

Fig. S1

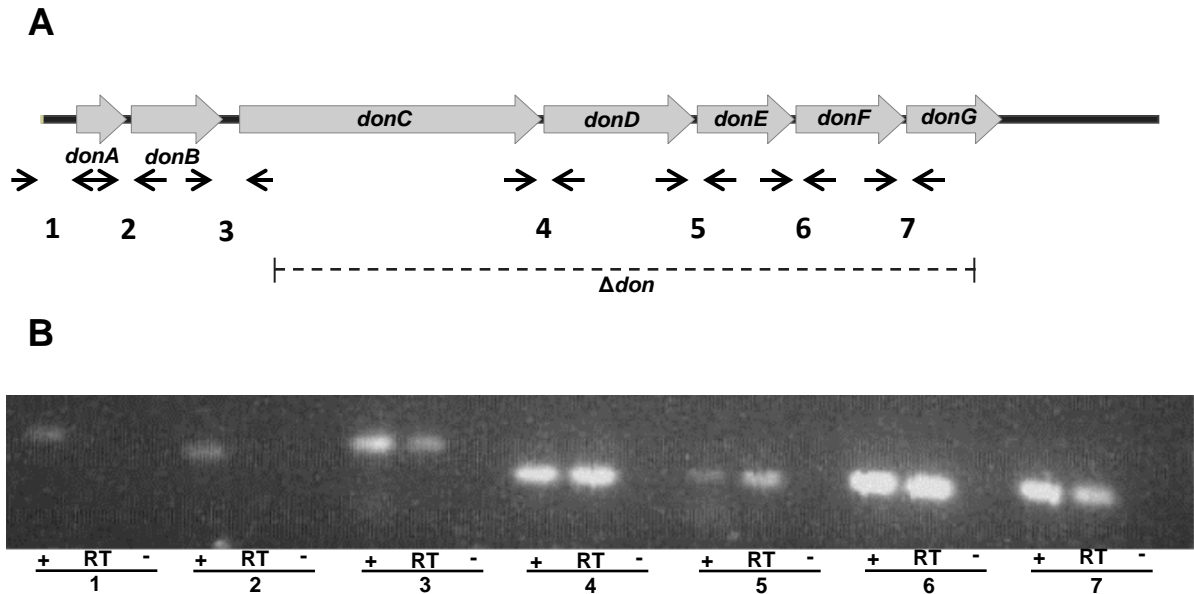


Fig. S1. Genetic map and RT-PCR analysis of the *donABCDEF G* genes (BF638R3437-43). **A)** Schematic diagram of the predicted *donABCDEF G* locus. The location of the Δdon deletion is shown with the dashed line. *donAB* are regulatory genes coding an ECF sigma factor and antisigma factor respectively. *donC* and *donD* code for the SusC and SusD paralogues respectively. The *donEFG* genes code for an endo-S-like endo-N-acetyl- β -D-glucosaminidase, a Concanavalin A lectin-like product with an α -N-acetylglucosaminidase domain, and a protein with a DUF 1735 domain common to acylhydrolases respectively. **B)** Agarose gel visualization of RT-PCR products performed with oligonucleotide primers specific for each intergenic region of the *donABCDEF G* locus. RNA was isolated from mid-logarithmic phase cultures of IB101 grown in defined media with mucin glycans as the carbon/energy source. Arrows below the schematic in panel A indicate the relative positions of primers used to amplify each intergenic region. “+”, positive control using genomic DNA; “RT”, complete RT-PCR reaction containing reverse transcriptase; “-”, negative control with reverse transcriptase omitted.

The RT-PCR results show a strong signal and linkage between the *donCDEF G* genes. There also is a weaker linkage between the antisigma factor gene, *donB*, and *donC* which are separated by 182 bp. Although these two genes are linked, based on the microarray data and regulatory models for other PULs it is likely a second promoter is located upstream *donCDEF G* that would act to amplify the expression of these structural genes under inducing conditions. Regulation of this locus is currently under investigation.

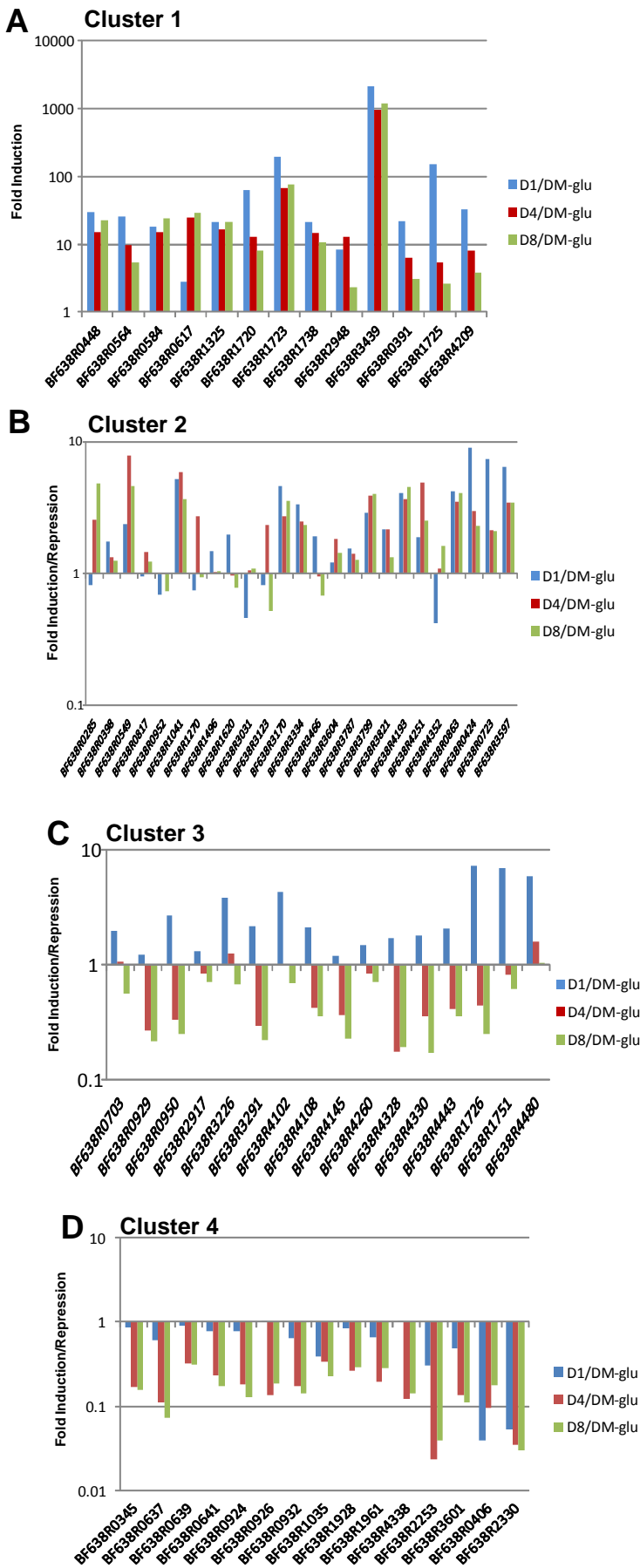


Fig. S2. Induction/repression patterns of SusC orthologues expressed during growth *in vivo*. Expression microarray data were used to determine induction or repression of all *susC*-like genes during growth of *B. fragilis* 638R in the rat tissue cage model relative to DM-glucose. Four induction patterns (Clusters) were identified by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. The four clusters are shown in panels A, B, C and D respectively. The fold-induction for *in vivo* growth relative to mid-logarithmic phase growth *in vitro* for 1, 4, and 8 days post-inoculation are shown by the blue, red, and green bars respectively.

