UC Riverside
UC Riverside Previously Published Works

Title

A Novel Family of RNA-Binding Proteins Regulate Polysaccharide Metabolism in Bacteroides thetaiotaomicron.

Permalink

https://escholarship.org/uc/item/41q287k6

Journal

Journal of Bacteriology, 203(21)

Authors

Adams, Amanda
Azam, Muhammad
Costliow, Zachary
et al.

Publication Date

2021-10-12

DOI

10.1128/JB.00217-21

Peer reviewed

A Novel Family of RNA-Binding Proteins Regulate Polysaccharide Metabolism in *Bacteroides thetaiotaomicron*

Amanda N. D. Adams,[a] Muhammad S. Azam,[a*] Zachary A. Costliow,[a] Xiangqian Ma,[a] Patrick H. Degnan,[b] Carin K. Vanderpool[a]

aDepartment of Microbiology, University of Illinois Urbana-Champaign, Urbana, Illinois, USA
bDepartment of Microbiology and Plant Pathology, University of California-Riverside, Riverside, California, USA

ABSTRACT Human gut microbiome composition is constantly changing, and diet is a major driver of these changes. Gut microbial species that persist in mammalian hosts for long periods of time must possess mechanisms for sensing and adapting to nutrient shifts to avoid being outcompeted. Global regulatory mechanisms mediated by RNA-binding proteins (RBPs) that govern responses to nutrient shifts have been characterized in *Proteobacteria* and *Firmicutes* but remain undiscovered in the *Bacteroidetes*. Here, we report the identification of RBPs that are broadly distributed across the *Bacteroidetes*, with many genomes encoding multiple copies. Genes encoding these RBPs are highly expressed in many *Bacteroides* species. A purified RBP, RbpB, from *Bacteroides thetaiotaomicron* binds to single-stranded RNA *in vitro* with an affinity similar to other characterized regulatory RBPs. *B. thetaiotaomicron* mutants lacking RBPs show dramatic shifts in expression of polysaccharide utilization and capsular polysaccharide loci, suggesting that these RBPs may act as global regulators of polysaccharide metabolism. A *B. thetaiotaomicron* Δ*rbpB* mutant shows a growth defect on dietary sugars belonging to the raffinose family of oligosaccharides (RFOs). The Δ*rbpB* mutant had reduced expression of *BT1871*, encoding a predicted RFO-degrading melibiase, compared to the wild-type strain. Mutation of *BT1871* confirmed that the enzyme it encodes is essential for growth on melibiose and promotes growth on the RFOs raffinose and stachyose. Our data reveal that RbpB is required for optimal expression of *BT1871* and other polysaccharide-related genes, suggesting that we have identified an important new family of global regulatory proteins in the *Bacteroidetes*.

IMPORTANCE The human colon houses hundreds of bacterial species, including many belonging to the genus *Bacteroides*, that aid in breaking down our food to keep us healthy. *Bacteroides* have many genes responsible for breaking down different dietary carbohydrates, and complex regulatory mechanisms ensure that specific genes are only expressed when the right carbohydrates are available. In this study, we discovered that *Bacteroides* use a family of RNA-binding proteins as global regulators to coordinate expression of carbohydrate utilization genes. The ability to turn different carbohydrate utilization genes on and off in response to changing nutrient conditions is critical for *Bacteroides* to live successfully in the gut, and thus the new regulators we have identified may be important for life in the host.

KEYWORDS RRM-1, RNA-binding protein, Hfq, melibiose, PUL, CPS, capsular polysaccharide

Address correspondence to Patrick H. Degnan, patrickd@ucr.edu, or Carin K. Vanderpool, cvanderp@illinois.edu.

* Present address: Muhammad S. Azam, Department of Microbiology, University of Chicago, Chicago, Illinois, USA.

The human gut microbiome is an important player in host health, with diet being one of the principal drivers of gut microbial composition and function (1–3). Dietary carbohydrates, including complex polysaccharides and oligosaccharides, are not readily absorbed by the host and reach the distal gut where they are broken down and metabolized by a consortium of microbes with diverse enzymatic capabilities (4). Members of the dominant bacterial phylum *Bacteroidetes* can readily switch between

carbohydrate types as they become available due to dozens of substrate-specific poly-saccharide utilization loci (PULs) that encode proteins responsible for sensing and catabolizing diverse polysaccharides (5–7). Characterized PULs are tightly regulated by several distinct families of transcriptional regulators so that they are only abundantly expressed when their substrates are available (5, 8–17). However, accumulating evidence suggests that posttranscriptional regulation also plays an important part in gut colonization and preferential use of carbohydrates through control of PULs (18–20).

Posttranscriptional regulation can be mediated by multiple regulators. In *Bacteroides* species, the roles of small RNA (sRNA) regulators (19, 21) and other RNA regulatory elements like riboswitches (22–24) in control of carbohydrate and vitamin metabolism are beginning to be recognized. In well-studied *Firmicutes* and *Proteobacteria*, posttranscriptional regulation of carbon metabolism and other systems often occurs through the actions of sRNAs and their helper RNA chaperones (25–28). Three of the most well-studied RNA chaperones include Hfq, CsrA, and ProQ. Collectively, these three RNA-binding proteins (RBPs) regulate the bulk of the RNA regulatory interactome in organisms such as *Escherichia coli* and *Salmonella enterica*, and each RBP has its own distinct RNA targets (27, 29–31). Hfq in particular functions as a global posttranscriptional regulator of gene expression (32, 33). It binds to both mRNAs and sRNAs, facilitating their interactions through short stretches of complementarity (34, 35). These interactions result in a variety of different regulatory outcomes, primarily resulting from changes in translation initiation or mRNA stability (32). In many organisms, mutation of *hfq* causes global changes in gene expression and pleiotropic phenotypes (36–38).

Though regulatory RNA chaperones have not been characterized in the *Bacteroidetes*, posttranscriptional regulation has been implicated in control of gene expression in *Bacteroides* species (21, 39), including regulation of PUL expression (19). In particular, the *cis*-antisense PUL-associated sRNA DonS in *B. fragilis* and several PUL-associated *cis*-antisense sRNAs in *B. thetaiotaomicron* (19, 21) have been implicated in the modulation of PUL function through repression of carbohydrate transporter gene expression. A recent study (21) identified dozens of sRNAs encoded throughout the genome of *B. thetaiotaomicron* with many being PUL-associated and a subset having a DonS-like antisense orientation to *susC* genes. This evidence of widespread sRNA expression suggests that sRNA-mediated regulation of PUL function may be a common phenomenon in *Bacteroides* (21). In addition, there are a growing number of examples of regulatory effects in *Bacteroides* mediated by sequences in mRNA untranslated regions (UTRs) (18, 40), and these may be mediated by as-yet-unidentified sRNAs or RNA chaperones. To better understand the scope of RNA-mediated regulatory mechanisms in the *Bacteroidetes*, we sought to identify and characterize RBPs that may act as regulatory RNA chaperones.

Here, we report the identification of a family of genes commonly found in *Bacteroidetes* genomes, which encode RBPs with a single RNA recognition motif 1 (RRM-1) domain. These genes are conserved, often exist in multiple copies per genome and are highly expressed in many human gut *Bacteroides* isolates. We demonstrate that a member of this family, RbpB, is a single-stranded RNA (ssRNA)-binding protein that binds with some specificity and affinities similar to other characterized RNA chaperones. *B. thetaiotaomicron* mutants lacking one or more of these RBPs have large-scale changes to their transcriptomes compared to the wild-type strain, with genes belonging to PUL and capsular polysaccharide (CPS) loci being the most differentially regulated. *B. thetaiotaomicron rbpB* mutants have growth defects on the common dietary plant sugars raffinose family oligosaccharides due to decreased expression of *BT1871*, an essential melibiase encoded in PUL24. Our findings suggest that this family of RBPs plays an important role in global regulation of polysaccharide metabolism in *Bacteroides*.

RESULTS

Identification of a conserved family of RNA-binding proteins in the phylum *Bacteroidetes*. To identify putative RNA-binding proteins that may act as global regulators in the *Bacteroidetes*, we compared a set of 313 human gut-associated microbial genomes representing major phyla commonly found in gut microbial communities
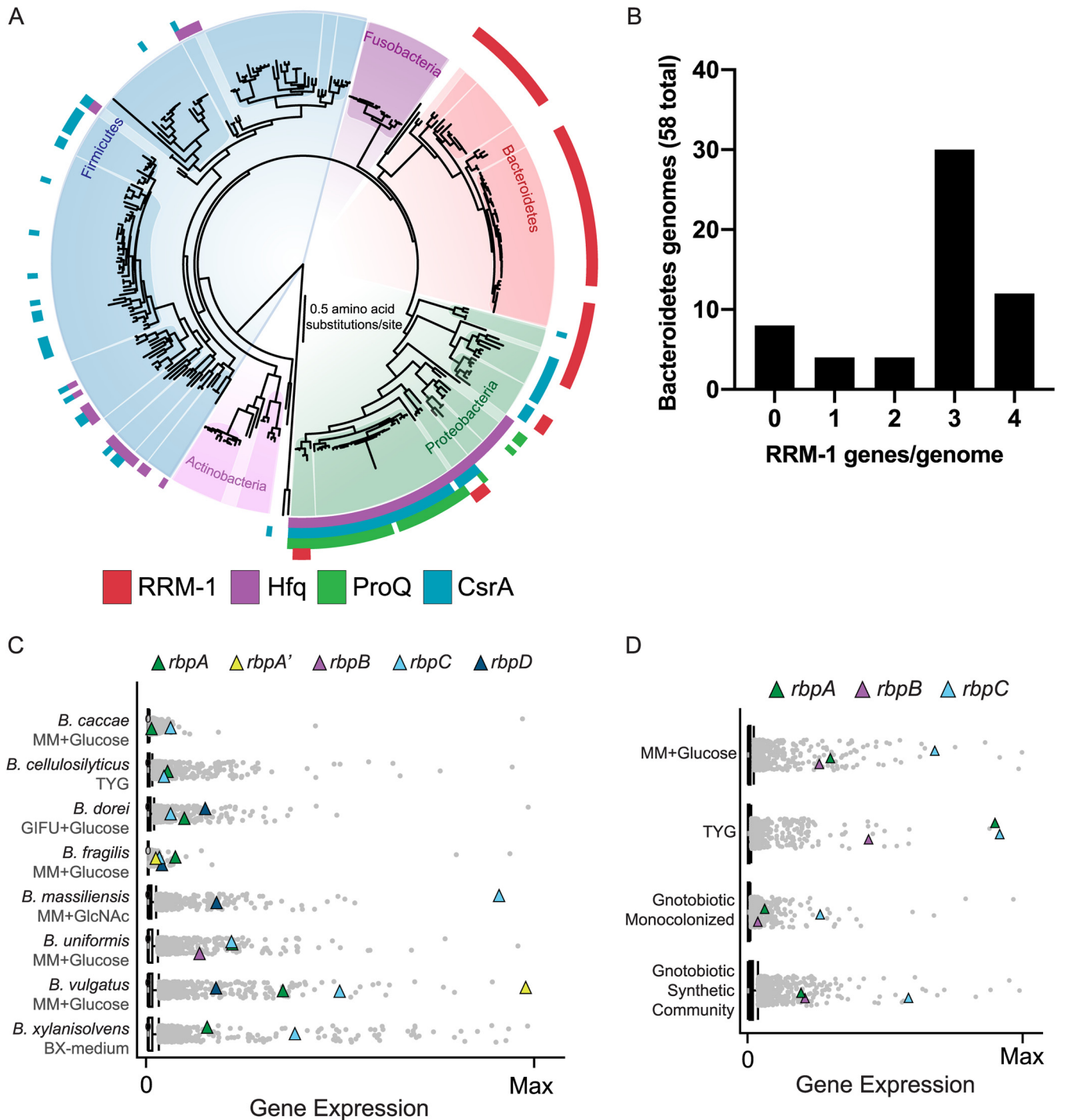
FIG 1 RRM-1 is a conserved, abundantly expressed RNA-binding domain in gut bacteria. (A) Maximum-likelihood phylogenetic species tree of 313 human gut-associated microbial genomes. Colored bars indicate the presence of at least one copy of the indicated RNA regulatory proteins in a given genome. (B) Histogram of total RRM-1 genes per genome in the 58 *Bacteroidetes* genomes represented in panel A. (C and D) RNA-seq expression plots of all genes in publicly available transcriptomes for various *Bacteroides* species (C) or *B. thetaiotaomicron* only in various growth conditions (D). Dots represent a single gene, and triangles represent *rbp* genes. The top 10% of expressed genes lie above the whiskers.

