

SUPPLEMENTAL DATA

SUPPLEMENTAL RESULTS AND DISCUSSION

Disruption of a single trans-envelope signaling system during growth in a condition where multiple systems are simultaneously activated—Multiple Sus-like systems involved in host glycan catabolism, including the *BT1032-53* system, are simultaneously expressed both *in vitro* during growth on *O*-glycans and *in vivo* in the distal mouse gut (1). The genetic experiments described in the main text demonstrate that the trans-envelope signaling switch encoded within the *BT1032-53* PUL is essential for its regulation. We also wished to determine if loss of the positive-acting components of this particular trans-envelope signaling switch (*BT1042* and *BT1053*) influenced other cellular responses to *O*-glycans, such as expression of other PULs. To address this question, we obtained whole genome transcriptional profiles for the $\Delta BT1053$ and $\Omega BT1042$ mutants grown to mid-exponential phase in MM-neutral *O*-glycans and compared them to our previously published GeneChip dataset for wild-type *B. thetaiotaomicron* grown in the same medium (n=2 biological replicates performed/strain). Differences in gene expression were determined by referencing each strain's transcriptional profile to wild-type *B. thetaiotaomicron* grown in MM-glucose and identifying genes with ≥ 5 -fold expression changes. By comparing each of the three MM-neutral *O*-glycan datasets to this common MM-glucose reference, we were able to directly compare all three datasets. A total of 223 genes exhibited ≥ 5 -fold changes in their expression: 195 of these (87.4%) were up-regulated relative to MM-glucose (see Table S5 for a complete list with fold-changes). A Venn diagram comparison of these three strains' responses to neutral *O*-glycans revealed genes present in all 7 sectors (groups) defined in Fig. S4A. However, the largest group of regulated genes (130 up-regulated and 15 down-regulated; 65% of all regulated genes) were contained in Group 5, indicating that their expression changes were similar among all three datasets and therefore not perturbed in either mutant. Notably, this group of shared responses encompasses 8 out of 9 PULs (blue-bracketed genes in Fig. S4B) that were upregulated ≥ 10 -fold during growth in MM-neutral *O*-glycans versus MM-glucose (1), and only excludes the *BT1032-53* PUL which belonged to Group 1. Thus, loss of the trans-envelope signaling switch linked to the *BT1032-53* locus does not cause noticeable alterations in expression of other PULs. This notion is consistent with the conclusions of our yeast two-hybrid analysis, which suggest that the transcription factors that activate PUL expression function primarily as specific local activators of their adjacent PUL genes.

Two additional groups delineated in Fig. S4A (1 and 3) include particularly interesting sets of genes. Group 1 includes genes that are induced in the wild-type strain, but that lose expression in both the $\Omega BT1042$ and $\Delta BT1053$ mutants. All except three genes in this group belong to the *BT1032-53* PUL (blue bracketed genes in the Group 1 section of Fig. S4B), confirming that this system's trans-envelope signaling switch primarily effects the adjacent PUL genes, and suggesting that regulators from the other 8 systems do not compensate for this loss of expression. However, three additional genes encoding proteins of unknown function (*BT3221-23*) are also in Group 1 and are therefore dependent on the *BT1032-53* signaling switch for expression in MM-neutral *O*-glycan medium (red bracketed genes in the Group 1 section of Fig. S4B). Each of these genes encodes a protein with a predicted signal peptide, suggesting that their products are secreted and possibly play a role in assimilating certain *O*-glycan-derived sugars that are specifically imported via the *BT1032-53* system. If this were the case, they may be indirectly dependent on ECF- σ^{BT1053} for activation because they would respond to products generated by the ECF- σ^{BT1053} -activated *BT1032-53* Sus-like system only when its *O*-glycan substrate(s) was present. Consistent with this notion, *BT3221-23* do *not* exhibit any transcriptional activation during artificial de-repression of ECF- σ^{BT1053} in MM-glucose (Table S4), suggesting that their transcription is not directly dependent on this ECF- σ factor.

Also of interest are 22 of the 24 up-regulated genes belonging to Group 3 (*i.e.*, those up-regulated in response to neutral *O*-glycans but *only* in the $\Delta BT1053$ mutant), which are encompassed within one of *B. thetaiotaomicron*'s eight *CPS* loci, *CPS8* (green bracketed genes in the Group 3 section of Fig. S4B). This

result is notable in light of the results presented in the main text that link four different trans-envelope signaling switches, including the one associated with the *BT1032-50* PUL, with regulation of various *CPS* loci (Fig. 4; note that *CPS8* was induced during de-repression of anti- σ^{BT1052}). In contrast to the anti- σ de-repression experiments described in the main text, which yielded transcriptional responses from ECF- σ -dependent genes in minimal medium lacking complex glycans, specific induction of *CPS8* in the $\Delta BT1053$ mutant reveals that shifting the repertoire of PULs that are simultaneously expressed by the cell can indeed result in alterations in capsule expression. The mechanism through which loss of BT1053, but not BT1042, uniquely results in increased *CPS8* expression remains unclear. One possibility is that ECF- σ^{BT1053} is essential for priming expression of some genes (*i.e.*, those in the *BT1032-53* PUL), whereas the SusC-like transporter BT1042 is only essential for subsequent amplification of PUL gene expression upon *O*-glycan recognition. Thus, loss of these two factors could result in slightly different effects on protein expression: *e.g.*, complete lack of *BT1032-53* expression in the $\Delta BT1053$ mutant versus low-level expression that cannot be increased subsequently in response to glycan cues in the $\Omega BT1042$ mutant. Because at least 8 other Sus-like systems are subsequently highly expressed in the same growth condition (MM-neutral *O*-glycans), the resulting differential in *BT1032-53* expression may result in the observed effect on *CPS8*. Future experiments that explore the pathways linking CPS and PUL gene expression will be needed to obtain additional insights into this phenomenon.

SUPPLEMENTAL REFERENCES

1. Martens, E. C., Chiang, H. C., and Gordon, J. I. (2008) *Cell Host & Microbe* **4**(5), 447-457
2. Sonnenburg, E. D., Sonnenburg, J. L., Manchester, J. K., Hansen, E. E., Chiang, H. C., and Gordon, J. I. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**(23), 8834-8839
3. D'Elia, J. N., and Salyers, A. A. (1996) *J. Bacteriol.* **178**(24), 7180-7186
4. Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V., and Gordon, J. I. (2003) *Science* **299**(5615), 2074-2076
5. Koropatkin, N. M., Martens, E. C., Gordon, J. I., and Smith, T. J. (2008) *Structure* **16**, 1-11
6. Salyers, A.A., Bonhoyo, G., and Shoemaker, N.B. (2000) *Methods* **20**, 35-46
7. Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D., and Gordon, J. I. (2005) *Science* **307**(5717), 1955-1959
8. Bjursell, M. K., Martens, E. C., and Gordon, J. I. (2006) *J. Biol. Chem.* **281**(47), 36269-36279

