*ungD2*ΔPSH compared with wild type

Gene	Characterization
Differentially expressed up	
COG category J (function: translation) (total in category: 1)	
BF1923	hypothetical protein
COG category K (function: transcription) (total in category: 2)	
BF0581	putative AraC family transcriptional regulatory protein
BF2332	hypothetical protein
COG category L (function: replication, recombination and repair) (total in category: 2)	
BF4220	putative histone-like DNA-binding protein HU3
BF4226	putative histone-like DNA-binding protein HU4
COG category M (function: cell wall/membrane biogenesis) (total in category: 6)	
BF1988	putative outer membrane protein
BF3193	putative lipoprotein
BF3451	putative LPS biosynthesis related conserved hypothetical protein
BF3452	putative LPS biosynthesis related dTDP-4-dehydrorhamnose 3,5-epimerase
BF3453	putative LPS biosynthesis related glucose-1-phosphate thymidylyltransferase
BF3944	putative TonB-family outer membrane receptor protein
COG category U (function: intracellular trafficking and secretion) (total in category: 3)	
BF3945	hypothetical protein
BF3946	hypothetical protein
BF3947	putative transmembrane MotA/TolQ/ExbB proton channel family protein
COG category O (function: posttranslational modification, protein turnover, chaperones) (total in category: 1)	
BF1279	putative thioredoxin
COG category C (function: energy production and conversion) (total in category: 3)	
BF1279	putative thioredoxin
BF2351	succinyl-CoA synthetase beta chain
BF3462	hypothetical protein
COG category G (function: carbohydrate transport and metabolism) (total in category: 1)	
BF3451	putative LPS biosynthesis related conserved hypothetical protein
COG category E (function: amino acid transport and metabolism) (total in category: 5)	
BF0393	putative glutamate decarboxylase
BF0394	glutaminase
BF2073	putative D-3-phosphoglycerate dehydrogenase
BF2665	diaminopimelate epimerase
BF2666	L,L-diaminopimelate aminotransferase
COG category H (function: coenzyme transport and metabolism) (total in category: 1)	
BF2073	putative D-3-phosphoglycerate dehydrogenase
COG category I (function: lipid transport and metabolism) (total in category: 1)	
BF3320	putative 7-alpha-hydroxysteroid dehydrogenase (bile acid catabolism)
COG category Q (function: secondary metabolites biosynthesis, transport, and catabolism) (total in category: 1)	
BF3320	putative 7-alpha-hydroxysteroid dehydrogenase (bile acid catabolism)
COG category R (general function prediction only) (total in category: 6)	
BF2334	hypothetical protein
BF3032	putative Pirin-like protein
BF3271	putative bacterioferritin-related protein
BF3320	putative 7-alpha-hydroxysteroid dehydrogenase (bile acid catabolism)
BF3463	putative LPS biosynthesis related polysaccharide transporter
BF3942	putative TPR-repeat family protein
COG category S (function unknown) (total in category: 2)	(*Table continues*)