Fig. S3

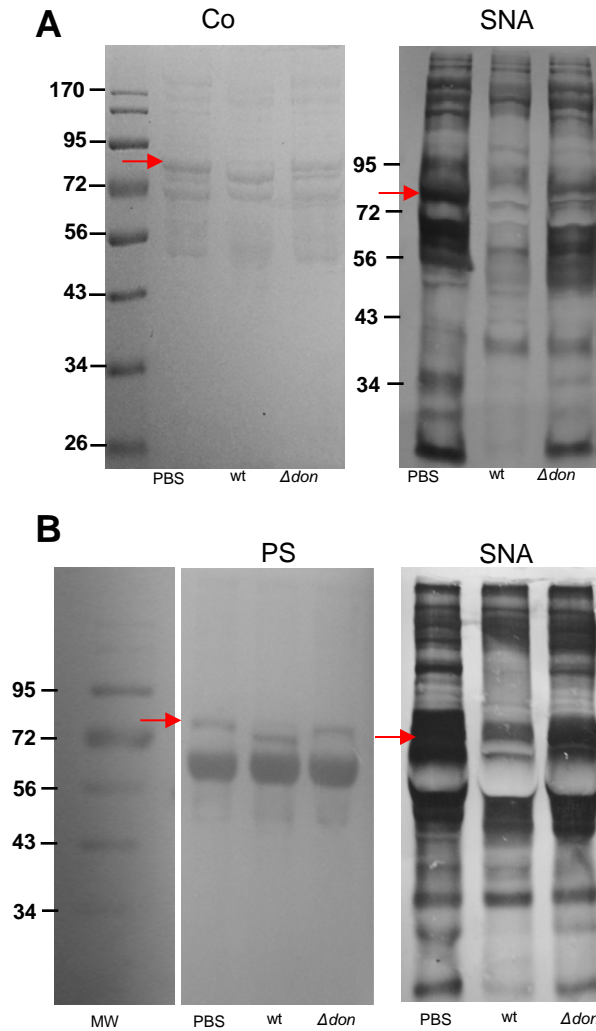


Fig. S3. Deglycosylation assays using total serous fluid proteins with serum albumin removed (**A**) or with serum albumin present (**B**). Standard deglycosylation analysis (SNA) was performed with wild type (wt) or Δdon strains for 3 h. Samples were analyzed by SDS-PAGE followed by protein staining with either Coomassie blue staining (Co) or Ponceau S (PS) or glycan staining SNA lectin. The arrows indicate the transferrin bands. Serum albumin was removed with the Qproteome Murine Albumin Depletion Kit, (Qiagen, Inc.).

Fig. S4

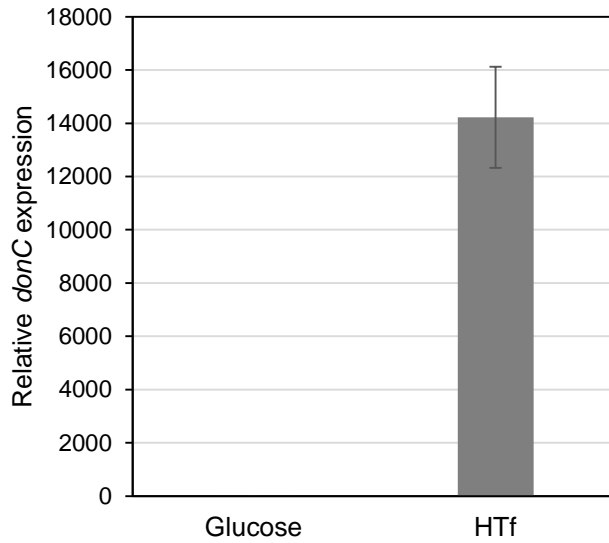


Fig. S4. DonC expression during growth in defined medium with human transferrin as the sole carbon/energy source. Expression of *donC* was determined by qRT-PCR for samples from mid-logarithmic cultures grown in DM-glucose (Glu) or DM-transferrin (HTf). The data were normalized to the amount of 16s RNA and the relative expression of transferrin grown cells to glucose grown cells was determined. The results are the average of triplicate samples.

Fig. S5

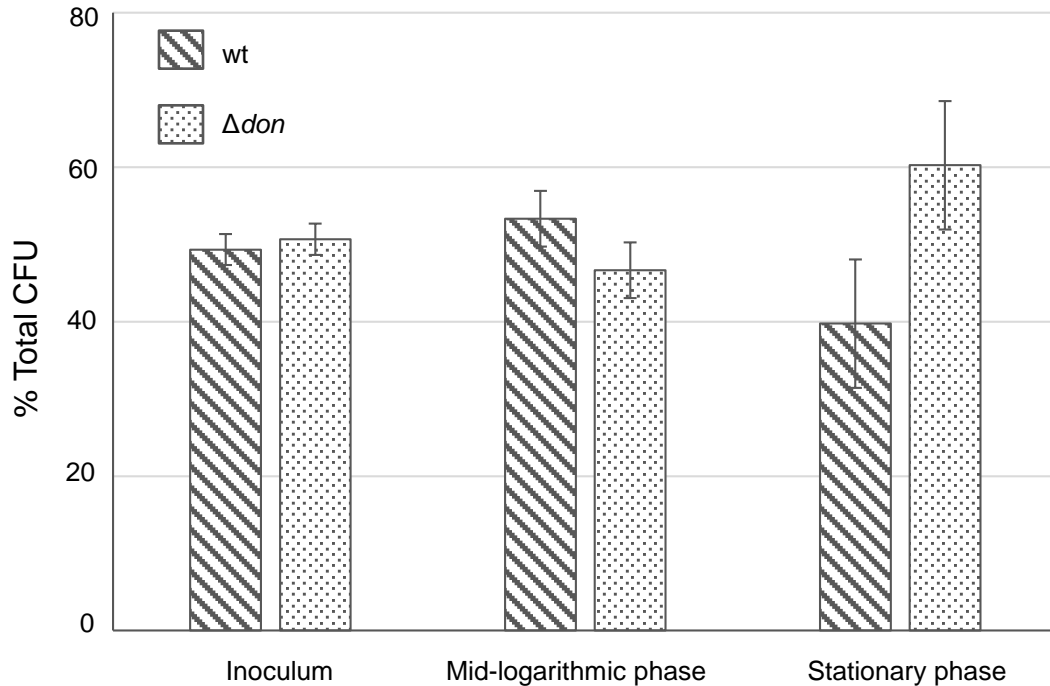


Fig. S5. *In vitro* competition assays in defined media with glucose. Mixtures containing a total of 4 ml of 10^5 CFU/ml of wild type and Δdon cells in a 1:1 ratio were inoculated into 25 mL of defined media with glucose. Aliquots of the cell mixture were removed from the inoculum, mid-logarithmic phase cultures and the stationary phase of the cultures, diluted in PBS and plated on BHIS for total cell counts. The percentage of Δdon among the total population was determined from the percentage of tetracycline resistant CFU counts among the total CFU counts.