SUPPLEMENTAL FIGURE LEGENDS

Fig. S1. (A) Housekeeping (σ^{AB}) promoters upstream of ECF- σ genes. Twenty-four out of 26 ECF- σ genes that are coupled to adjacent anti- σ genes are located downstream from putative σ^{AB} promoters. Two others, *BT3993* and *BT2169* appear to be internal genes in operons for which we could not discern potential promoter sequences. The transcription start site for *BT1053* was confirmed using 5'-RACE. Asterisks in *BT1053* mark the two potential initiating bases for this promoter. The nucleotide at which transcription starts is ambiguous because a 3' poly-dC tailing strategy was used during mapping; therefore, we cannot determine if the upstream dC that is marked represents the actual 5' end of the transcript, or the result of enzymatic addition of a dC residue to a transcript end beginning with the marked dA. **(B)** Housekeeping (σ^{AB}) promoters upstream of other classes of PUL regulators. Putative σ^{AB} promoters located upstream of three major classes of inner membrane-spanning transcription factors commonly associated with *B. thetaiotaomicron* PULs: hybrid two-component systems (HTCS); SusR-like regulators; and SARP-family regulators (1-3). A subset of HTCS regulators appears to be subject to auto-regulation in response to glycans such as α -mannan (1). Thus, five representative HTCS regulators were chosen from the 36 present in the *B. thetaiotaomicron* VPI-5482 genome based on the following two

criteria: (i) they are associated with a PUL that was induced in one or more growth conditions; and (ii) they do not exhibit noticeable auto-regulation during PUL induction. The *B. thetaiotaomicron* VPI-5482 genome encodes four SusR-like proteins; one of these (BT3309) is not shown because it is an internal component of an operon. All three SARP-family regulators present in the VPI-5482 strain are presented.

Fig. S2. Signaling domains used in yeast two-hybrid analysis. **(A)** Amino acid sequence alignment of the 5 signal transducing SusC-like transporters analyzed by yeast two-hybrid assay. Major functional domains are color coded as indicated by the legend at the upper right. Sequences highlighted in red ('transducing domain') correspond to the precise peptide segments used for fusions with the GAL4 binding domain. **(B)** Amino acid sequence alignment of the 5 anti- σ factor proteins analyzed in the yeast two-hybrid assay. Major functional domains are color coded as indicated by the legend at the upper right. Sequences highlighted in green ('cytoplasmic domain') and yellow ('periplasmic domain') correspond to the precise peptide sequences fused to the GAL activating domain. Residues highlighted in blue boxes indicate the last residue retained in each anti- σ gene truncation resulting from construction of plasmid insertion mutants. An alignment of the ECF- σ signaling partners is not shown because the entire predicted coding region for each gene was used for our yeast two-hybrid analysis.

Fig. S3. Only BT1042 and not functions encoded by downstream genes are required for trans-envelope signaling. A histogram showing transcriptional activation of three genes (*BT1040*, *BT1043* and *BT1046*) from three different operons in the *BT1032-53* PUL during growth in MM-neutral *O*-glycans. Both wild-type *B. thetaiotaomicron* (yellow bars) and the Ω *BT1043* mutant (orange bars) activate gene expression in response to *O*-glycans, suggesting that the signal transducing SusC-like transporter BT1042 is the only protein encoded by the *BT1042-46* operon that is required for trans-envelope signaling (note that *BT1043* is not expressed in the Ω *BT1043* mutant due to disruption). Data are based on qRT-PCR-based quantification of the expression levels of the three genes indicated. Bars indicate fold-changes in expression during growth in MM-neutral *O*-glycans relative to each strain grown in MM-glucose (n=3 biological replicates per strain in each growth condition; error bars represent one standard deviation).

Fig. S4. Global effects of mutations in the signaling components of a single trans-envelope system. **(A)** Venn diagram showing the distribution of genes with altered expression in response to growth in MM-neutral *O*-glycans for wild-type, Δ *BT1053* and Ω *BT1042* strains (fold-changes for individual genes are summarized by group in Table S5). Note that most genes with altered expression are up regulated and reside in Group 5; these genes are shared among all data sets and indicate that loss of either trans-envelope signaling factor (ECF- σ^{BT1053} or SusC-like transporter BT1042) does not result in global changes in the transcriptional response to neutral *O*-glycans. **(B)** Heatmap illustration of upregulated genes according to the Venn groups illustrated in panel A (gene expression is normalized to the color bars at the bottom of each panel; green represents lower expression and red higher expression). Vertical columns represent the strain and growth conditions listed along the top. Data from duplicate GeneChip experiments are shown for each condition (total of 8 GeneChips from 4 conditions, including the MM-glucose reference). Several gene clusters of interest are labeled along the right of the heatmap (these genes are correspondingly labeled with the same color in Table S5). Labeled genes include: (i) 9 different PULs where the included genes exhibit an average fold-change in response to neutral *O*-glycans that is ≥ 10 -fold [blue brackets and gene labels; (1); note that only expression of the *BT1032-53* PUL (Group 1) is compromised in the Δ *BT1053* and Ω *BT1042* mutants, indicating that loss of this trans-envelope switch only effects its local PUL and not the 8 others]; (ii) a cluster of three genes with no known function (*BT3221-23*; red brackets and gene labels) that also belong to Group 1, indicating that they are induced by neutral *O*-glycans in a Δ *BT1053*/ Ω *BT1042*-dependent manner; and (iii) genes from *CPS8* (green brackets and gene labels), which exhibits increased expression in response to *O*-glycans, but only in the Δ *BT1053* mutant. **(C)** A heatmap similar to the one shown in panel B representing genes with higher expression in MM-glucose relative to MM- neutral *O*-glycans (for a list of these genes, see Table S5).

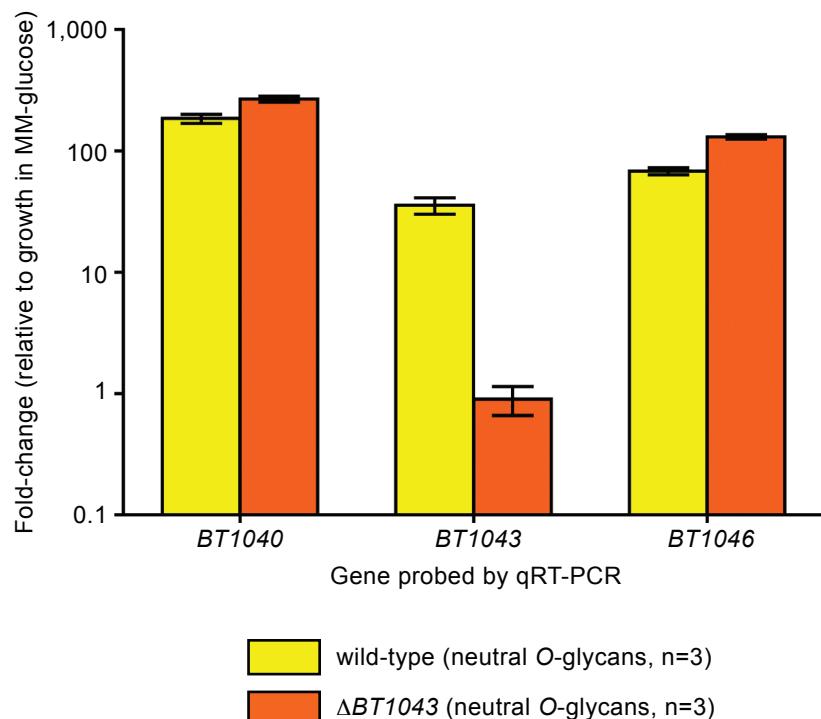
Fig. S5. Increased capsule thickness in the $\Omega BT3992$ anti- σ mutant. Representative images of India ink-stained *B. thetaiotaomicron* strains: wild-type, $\Omega BT3992$ and $\Omega BT3992/\Delta CPS2$. Bacterial cells appear as a dark shape surrounded by a bright capsular layer, which excludes the dark and particulate India ink. Magnification is 630X; Bar = 2 μm .