(41). We first searched for canonical RNA chaperones—Hfq, ProQ, and CsrA—which are involved in the posttranscriptional regulation of gene expression in *Proteobacteria*. Using hidden Markov models (HMMs) with trusted cutoffs, we identified Hfq in 23% (72/313) of the genomes (a total of 79 Hfq homologs) mostly in the *Proteobacteria*, although there were some identified in *Firmicutes* genomes (Fig. 1A; see also Data Set

S1). ProQ homologs (a total of 40 across 36 genomes) were entirely restricted to the *Proteobacteria*, with the majority being found in gammaproteobacterial genomes. CsrA homologs were identified in 25% (79/313) of genomes with a total of 93 CsrA homologs distributed across the *Proteobacteria* and *Firmicutes*. We did not identify any Hfq, ProQ, or CsrA homologs among *Bacteroidetes* genomes suggesting that if RNA chaperone regulators are present in this phylum, they do not belong to these canonical families.

To identify other putative RNA chaperones in our model organism, we searched the *B. thetaiotaomicron* VPI-5482 genome for proteins with conserved RNA-binding domains. This yielded three intriguing candidates comprised of a single RNA Recognition Motif 1 (RRM-1; PF00076) domain, here named RbpA (BT0784), RbpB (BT1887), and RbpC (BT3840). The small, ~70-amino-acid (aa) RRM-1 domain is one of the most common RNA-binding domains in eukaryotes, where it is typically found in multidomain proteins involved in posttranscriptional RNA processing events, including regulation of RNA stability, translation, and turnover (42). Though poorly characterized, many bacterial genomes appear to encode RRM-1 domain-containing proteins (43, 44). Given the characterized roles of RRM-1 domain-containing proteins in posttranscriptional RNA regulatory processes, we chose to focus on these homologs for further characterization.

Expanding our search for RRM-1 domain-containing proteins to our larger set of gut microbial genomes identified homologs of *B. thetaiotaomicron* RBPs in 69 of 313 genomes (Fig. 1A; see also Data Set S1 in the supplemental material). These proteins were widely distributed among *Bacteroidetes* genomes accounting for 86% (149/174) of the total number of RRM-1 domain proteins identified. We also identified homologs in a small subset of proteobacterial genomes (Fig. 1A; see also Data Set S1). In contrast to eukaryotes, the bacterial RRM-1 proteins we identified are small, single-domain proteins ranging in size from 60 to 132 aa, with the majority being 80 to 100 aa. Each protein contains a single ferredoxin-like fold RRM-1 motif, followed by predicted disordered C termini of various lengths. This structure is reminiscent of the disordered C termini of Hfq and other RNA chaperones that plays a role in RNA-binding and cycling among various binding partners (45–49). CsrA, Hfq, and ProQ homologs were largely encoded in single copy, with only a few instances of more than one copy in individual genomes (see Data Set S1). In contrast, RRM-1 genes frequently occurred in multiple copies per genome in *Bacteroidetes* genomes (Fig. 1B; see also Table S1). Of the *Bacteroidetes* genomes we analyzed, 50 of 58 contained one to four copies of genes encoding RRM-1 domain proteins, with the majority of genomes containing three (Fig. 1B). Eight strains did not contain any RRM-1 containing genes, including *Prevotella copri* DSM 18205, *Paraprevotella clara* YIT 11840, and *Bacteroides plebeius* DSM 17135 (see Data Set S1 for a full list).

To analyze phylogenetic relationships among novel RRM-1 domain proteins found in *Bacteroidetes* genomes, we used MCL (Markov cluster algorithm) with a 70% amino acid identity cutoff and compared the resulting clusters to a species phylogeny of the *Bacteroidetes* (see Data Set S1). Clustering was chosen because the extent of divergence among the homologs and their short lengths makes phylogenetic reconstruction unreliable. The clustering revealed a complicated history of divergence and duplication resulting in 10 clusters designated *rbpA to rbpJ*. It is notable that the three loci represented in *B. thetaiotaomicron*—*rbpA*, *rbpB*, and *rbpC*—represent the most widespread clusters. Even within these clusters we identified evidence of likely duplications or horizontal gene transfer among particular lineages, resulting in genomes that encode two genes belonging to a single cluster. For example, the *Bacteroides fragilis* 3_1_2 genome contains *rbpA* and *rbpA'* which share 84% amino acid identity.

The well-characterized RNA chaperone Hfq is an abundant mRNA and protein in *Proteobacteria* (50, 51). To determine whether genes encoding *Bacteroides* RBPs show similarly high levels of expression, we analyzed available RNA-seq data for *B. thetaiotaomicron* (generated by us [see Materials and Methods] and others [23]) and eight additional species—*Bacteroides caccae* (52), *Bacteroides cellulosyliticus* (53), *Bacteroides*
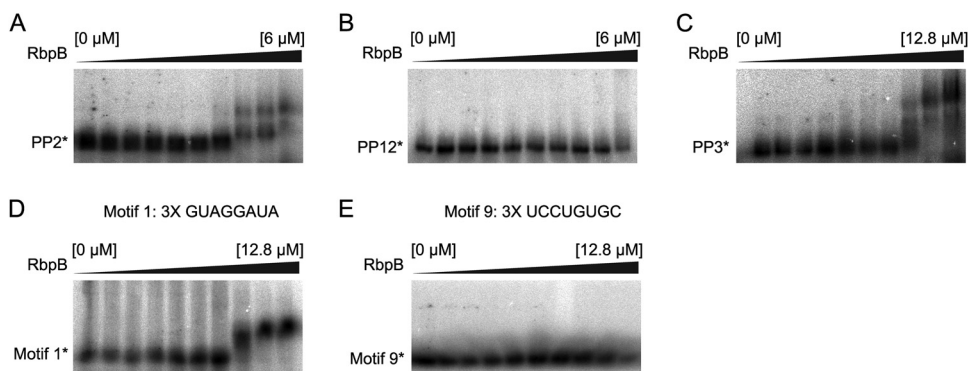
A

RbpB [0 µM]   [6 µM]

PP2*

B

RbpB [0 µM]   [6 µM]

PP12*

C

RbpB [0 µM]   [12.8 µM]

PP3*

D Motif 1: 3X GUAGGAUA

RbpB [0 µM]   [12.8 µM]

Motif 1*

E Motif 9: 3X UCCUGUGC

RbpB [0 µM]   [12.8 µM]

Motif 9*

FIG 2 RbpB is an ssRNA-binding protein. (A and B) RbpB-pentaprobe EMSAs were performed for pentaprobes 2 (A) and 12 (B) (PP2 and PP12, respectively). RbpB (µM) increases from left to right as follows: 0, 0.02, 0.05, 0.09, 0.19, 0.38, 0.75, 1.50, 3.00, and 6.00. (C) A pentaprobe 3 repeat EMSA with RbpB (µM) was performed, increasing from left to right in the gel as follows: 0, 0.05, 0.10, 0.20, 0.40, 0.80, 1.60, 3.20, 6.40, and 12.80. (D and E) EMSAs of a 3× repeat of MEME motif 1 (D) or motif 9 (E) were performed, with RbpB increasing from left to right as in panel C. *, unbound radiolabeled pentaprobe.
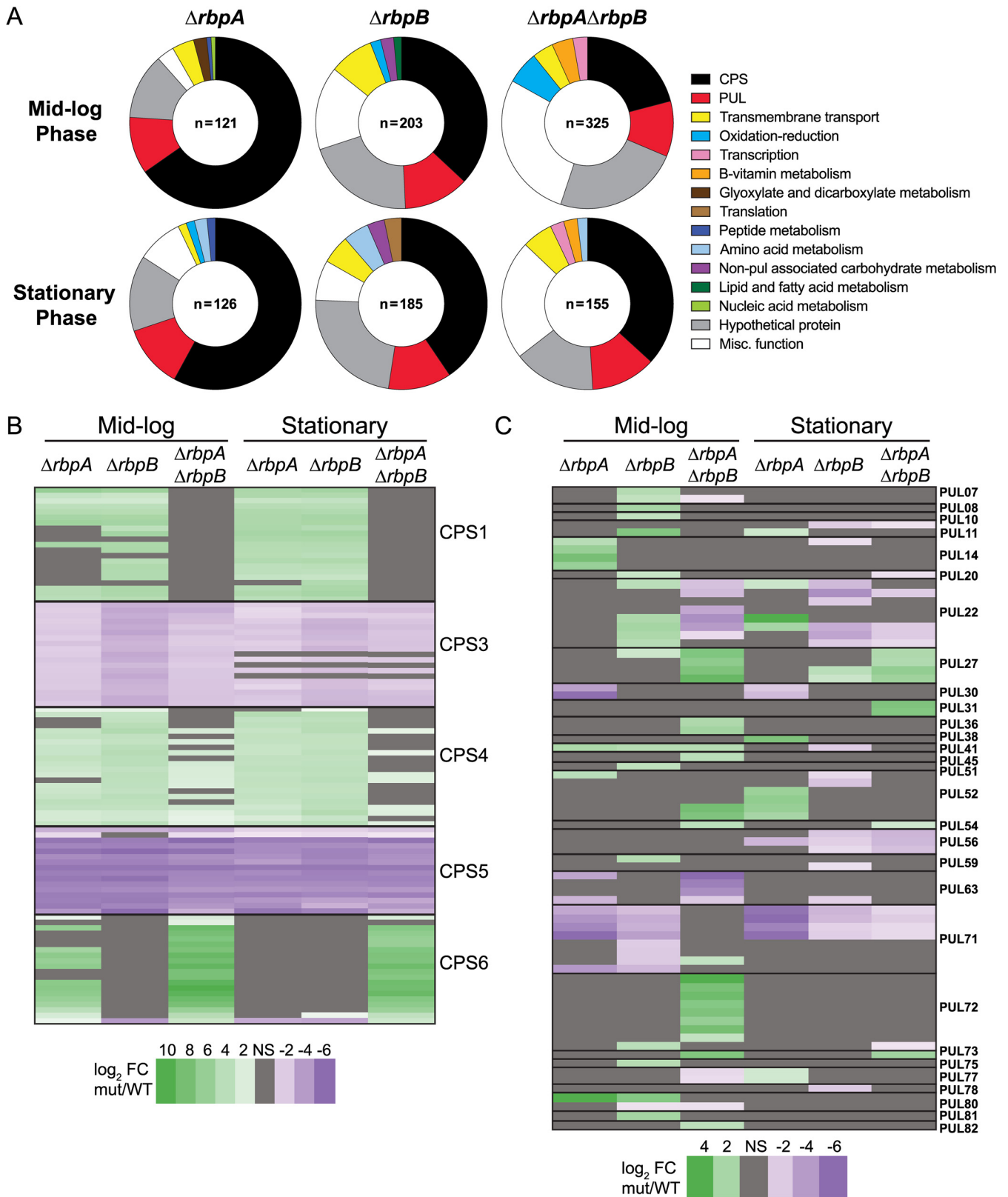
dorei (54), *Bacteroides massiliensis* (18), *Bacteroides uniformis* (24), *Bacteroides vulgatus* (24), and *Bacteroides xylanisolvens* (55) that were grown under a variety of *in vitro* conditions. Virtually all of the genes encoding RBP homologs were highly expressed in these data sets. Most *Bacteroides rbp* genes (represented by triangles in Fig. 1C and D) were expressed at levels placing them among the top 10% of most highly expressed genes (represented by dots in Fig. 1C and D). Looking specifically at *B. thetaiotaomicron rbpA*, *rbpB*, and *rbpC*, we observed that these genes are highly expressed both *in vitro* (minimal medium with glucose and TYG medium) and *in vivo* in monocolonized mice (56) or mice colonized with a synthetic consortium (52) (Fig. 1D).