Fig. S6

A) DonE

B) DonF

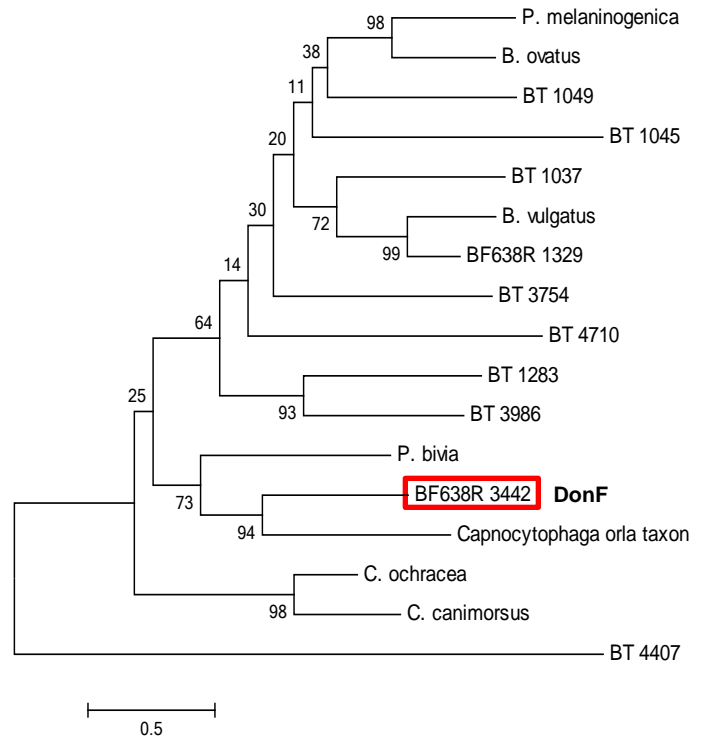
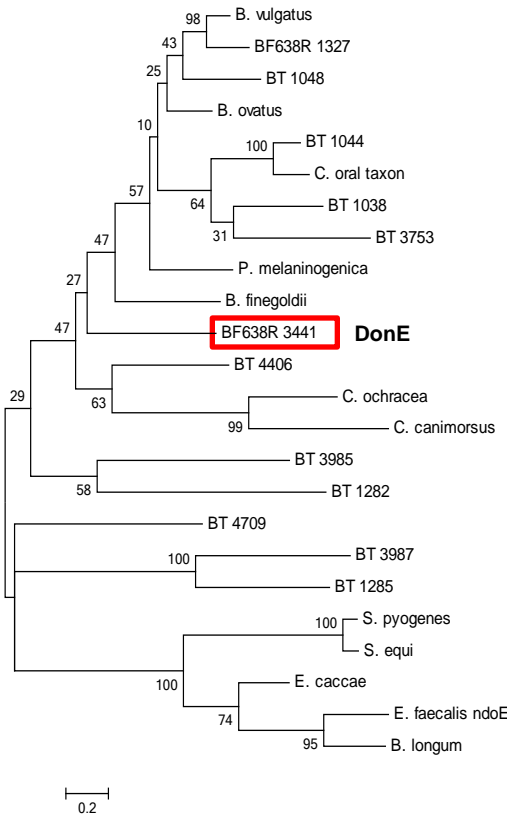


Fig. S6. Phylogenetic analysis of the *donE*, *F* and *G* gene products coding for an Endo-S-like endo-N-acetyl- β -D-glucosaminidase, a Concanavalin A-like lectin/glucanase, and an uncharacterized sugar binding protein respectively. **A)** DonE (BF638R3441) **B)** DonF (BF638R3442) **C)** DonG (BF638R3443). The evolutionary trees were determined using the Maximum Likelihood method as implemented in the MEGA6 software. The trees with the highest likelihood are shown. Numbers at the branches are the percentage of trees in which the associated taxa clustered together. Initial trees were obtained by the Neighbor-Joining method. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The protein sequences were obtained by using NCBI BLAST to identify similar sequences in the database. Most organisms are indicated by the species name. *B. thetaiotaomicron* (strain 5482) had many similar sequences and these were designated BT followed by the gene number.

Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

Fig. S6 (continued)

C) DonG

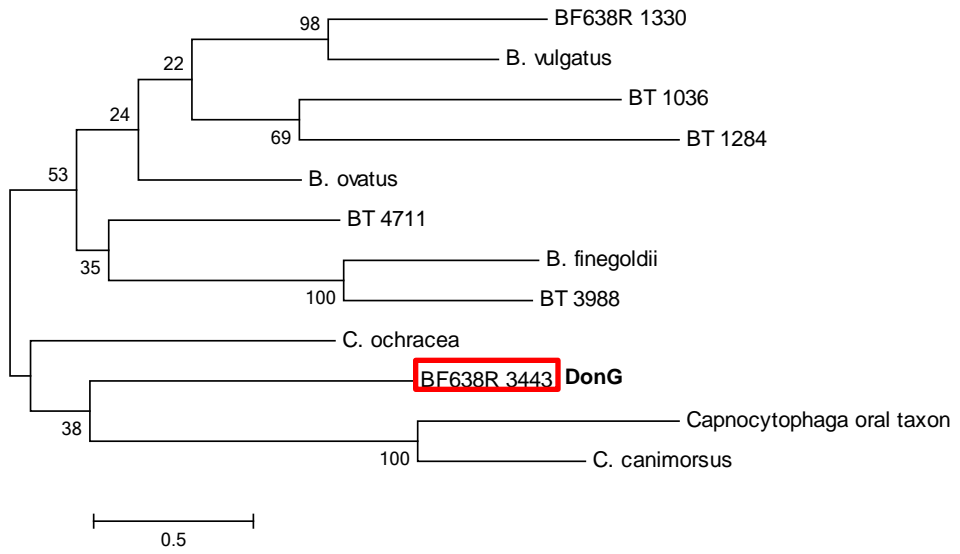


Fig. S7

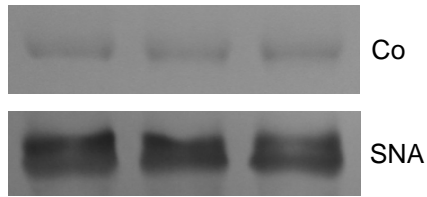


Fig. S7. Deglycosylation analysis of IgG. Deglycosylation assay of human IgG by *B. fragilis*. Commercially prepared human IgG was mixed and incubated anaerobically at 37°C for 3 h with PBS or wild type or Δdon cells grown in an inducing DM-mucin glycan medium. Samples were analyzed by SDS-PAGE and subjected to Coomassie blue staining (Co) or SNA glycan-staining (SNA).

Fig. S8

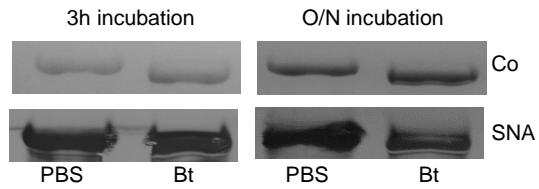


Fig. S8. Deglycosylation analysis of transferrin by *B. thetaiotaomicron* strain VPI5482. Commercially prepared human transferrin was incubated with PBS or wild type *B. thetaiotaomicron* cells (Bt) for 3 h or overnight (O/N) in standard deglycosylation assays. Cells used in the assays were grown to mid-logarithmic phase in DM-mucin glycan medium. Samples were analyzed by SDS-PAGE followed by Coomassie blue staining (Co) or SNA glycan staining (SNA).

Table S1. Induction and repression of *Bacteroides fragilis* BF638R PUL genes in vivo and in vitro.