Fig. S6. Representative HPAEC-PAD traces. A single HPAEC-PAD trace is shown for each strain and sugar analysis condition presented in Fig. 5C of the main text (note that each experiment was performed in triplicate with similar results): (A) wild-type, neutral monosaccharides; (B) $\Omega BT3992$, neutral monosaccharides; (C) $\Omega BT3992/\Delta CPS2$ neutral monosaccharides; (D) wild-type, acidic monosaccharides; (E) $\Omega BT3992$, acidic monosaccharides; (F) $\Omega BT3992/\Delta CPS2$, acidic monosaccharides. Monosaccharide abbreviations are identical to those used in Fig. 5C of the main text: Fucose, Fuc; N-acetyl-galactosamine, GalNAc; N-acetyl-glucosamine, GlcNAc; galactose, Gal; glucose, Glc; mannose, Man; galacturonic acid, GalA; glucuronic acid, GlcA; and iduronic acid, IdoA (Note the presence of GalNAc and GlcNAc, which are de-acetylated during acid hydrolysis, is inferred from galactosamine and glucosamine standards, respectively). Sugars labeled in red text in panel E correspond to the two unknown sugars that become abundant upon *BT3992* disruption. Note the proximity of ‘unknown 4’ to the peaks corresponding to the acidic monosaccharides GalA and GlcA; this unknown sugar is flanked by these standards, but was clearly distinguishable based on retention time, in each experiment.