B. thetaiotaomicron RbpB is a single-stranded RNA-binding protein. To test the RNA-binding activity of a representative of this family of *Bacteroides* RBPs, we conducted electrophoretic mobility shift assays (EMSAs). We overexpressed and purified *B. thetaiotaomicron* RbpB and tested binding to a series of *in vitro* transcribed single-stranded RNA (ssRNA) "pentaprobes" (57). The 12 ssRNA pentaprobes are each 100 nucleotides (nt) in length and collectively contain all possible 5-nt sequence combinations (see Table S1). RbpB shifted 10 of the 12 probes to various degrees, indicating that RbpB binds ssRNA *in vitro* and suggesting that it does so with some degree of sequence specificity (Fig. 2A; see also Fig. S1). RbpB showed no evidence of binding to two of the pentaprobes PP6 and PP12 (see Fig. S1A and 2B). Probes PP2 and PP3 shifted at lower concentrations of RbpB compared to other probes (Fig. 2A; see also Fig. S1A). The calculated K_d of RbpB binding to PP3 is 10.5 µM (Fig. 2C), a dissociation constant similar to those previously reported for RRM domains (42, 58). RRM domains can interact with a variable number of nt in the binding pocket, with binding motifs that are typically 5 to 8 nt in length (59). To identify candidate RbpB binding motifs in the pentaprobes, we used MEME motif discovery tool (60) to identify sequence motifs (<9 nt in length) that occurred in RbpB-binding pentaprobes but were absent in nonbinding pentaprobes (see Fig. S1B). MEME identified 10 such motifs in RbpB-binding pentaprobe sequences. Motif 1—comprised of G, U, and A residues—was the most common motif found exclusively in the bound pentaprobes (Fig. 2D; see also Fig. S1B). A C/U-rich motif, motif 9 (Fig. 2E; see also Fig. S1B), was present in probes PP2 and PP3, which bound RbpB with higher affinity than other pentaprobes (Fig. 2A and C; see also Fig. S1A). To test whether RbpB would bind specifically to motif 1 (5′-GUAGGAUA-3′) or motif 9 (5′-UCCUGUGC-3′), we conducted EMSAs using new RNA oligonucleotide probes containing three repeats of each motif. RbpB shifted the probe containing three copies of motif 1 with a K_d of 5.1 µM (Fig. 2D). In contrast, RbpB did not shift a probe containing three copies of motif 9 (Fig. 2E). Overall, these results demonstrate that RbpB binds ssRNA

with some degree of specificity at affinities comparable to known RNA-binding proteins (61–63).

Loss of RBPs leads to altered expression of PUL and CPS loci. To assess possible functions of RBPs in *B. thetaiotaomicron*, we made mutant strains lacking *rbpA* and *rbpB*. We generated three strains (Δ*rbpA* [*BT0784*], Δ*rbpB* [*BT1887*], and Δ*rbpA* Δ*rbpB*). We were unable to generate a Δ*rbpC* (BT3840) mutant. We performed RNA-seq on RNA samples from Δ*rbpA*, Δ*rbpB*, and Δ*rbpA* Δ*rbpB* strains grown to mid-log phase or stationary phase in rich tryptone-yeast extract-glucose medium (TYG). Among protein coding genes, 12.3% (587/4,778) were significantly differentially regulated ($q < 0.06$, $\log_2$ fold change [$\log_2$FC] of $\geq +1$ or ≤ -1) in at least one condition (see Data Set S2A), with the Δ*rbpA* Δ*rbpB* mutant having the greatest number of differentially regulated genes among the three mutants. To identify functional classes of differentially expressed genes, gene set enrichment analysis (GSEA) (64) was used with *B. thetaiotaomicron*-specific custom gene sets (see Materials and Methods). Differentially regulated genes that were not categorized in GSEA were further grouped according to gene ontology (see Data Set S2A and Materials and Methods). Considering all differentially regulated genes across all strains, enriched functional groups included CPS loci, PULs, hypothetical proteins, transmembrane transport, redox activities, B-vitamin metabolism, transcription, translation, and a variety of other metabolic pathways (Fig. 3A). The largest functional group of differentially regulated genes was CPS genes, accounting for 17% (98/587) of differentially regulated genes across all six wild-type to mutant comparisons (Fig. 3A and B). CPS loci encode functions that produce the polysaccharide coats that surround the *Bacteroides* cell surface (10, 17, 65). Of the eight CPS loci in *B. thetaiotaomicron* VPI-5482, five were differentially regulated across the three mutant strains, including CPS1, CPS3, CPS4, CPS5, and CPS6 loci (Fig. 3B). CPS3 and -5 loci were downregulated in all three mutants compared to wild type in both conditions, whereas CPS1, -4, and -6 loci were upregulated in some mutants compared to the wild type in a subset of conditions. CPS1 was upregulated in both Δ*rbpA* and Δ*rbpB* mutants in mid-log phase and stationary phase but was unchanged in the Δ*rbpA* Δ*rbpB* mutant in either growth condition. Expression patterns for all mutants were similar between mid-log-phase and stationary-phase conditions, except for the CPS6 locus. CPS6 was upregulated in Δ*rbpA* and Δ*rbpA* Δ*rbpB* mutants compared to wild-type in mid-log-phase cells, but only the Δ*rbpA* Δ*rbpB* mutant showed a difference from the wild type in stationary phase.

PULs were the second most abundantly represented functional group among differentially regulated genes (Fig. 3A). Of the 88 annotated PULs (5), 29 (33%) had at least one differentially regulated gene in *rbp* mutant strains compared to the wild type, accounting for 29% (75/263) of genes across the 29 PULs (Fig. 3C). In contrast to CPS expression, PUL expression differences in mutant strains frequently varied according to growth phase. PUL56 was downregulated in all three mutant strains exclusively during the stationary phase. In contrast, PUL71 was downregulated in all three mutant strains during stationary phase but in mid-log phase was only downregulated in Δ*rbpA* and Δ*rbpB* single mutants (Fig. 3C). Similar to the CPS loci, several PULs demonstrated expression patterns indicative of interactions between *rbpA* and *rbpB*, including PUL22. PUL22 was upregulated in the Δ*rbpB* mutant but downregulated in Δ*rbpA* Δ*rbpB* double mutant during mid-log-phase growth. In contrast, in the stationary phase, PUL22 was upregulated in Δ*rbpA* mutant and downregulated in the Δ*rbpB* and Δ*rbpA* Δ*rbpB* mutant strains. Several PULs, including PUL08, -10, -14, -51, -59, -75, -80, and -81, were differentially regulated in specific single mutant strains but not differentially regulated in the Δ*rbpA* Δ*rbpB* double mutant. We also saw some expression patterns that may be indicative of redundant regulation by RbpA and RbpB; PUL36, -45, -54, -72, -73, and -82 were not differentially expressed in the single deletion mutants but were differentially expressed in the Δ*rbpA* Δ*rbpB* double mutant. Collectively, these results suggest that RbpA and RbpB play global roles in *B. thetaiotaomicron* gene expression and in

FIG 3 Loss of RBPs leads to altered expression of PULs and CPS loci. (A) Functional categories enriched in differentially regulated genes in rich-medium RNA-seq (genes with a log$_2$FC ≥ +1 or ≤ −1 with a *q* value of <0.06). (B and C) Differentially regulated genes for PULs and CPS loci. Only genes that were significantly differentially regulated are shown. Gene names are listed in Data Set S2A.

particular suggest that they coordinate capsular polysaccharide production and carbohydrate utilization through control of CPS and PUL genes, respectively.

The ΔrbpB mutant is defective for growth on raffinose family oligosaccharides. To determine whether *B. thetaiotaomicron* RBPs are required for growth on specific carbohydrates, we carried out an initial screen of ΔrbpA, ΔrbpB, and ΔrbpA ΔrbpB strains for growth defects on Biolog plates containing a variety of carbon sources (see Fig. S2A and Data Set S3). All three strains were defective for utilization of a number of dietary and host-associated carbohydrates (see Fig. S1A and Data Set S3). As observed for the transcriptome, phenotypes for the double mutant ΔrbpA ΔrbpB strain did not recapitulate all growth defects observed in single mutant ΔrbpA or ΔrbpB strains, implying a genetic interaction between *rbpA* and *rbpB*. For example, the ΔrbpA strain showed faster growth than the wild-type strain on maltotriose, α-methyl-D-galactoside, α-D-lactose, and lactulose while showing slower growth on sucrose, D-trehalose, turanose, D-mannose, and palatinose compared to the wild type. Defects on turanose, D-trehalose, and palatinose were recapitulated in the ΔrbpA ΔrbpB strain, but the other growth changes seen in the ΔrbpA strain were not observed in the ΔrbpA ΔrbpB strain. The ΔrbpB strain grew slower than did the wild type on D-melibiose, β-methyl-D-galactoside, palatinose, and mannan, and defects on β-methyl-D-galactoside, palatinose, and mannan were also observed for the ΔrbpA ΔrbpB strain. Interestingly, all three strains were defective for growth on palatinose and a methylated galactoside. Unique to ΔrbpA ΔrbpB strain was slow growth on gentiobiose and N-acetyl-D-galactosamine. Overall, these results are consistent with transcriptome results that suggest that both *rbpA* and *rbpB* play a role in regulation of carbohydrate utilization.