Gene	FUNCTION	*DM-M/DM-G	D1/DM-G	D4/DM-G	D8/DM-G
BF638R0283	exported alpha-galactosidase	0.84	0.90	1.70	2.85
BF638R0284	conserved hypothetical exported protein	0.84	0.60	1.28	2.08
BF638R0285	SusC-like	0.61	0.82	2.58	4.84
BF638R0286	regulatory protein	0.80	0.61	19.93	15.68
BF638R0287	ECF-type sigma factor	1.10	0.74	0.83	1.09
BF638R0288	exported alpha-galactosidase	1.63	0.94	1.18	0.86
BF638R0345	SusC-like	2.29	0.85	0.17	0.16
BF638R0346	lipoprotein	1.96	0.70	0.38	0.34
BF638R0347	exported protein	2.14	0.40	0.23	0.24
BF638R0348	exported protein	2.18	0.91	1.03	0.75
BF638R0349	exported endo-arabinase	0.79	0.69	0.75	0.65
BF638R0384	conserved hypothetical exported protein	1.99	1.61	0.40	0.33
BF638R0385	beta-galactosidase	2.23	2.29	0.24	0.18
BF638R0386	exported beta-hexosaminidase	5.52	3.61	0.35	0.32
BF638R0387	conserved hypothetical exported protein	5.99	4.49	0.44	0.78
BF638R0388	exported hydrolase	5.17	4.79	0.86	0.80
BF638R0389	conserved hypothetical protein	4.17	6.56	4.08	3.39
BF638R0390	exported protein	20.14	12.38	4.37	3.01
BF638R0391	SusC-like	37.20	21.85	6.25	3.07
BF638R0392	two-component system sensor histidine ki	1.22	3.45	3.54	2.64
BF638R0393	two-component sensor histidine kinase	1.78	11.73	3.31	2.65
BF638R0394	conserved hypothetical exported protein	1.41	8.51	4.32	3.41
BF638R0395	conserved hypothetical exported protein	1.27	2.73	1.72	1.52
BF638R0397	exported protein	1.28	1.92	1.39	1.40
BF638R0398	SusC-like	1.22	1.74	1.33	1.26
BF638R0406	SusC-like	0.03	0.04	0.10	0.18
BF638R0407	outer membrane protein	0.06	0.10	0.21	0.30
BF638R0421	exported protein	4.32	3.75	1.32	1.30
BF638R0422	exported protein	5.70	4.33	1.56	1.23
BF638R0423	outer membrane protein	10.43	6.80	1.52	1.50
BF638R0424	SusC-like	15.08	8.96	2.96	2.29

BF638R0425	two-component system sensor histidine ki	1.82	16.18	2.42	2.01
BF638R0444	possible exported xanthan lyase/N-acetylm	4.30	2.57	0.83	0.69
BF638R0445	cation symporter	4.23	18.45	12.65	12.29
BF638R0446	conserved hypothetical exported protein	4.93	6.79	11.61	14.21
BF638R0447	conserved hypothetical lipoprotein	44.85	25.01	10.22	16.46
BF638R0448	SusC-like	55.44	30.24	14.86	22.53
BF638R0546	melibiase	1.65	0.86	1.79	1.94
BF638R0547	glycosyl hydrolase, alpha-xylosidase	2.40	0.86	0.91	0.89
BF638R0548	outer membrane protein	2.16	0.99	1.19	0.85
BF638R0549	SusC-like	2.97	2.36	7.81	4.60
BF638R0550	anti-sigma factor	1.40	0.73	18.21	6.61
BF638R0551	ECF-type RNA polymerase sigma factor	1.42	1.62	0.50	0.63
BF638R0555	ATP-binding component of ABC transporte	0.42	0.23	0.13	0.27
BF638R0556	ABC transporter	0.47	0.32	0.25	0.51
BF638R0564	SusC-like	21.32	25.78	9.83	5.33
BF638R0565	conserved hypothetical protein	9.58	13.48	4.87	3.46
BF638R0566	arylsulfatase	6.60	6.71	2.58	1.69
BF638R0582	RNA polymerase ECF sigma factor	0.91	0.66	0.92	1.08
BF638R0583	anti-sigma factor	1.18	0.79	2.15	3.34
BF638R0584	SusC-like	3.30	18.29	14.87	23.79
BF638R0585	outer membrane protein	2.65	8.11	8.98	14.03
BF638R0586	conserved hypothetical protein	2.50	5.69	4.39	6.85
BF638R0614	RNA polymerase ECF-type sigma factor	1.54	1.30	2.48	3.91
BF638R0615	anti-sigma factor	1.33	0.43	1.62	1.24
BF638R0616	conserved hypothetical protein	1.32	1.77	6.72	6.50
BF638R0617	SusC-like	2.19	2.80	24.96	29.00
BF638R0618	outer membrane protein	2.06	2.16	9.34	14.91
BF638R0619	beta-galactosidase	1.44	0.71	1.71	3.31
BF638R0620	glycosyl hydrolase	1.32	0.64	0.84	1.39
BF638R0621	glycosyl hydrolase/xylanase	1.30	0.83	0.86	1.32
BF638R0636	conserved hypothetical protein	0.52	0.64	0.12	0.08
BF638R0637	SusC-like	0.57	0.60	0.11	0.07
BF638R0638	membrane protein	0.53	0.99	0.33	0.29
BF638R0639	SusC-like	0.58	0.91	0.32	0.31

BF638R0641	SusC-like	0.70	0.77	0.23	0.17
BF638R0642	lipoprotein	0.65	0.85	0.30	0.23
BF638R0643	GTP-binding protein	0.69	0.30	0.48	0.38
BF638R0644	conserved hypothetical protein	0.62	0.52	0.28	0.24
BF638R0645	aminopeptidase	0.96	0.41	0.16	0.16
BF638R0703	SusC-like	3.60	1.94	1.04	0.55
BF638R0704	conserved hypothetical protein	2.61	1.39	0.66	0.36
BF638R0705	conserved hypothetical exported protein	2.20	1.31	0.80	0.59
BF638R0706	conserved hypothetical exported protein	1.66	0.94	0.43	0.23
BF638R0722	conserved hypothetical lipoprotein	3.17	6.13	1.76	1.95
BF638R0723	SusC-like	3.85	7.41	2.15	2.09
BF638R0724	alpha-1,2-mannosidase precursor	2.84	6.11	2.05	1.62
BF638R0725	alpha-1,2-mannosidase precursor	3.03	7.13	3.62	3.16
BF638R0726	anti-sigma factor	0.77	2.23	3.46	4.25
BF638R0727	RNA polymerase ECF-type sigma factor	0.59	1.17	0.65	0.43
BF638R0728	alpha-1,2-mannosidase precursor	0.64	1.17	0.96	0.85
BF638R0816	glycosyl hydrolase	1.68	0.74	2.88	3.75
BF638R0817	SusC-like	1.96	0.96	1.47	1.24
BF638R0818	outer membrane protein	1.13	0.69	1.06	0.83
BF638R0819	lipoprotein	1.08	0.51	1.00	0.97
BF638R0858	secreted sulfatase	1.50	2.38	0.31	0.26
BF638R0859	possible alpha-galactosidase	1.34	2.98	0.32	0.24
BF638R0860	conserved hypothetical protein	2.02	2.38	0.66	0.43
BF638R0861	sulfatase	4.55	2.10	2.88	2.00
BF638R0862	conserved hypothetical protein	3.68	3.97	3.52	4.19
BF638R0863	SusC-like	4.95	4.21	3.53	4.12
BF638R0864	two-component system sensor histidine ki	0.99	0.87	0.57	0.43
BF638R0920	beta-glucosidase	4.59	1.69	1.59	1.37
BF638R0921	cytochrome c binding protein	2.95	0.71	0.47	0.40
BF638R0922	transmembrane protein	2.41	3.45	1.04	0.81
BF638R0923	outer membrane protein	6.39	1.63	0.80	0.62
BF638R0924	SusC-like	7.44	0.78	0.18	0.13
BF638R0925	conserved hypothetical protein	1.14	1.45	0.32	0.38
BF638R0926	SusC-like	1.16	1.01	0.14	0.18