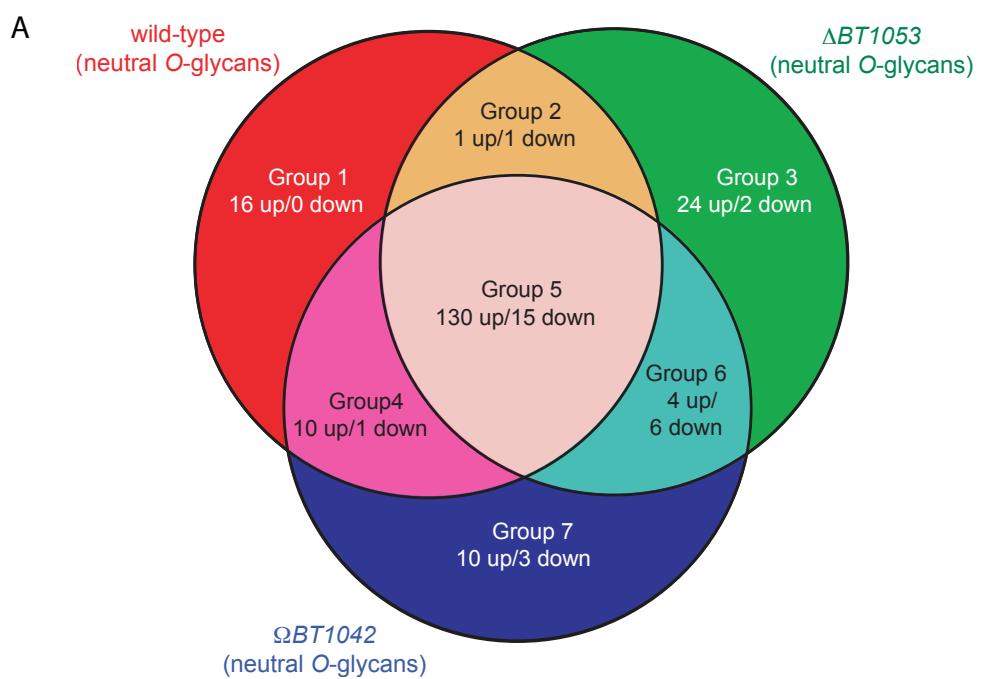
Martens *et al.* Figure S1B

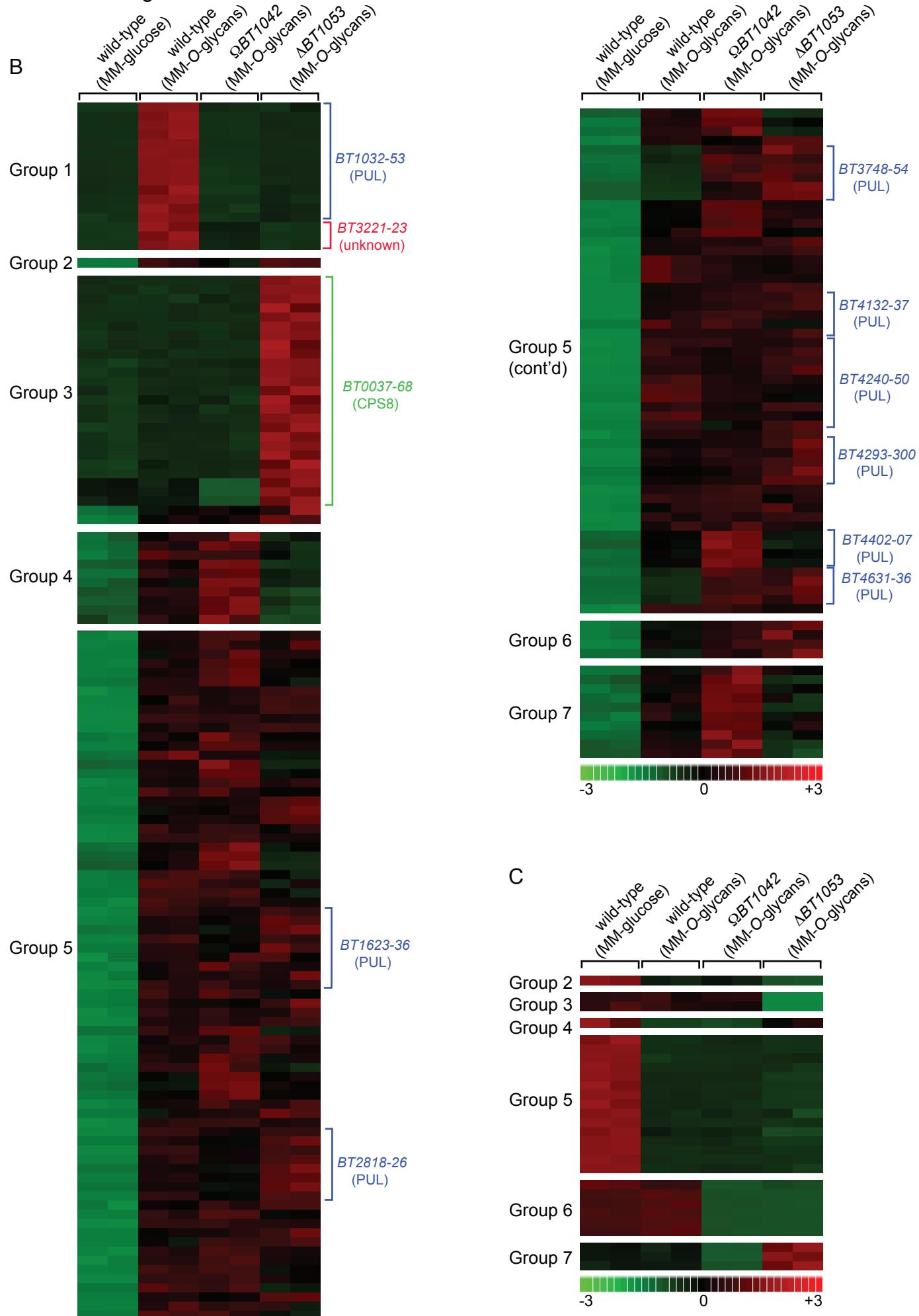
| Gene | -33 | -7 | start codon | -33/-7 spacing |
|---|--------------------------------|--|--------------------|-----------------------|
| Consensus: tttg-----ta--tttg-----atg | | | | |
| HTCS regulators: | | | | |
| BT2391: | ttcg-cctttctgttataatattgat--- | tat tttg --tgacactaaactaatatgaatatata----- | atg | 21 |
| BT2826: | ttta-aatagaatttcctataattta--- | cac tttg --tacagtattatgtggctattt----- | atg | 20 |
| BT3334: | ttcc-cgaaaatcaaataaaaacaa--- | tat tttg --ccattagtaataccctaaacaaccgaatta----- | gtg | 20 |
| BT4137: | tctg-gtaatatttctcttaaatgttag- | tat tttg --aaaactgaatcttataatctgactata----- | atg | 23 |
| BT4663: | catg-gggatttaaggaaattac--- | tat tttg --gcgtcattattaatctcgtaactaata----- | atg | 20 |
| SusR-like regulators: | | | | |
| SusR: | tttc-gtttttattactataatcat--- | tat tttg --tcacatctaaaaagcacctt----- | atg | 20 |
| BT3091: | gctg-tcggtttctgaaagacttt--- | tat tttg --cagatataatacatatggcgct----- | atg | 19 |
| BT4069: | tttg-acgaattacccaaagaaaattca- | ttt tttg --caaattgtatataacacattgat----- | atg | 23 |
| SARP-family regulators: | | | | |
| BT1770: | tatg-aatatcatattgtttcg---- | tac tttg --ccacataggaacagattcaagtagact----- | atg | 19 |
| BT2204: | ttag-gcggttatttaaaaataatcctg-- | tat tttg --tgataataatcgaaagact----- | atg | 22 |
| BT3853: | ttgg-tcttagacataagttct---- | tac tttg --aacaaggattccgggttt----- | atg | 18 |