Gene	Characterization
BF2074	hypothetical protein
BF3461	hypothetical protein
Total genes represented once or more: 35	
Ups not in COGs (total in category: 32)	
BF0322	hypothetical membrane protein
BF0468	hypothetical protein
BF0470	hypothetical protein
BF0567	hypothetical protein
BF1047	hypothetical protein
BF1158	putative alpha-glucosidase
BF1567	hypothetical protein
BF1629	hypothetical protein
BF1639	hypothetical protein
BF1733	hypothetical protein
BF2113	hypothetical protein
BF2162	hypothetical protein
BF2168	hypothetical protein
BF2257	hypothetical protein
BF2347	hypothetical protein
BF2349	hypothetical protein
BF2377	hypothetical protein
BF2441	hypothetical protein
BF2442	hypothetical protein
BF2443	hypothetical protein
BF2445	hypothetical protein
BF2496	putative DNA-binding protein
BF2582	hypothetical protein
BF2690	hypothetical protein
BF2912	hypothetical protein
BF2957	hypothetical protein
BF3394	hypothetical protein
BF3460	putative LPS biosynthesis related glycosyltransferase
BF3789	hypothetical protein
BF3930	hypothetical protein
BF4230	hypothetical protein
pBF9343.11	putative integral membrane protein
Differentially expressed down	
COG category M (function: cell wall/membrane biogenesis) (total in category: 7)	
BF1551	putative glucose-1-phosphate thymidyl transferase
BF1552	putative dTDP-4-dehydrorhamnose 3,5 epimerase
BF1554	putative nucleotide-sugar dehydrogenase
BF1555	putative dNTP-hexose dehydratase-epimerase
BF1557	putative glucosyltransferase
BF1560	putative glycosyltransferase
BF2848	putative capsular polysaccharide biosynthesis protein
COG category G (function: carbohydrate transport and metabolism) (total in category: 2)	
BF1555	putative dNTP-hexose dehydratase-epimerase
BF2848	putative capsular polysaccharide biosynthesis protein
COG category R (general function prediction only) (total in category: 1)	
BF1556	putative polysaccharide transporter/flippase
Total genes represented once or more: 10	
Downs not in COGs (total in category: 21)	
BF0230	hypothetical protein
BF0888	hypothetical protein
BF1260	hypothetical protein
BF1324	hypothetical protein
BF1355	hypothetical protein
BF1549	putative transcriptional regulator
BF1550	putative transcriptional regulator
BF1553	hypothetical protein
BF1558	putative glycosyltransferase protein
BF1559	hypothetical protein
BF2347B	hypothetical protein

Gene	Characterization
BF2995	hypothetical protein
BF2999	hypothetical protein
BF3000	hypothetical protein
BF3001	hypothetical protein
BF3002	hypothetical protein
BF3003	putative lipoprotein
BF3004	hypothetical protein
BF3427	hypothetical protein
BF3742	hypothetical protein
BF3949	putative transmembrane protein

Abbreviations: COG, cluster of orthologous groups.

Table S3. Primers used in this study

Primer sequence (5′->3′)	5′ addition	Purpose
ATTTCCGAAAGCATTGAAGTCTAA	AA GAGCTC (*Sst* I)	Delete *ungD1*, upstream flank
GCAACTCGCGATGAAAGATACTTA	AT CTGCAG (*Pst* I)	
TTTGTAAGCAAGAATTCGGAGTTT	TT CTGCAG (*Pst* I)	Delete *ungD1*, downstream flank
AATCCTGATCACCTCCTACAAGAA	AA GAGCTC (*Sst* I)	
AGAGGTTCAAAGAACAGACGATG	GT GGATCC (*Bam* HI)	Delete *ungD2*, upstream flank
AGTACCTTCGCAGAAAGATACCG	GC GAATTC (*Eco* RI)	
GTGAGCAAAAACTCTTGTTTTGAG	GA GAATTC (*Eco* RI)	Delete *ungD2*, downstream flank
AAATATCAACCACTTCCGTCAGTT	TT GGATCC (*Bam* HI)	
CGTTGAAACTAGCATCTCGTACTC	CTCA GGATCC (*Bam* HI)	Delete PSH, upstream flank
TTATGAAGTTTAAAGGGGTCAAGC	CGAA ACGCGT (*Mlu* I)	
TGCAAAGAAAAGATGTCCATTAAA	AAAT ACGCGT (*Mlu* I)	Delete PSH, downstream flank
CTTTCAGAAAAGAAAAAGCATTCC	CCTT GGATCC (*Bam* HI)	
TGAAATGTTTATAAAATAAAAAGGTTAGGG	TA GGATCC (*Bam* HI)	Clone *ungD1*
CATTTATTTCATTCCTAAGTTTTCG	GC GGATCC (*Bam* HI)	
ACCGCCTAAAACAAAAACTGTAAA	AA GGATCC (*Bam* HI)	Clone *ungD2*
ACGATCAGTCTTTCTTATCCAACG	GC GGATCC (*Bam* HI)	
GTCAAATAAAAACGCACGTGTTAG		Detect *ungD2* deletion
AACGTTTCAAGTTACACGAACAAA		
AAAAAGAAGATGACAAGCTCCCTAT		PSH promoter orientation
TGGGATCCTCCTGTCACACGCACACGCTG		