BF638R0927	possible outer membrane protein	2.34	1.34	0.31	0.37
BF638R0929	SusC-like	2.80	1.21	0.27	0.22
BF638R0930	RNA polymerase ECF-type sigma factor	0.62	0.58	0.52	0.63
BF638R0931	anti-sigma factor	5.77	0.82	11.63	6.73
BF638R0932	SusC-like	3.68	0.64	0.17	0.14
BF638R0933	possible outer membrane protein	3.76	0.66	0.40	0.33
BF638R0949	conserved hypothetical protein	3.16	2.27	0.54	0.35
BF638R0950	SusC-like	4.16	2.69	0.33	0.25
BF638R0951	conserved hypothetical protein	1.43	0.82	1.04	1.61
BF638R0952	SusC-like	2.06	0.69	1.01	0.74
BF638R1032	RNA polymerase ECF-type sigma factor	0.80	1.33	2.93	2.41
BF638R1033	anti-sigma factor	1.65	1.13	5.94	3.39
BF638R1035	SusC-like	1.12	0.39	0.34	0.23
BF638R1036	conserved hypothetical protein	1.10	0.37	0.12	0.10
BF638R1037	endonuclease/exonuclease/phosphatase f	1.25	0.47	0.18	0.16
BF638R1038	exported protein	1.93	0.39	0.32	0.37
BF638R1039	RNA polymerase ECF-type sigma factor	1.47	1.12	3.34	2.11
BF638R1040	anti-sigma factor	7.58	0.94	15.74	9.82
BF638R1041	SusC-like	29.69	5.19	5.89	3.69
BF638R1042	outer membrane protein	18.88	3.83	3.75	2.63
BF638R1043	arylsulfatase precursor	14.97	4.08	4.27	2.98
BF638R1045	arylsulfatase	1.77	0.48	1.58	1.74
BF638R1264	alpha-glucosidase, glycosylhydrolase	2.19	1.49	1.18	1.11
BF638R1265	ECF-type RNA polymerase sigma factor	1.17	0.34	0.22	0.23
BF638R1266	xylosidase/arabinosidase	2.18	0.16	0.33	0.39
BF638R1267	beta-lactamase	0.73	0.49	0.68	1.14
BF638R1268	hypothetical protein	0.80	0.46	0.60	0.53
BF638R1269	anti-sigma factor	0.56	0.39	3.33	2.04
BF638R1270	SusC-like	0.79	0.74	2.71	0.93
BF638R1271	conserved hypothetical protein	0.77	0.68	1.24	0.51
BF638R1272	phosphohydrolase, lcc family	1.14	1.23	1.33	1.24
BF638R1273	possible phosphodiesterase/nucleotide pyr	1.67	0.51	1.08	0.87
BF638R1274	endonuclease/exonuclease/phosphatase f	1.94	0.62	0.43	0.37
BF638R1323	RNA polymerase ECF-type sigma factor	0.76	2.33	5.17	7.21

BF638R1324	anti-sigma factor	2.00	2.03	13.65	13.92
BF638R1325	SusC-like	20.82	21.43	16.60	20.89
BF638R1326	conserved hypothetical protein	20.04	17.86	12.88	17.23
BF638R1327	possible endo-beta-N-acetylglucosaminidas	19.35	17.54	11.15	16.70
BF638R1329	conserved hypothetical protein	20.68	15.05	9.62	13.49
BF638R1330	conserved hypothetical protein	13.81	13.37	8.34	10.29
BF638R1494	RNA polymerase ECF-type sigma factor	1.07	7.96	6.75	13.31
BF638R1495	anti-sigma factor	1.20	1.83	11.32	4.29
BF638R1496	SusC-like	1.07	1.47	1.04	1.05
BF638R1497	outer membrane protein	0.88	0.54	0.37	0.34
BF638R1498	hypothetical protein	0.96	1.08	0.97	0.79
BF638R1499	alpha-galactosidase	1.03	1.68	0.99	1.04
BF638R1620	SusC-like	1.14	1.99	0.97	0.79
BF638R1621	conserved hypothetical protein	1.25	3.01	1.24	1.05
BF638R1715	NanR, ROK family transcriptional regulator	1.40	2.42	3.22	3.16
BF638R1716	NanL, NANA aldolase	4.30	16.33	7.74	6.73
BF638R1717	NanE, N-acetylglucosamine 2-epimerase	5.01	26.95	9.11	8.42
BF638R1718	NanT, transport protein	9.54	64.83	15.77	13.94
BF638R1719	hypothetical protein	5.86	43.75	8.59	9.89
BF638R1720	SusC-like	5.94	62.69	12.81	8.02
BF638R1721	conserved hypothetical protein	5.43	74.43	16.70	11.68
BF638R1722	conserved hypothetical protein	21.28	188.38	63.00	75.48
BF638R1723	SusC-like	47.55	195.37	67.18	76.87
BF638R1724	conserved hypothetical protein	13.49	163.55	7.43	4.49
BF638R1725	SusC-like	16.81	152.89	5.42	2.61
BF638R1726	SusC-like	3.38	7.18	0.44	0.25
BF638R1727	exported protein	2.52	7.96	0.71	0.52
BF638R1728	NanH, neuraminidase precursor	9.81	37.60	12.63	12.98
BF638R1729	NahA, beta-N-acetylhexoosaminidase	19.66	128.60	15.90	15.22
BF638R1730	sialate-O-acetyltransferase	11.99	75.34	4.85	7.61
BF638R1731	sialate O-acetylesterase	2.60	3.43	1.19	1.28
BF638R1732	beta-mannosidase	1.38	1.84	0.45	0.61
BF638R1733	beta-N-acetylhexosaminidase	1.63	2.12	0.42	0.48
BF638R1734	lipoprotein	1.29	2.37	0.27	0.40