Martens *et al.* Figure S3

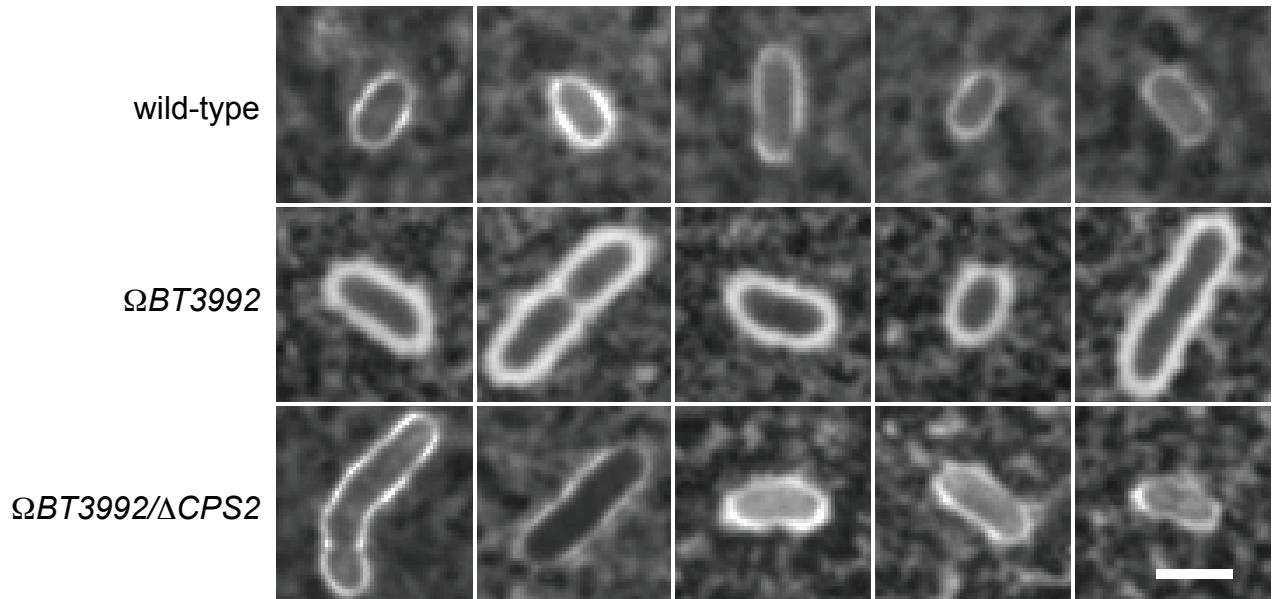


Martens *et al.* Figure S4A

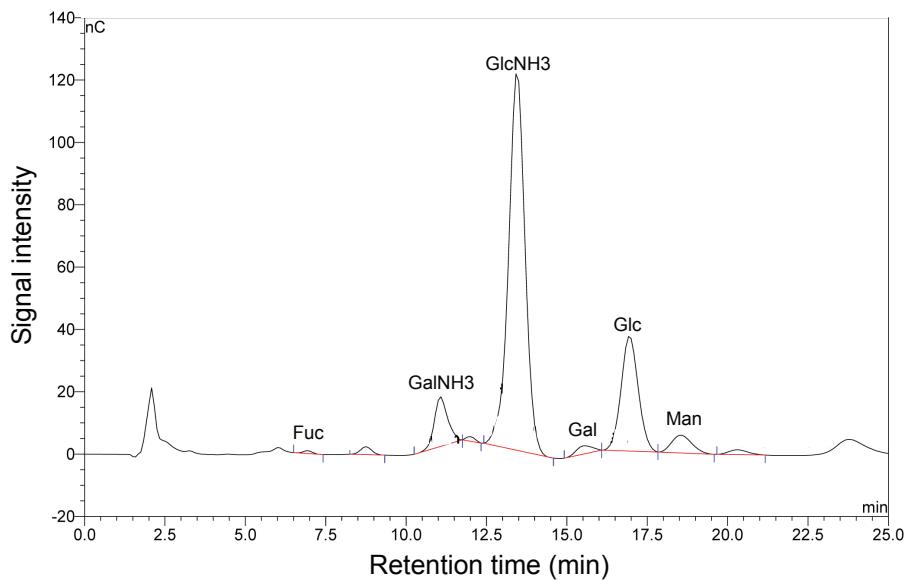




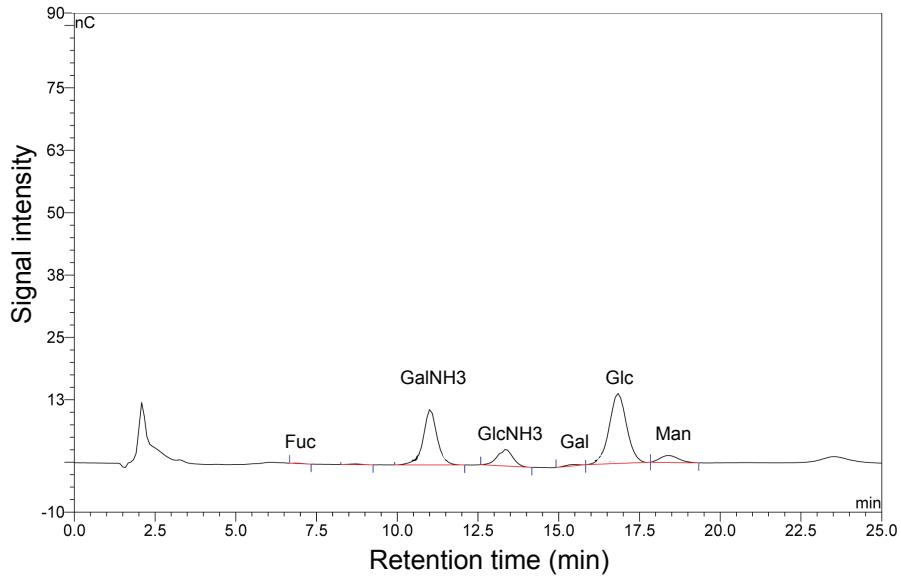
Martens et al. Figure S5



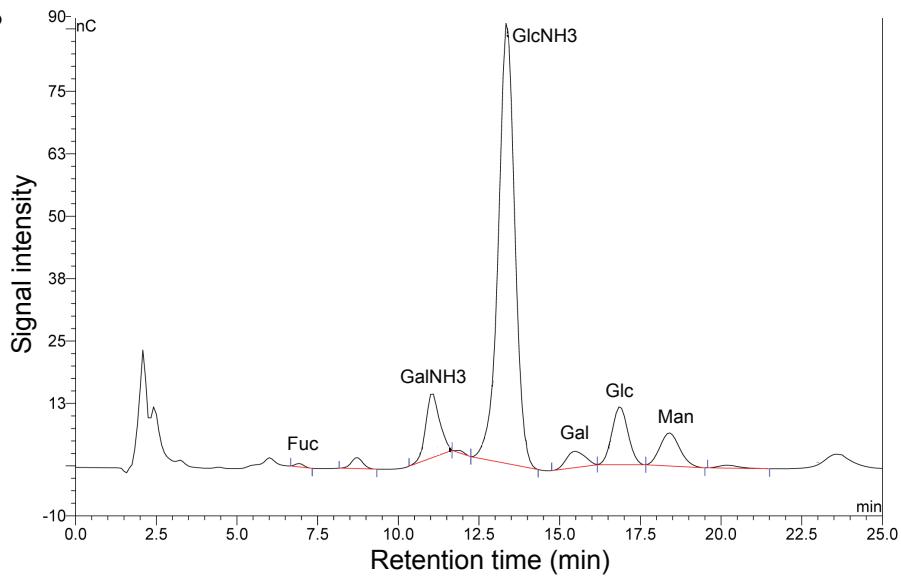
A wild-type



B $\Omega BT3992$

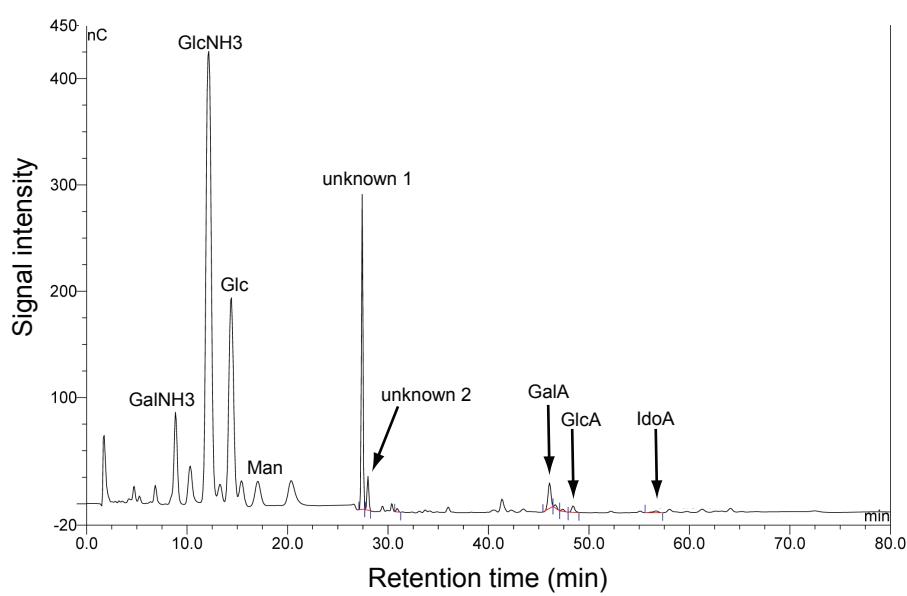


C $\Omega BT3992/\Delta CPS2$

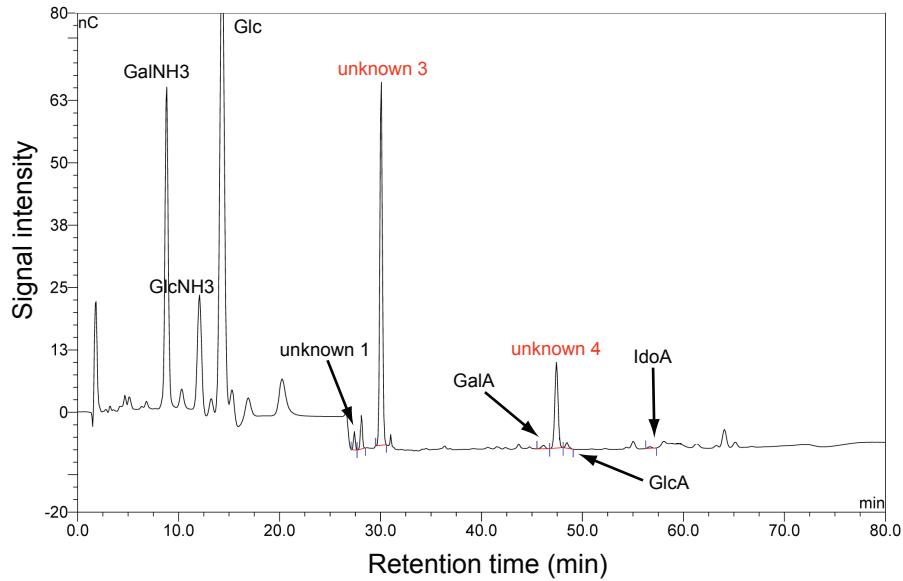


Martens et al. Figure S6D-F

D wild-type



E $\Omega BT3992$



F $\Omega BT3992/\Delta CPS2$

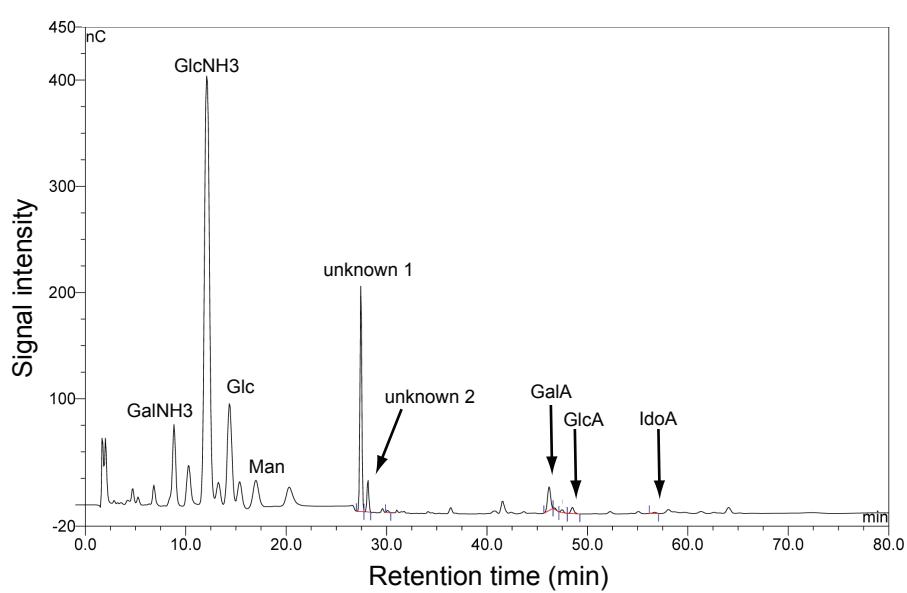


Table S1. Strains and plasmids used in this study

| Strain or plasmid | Relevant feature(s) | Use | Source or reference |
|-------------------------------------|--|---|---------------------|
| B. thetaiotaomicron strains: | | | |
| ATCC 29148 | wild-type strain; VPI-5482 | wild-type strain | (4) |
| ATCC 29148 <i>tdk</i> | Tdk ^R , FUdR ^R | "wild-type" strain for gene deletion by allelic exchange | (5) |
| ΩBT1052 | Erm ^R , lacks anti- σ BT1052 | ATCC 29148 with pGERM insertion in BT1052 | This study |
| ΩBT1052::NBU2-BT1052 | Erm ^R , Tet ^R | ΩBT1052 with complementing copy of BT1052 inserted into att NBU2 site 1 via pNBU2-bla-tetQb | This study |
| ΩBT1618 | Erm ^R , lacks anti- σ BT1618 | ATCC 29148 with pGERM insertion in BT1618 | This study |
| ΩBT3992 | Erm ^R , lacks anti- σ BT3992 | ATCC 29148 with pGERM insertion in BT3992 | This study |
| ΩBT4249 | Erm ^R , lacks anti- σ BT4249 | ATCC 29148 with pGERM insertion in BT4249 | This study |
| ΩBT4356 | Erm ^R , lacks anti- σ BT4356 | ATCC 29148 with pGERM insertion in BT4356 | This study |
| ΩBT1042 | Erm ^R , lacks SusC-like protein BT1042 | ATCC 29148 with pGERM insertion in BT1042 | This study |
| ΩBT1043 | Erm ^R , lacks SusD-like protein BT1043 | ATCC 29148 with pGERM insertion in BT1043 | This study |
| ΔBT1053 | Tdk ^R , FUdR ^R , lacks ECF- σ BT1053 | ATCC 29148 tdk with deletion of BT1053 | This study |
| ΩBT3992/ΔBT3993 | Tdk ^R , FUdR ^R , Erm ^R , lacks ECF- σ BT3993 and anti- σ BT3992 | ATCC 29148 tdk with deletion of BT3993 and pGERM insertion in BT3992 | This study |