One of the carbohydrates on which the ΔrbpB mutant alone had substantial growth defects was D-melibiose, a subunit of the raffinose family oligosaccharides (RFOs) (see Fig. S2A). Since RFOs are prevalent in the human diet and are available to organisms that can metabolize them in the distal gut, we chose this phenotype for further evaluation. RFOs consist of the disaccharide sucrose [glucose-(α-1,2)-fructose] bound to repeating α-1,6-galactosyl residues, producing the trisaccharide raffinose and the tetrasaccharide stachyose (Fig. 4A), along with the pentasaccharide verbascose. In addition to sucrose, the galactose-(α-1,6)-glucose disaccharide melibiose is an RFO subunit. When grown in minimal medium with RFOs or their subunits as the sole carbon source, the ΔrbpB strain displayed growth defects on melibiose, raffinose, and stachyose (Fig. 4B). The doubling time of the wild-type strain on minimal medium with melibiose as the sole carbon source was 2.56 h compared to approximately twice that for the ΔrbpB strain (5.12 h) (see Fig. S2B). The differences in growth on raffinose and stachyose were less pronounced: the ΔrbpB mutant's doubling time was ~1.4-fold greater than that of the wild-type strain. The ΔrbpA and ΔrbpA ΔrbpB strains did not have growth defects on these substrates, again consistent with possible interactions between RbpA and RbpB with respect to growth on RFOs. ΔrbpA, ΔrbpB, and ΔrbpA ΔrbpB strains showed no growth defects on monosaccharide subunits of RFOs, including glucose, galactose, and fructose, or on the disaccharide sucrose (Fig. 4B and C), suggesting that the ΔrbpB growth defect is due to the inability of this strain to utilize sugars containing the galactose-α-1,6-glucose linkage.

Complementation of the ΔrbpB strain was attempted with two different constructs (see Fig. S3A). Neither complementation construct restored growth of the ΔrbpB mutant on RFOs (see Fig. S3B). We measured *rbpB* mRNA levels from wild-type (*rbpB*⁺), ΔrbpB, and both complementation strains and found that levels of *rbpB* mRNA in complementation strains were significantly lower than in the wild-type strain (see Fig. S3C), which may account for the inability to restore growth on melibiose.

Upon further inspection of the *rbpB* (*BT1887*) native locus in our TYG RNA-seq data, we noticed reduced expression of the immediately adjacent genes *BT1886*, *BT1885*, and *BT1884* in the ΔrbpB strain, especially in mid-log phase, suggesting a possible polar effect of the *rbpB* mutation on *BT1886-BT1884* (see Fig. S4A). We also observed a single transcription start site upstream of *rbpB-BT1884* in TYG (21), a terminator prediction
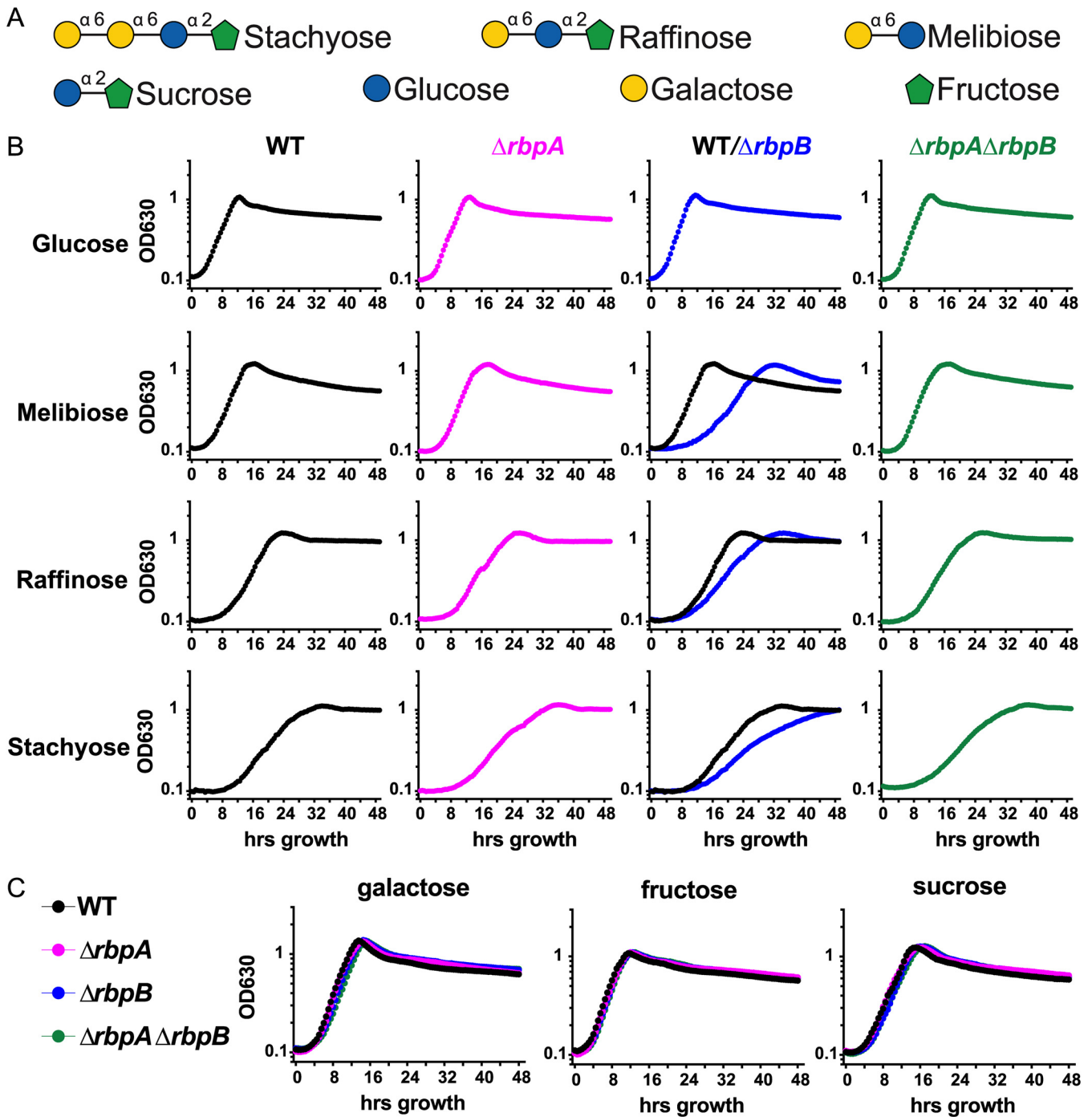
FIG 4 Δ*rbpB* is defective for growth on raffinose family oligosaccharides. (A) Raffinose family oligosaccharides and their subunits. (B and C) Representative growth curves in minimal media from a single biological replicate ($n = 3$) for wild-type (WT) and *rbp* mutants with optical densities (OD_{630}) recorded every 30 min. WT growth curves are repeated in the Δ*rbpB* graphs for melibiose, raffinose, and stachyose for comparison with the Δ*rbpB* mutant.

after the *rbpB* open reading frame (ORF), and a terminator prediction after *BT1884* (21), suggesting *rbpB* may be expressed as both a monocistronic mRNA and polycistronic with *BT1886-BT1884* (see Fig. S4B). To determine whether we could detect *rbpB* cotranscription with *BT1886-BT1884* under conditions relevant to the *rbpB* mutant phenotype, we conducted RT-PCR on RNA samples harvested from wild-type (*rbpB*⁺) cells grown to mid-log phase on minimal media with glucose or melibiose. Primer sets spanning junctions between each gene in the putative operon yielded PCR products (see Fig. S4C), suggesting that *rbpB* and *BT1886-BT1884* are expressed as an operon. *BT1886*,

BT1885, and *BT1884* encode a putative RhlE DEAD box RNA helicase, a hypothetical protein, and a cold shock domain-containing protein, respectively. The operon structure suggests the functions of these proteins are linked. One other possibility that may explain the inability to complement the *rbpB* mutant melibiose growth phenotype is that the appropriate stoichiometry of these proteins was not restored by the complementation constructs.

Loss of RbpB leads to decreased expression of an essential melibiase in PUL24. Given the inability of Δ*rbpB* to utilize α-1,6-linked RFOs, we hypothesized that a gene (or genes) encoding an α-galactosidase would be differentially regulated in the Δ*rbpB* mutant strain compared to the wild type. Although there are several α-galactosidases annotated in the genome (7), none of them were significantly differentially regulated in our TYG RNA-seq, suggesting differential regulation may be specific to growth in minimal medium with melibiose. We therefore performed more RNA-seq to identify candidate genes responsible for this phenotype. We compared transcriptome profiles of wild-type and Δ*rbpB* strains grown in minimal media with glucose or melibiose. To identify genes that are uniquely transcriptionally responsive to the α-1,6 linkage in melibiose, we also compared the glucose- and melibiose-grown cell transcriptomes to those of cells grown in minimal medium with a 1:1 mixture of glucose and galactose, the monosaccharides that make up melibiose.

Comparing wild-type and Δ*rbpB* transcriptomes in all three media, we identified genes in PUL24 that were strongly differentially regulated (Fig. 5A; see also Data Set S2B). PUL24 (genes *BT1871* to *BT1878*) contains a SusC/ᴅ-like pair (BT1874 and BT1875), a σ/anti-σ factor pair (BT1876 and BT1877), and four putative glycosyl hydrolases belonging to families GH3 (BT1872), GH43 (BT1873), GH76 (BT1878), and GH97 (BT1871). *BT1871* and *BT1872* were highly expressed in the wild-type strain growing in glucose and melibiose (Fig. 5A) (and the glucose-galactose mixture [see Data Set S2B]) but were expressed at barely detectable levels in the Δ*rbpB* strain. In contrast, *BT1871* and *BT1872* were not differentially expressed between wild-type and Δ*rbpB* strains grown in TYG (see Data Set S2A and Fig. 5A). Genes *BT1873* to *BT1878* were expressed at very low levels in wild-type and Δ*rbpB* strains in all of the conditions we tested (Fig. 5A; see also Data Set S2B), suggesting that despite being located in the same PUL as highly expressed *BT1871* and *BT1872*, these genes are not involved in glucose, galactose, or melibiose metabolism.

Previous work showed that BT1871 has *in vitro* melibiase activity (66), and transposon insertions in *BT1871* led to decreased fitness in melibiose in a carbohydrate utilization screen (67). To confirm that BT1871 was important for *B. thetaiotaomicron* utilization of RFOs, including melibiose as a sole carbon source, we deleted *BT1871* and cultured the Δ*BT1871* and wild-type strains in minimal media with melibiose, raffinose, stachyose, and sucrose (Fig. 5B and C). The Δ*BT1871* strain showed no growth defect on sucrose compared to the wild-type strain, which is expected based on its predicted melibiase activity. In contrast, the Δ*BT1871* mutant could not grow on melibiose, indicating *BT1871* is essential for melibiose utilization. In addition, the Δ*BT1871* mutant showed reduced growth on raffinose and stachyose compared to wild-type (Fig. 5B and C), indicating that *BT1871* is required for metabolism of RFOs in general. Residual growth of the Δ*BT1871* mutant on raffinose and stachyose is presumably due to the ability to utilize fructose from the α-1,2 sucrose linkage.