BF638R1735	beta-N-acetylhexosaminidase	1.28	1.18	0.52	0.43
BF638R1736	outer membrane protein	1.09	1.27	0.37	0.38
BF638R1737	beta-galactosidase	0.99	1.32	0.43	0.50
BF638R1738	SusC-like	6.81	21.16	14.41	10.74
BF638R1739	conserved hypothetical protein	5.40	6.36	3.37	1.96
BF638R1740	outer membrane protein	4.59	3.95	2.57	2.02
BF638R1751	SusC-like	23.68	6.85	0.81	0.62
BF638R1752	outer membrane protein	19.34	6.23	0.68	0.67
BF638R1753	alpha-1,2-mannosidase precursor	13.55	3.03	1.06	1.28
BF638R1754	endonuclease/exonuclease/phosphatase f	4.03	1.65	0.55	0.58
BF638R1755	Alkaline phosphatase	2.41	0.70	0.23	0.24
BF638R1756	conserved hypothetical protein	2.22	1.03	0.29	0.29
BF638R1757	exported protein	2.47	0.71	0.21	0.21
BF638R1758	exported protein	1.83	0.92	0.23	0.18
BF638R1928	SusC-like	1.01	0.85	0.26	0.29
BF638R1929	conserved hypothetical protein	1.01	1.10	0.32	0.36
BF638R1960	TonB-dependent outer membrane recepto	0.53	0.31	0.44	0.39
BF638R1961	SusC-like	0.34	0.65	0.19	0.28
BF638R1962	conserved hypothetical protein	0.32	0.42	0.09	0.15
BF638R1963	conserved hypothetical protein	0.33	0.34	0.06	0.07
BF638R1964	hypothetical protein	0.33	0.42	0.06	0.10
BF638R1965	hypothetical protein	0.37	0.54	0.09	0.14
BF638R1966	hypothetical protein	0.43	0.59	0.14	0.17
BF638R1967	conserved hypothetical protein	0.42	0.46	0.10	0.14
BF638R1968	conserved hypothetical protein	0.46	0.51	0.07	0.09
BF638R1969	polysialic acid capsule transport protein	0.52	0.82	0.16	0.21
BF638R2253	SusC-like	0.13	0.31	0.02	0.04
BF638R2254	lipoprotein	0.15	0.28	0.04	0.07
BF638R2329	lipoprotein	0.07	0.08	0.06	0.07
BF638R2330	SusC-like	0.07	0.05	0.03	0.03
BF638R2916	conserved hypothetical protein	2.63	1.04	0.73	0.95
BF638R2917	SusC-like	3.23	1.30	0.83	0.70
BF638R2918	membrane protein	3.26	0.78	1.83	1.64
BF638R2920	ECF sigma factor	1.90	3.73	13.53	10.60

BF638R2946	exported protein	15.80	3.53	8.07	1.32
BF638R2947	membrane protein	18.11	3.53	7.19	1.50
BF638R2948	SusC-like	34.43	8.37	12.90	2.29
BF638R2949	conserved hypothetical protein	15.35	1.25	12.24	3.46
BF638R3029	carbon-nitrogen hydrolase	0.64	1.25	1.39	0.97
BF638R3030	exported protein	0.39	1.59	1.85	0.87
BF638R3031	SusC-like	0.37	0.46	1.07	1.09
BF638R3032	exported protein	0.50	0.21	0.40	0.53
BF638R3033	exported protein	0.67	0.77	1.70	1.00
BF638R3120	sulfatase	1.48	0.63	0.39	0.24
BF638R3121	exported hydrolase	1.64	0.50	0.60	0.33
BF638R3122	lipoprotein	1.47	0.34	0.59	0.24
BF638R3123	SusC-like	1.20	0.82	2.34	0.52
BF638R3124	anti-sigma factor	1.81	1.40	3.64	2.52
BF638R3125	sigma factor	1.15	1.73	0.78	1.22
BF638R3167	conserved hypothetical lipoprotein	8.14	7.10	2.76	4.31
BF638R3168	lipoprotein	2.90	3.23	1.60	2.29
BF638R3169	membrane protein	4.07	3.93	1.86	3.07
BF638R3170	SusC-like	4.23	4.61	2.72	3.57
BF638R3226	SusC-like	5.03	3.82	1.24	0.67
BF638R3227	lipoprotein	3.44	3.90	1.43	0.81
BF638R3228	lipoprotein	3.73	5.26	3.18	2.70
BF638R3290	exported protein	7.72	1.67	0.27	0.20
BF638R3291	SusC-like	8.76	2.17	0.29	0.22
BF638R3331	hypothetical protein	0.90	0.56	0.27	0.32
BF638R3332	lipoprotein	1.08	0.42	0.53	0.44
BF638R3333	exported protein	3.34	2.11	1.27	0.87
BF638R3334	SusC-like	5.06	3.38	2.48	2.34
BF638R3437	donA, ECF sigma factor	3.60	23.33	14.85	22.79
BF638R3438	donB, antisigma factor	33.84	223.63	597.79	532.11
BF638R3439	donC, SusC-like	559.54	2146.72	933.46	1161.15
BF638R3440	donD, SusD-like	505.26	1907.71	700.53	918.12
BF638R3441	donE, endo-S-like	334.26	1494.32	443.63	631.07
BF638R3442	donF, lectin-like	233.61	765.28	367.94	438.35

BF638R3443	donG, exported protein	90.69	1227.85	235.37	363.24
BF638R3466	SusC-like	0.50	1.92	0.96	0.68
BF638R3467	conserved hypothetical protein	0.50	1.90	0.28	0.38
BF638R3468	lipoprotein	0.51	0.82	0.08	0.12
BF638R3469	peptidase	0.51	0.88	0.10	0.10
BF638R3470	hypothetical protein	0.41	0.70	0.12	0.15
BF638R3593	exported phosphatase	1.48	1.37	1.03	1.13
BF638R3594	exported protein	1.69	0.99	0.49	0.38
BF638R3595	exported phosphoesterase protein	1.81	0.89	0.72	0.48
BF638R3596	conserved hypothetical protein	6.21	4.97	3.34	3.16
BF638R3597	SusC-like	7.16	6.41	3.45	3.47
BF638R3598	anti sigma factor	2.46	1.60	17.35	14.63
BF638R3599	RNA polymerase sigma factor	1.76	1.46	0.82	0.69
BF638R3600	lipoprotein	8.64	0.67	0.20	0.15
BF638R3601	SusC-like	9.98	0.49	0.14	0.11
BF638R3602	conserved hypothetical protein	6.63	0.82	1.21	1.83
BF638R3603	membrane protein	10.52	0.87	1.31	1.01
BF638R3604	SusC-like	17.96	1.22	1.82	1.44
BF638R3605	regulatory protein (possible anti-sigma fact	5.53	0.74	10.18	7.99
BF638R3606	RNA polymerase sigma factor	0.87	0.49	0.83	1.81
BF638R3783	transport related, membrane protein	2.67	1.65	1.55	1.22
BF638R3784	sialidase	3.29	1.67	0.67	0.60
BF638R3785	hypothetical protein	3.64	2.55	2.00	1.42
BF638R3786	lipoprotein	2.56	1.55	1.31	1.32
BF638R3787	SusC-like	2.70	1.54	1.41	1.28
BF638R3788	N-acetylneuraminase lyase (sialic acid lyase	1.90	2.57	1.51	1.69
BF638R3789	GntR-family regulatory protein	1.29	0.38	1.12	1.22
BF638R3796	exported sulfatase	1.59	0.76	1.08	0.84
BF638R3797	exported uslfatase	2.21	1.11	1.20	0.82
BF638R3798	conserved hypothetical protein	2.31	1.70	2.66	2.92
BF638R3799	SusC-like	2.99	2.90	3.90	4.01
BF638R3800	membrane protein	1.37	0.66	15.37	9.87
BF638R3801	RNA polymerase sigma factor	0.59	2.08	3.20	4.32
BF638R3819	RNA polymerase sigma factor	1.95	2.42	3.01	3.69