| ΩBT3992/ΔBT3983-88 | Tdk ^R , FUdR ^R , Erm ^R , lacks Sus-like system encoded by BT3983-88 and anti- σ BT3992 | ATCC 29148 tdk with deletion of BT3983-88 and pGERM insertion in BT3992 | This study |
| ΩBT3992/ΔCPS2 | Tdk ^R , FUdR ^R , Erm ^R lacks CPS2 genes and anti- σ BT3992 | ATCC 29148 tdk with deletion of CPS2 (BT0462-82) and pGERM insertion in BT3992 | This study |
| Plasmids: | | | |
| pGERM | <i>Bacteroides</i> spp. suicide vector | construction of insertion-duplication mutants | (6) |
| pExchange- <i>tdk</i> | Derivative of pKNOCK-bla-ermGb carrying cloned <i>tdk</i> for counter-selection | construction of gene deletions | (5) |
| pNBU2-bla-tetQb | Targets cloned fragments to NBU2 att1 and/or att2 sites, Tet ^R , Amp ^R | NBU2-based complementation and tagging vector | (1) |

| | | | | | | | |
|-------------------------------|-----------|---------|------------------------|---|----------|--------------|-----------|
| GSE14592 | GSM364898 | GPL1821 | <i>QB</i> <i>T1042</i> | SusC-like transporter <i>BT1042</i> mutant grown in MM-neutral O- | Tube/5ml | mid-log/0.62 | This work |
| GSE14592 | GSM364899 | GPL1821 | <i>ΔBT1053</i> | ECF- σ <i>BT1053</i> mutant grown in MM-neutral O- glycans | Tube/5ml | mid-log/0.6 | This work |
| GSE14592 | GSM364900 | GPL1821 | <i>ΔBT1053</i> | ECF- σ <i>BT1053</i> mutant grown in MM-neutral O- glycans | Tube/5ml | mid-log/0.63 | This work |
| Published microarrays: | | | | | | | |
| GSE11980 | GSM40886 | GPL1821 | wild-type | Cecal bacterial population from adult NMRI inbred mouse fed simple sugar diet; replicate 1/3 | N/A | N/A | (7) |
| GSE11980 | GSM40887 | GPL1821 | wild-type | Cecal bacterial population from adult NMRI inbred mouse fed simple sugar diet; replicate 2/3 | N/A | N/A | (7) |
| GSE11980 | GSM40888 | GPL1821 | wild-type | Cecal bacterial population from adult NMRI inbred mouse fed simple sugar diet; replicate 3/3 | N/A | N/A | (7) |
| GSE11980 | GSM119522 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day- old NMRI inbred mice suckling on mother's milk; replicate 1/6 | N/A | N/A | (8) |
| GSE11980 | GSM119523 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day- old NMRI inbred mice suckling on mother's milk; replicate 2/6 | N/A | N/A | (8) |