We constructed three different complementation strains to confirm that BT1871 is responsible for the melibiose growth defect. There is a single predicted promoter upstream of *BT1872* (21) and the *BT1872* and *BT1871* open reading frames are separated by only 32 bp, suggesting that they are coexpressed. Complementation strain 1 (compl1 [Fig. 5D]) carried the native promoter upstream of *BT1872*, followed by a deletion of the *BT1872* ORF and the intact *BT1871* gene. Complementation strain 2 (compl2 [Fig. 5D]) carried the intact promoter and *BT1872* and *BT1871* genes. Complementation strain 3 (compl3 [Fig. 5D]) carried the promoter and *BT1872* only. The compl1 construct partially restored growth on melibiose, raffinose, and stachyose (Fig. 5E). The compl2
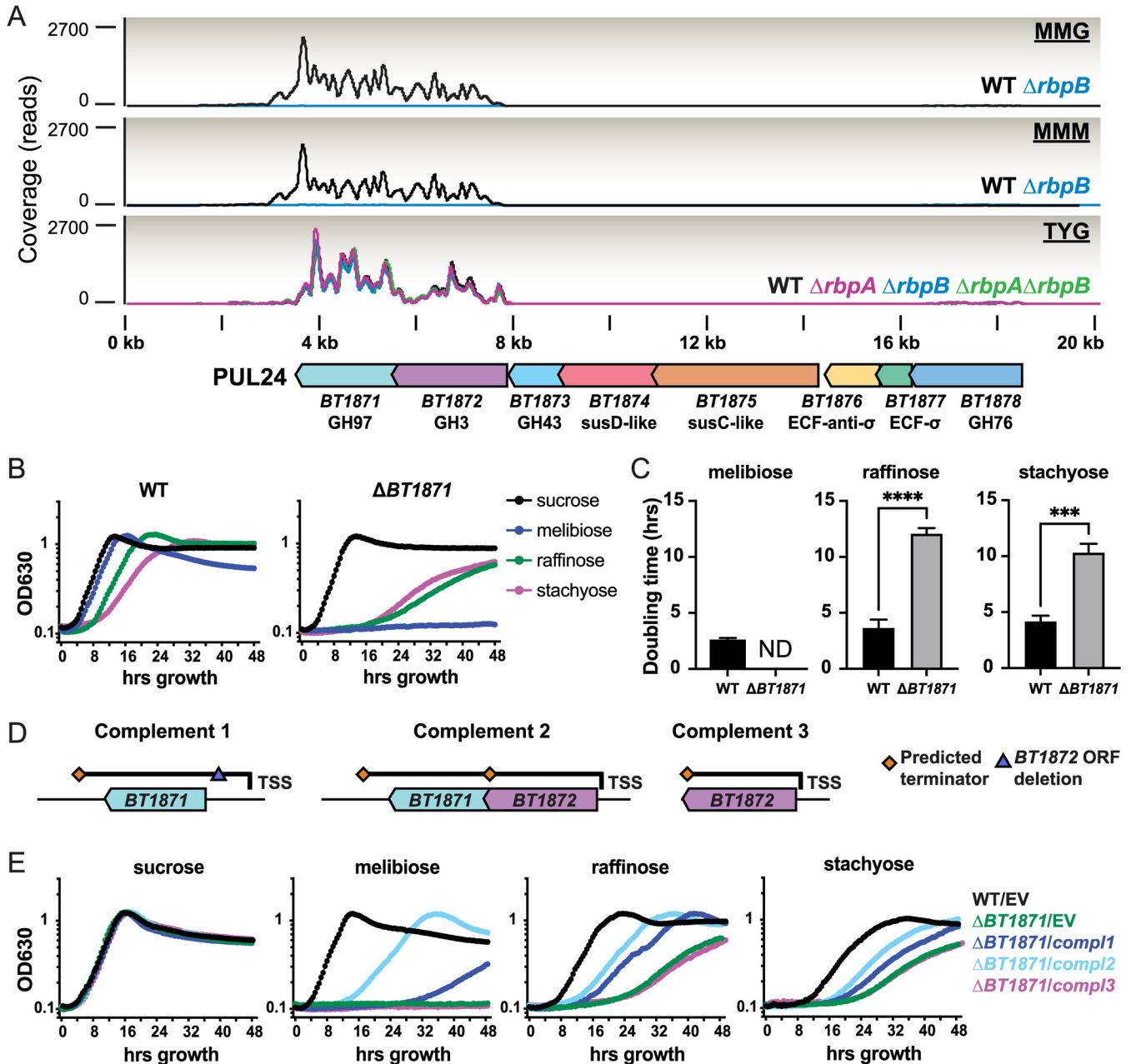
FIG 5 Loss of *rbpB* leads to loss of expression of an essential melibiase in PUL24. (A) Normalized expression coverage curves for mid-log-phase cultures in PUL24 with putative gene annotations. (B and E) Representative growth curves in minimal media from a single biological replicate ($n = 3$) with optical densities (OD_{630}) recorded every 30 min. (C) Average doubling times shown with standard deviations ($n = 3$) of WT and $\Delta BT1871$ strains grown on melibiose, raffinose, and stachyose (nonparametric t test significance: ***, $P < 0.001$; ****, $P < 0.0001$). ND, not determined. (D) Genomic regions inserted into pNBU2 vectors for complementation of *BT1871* and *BT1872* shown in panel E. (E) Empty vector (EV) controls contain integrated pNBU2 without an insert.

construct improved growth on RFOs compared to the compl1 construct, whereas the compl3 construct (*BT1872* alone) failed to complement (Fig. 5E). Taken together, these results indicate that *BT1871* is an essential melibiase required for RFO utilization and that the decrease in *BT1871* mRNA in the $\Delta rbpB$ strain is responsible for the melibiose growth defect.

DISCUSSION

Though it is well established that rapid nutrient shifts affect the composition and metabolic activities of gut microbes (52, 68–70), the regulatory mechanisms that allow
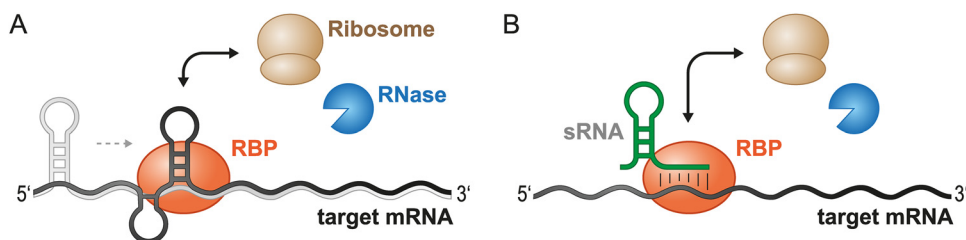
FIG 6 Model for possible mechanisms of RBP-mediated regulation. (A) RBP binding directly to mRNAs could promote structural changes (represented by the transition from light gray to dark gray conformation) that alters access of ribosomes or RNases to change translation or mRNA stability. (B) RBPs could facilitate sRNA binding to mRNA targets and alter the access of ribosomes or RNases to change translation or mRNA stability.

them to sense and rapidly adapt to use of different nutrient sources are poorly understood. Canonical mechanisms for global transcriptional regulation of carbon source utilization in model organisms from the phyla *Proteobacteria* and *Firmicutes* (71) are absent in the *Bacteroidetes* (72–74). Likewise, RNA chaperones and RNA-mediated posttranscriptional regulatory mechanisms that coordinate metabolism and responses to changing environmental conditions (28, 32, 75–77) are commonly found in *Proteobacteria* and *Firmicutes* but have not been described in *Bacteroidetes*. In this study, we identify a family of conserved RNA-binding proteins that is broadly distributed among members of the *Bacteroidetes* and some *Proteobacteria* that lack canonical RNA chaperone regulators. These RBPs occur in multiple copies in a given genome and are highly expressed in a number of *Bacteroides* species from the human gut in culture and in mouse models. At least one of these proteins from *B. thetaiotaomicron*, RbpB, is able to bind ssRNA *in vitro* in a sequence-specific manner. Deletion of *rbpA* and *rbpB* in *B. thetaiotaomicron* leads to global alterations in expression of CPS loci and PULs and perturbed growth on a variety of carbohydrate sources. Overall, these results suggest that this family of RBPs may play global regulatory roles in carbohydrate metabolism in the *Bacteroides*.

While our study provides strong evidence for the importance of these RBPs in global regulation of gene expression, the mechanisms by which RBPs mediate these effects are still unknown. We hypothesize that like canonical RNA chaperones in *Proteobacteria*, RBPs act as RNA chaperones that control mRNA stability and translation by binding to target mRNAs and modulating ribosome association or access of RNases (Fig. 6). RBP modulation of mRNA translation or stability may be through direct interaction of RBPs with target mRNAs (Fig. 6A) or through facilitating base pairing of sRNAs to mRNAs (Fig. 6B), either of which could result in changes to mRNA structure that alter accessibility to ribosomes or RNases. Little is known about RNA chaperone function in the *Bacteroidetes*, but in the case of RbpB, its role in RNA metabolism is supported by its genomic location. Annotations for BT1885 (DEAD box RNA helicase) and BT1884 (cold shock protein) suggest that *rbpB-BT1884* may be an RNA metabolism operon. To date, we do not have evidence supporting or refuting a role for RBPs in modulation of sRNA function. However, recent literature suggests that sRNAs may play an important role in modulation of carbohydrate metabolism in *Bacteroides* (19, 21). One recent study described an *N*-acetyl-ᴅ-glucosamine-inducible sRNA called GibS that binds *in vitro* to mRNAs involved in carbohydrate metabolism. Mutant strains lacking GibS had nine differentially regulated genes compared to the wild-type parent strain. Two of these were *BT1871* and *BT1872*, where expression was reduced in the Δ*gibS* mutant compared to the wild-type strain. GibS binding to *BT1871* mRNA was predicted *in silico* but could not be demonstrated *in vitro*. The authors of that study speculated that GibS binding to *BT1871* mRNA required an unidentified RNA chaperone. To test whether RbpB facilitates RFO utilization by a GibS-dependent mechanism, we generated Δ*gibS* and Δ*rbpB* Δ*gibS* deletion mutants and grew these strains alongside the wild type in the presence of melibiose (see Fig. S5A). The Δ*gibS* mutant grew similarly to the wild type, and the Δ*rbpB* Δ*gibS* mutant

grew similarly to the Δ*rbpB* parent strain on melibiose (see Fig. S5A). RT-qPCR showed that levels of *BT1871* mRNA were also similar between wild-type and Δ*gibS* strains (see Fig. S5B), suggesting that GibS does not play a major role in modulating *BT1871* mRNA levels in our growth conditions. Overall, these results suggest that under our growth conditions, RbpB regulates *BT1871* independently of GibS.

We have yet to explore the role of RBPs in helping *B. thetaiotaomicron* colonize or be maintained in the host gut. In a recent study (67) screening transposon (Tn) mutants for a wide variety of *in vitro* and *in vivo* phenotypes, there were no reported insertions in *rbpA* or *rbpB*. Insertions in *rbpC* led to reduced fitness on glucose-containing media (67), possibly explaining our inability to generate Δ*rbpC* mutants in our standard glucose-rich media. The *rbpC* mutants also had an increased growth on melibiose, suggesting that *rbpC* also plays a role in utilization of RFOs. In the same study (67), colonization of germfree mice fed a plant polysaccharide-rich diet with the *B. thetaiotaomicron* Tn-mutant pool led to increased fitness of *rbpC* mutants. *BT1871* mutants showed decreased fitness over time, whereas Tn-insertions into several other PUL24 genes led to increased fitness *in vivo*. Overall, these data indicate that RBPs and the PULs they regulate may be important for *in vivo* fitness.