BF638R3820	membrane protein	4.68	4.53	12.17	6.41
BF638R3821	SusC-like	2.13	2.18	2.17	1.33
BF638R3822	lipoprotein	1.80	0.94	0.60	0.35
BF638R3823	lipoprotein	0.95	1.39	1.07	0.74
BF638R3824	exported protein	0.08	0.12	0.07	0.06
BF638R4102	SusC-like	4.38	4.27	0.97	0.69
BF638R4103	outer membrane protein	3.60	3.57	0.51	0.37
BF638R4104	endo-beta-galactosidase	3.00	3.34	0.74	0.66
BF638R4105	beta-hexosaminidase precursor	2.15	1.40	0.18	0.17
BF638R4106	beta-glucanase	1.99	1.88	0.49	0.44
BF638R4107	beta-galactosidase	1.92	2.37	1.23	0.59
BF638R4108	SusC-like	2.63	2.12	0.42	0.36
BF638R4109	outer membrane protein	2.44	1.50	0.46	0.42
BF638R4142	glycosyl hydrolase	2.02	1.68	0.52	0.52
BF638R4143	conserved hypothetical protein	1.71	0.96	0.30	0.28
BF638R4144	outer membrane protein	2.27	0.84	0.48	0.42
BF638R4145	SusC-like	2.75	1.19	0.37	0.23
BF638R4146	anti-sigma factor	1.58	1.07	1.26	0.75
BF638R4147	ECF-type RNA polymerase sigma factor	0.80	0.72	0.33	0.36
BF638R4192	Transcriptional regulator	1.02	3.06	3.10	2.59
BF638R4193	SusC-like	42.41	4.08	3.66	4.58
BF638R4194	lipoprotein	32.61	3.23	1.83	2.95
BF638R4207	RNA polymerase ECF-type sigma factor	2.17	7.16	12.22	11.38
BF638R4208	anti-sigma factor	8.54	2.29	5.73	2.32
BF638R4209	SusC-like	57.11	32.97	8.14	3.80
BF638R4210	conserved hypothetical protein	24.09	9.37	2.59	2.31
BF638R4247	conserved hypothetical protein	1.41	2.26	2.58	2.10
BF638R4248	possible xylosidase/arabinoxidase	1.71	2.17	2.75	2.22
BF638R4249	conserved hypothetical exported protein	2.03	1.70	4.03	4.00
BF638R4250	outer membrane protein	1.59	1.20	3.79	2.11
BF638R4251	SusC-like	2.29	1.88	4.90	2.51
BF638R4258	hypothetical protein	1.46	0.52	0.20	0.19
BF638R4259	outer membrane protein	2.76	1.06	0.73	0.46
BF638R4260	SusC-like	3.79	1.46	0.83	0.70

BF638R4261	two-component system sensor kinase/resp	0.89	5.25	3.59	2.53
BF638R4328	SusC-like	1.91	1.69	0.17	0.19
BF638R4329	conserved hypothetical protein	1.69	1.29	0.18	0.19
BF638R4330	SusC-like	8.45	1.79	0.35	0.17
BF638R4331	outer membrane protein	6.52	1.39	0.40	0.17
BF638R4332	hypothetical protein	5.68	1.11	1.07	1.36
BF638R4338	SusC-like	2.02	1.03	0.12	0.14
BF638R4339	conserved hypothetical protein	2.18	1.09	0.12	0.14
BF638R4348	conserved hypothetical protein	0.93	0.86	1.30	1.18
BF638R4349	arylsulfatase	1.14	0.52	1.32	1.63
BF638R4350;BF6:	conserved hypothetical protein; lipoprotein	1.24	0.53	1.10	1.81
BF638R4352	SusC-like	1.24	0.42	1.10	1.61
BF638R4353	anti-sigma factor	1.21	0.28	12.51	9.78
BF638R4354	ECF-type RNA polymerase	0.95	1.05	2.02	3.62
BF638R4442	outer membrane protein	1.00	1.00	1.00	1.00
BF638R4443	SusC-like	2.59	2.06	0.41	0.36
BF638R4444	lipoprotein	0.65	0.46	0.15	0.21
BF638R4478	hypothetical protein	2.23	2.84	1.45	2.03
BF638R4479	outer membrane protein	4.48	4.32	0.95	0.72
BF638R4480	SusC-like	4.99	5.82	1.57	1.04
BF638R4481	hypothetical protein	2.09	1.35	1.07	0.70
BF638R4482	transmembrane sugar transporter	2.19	0.83	2.38	1.95

*All values were calculated from the expression microarray data deposited at the NCBI GEO, accession # GSE53883. *B. fragilis* 638R cultures were grown in vitro in defined medium with either 0.4% glucose (DM-G) or 2% mucin glycans (DM-M), or grown in vivo in the rat tissue cage for 1 (D1), 4 (D4), or 8 (D8) days. Values given in columns C-F are the ratios (induction or repression) of the RMA normalized Log2 expression values obtained from at least two biological replicates. The table rows are color coded according to which expression cluster the PUL is associated (see below). Genes associated with sialic acid utilization but not part of a recognized PUL are not color coded and remain uncolored.

Cluster 1: highly induced

Cluster 2: moderately induced or no change

Cluster 3: repressed following moderate induction on day 1

Cluster 4: repressed

Table S2. Proteomics Analysis of Rat Serous Fluid: Estimate of Relative Protein Abundance¹

Identified Proteins ²	Mol Wt	% of Total Protein ³	g/dL ⁴	Glycosylation sites ⁵
Serum albumin	69 kDa	26%	1.69	0
Serotransferrin	76 kDa	7.00%	0.45	1
Murinoglobulin-2	162 kDa	1.96%	0.32	11
Alpha-1-inhibitor 3	164 kDa	1.50%	0.10	13
Protein LOC297568	164 kDa	1.50%	0.10	0
Murinoglobulin-1	165 kDa	1.40%	0.09	13
Hemopexin	51 kDa	1.30%	0.08	5
Complement C3	186 kDa	1.30%	0.08	2
Alpha-1-antiproteinase	46 kDa	1.20%	0.08	3
Serine protease inhibitor A3I	47 kDa	1.20%	0.08	4
Protein LOC299282	68 kDa	1.10%	0.07	0
Alpha-1-macroglobulin	167 kDa	1.00%	0.06	11
Serine protease inhibitor A3L	46 kDa	1.00%	0.06	4
Complement C4	192 kDa	0.65%	0.04	5
Ceruloplasmin	121 kDa	0.64%	0.04	6

¹ Determined by Q-Exactive Orbitrap Mass Spectrometry at the UC Davis Proteomics Core Facility. Uninfected serous fluid was subject to SDS-PAGE, digested with trypsin (in-gel), extracted and analyzed by LC-MS/MS.

² The 15 most abundant proteins are shown and account for about 50% of total protein spectra observed.

³ Estimated from the % of the total number of spectra observed for each protein

⁴ Protein concentrations were estimated based on their respective percentage of total spectra relative to the total amount of protein concentration in the serous fluid (6.5 g/dL).

⁵ The number of annotated N-glycosylations sites. These proteins have no annotated or predicted O-linked glycosylation.

Table S3. Bacterial strains and plasmids used in this study.