| | | | | | | | |
|----------|-----------|---------|-----------|--|------------------|--|-----|
| GSE11980 | GSM119524 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day-old NMRI inbred mice suckling on mother's milk; replicate 3/6 | N/A | N/A | (8) |
| GSE11980 | GSM119525 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day-old NMRI inbred mice suckling on mother's milk; replicate 4/6 | N/A | N/A | (8) |
| GSE11980 | GSM119526 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day-old NMRI inbred mice suckling on mother's milk; replicate 5/6 | N/A | N/A | (8) |
| GSE11980 | GSM119527 | GPL1821 | wild-type | Pooled cecal bacterial populations from 17 day-old NMRI inbred mice suckling on mother's milk; replicate 6/6 | N/A | N/A | (8) |
| GSE11980 | GSM301635 | GPL1821 | wild-type | Minimal medium plus glucose 0.5% (w/v); replicate 1/3; reference for anti- α mutant | Chemostat/ 800ml | middle logarithmic growth phase/ 0.645 | (1) |
| GSE11980 | GSM301637 | GPL1821 | wild-type | Minimal medium plus glucose 0.5% (w/v); replicate 2/3; reference for anti- α mutant | Chemostat/ 800ml | middle logarithmic growth phase/ 0.62 | (1) |
| GSE11980 | GSM301639 | GPL1821 | wild-type | Minimal medium plus glucose 0.5% (w/v); replicate 3/3; reference for anti- α mutant | Chemostat/ 800ml | middle logarithmic growth phase/ 0.635 | (1) |
| GSE11980 | GSM301720 | GPL1821 | wild-type | Minimal medium plus glucose 0.5% (w/v); replicate 1/2, reference for BT1042 and BT1053 mutant profiles | Tube/5ml | middle logarithmic growth phase/0.62 | (1) |
| GSE11980 | GSM301721 | GPL1821 | wild-type | Minimal medium plus glucose 0.5% (w/v); replicate 2/2, reference for BT1042 and BT1053 mutant profiles | Tube/5ml | middle logarithmic growth phase/0.64 | (1) |

| | | | | | | | |
|----------|-----------|---------|-----------|--|----------|---|-----|
| GSE11980 | GSM301722 | GPL1821 | wild-type | Minimal medium plus 0.5% (w/v) neutral PMG glycans; replicate 1/2 | Tube/5ml | middle logarithmic growth phase/0.59 | (1) |
| GSE11980 | GSM301731 | GPL1821 | wild-type | Minimal medium plus 0.5% (w/v) neutral PMG glycans; replicate 2/2 | Tube/5ml | middle logarithmic growth phase/0.63 | (1) |
| GSE11980 | GSM302149 | GPL1821 | wild-type | Minimal medium plus 0.5% (w/v) 100mM fraction PMG glycans; earlt phase, replicate | Tube/5ml | mid 1st log phase/0.3 | (1) |
| GSE11980 | GSM302150 | GPL1821 | wild-type | Minimal medium plus 0.5% (w/v) 100mM fraction PMG glycans; early phase, replicate | Tube/5ml | mid 1st log phase/0.3 | (1) |

BT3992n YTH rev CDSIII tctagaggccgaggcgccgacatgtcatctcatccagcgatgcataatgc *BT3992* N-terminal domain
BT3992n YTH fwd SM3 aagcagtggtatcaacgcagactggccattatggccgggatggaggtaaa *BT3992* N-terminal domain
tgccctggatca

Gene deletions:

| | | |
|---|--|---|
| <i>BT3983-88</i> upstream (<i>Sall</i>) | gcggtc <u>gacgg</u> gaagatgaactgaatacggg | <i>BT3983-88</i> deletion, upstream flank |
| <i>BT3983-88</i> 5' out | cattagtttatataaaatattagaatcg | <i>BT3983-88</i> deletion, upstream flank |
| <i>BT3983-88</i> 3' out | <u>cgattctaata</u> tatttataaaaaactaa <u>atgaatt</u> taataatcaagataggat | <i>BT3983-88</i> deletion, downstream flank |
| <i>BT3983-88</i> downstream (<i>SmaI</i>) | gcccgggtattcactactttgccggac | <i>BT3983-88</i> deletion, downstream flank |
| CPS2 upstream (<i>Sall</i>) | gcggtc <u>gacag</u> ctaaaaagaactccatacag | CPS2 deletion, upstream flank |
| CPS2 5' out | cattatcacccattacccttg | CPS2 deletion, upstream flank |
| CPS2 3' out | <u>caagggttaatgggt</u> ataatgtaatcgtaatccggttctaag | CPS2 deletion, downstream flank |
| CPS2 downstream (<i>XbaI</i>) | gcgtct <u>tagact</u> tgctttaccgctcatcc | CPS2 deletion, downstream flank |

Complementation alleles:

| | | |
|--|---|-------------------------------|
| <i>BT1052</i> comp fwd (<i>XbaI</i>) | gcgtct <u>tagat</u> gttaatatcttaggtatataaggag | <i>BT1052</i> complementation |
| <i>BT1052</i> comp rev (<i>Sall</i>) | gcgg <u>tcgac</u> acatacccttttaacattctc | <i>BT1052</i> complementation |

qRT-PCR primers:

| | |
|-------------------|-------------------------------------|
| <i>BT1040</i> fwd | tcagcgttgcgtcagtctcctaa |
| <i>BT1040</i> rev | attccatcccttccaagcaacactac |
| <i>BT1043</i> fwd | gtaaacc <u>accgg</u> actcactg |
| <i>BT1043</i> rev | ttttgcgtataataatttctgtactg |
| <i>BT1046</i> fwd | ctaccggacggatacgtacgacga |
| <i>BT1046</i> rev | cagtac <u>aggcc</u> ataagccgacaga |
| <i>BT3983</i> fwd | gaaatggctgtacgcggacctat |
| <i>BT3983</i> rev | ttacggc <u>cgtcc</u> aaactgtgaag |
| <i>BT3984</i> fwd | ggggtg <u>cagacgg</u> gtgtgga |
| <i>BT3984</i> rev | cgataatgcgttcttgc <u>ttcttct</u> |
| <i>BT0463</i> fwd | ctatttcgtattgtatggctggta |
| <i>BT0463</i> rev | tctccgataataatgt <u>ctgg</u> ctaatt |
| <i>BT0482</i> fwd | attaacc <u>cgaca</u> gagaacgaaaaaga |
| <i>BT0482</i> rev | t <u>gatgg</u> ctaaattggcggagataa |

BT0206 fwd

BT0206 rev

BT4404 fwd

BT4404 rev

BT3992 fwd

BT3992 rev

gctgaaagtggcacgaatacaat

acataaagcgtgaaccggaaatagg

ggtcgctggcaagaggctaca

accgggagttccagtcattacga

gcggtcgacgcatgaaagaaatttatcatctgg

gcgggtaccgatacacattacgttcttttcgg

Table S5. Neutral O-glycan-responsive genes in wild-type, $\Delta BT1053$ and $\Omega BT1042$ mutants