Regulation by RBPs may represent a critical mechanism for coordination of carbohydrate utilization and production of cell surface capsular polysaccharides. Differential regulation of PULs and CPS loci in *rbp* deletion strains is consistent with several reports indicating a regulatory link between these polysaccharide metabolic processes in *B. thetaiotaomicron* (17, 78, 79). Our RNA-seq data showed that deletion of RBPs leads to reduced expression of CPS3 and CPS5 and increased expression of CPS4. In *B. thetaiotaomicron*, CPS3 is normally the most highly expressed locus *in vitro*, but CPS3 dominance declines after colonization when CPS4 becomes the most highly expressed CPS locus in mouse models when dietary glycans are present (69, 79, 80). Disruption of CPS4 expression leads to decreased fitness in mouse competitions. Strains exclusively expressing CPS5 usually outcompete strains expressing other individual CPS loci (79, 81, 82). In a study monitoring CPS expression in a mouse model over time, it was observed that even when the *B. thetaiotaomicron* inoculum expressed one dominant CPS locus, expression of different loci over time varied between mice (79). While CPS4 was most often highly expressed in mice fed a high-fiber diet, mice on fiber-free diets typically expressed CPS5 or CPS6. *B. thetaiotaomicron* mutants that could only express a single CPS locus had a decreased ability to recover from antibiotic-induced stress (79). These studies, along with our present work, collectively suggest that the ability to shift among different CPS types is advantageous in the host and that this regulation may be mediated in part by RBPs. Further characterization of the RBPs and their regulatory mechanisms may provide critical insight into how *Bacteroides* coordinately control carbohydrate availability with cell surface properties. This could reveal key principles governing mechanisms in host dynamics.

MATERIALS AND METHODS

Bacterial culturing and genetic manipulation. *B. thetaiotaomicron* VPI-5482 strains were grown anaerobically in a Coy Laboratory Products vinyl anaerobic chamber with an input gas of 20% CO_2, 10% H_2, and 70% N_2 balance. Routine culturing of *B. thetaiotaomicron* was done in TYG (tryptone, yeast extract, and glucose) (83) broth and on Difco brain heart infusion (BHI) agar plates with 10% defibrinated horse blood (HB; Quad Five) at 37°C. *Escherichia coli* strains were grown aerobically at 37°C on BHI–10% HB for conjugations and in Luria broth for all other applications. Minimal medium (22) was supplemented with B_{12} (3.75 nM, final; Sigma) and carbohydrates as needed at the following final wt/vol concentrations unless otherwise indicated: 4.0% stachyose (Sigma), 2.0% D-(+)-raffinose (Sigma), 0.5% D-(+)-melibiose (Sigma), 0.5% α-D-glucose (Sigma), 0.5% D-(+)-galactose (Sigma), 0.5% β-D-(–)-fructose (MP Biomedicals), and 0.5% sucrose (MP Biomedicals). When needed, antibiotics were added at the following final concentrations: 100 μg/ml ampicillin (Sigma), 200 μg/ml gentamicin (Goldbio), 25 μg/ml erythromycin (VWR), 200 μg/ml 5′-fluoro-2′-deoxyuridine (VWR), 100 ng/ml anhydrotetracycline (Sigma), and 25 μg/ml kanamycin (Fisher). All strains, vectors, and primers are listed in Table S1 in the supplemental material. For all experiments, wild-type *B. thetaiotaomicron* is the Δ*tdk* strain (strain AA0014 in Table S1).

Markerless deletions were made in *B. thetaiotaomicron* using the pExchange_*bla_tdk_ermGb* (84) and the pLGB13_*bla_ermG* (85) suicide vector-based allelic exchange methods. Upstream and downstream regions of the gene to be deleted were amplified using Kappa HiFi (Kappa Biosystems) and

cloned into pExchange_*bla_tdk_ermGb* using standard restriction digestion and ligation methods and splicing by overlap exchange (SOE) (22). Alternatively, inserts were cloned into Q5 (NEB) amplified pExchange_*bla_tdk_ermGb* using restriction digest and ligation of a gBlock insert (IDT). GibS flanks were Q5 amplified and ligated to restriction-digested pLGB13_*bla_ermG*. Complete vectors were conjugated into *B. thetaiotaomicron* with *E. coli* S17 λ-pir using established methods (22). Complementation pNBU2_*bla_ermGb* vectors (5, 86) were cloned using standard restriction digest and ligation methods and conjugated into *B. thetaiotaomicron*, as done with pExchange. pNBU2_*bla_ermGb* vectors were PCR screened for insertion into a single attachment site as done previously (5). The pET-28a-*rbpB* protein expression vector was generated by inserting the *rbpB* (*BT1887*) ORF 5' to the thrombin cleavage site and 6×His tag in the pET28a backbone. pET-28a-*rbpB* was cloned using Q5 PCR amplification and NEBuilder assembly (NEB) in *E. coli* XL10-Gold competent cells (Agilent) before being moved into *E. coli* BL21(DE3) for protein expression.

Computational identification of RNA regulators in human gut-associated microbial genomes. To identify genomes containing CsrA, ProQ, RRM-1, and Hfq in the human gut microbiome, we utilized a custom database of 313 human gut-associated microbial genomes containing a single representative genome for a species (22, 41). Candidate RNA regulator genes were identified using hmmer v3.3 (hmmer.org) with trusted cutoffs, and the individual hidden Markov model from each protein was queried: Hfq, PF17209.4; CsrA, PF02599.17; ProQ, PF04352.14; and RRM-1, PF00076.23 (43). The resulting gene list was then run against Pfam-A.hmm version 33.1 using hmmer to verify that the query PFAM was the top hit for the target domain using trusted cutoff values. ProQ PF04352.14 gene hits that also contained an N-terminal FinO_N domain (PF12602.9) were removed from the final annotation list. RRM-1 PF00076.23 gene hit list was limited to fewer than 150 aa to remove a few genes containing transmembrane domains. The maximum-likelihood phylogenetic tree in Fig. 1A was built using a multisequence alignment of 13 conserved core genes (AspS, Ffh, FusA, GltX, InfB, LeuS, RplB, RpsE, RpsH, RpsK, TopA, TufA, and RpoB) identified and described previously (87). Briefly, protein sequences for each group of orthologs were individually aligned with MUSCLE (88), concatenated, and subjected to phylogenetic reconstruction with RAxML (89). The phylogeny was visualized using FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

RBP expression in publicly available RNA-seq data sets. Publicly available RNA-seq data sets were downloaded from NCBI (see Table S2 for sample IDs). RNA-seq reads were quality filtered with Trimmomatic v0.36 (90). Read mapping and sample normalization was calculated with Rockhopper v2.03 (91, 92), and normalized expression values were graphed using JMP v15 (93).

RbpB EMSAs and motif identification. (i) Purification of RbpB. *E. coli* BL21(DE3) cells with the pET-*rbpB* vector were grown to late exponential phase (optical density at 600 nm [OD_{600}] of 0.6 to 0.8, as measured on an Ultraspec 2100 Pro [Amersham]), and protein expression was induced with 1 mM final IPTG (isopropyl-β-D-thiogalactopyranoside; Goldbio) for 4 h at 37°C. Cells were harvested by centrifugation and pellets resuspended in 30 ml of extraction buffer (1× phosphate-buffered saline [PBS], 0.5 M NaCl [pH 7.2]) before being lysed in a French press. Supernatant was collected after centrifugation at 16,000 × *g* for 10 min at 4°C. The supernatant was then fractionated using a HiTrap Ni^{2+} column (GE Healthcare) according to the manufacturer's instructions. Fractions containing RbpB were dialyzed overnight in TGED buffer (10 mM Tris-HCl [pH 8], 5% glycerol, 0.1 mM EDTA, and dithiothreitol [DTT] at 0.015 mg/ml) and loaded onto a HiTrap-Q column (GE Healthcare). The column was washed with TGED buffer, and protein was eluted with a linear gradient of NaCl (0.1 to 1 M) in TGED buffer. The fractions containing the protein were pooled, dialyzed, and concentrated using Centricon 10 concentrators (Millipore-Sigma), mixed with an equal volume of 100% glycerol, and stored at −20°C.

(ii) Radiolabeled pentaprobe synthesis. Twelve pentaprobes containing all the possible 5-nt combinations were prepared based on a published protocol with some modifications (57). All oligonucleotides used in pentaprobe synthesis are listed in Table S1. Single-stranded oligonucleotides for PP1-PP6 were Q5 PCR amplified with a 5' T7 promoter for either the Watson strand or the Crick strand, generating 12 dsDNA templates with a single T7 site, two each for PP1 to PP6. The Watson strand of PP1 dsDNA is identical to the coding strand of the PP1 pentaprobe, and the Crick strand is identical to the coding strand of the PP7 pentaprobe. These dsDNA fragments were then used to produce 12 different ssRNA pentaprobes by *in vitro* transcription from the T7 promoter with a MEGAscript T7 kit (Ambion). Transcribed RNA fragments were 5' end labeled with [γ^{32}P]ATP (Perkin-Elmer) using the KinaseMax kit (Ambion) according to the manufacturer's protocol.

(iii) Electrophoretic mobility shift assay and motif prediction. RNA-protein gel electrophoretic mobility shift assays were performed using 0.01 pmol of ^{32}P-labeled pentaprobe RNA and the indicated amounts of RbpB in binding buffer (10 mM Tris-HCl [pH 8.0], 0.5 mM DTT, 0.5 mM $MgCl_2$, 10 mM KCl, and 5 mM Na_2HPO_4–NaH_2PO_4 [pH 8.0]). The mixture was incubated at 37°C for 30 min, and nondenaturing loading buffer (50% glycerol and 0.1% bromophenol blue) was added. The samples were resolved on a 4.6% native polyacrylamide gel for 1.5 h at 10 mA. The fraction of RbpB bound was determined using a FLA-3000 fluorescent image analyzer (Fujifilm) to quantify the band intensities. K_D values were calculated using Sigmaplot software based on a published method (94). The MEME program (60) was used to predict conserved motifs for the positive pentaprobe sequences with the following parameters: maximum number of motifs, 10; minimum motif width, 4; and maximum motif width, 8. The K_D was calculated for RbpB binding to PP3 and motif 1 using three technical replicates as done previously (61).

RNA sequencing sample prep and processing. For rich-medium RNA-seq, strains were cultured in 5 ml of TYG in biological triplicate to stationary phase overnight. Each culture was then subcultured 1:100 into two 5-ml TYG cultures. One tube was cultured to mid-log phase (0.35 to 0.6 OD_{600}), and the

second tube was cultured to early stationary phase (1.2 to 1.4 OD_{600}) as measured in a Thermo Spectronic 200 (Thermo Fisher Scientific; referred to as ThermoSpec below). Then, 500 μl of cells was then spun down at 7,500 $\times$ g for 3 min at room temperature, the supernatant was removed, and the pellets were resuspended in 600 μl of TriReagent (Sigma). RNA was then isolated from the resuspensions using the Zymo Direct-Zol RNA Mini-Prep kit (Zymo), which includes on-column DNase I treatment. RNA quality was evaluated using a Qubit 2.0 fluorometer (Invitrogen) and Agilent 2100 Bioanalyzer (UIUC Biotechnology Center). Total RNA was then submitted to the W. M. Keck Center for Comparative and Functional Genomics at UIUC for rRNA depletion, library construction, and sequencing. Briefly, ribosomal RNAs were removed from total RNA with an Illumina Ribo-Zero rRNA removal kit for bacteria. RNA-seq libraries were produced with a ScriptSeq v2 kit (Illumina) and cleaned with AMPure beads (Beckman Coulter) to remove any fragments <80 nt. Libraries were sequenced on an Illumina HiSeq2500 using HiSeq SBS sequencing kit v4 to give 160-nt single-end reads. Results were demultiplexed with bcl2fastq v2.17.1.14 (Illumina). Reads were quality filtered and trimmed using Bioconductor package ShortRead (95) to first remove reads with >1 N or, if ≥75% of a read is a single nucleotide, and then the first 2 nt were removed from each sequence read. Sequencing adapters were removed with fastx_clipper (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Residual rRNAs were removed using bowtie2 (96, 97), and final reads were mapped to the genome and analyzed using Rockhopper v2.03. All raw Rockhopper calculated expression values were increased by 1, and fold changes (FC) were calculated as the $\log_2$(mutant expression value + 1/wild-type expression value + 1). RNA-seq processing statistics are summarized in Data Set S2C.