Bacterial strain or plasmid	^a Description	Reference or source
<i>Bacteroides</i> strains		
IB101	<i>B. fragilis</i> , 638R, clinical isolate, Rf ^f	(1)
IB102	<i>Bacteroides uniformis</i> , VPI strain 006-1 (ATCC8492), Rf ^f	(2)
IB103	<i>Bacteroides ovatus</i> , VPI 0038 (ATCC 8483)	(2)
IB114	<i>B. fragilis</i> , ATCC 25285, clinical isolate, Rf ^f	(3)
IB116	<i>Bacteroides thetaiotaomicron</i> , VPI strain 2302, Rf ^f	(2)
IB351	<i>Bacteroides vulgatus</i> , ATCC 8482	(2)
IB555	638R Δ don, BF638R3439-3443 genes replaced with a <i>tetQ</i> cassette, Tc ^r , Rf ^f	This study
BT5482R	<i>Bacteroides thetaiotaomicron</i> , VPI strain 5482 (ATCC 29148), Rf ^f	(4)
BER37	<i>Parabacteroides distasonis</i> , clinical isolate CLA348	(5)
BER39	<i>Parabacteroides merdae</i> , ATCC 43185	
<i>E. coli</i> strains		
DH10B	<i>E. coli</i> , F ⁻ <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74</i> <i>recA1 endA1 araD139</i> Δ (<i>ara leu</i>) 7697 <i>galJ galK rpsL nupG</i> λ ⁻	Invitrogen
HB101::RK231	<i>E. coli</i> , HB101 containing RK231, Kn ^r , Tc ^r , Sm ^r	(6)
Plasmids		
pFD516	<i>Bacteroides</i> suicide vector, 7.7 kb, Sp ^r in <i>E. coli</i> , Em ^r in <i>Bacteroides</i>	(7)
pFD516/omp117 \pm 2K/tetQ	pFD516 containing the BF638R3439-3443 deletion construct with the <i>tetQ</i> gene cassette, 14.6 kb, Sp ^r in <i>E. coli</i> , Em ^r and Tc ^r in <i>Bacteroides</i> .	This study

^a Rf, rifampicin; Tc, tetracycline; Sp, spectinomycin; Em, erythromycin; Kn, kanamycin; Sm, streptomycin.

REFERENCES

1. Privitera G, Dublanchet A, & Sebald M (1979) Transfer of multiple antibiotic resistance between subspecies of *Bacteroides fragilis*. *J.Infect.Dis.* 139:97-101.
2. Johnson JL (1978) Taxonomy of the bacteroides. I. Deoxyribonucleic acid homologies among *Bacteroides fragilis* and other saccharolytic *Bacteroides* species. *Int.J.Syst.Bacteriol.* 28:245-256.
3. Sund CJ, et al. (2008) The *Bacteroides fragilis* Transcriptome Response To Oxygen and H2O2: The Role Of OxyR And Its Effect On Survival And Virulence. *Mol.Microbiol.* 67:129-142.
4. Bedzyk L, Shoemaker NB, Young KE, & Salyers AA (1992) Insertion and excision of *Bacteroides* conjugative chromosomal elements. *J.Bacteriol.* 174:166-172.
5. Jacobs MR, Spangler SK, & Appelbaum PC (1990) Beta-lactamase production, beta-lactam sensitivity and resistance to synergy with clavulanate of 737

Bacteroides fragilis group organisms from thirty-three US centres. *J.Antimicrob. Chemother.* 26(3):361-370.

6. Guiney DG, Hasegawa P, & Davis CE (1984) Plasmid transfer from *Escherichia coli* to *Bacteroides fragilis*: differential expression of antibiotic resistance phenotypes. *Proc.Natl.Acad.Sci.USA* 81:7203-7206.
7. Smith CJ, Rollins LA, & Parker AC (1995) Nucleotide sequence determination and genetic analysis of the *Bacteroides* plasmid, pBI143. *Plasmid* 34:211-222.

Table S4. Preparation of porcine gastric mucin glycans.

Procedure ^{a,b}
1. Suspend 2.5% (w/v) of Porcine gastric mucin (Type II, Sigma-Aldrich, Inc.) in 100 mM Tris Buffer (pH7.4) and autoclave for 5 min.
2. Cool mixture to 65 °C and incubate with proteinase K (0.1mg/ml) at 65 °C for 16 to 20 h.
3. Centrifuge mixture at 21,000 g for 30 min at 4 °C.
4. Collect supernatant, add NaOH and NaBH ₄ to final concentrations of 0.1 M and 1.0 M respectively.
5. Incubate reaction mixture at 55 °C for 20 hours.
6. Neutralize mixture to pH7.0 by the addition of HCl.
7. Centrifuge at 21,000 g for 30 min at 4 °C.
8. Collect supernatant and filter through a 0.22 µm filter. Extensively dialyze filtrate against deionized distilled water with 1 kDa cutoff dialysis tubing.
9. Lyophilize and store dry until needed.

^aAdapted from the Supplemental Experimental Procedures of Martens et al. 2008, Cell Host & Microbe 4:447-457

^bAlthough optimized for the release of O-linked glycans, the procedure does liberate some N-linked glycans as shown previously (Supplemental Data Table S1, Martens et al. 2008, Cell Host & Microbe 4:447-457).

Table S5. Oligonucleotides used in this study.

Name	Sequence (5' → 3')	Tag	Description
UF	AGTCCTGCAGCAAACAGATGCCTTATCAGTTT	PstI	Designed to amplify 2 kb upstream of <i>donC</i>
UR	AGTCGGATCCCAGTGTCTTATCCAGAGATTCA	BamHI	
DF	AGTCGGATCCCTGGACAATATCGATTCAAGCA	BamHI	Designed to amplify 2 kb downstream for <i>donG</i>
DR	AGTCGAGCTCATAGCGACCAAAACCC TTCT	SacI	
Omp117rtL	GGTGAAGGCATTTCCGACTT		Designed to amplify a 140 bp fragment of <i>donC</i> gene for quantitative PCR
Omp117rtR	TTGCCTTCCTGCCCTTTCTT		
16srL	GATGCGTTCCATTAGGTTGTTG		Designed to amplify a 127 bp fragment of 16s ribosomal RNA gene for quantitative PCR
16srR	CACTGCTGCCTCCCGTAG		
PsigOKL2	AATCACAATCAGCCTTATATTCTAC		Designed to amplify a 375 bp fragment spanning <i>donA</i> and its potential TIS region
PsigOKR	CATTGCTCCTGATAGGTCTG		
PantisigOKL	GGATATTTTCGATCCGCACTG		Designed to amplify a 298 bp fragment spanning the intergenic region of <i>donA</i> and <i>donB</i>
PantisigOKR	TTTCCTTCGGTTTCGTTCCA		
Pomp117L	AGATATCCTGAACATTATGCAGG		Designed to amplify a 339 bp fragment spanning the intergenic region of <i>donB</i> and <i>donC</i>
Pomp117R	GCCGTTTGTAGAATAAAACAGC		
PsusDL	AATGACCTCTTCGACAGGTA		Designed to amplify a 220 bp fragment spanning the intergenic region of <i>donC</i> and <i>donD</i>
PsusDR	GGTTCCGGTTGATATCTTCG		
PendoSL	ATTGATCGGAAACGGATACG		Designed to amplify a 240 bp fragment spanning the intergenic region of <i>donD</i> and <i>donE</i>
PendosR	CTACCAACAATGCTACTCCG		
PsusEL	GAAGACGCATTTGAAGACGA		Designed to amplify a 261 bp fragment spanning the intergenic region of <i>donE</i> and <i>donF</i>
PsusER	GCTCTAAAGTGAAGTTGACGA		
PsusFL	GAGACCAAAGACCTGAAAGC		Designed to amplify a 288 bp fragment spanning the intergenic region of <i>donF</i> and <i>donG</i>
PsusFR	GAGTTGCTCACGGATATCTT		