Up-regulated genes:

| Gene | Description | wild-type fold-change | $\Delta BT1053$ fold-change | $\Omega BT1042$ fold-change |
|----------------|--|-----------------------|-----------------------------|-----------------------------|
| Group 1 | | | | |
| BT1035 | hypothetical protein | +8.9 | | |
| BT1036 | hypothetical protein | +37.7 | | |
| BT1037 | hypothetical protein | +54.4 | | |
| BT1038 | hypothetical protein | +62.0 | | |
| BT1039 | susD-like | +37.2 | | |
| BT1040 | susC-like | +48.0 | | |
| BT1042 | susC-like | +20.4 | | |
| BT1043 | susD-like | +17.8 | | |
| BT1044 | Glycoside Hydrolase Family 18 | +20.7 | | |
| BT1045 | hypothetical protein | +13.1 | | |
| BT1046 | susC-like | +78.3 | | |
| BT1048 | Glycoside Hydrolase Family 18 | +11.5 | | |
| BT1049 | putative patatin-like protein | +20.5 | | |
| BT3221 | hypothetical protein | +27.0 | | +5.1 |
| BT3222 | hypothetical protein | +15.9 | | |
| BT3223 | hypothetical protein | +10.4 | | |
| Group 2 | | | | |
| p5482_18 | VPI-5482 plasmid ORF_18 | | +5.0 | |
| Group 3 | | | | |
| BT0037 | putative transcriptional regulatory protein (CPS8) | | +8.3 | |
| BT0038 | hypothetical protein (CPS8) | | +7.1 | |
| BT0040 | hypothetical protein (CPS8) | | +5.8 | |
| BT0041 | F420H2:quinone oxidoreductase (CPS8) | | +6.3 | |
| BT0044 | putative glycosyltransferase (CPS8) | | +6.3 | |
| BT0048 | hypothetical protein (CPS8) | | +6.1 | |

| | | | |
|--------|--|---------|---------|
| BT3736 | hypothetical protein | +5.0 | +6.7 |
| BT3737 | hypothetical protein | +5.4 | +7.1 |
| BT3739 | Na+/dicarboxylate or sulfate symporter | +6.0 | |
| BT3749 | putative anti-sigma factor | +8.6 | +6.2 |
| BT3750 | susC-like | +132.0 | +302.5 |
| BT3751 | outer membrane protein | +15.8 | +44.6 |
| BT3752 | susD-like | +19.1 | +60.1 |
| BT3753 | endo-beta-N-acetylglucosaminidase F2 precursor (mannosyl-glyco | +12.7 | +60.8 |
| BT3754 | hypothetical protein | +13.1 | +68.6 |
| BT3787 | hypothetical protein | +5.5 | +6.3 |
| BT3789 | susD-like | | +5.4 |
| BT3791 | hypothetical protein | | +5.3 |
| BT3792 | Glycoside Hydrolase Family 76 | | +6.1 |
| BT3868 | Glycoside Hydrolase Family 20 | +7.1 | +9.4 |
| BT3990 | Glycoside Hydrolase Family 92 | | +5.7 |
| BT4038 | susD-like | +8.4 | +6.5 |
| BT4039 | susC-like | +9.3 | +7.2 |
| BT4040 | putative galactose oxidase precursor | +8.7 | +6.6 |
| BT4050 | Glycoside Hydrolase Family 2 | +9.2 | +11.9 |
| BT4132 | putative chitobiase | +100.7 | +129.5 |
| BT4133 | hypothetical protein | +395.8 | +518.5 |
| BT4134 | susD-like | +303.6 | +309.3 |
| BT4135 | susC-like | +723.8 | +415.0 |
| BT4136 | Glycoside Hydrolase Family 29 | +1136.0 | +1275.0 |
| BT4240 | conserved hypothetical protein, with a phosphotransferase enzyme | +8.5 | +7.8 |
| BT4241 | Glycoside Hydrolase Family 2 | +7.7 | +8.2 |
| BT4242 | putative transporter | +7.3 | +7.9 |
| BT4243 | putative oxidoreductase (putative secreted protein) | +7.8 | +9.2 |
| BT4244 | hypothetical protein | +49.3 | +40.3 |
| BT4245 | hypothetical protein | +36.0 | +24.6 |
| BT4246 | susD-like | +38.9 | +23.0 |
| BT4247 | susC-like | +36.7 | +40.4 |
| BT4248 | putative anti-sigma factor | +34.6 | +24.6 |
| BT4249 | hypothetical protein | +12.9 | +14.3 |
| BT4251 | hypothetical protein | +6.2 | +7.2 |
| BT4294 | hypothetical protein | +24.3 | +37.3 |

Down-regulated genes:

| Gene | Description | wild-type fold-change | $\Delta BT1053$ fold-change | $\Omega BT1042$ fold-change |
|----------------|--|-----------------------|-----------------------------|-----------------------------|
| Group 1 | | | | |
| Group 2 | | | | |
| BT1915 | pyruvate carboxylase subunit A | | -5.1 | |
| Group 3 | | | | |
| BT1053 | RNA polymerase ECF-type sigma factor | | -21.2 | |
| BT2275 | thymidine kinase | | -134.2 | |
| Group 4 | | | | |
| BT1886 | hypothetical protein | -5.9 | | -7 |
| Group 5 | | | | |
| BT0490 | hypothetical protein | -9.4 | -9.7 | -6.5 |
| BT0970 | haloacid dehalogenase-like hydrolase | -6.6 | -7.5 | -6.1 |
| BT1419 | hypothetical protein | -73.5 | -11 | -17.3 |
| BT1420 | hypothetical protein | -12.5 | -20.7 | -11.6 |
| BT1563 | hypothetical protein | -6.7 | -20 | -7.3 |
| BT1564 | hypothetical protein | -6.6 | -12.3 | -7.4 |
| BT1566 | putative aluminum resistance protein | -5.1 | -9.1 | -5.6 |
| BT1567 | hypothetical protein | -5.6 | -11 | -6.9 |
| BT1999 | anaerobic ribonucleoside-triphosphate reductase activating protein | | -5.5 | |
| BT2156 | putative sugar phosphate isomerase/epimerase | -5.4 | -6.8 | -6 |
| BT2158 | putative dehydrogenases and related proteins | | -5.4 | |
| BT2159 | putative oxidoreductase | -5.1 | -7.6 | -5.2 |
| BT2178 | hypothetical protein | -9.7 | -11.8 | -7.9 |
| BT3113 | putative transmembrane efflux protein | -11.9 | -14 | -14.7 |
| BT4499 | conserved hypothetical protein, putative membrane protein | -10.6 | -28.7 | -7.5 |

Group 6

| | | | |
|--------|----------------------|-------|-------|
| BT2675 | hypothetical protein | -5.4 | -7.5 |
| BT3240 | susC-like | -11.2 | -10.1 |
| BT3241 | susD-like | -12.7 | -9.9 |
| BT3242 | hypothetical protein | -13.6 | -12.6 |
| BT3243 | hypothetical protein | -13.1 | -11.4 |
| BT3244 | hypothetical protein | -10.2 | -9.4 |

Group 7

| | | |
|--------|----------------------|------|
| BT2987 | hypothetical protein | -8.8 |
| BT2988 | hypothetical protein | -8.5 |
| BT2989 | hypothetical protein | -9.6 |