For RNA-seq of cultures grown in minimal media, a single colony/strain was smeared onto half a 100-mm BHI-10%HB agar plate with a cotton swab and cultured for 24 h. Lawns were then resuspended in 5 ml of minimal medium plus glucose (MMG) and spun down at 4,000 $\times$ g for 5 min. Cell pellets were washed three times with 1 ml of MMG and then diluted to 0.07 OD_{630} in 200 μl of MMG, as measured on a BioTek Synergy HT plate reader (referred to as BioTek below). Cells were then diluted 1:1,000 in 25 ml of MMG and cultured overnight to 0.35 to 0.50 OD_{600} (ThermoSpec). Cells were then spun down and resuspended in 1 ml of minimal medium without a carbon source per every 5 ml of culture. For each biological replicate, these suspensions were then diluted to 0.1 OD_{600} (ThermoSpec) in 5 ml of MMG, minimal medium plus melibiose (MMM), or minimal medium plus 0.25% (wt/vol) glucose plus 0.25% (wt/vol) galactose (MMGG). Cultures were grown to 0.45 to 0.65 OD_{600} (ThermoSpec) and then stabilized in Qiagen RNA protect. Briefly, 4 ml of culture was combined with 8 ml of RNA Protect, vortexed, and then incubated at room temperature for 5 min. Suspensions were then spun at 4,000 $\times$ g for 10 min at 4°C. Supernatants were decanted and pellets stored at −80°C. Cell pellets were thawed, and RNA was prepped by using a Qiagen RNeasy minikit. RNA was sequenced and analyzed, as done above with the RNA-seq for cells grown in TYG, with the exception that libraries were sequenced on an Illumina HiSeq4000 to produce 150-nt single-end reads.

RNA-seq protein functional and pathway analyses. RNA-seq data from each mutant strain (Δ*rbpA*, Δ*rbpB*, and Δ*rbpA* Δ*rbpB*) was compared to wild-type in rich medium at mid-log phase and stationary phase, yielding six comparisons. All genes with a $\log_2FC$ ≥+1 or ≤−1 with a q value of <0.06 were assigned to functional groups using the following framework. GSEA was used on each of the six differentially expressed gene sets individually to identify enriched functional groups. *B. thetaiotaomicron*-specific functional gene sets used included: KEGG pathways, CPS loci, PULs, and corrinoid transport (22). GSEA identified enriched gene sets corresponding to PULs, CPS loci, tricarboxylic acid (TCA) cycle (KEGG pathway bth00020), corrinoid transport, and microbial metabolism in diverse environments (bth01120). Since these were identified as enriched categories, if one of the significantly differentially regulated genes in Data Set S2A was part of one of these gene sets, it was assigned that as a category identifier with the exception of "microbial metabolism in diverse environments." GSEA-identified enriched genes in this KEGG category were further split into subcategories, including TCA cycle and glyoxylate and dicarboxylate metabolism (bth00630). If genes were associated with both TCA cycle and glyoxylate and dicarboxylate metabolism, they were assigned TCA cycle since this category was identified as enriched by GSEA directly, but glyoxylate and dicarboxylate metabolism was not. Since corrinoid transport was enriched in our differentially expressed genes, any gene associated with B_{12} metabolism that was in our differentially regulated genes was assigned a "B-vitamin metabolism" category. Remaining gene functions were assigned using gene ontology (GO) from QuickGO. Briefly, gene names were used to extract UniprotKB identifiers that were then used to pull GO biological process annotations from QuickGO when available. Only the first reported GO assignment was used for each gene. GO terms were further grouped into custom functional categories to make a more tractably sized list of functional categories for visualization. Genes without any of the above functional assignments were labeled as either "hypothetical protein" if they were annotated as such or "miscellaneous" if the gene had a putative functional annotation that was not captured by the other functional categories. All resulting functional groups are listed in Data Set S2A.

Biolog carbon utilization assays. Biolog carbon utilization assays were conducted according to manufacturer recommendations as follows. Biolog carbon source PM1 and PM2A MicroPlates were brought to room temperature to avoid condensation prior to opening the seals. Plates were then cycled into the anaerobic chamber and maintained in an anaerobic desiccant box for 24 h. Single colonies of each strain were swabbed onto BHI–10%HB plates and cultured overnight to produce a lawn of cells. Cells were aerobically suspended into 5 ml of reduced minimal medium without a carbon source to 40% turbidity (OD_{590}; ThermoSpec) using a cotton swab. Suspensions were cycled into the chamber, and 1.5 ml was combined with 22 ml of anoxic, reduced minimal medium without a carbon source. Each carbon source plate was then inoculated with 100 μl of diluted cells and statically incubated for 30 min at

room temperature to facilitate compound dissolution before measuring time point zero. Plates were statically incubated at 37°C with manual OD_{630} readings taken every hour in the plate reader for the first 11 h of growth. Plates were then left in the chamber overnight and optical density readings were resumed after 24 h of growth. Time points were then taken every 3 h to a final time point of 36 h of growth. Linear regression and prediction curves were calculated using Prism. Negative-control wells and Xylitol (PM2A) were removed from linear regression calculations. Xylitol was removed due to an unknown occlusion (potentially condensation or precipitation) causing transiently high OD_{630} readings. In the absence of these transient values, *B. thetaiotaomicron* could not grow on Xylitol as a sole carbon source in these experiments.

Minimal medium growth assays. Strains were cultured from a colony in 5 ml of TYG for 24 h and then subcultured at 1:1,000 into 5 ml of MMG for 24 h. Next, 1 ml of stationary-phase MMG cultures were spun down 4,000 $\times$ g for 10 min at room temperature. Supernatants were removed, and the pellets were resuspended in 1 ml of minimal medium without a carbon source. Next, 2 μl of cells was subcultured into 198 μl of minimal media containing carbon sources to appropriate final concentrations in flat-bottom, 96-well Corning Costar tissue culture-treated plates (Sigma). Plates were sealed with a Breathe-Easy gas permeable membrane (Sigma) and statically cultured in the BioTek plate reader for 48 h with the optical density recorded every 30 min. Doubling times were calculated using the least-squares method for growth between 0.2 and 0.4 OD_{630} ($n = 3$).

qRT-PCR of *rbpB* and *BT1871*. Strains were cultured in MMG to stationary phase overnight, subcultured 1:100 into 5 ml of MMG, and then cultured to mid-log phase (0.38 to 0.52 OD_{600}; ThermoSpec). All cultures for strains containing pNBU2_*ermGb* vectors contained erythromycin. Next, 4 ml of cells was pelleted at 4,000 $\times$ g for 10 min, supernatant decanted, and then RNA was isolated with a Qiagen RNeasy minikit. Residual DNA was degraded on-column using a Qiagen RNase-Free DNase set, and the RNA was cleaned with a Qiagen RNeasy minikit. First-strand cDNA synthesis was done with a SuperScript II RT kit (Invitrogen) and random hexamers (Invitrogen). After reverse transcription, the SuperScript reaction mixture was incubated with 27 μl of 1 N NaOH at 65°C for 30 min, neutralized with 27 μl of 1N HCl, and cleaned up with a Qiagen MinElute PCR purification kit. cDNA was diluted and *rbpB*, *BT1871*, and 16S rRNA copies amplified using 2$\times$ QX200 ddPCR EvaGreen Supermix (Bio-Rad) and quantified using the QX200 Droplet Digital PCR system (Bio-Rad) according to the manufacturer's instructions. All ddPCR consumables were supplied by Bio-Rad and Rainin (pipette tips only). The relative ratios were calculated by dividing the *rbpB* or *BT1871* counts by the 16S rRNA counts.

Determination of operon structure of *rbpB* in *B. thetaiotaomicron*. Strains were cultured in MMG to stationary phase overnight and then subcultured at 1:1,000 into 4 ml of MMM and at 1:10,000 into 4 ml of MMG. MMG and MMM cultures were grown to mid-log phase (OD_{630} 0.25 to 0.35; BioTek), pelleted at 4,000 $\times$ g for 10 min at 4°C, and the supernatant was removed. RNA and cDNA were prepped as done for qPCR, with the exception that residual DNA was degraded on beads using an Ambion nuclease-free DNase kit. Overlap endpoint PCR was done with KAPA HiFi (KAPA Biosystems). gDNA was prepped using a Qiagen DNeasy blood and tissue kit.

Data availability. All RNA-seq data sets corresponding to the samples listed in Data Set S2C in the supplemental material are publicly available from the NCBI under BioProject accession number PRJNA723047.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, PDF file, 2.4 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.04 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.5 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.02 MB.

REFERENCES

1. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, Henrissat B, Knight R, Gordon JI. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332:970–974. https://doi.org/10.1126/science.1198719.

2. Carmody RN, Gerber GK, Luevano JM, Gatti DM, Somes L, Svenson KL, Turnbaugh PJ. 2015. Diet dominates host genotype in shaping the murine gut microbiota. Cell Host Microbe 17:72–84. https://doi.org/10.1016/j.chom.2014.11.010.

3. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD. 2011. Linking long-term dietary patterns with gut microbial enterotypes. Science 334:105–108. https://doi.org/10.1126/science.1208344.

4. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B. 2013. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. Nat Rev Microbiol 11:497–504. https://doi.org/10.1038/nrmicro3050.

5. Martens EC, Chiang HC, Gordon JI. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe 4:447–457. https://doi.org/10.1016/j.chom.2008.09.007.

6. Rogers TE, Pudlo NA, Koropatkin NM, Bell JSK, Balasch MM, Jasker K, Martens EC. 2013. Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures. Mol Microbiol 88:876–890. https://doi.org/10.1111/mmi.12228.

7. Terrapon N, Lombard V, Drula É, Lapébie P, Al-Masaudi S, Gilbert HJ, Henrissat B. 2018. PULDB: the expanded database of polysaccharide utilization loci. Nucleic Acids Res 46:D677–D683. https://doi.org/10.1093/nar/gkx1022.

8. Pereira GV, Abdel-Hamid AM, Dutta S, D'Alessandro-Gabazza CN, Wefers D, Farris JA, Bajaj S, Wawrzak Z, Atomi H, Mackie RI, Gabazza EC, Shukla D, Koropatkin NM, Cann I. 2021. Degradation of complex arabinoxylans by human colonic *Bacteroidetes*. Nat Commun 12:459. https://doi.org/10.1038/s41467-020-20737-5.

9. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JI. 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol 9:e1001221. https://doi.org/10.1371/journal.pbio.1001221.

10. Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV, Gordon JI. 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science 299:2074–2076. https://doi.org/10.1126/science.1080029.

11. Xu J, Chiang HC, Bjursell MK, Gordon JI. 2004. Message from a human gut symbiont: sensitivity is a prerequisite for sharing. Trends Microbiol 12:21–28. https://doi.org/10.1016/j.tim.2003.11.007.

12. Cho KH, Cho D, Wang GR, Salyers AA. 2001. New regulatory gene that contributes to control of *Bacteroides thetaiotaomicron* starch utilization genes. J Bacteriol 183:7198–7205. https://doi.org/10.1128/JB.183.24.7198-7205.2001.

13. Schwalm ND, Townsend GE, Groisman EA. 2016. Multiple signals govern utilization of a polysaccharide in the gut bacterium *Bacteroides thetaiotaomicron*. mBio 7:e01342-16. https://doi.org/10.1128/mBio.01342-16.

14. Chang C, Tesar C, Li X, Kim Y, Rodionov DA, Joachimiak A. 2015. A novel transcriptional regulator of L-arabinose utilization in human gut bacteria. Nucleic Acids Res 43:10546–10559. https://doi.org/10.1093/nar/gkv1005.

15. Ravcheev DA, Godzik A, Osterman AL, Rodionov DA. 2013. Polysaccharides utilization in human gut bacterium *Bacteroides thetaiotaomicron*: comparative genomics reconstruction of metabolic and regulatory networks. BMC Genomics 14:873. https://doi.org/10.1186/1471-2164-14-873.

16. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL. 2010. Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. Cell 141:1241–1252. https://doi.org/10.1016/j.cell.2010.05.005.

17. Martens EC, Roth R, Heuser JE, Gordon JI. 2009. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. J Biol Chem 284:18445–18457. https://doi.org/10.1074/jbc.M109.008094.

18. Pudlo NA, Urs K, Kumar SS, German JB, Mills DA, Martens EC. 2015. Symbiotic human gut bacteria with variable metabolic priorities for host mucosal glycans. mBio 6:e01282-15–e01215. https://doi.org/10.1128/mBio.01282-15.

19. Cao Y, Förstner KU, Vogel J, Smith CJ. 2016. *cis*-encoded small RNAs, a conserved mechanism for repression of polysaccharide utilization in *Bacteroides*. J Bacteriol 198:2410–2418. https://doi.org/10.1128/JB.00381-16.

20. Townsend GE, Han W, Schwalm ND, Hong X, Bencivenga-Barry NA, Goodman AL, Groisman EA. 2020. A master regulator of *Bacteroides thetaiotaomicron* gut colonization controls carbohydrate utilization and an alternative protein synthesis factor. mBio 11:e03221-19. https://doi.org/10.1128/mBio.00301-20.

21. Ryan D, Jenniches L, Reichardt S, Barquist L, Westermann AJ. 2020. A high-resolution transcriptome map identifies small RNA regulation of metabolism in the gut microbe *Bacteroides thetaiotaomicron*. Nat Commun 11:3557. https://doi.org/10.1038/s41467-020-17348-5.

22. Degnan PH, Barry NA, Mok KC, Taga ME, Goodman AL. 2014. Human gut microbes use multiple transporters to distinguish vitamin B_{12} analogs and compete in the gut. Cell Host Microbe 15:47–57. https://doi.org/10.1016/j.chom.2013.12.007.

23. Costliow ZA, Degnan PH. 2017. Thiamine acquisition strategies impact metabolism and competition in the gut microbe *Bacteroides thetaiotaomicron*. mSystems 2:e00116-17. https://doi.org/10.1128/mSystems.00116-17.

24. Costliow ZA, Degnan PH, Vanderpool CK. 2019. Thiamine pyrophosphate riboswitches in *Bacteroides* species regulate transcription or translation of thiamine transport and biosynthesis genes. bioRxiv 867226. https://doi.org/10.1101/867226.

25. Papenfort K, Vogel J. 2014. Small RNA functions in carbon metabolism and virulence of enteric pathogens. Front Cell Infect Microbiol 4:1–12. https://doi.org/10.3389/fcimb.2014.00091.

26. Lay ND, Schu DJ, Gottesman S. 2013. Bacterial small RNA-based negative regulation: Hfq and its accomplices. J Biol Chem 288:7996–8003. https://doi.org/10.1074/jbc.R112.441386.

27. Bouloc P, Repoila F. 2016. Fresh layers of RNA-mediated regulation in Gram-positive bacteria. Curr Opin Microbiol 30:30–35. https://doi.org/10.1016/j.mib.2015.12.008.

28. Kavita K, de Mets F, Gottesman S. 2018. New aspects of RNA-based regulation by Hfq and its partner sRNAs. Curr Opin Microbiol 42:53–61. https://doi.org/10.1016/j.mib.2017.10.014.

29. Smirnov A, Schneider C, Hör J, Vogel J. 2017. Discovery of new RNA classes and global RNA-binding proteins. Curr Opin Microbiol 39:1–9.

30. Hör J, Gorski SA, Vogel J. 2018. Bacterial RNA biology on a genome scale. Mol Cell 70:785–799. https://doi.org/10.1016/j.molcel.2017.12.023.

31. Melamed S, Adams PP, Zhang A, Zhang H, Storz G. 2020. RNA-RNA interactomes of ProQ and Hfq reveal overlapping and competing roles. Mol Cell 77:411–425. https://doi.org/10.1016/j.molcel.2019.10.022.

32. Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. Nat Rev Microbiol 9:578–589. https://doi.org/10.1038/nrmicro2615.

33. Santiago-Frangos A, Woodson SA. 2018. Hfq chaperone brings speed dating to bacterial sRNA. Wiley Interdiscip Rev RNA 9:e1475. https://doi.org/10.1002/wrna.1475.

34. Iosub IA, van Nues RW, McKellar SW, Nieken KJ, Marchioretto M, Sy B, Tree JJ, Viero G, Granneman S. 2020. Hfq CLASH uncovers sRNA-target interaction networks linked to nutrient availability adaptation. Elife 9:e54655. https://doi.org/10.7554/eLife.54655.

35. Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L, Margalit H. 2016. Global mapping of small RNA-target interactions in bacteria. Mol Cell 63:884–897. https://doi.org/10.1016/j.molcel.2016.07.026.

36. Chao Y, Vogel J. 2010. The role of Hfq in bacterial pathogens. Curr Opin Microbiol 13:24–33. https://doi.org/10.1016/j.mib.2010.01.001.

37. Richards GR, Vanderpool CK. 2011. Molecular call and response: the physiology of bacterial small RNAs. Biochim Biophys Acta 1809:525–531. https://doi.org/10.1016/j.bbagrm.2011.07.013.

38. Bobrovskyy M, Vanderpool CK. 2014. The small RNA SgrS: roles in metabolism and pathogenesis of enteric bacteria. Front Cell Infect Microbiol 4:1–8. https://doi.org/10.3389/fcimb.2014.00061.

39. Waters JL, Salyers AA. 2012. The small RNA RteR inhibits transfer of the *Bacteroides* conjugative transposon CTnDOT. J Bacteriol 194:5228–5236. https://doi.org/10.1128/JB.00941-12.

40. Townsend GE, Han W, Schwalm ND, Raghavan V, Barry NA, Goodman AL, Groisman EA. 2019. Dietary sugar silences a colonization factor in a mammalian gut symbiont. Proc Natl Acad Sci U S A 116:233–238. https://doi.org/10.1073/pnas.1813780115.

41. Methé BA, Nelson KE, Pop M, Creasy HH, Giglio MG, Huttenhower C, Gevers D, Petrosino JF, Abubucker S, Badger JH, Chinwalla AT, Earl AM, FitzGerald MG, Fulton RS, Hallsworth-Pepin K, Lobos EA, Madupu R, Magrini V, Martin JC, Mitreva M, Muzny DM, Sodergren EJ, Versalovic J, Wollam AM, Worley KC, Wortman JR, Young SK, Zeng Q, Aagaard KM,

Abolude OO, et al. 2012. A framework for human microbiome research. Nature 486:215–221. https://doi.org/10.1038/nature11209.

42. Maris C, Dominguez C, Allain FHT. 2005. The RNA recognition motif, a plastic RNA-binding platform to regulate posttranscriptional gene expression. FEBS J 272:2118–2131. https://doi.org/10.1111/j.1742-4658.2005.04653.x.

43. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432. https://doi.org/10.1093/nar/gky995.

44. Herren CD, Rocha ER, Smith CJ. 2003. Genetic analysis of an important oxidative stress locus in the anaerobe Bacteroides fragilis. Gene 316:167–175. https://doi.org/10.1016/s0378-1119(03)00759-5.

45. Voss JE, Luisi BF, Hardwick SW. 2014. Molecular recognition of RhlB and RNase D in the Caulobacter crescentus RNA degradosome. Nucleic Acids Res 42:13294–13305. https://doi.org/10.1093/nar/gku1134.

46. Al-Husini N, Tomares DT, Bitar O, Childers WS, Schrader JM. 2018. α-Proteobacterial RNA degradosomes assemble liquid-liquid phase-separated RNP bodies. Mol Cell 71:1027–1039. https://doi.org/10.1016/j.molcel.2018.08.003.

47. Santiago-Frangos A, Kavita K, Schu DJ, Gottesman S, Woodson SA. 2016. C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. Proc Natl Acad Sci U S A 113:E6089–E6096. https://doi.org/10.1073/pnas.1613053113.

48. Večerek B, Rajkowitsch L, Sonnleitner E, Schroeder R, Bläsi U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. Nucleic Acids Res 36:133–143. https://doi.org/10.1093/nar/gkm985.

49. Santiago-Frangos A, Jeliazkov JR, Gray JJ, Woodson SA. 2017. Acidic C-terminal domains autoregulate the RNA chaperone Hfq. Elife 6:e27049. https://doi.org/10.7554/eLife.27049.

50. Sobrero P, Valverde C. 2012. The bacterial protein Hfq: much more than a mere RNA-binding factor. Crit Rev Microbiol 38:276–299. https://doi.org/10.3109/1040841X.2012.664540.

51. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JCD. 2013. An infection-relevant transcriptomic compendium for Salmonella enterica serovar Typhimurium. Cell Host Microbe 14:683–695. https://doi.org/10.1016/j.chom.2013.11.010.

52. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, Pudlo NA, Kitamoto S, Terrapon N, Muller A, Young VB, Henrissat B, Wilmes P, Stappenbeck TS, Núñez G, Martens EC. 2016. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. Cell 167:1339–1353. https://doi.org/10.1016/j.cell.2016.10.043.

53. Wu M, McNulty NP, Rodionov DA, Khoroshkin MS, Griffin NW, Cheng J, Latreille P, Kerstetter RA, Terrapon N, Henrissat B, Osterman AL, Gordon JI. 2015. Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. Science 350:aac5992. https://doi.org/10.1126/science.aac5992.

54. Huang Y, Sheth RU, Kaufman A, Wang HH. 2020. Scalable and cost-effective ribonuclease-based rRNA depletion for transcriptomics. Nucleic Acids Res 48:e20. https://doi.org/10.1093/nar/gkz1169.

55. Despres J, Forano E, Lepercq P, Comtet-Marre S, Jubelin G, Chambon C, Yeoman CJ, Miller MEB, Fields CJ, Martens E, Terrapon N, Henrissat B, White BA, Mosoni P. 2016. Xylan degradation by the human gut Bacteroides xylanisolvens XB1AT involves two distinct gene clusters that are linked at the transcriptional level. BMC Genomics 17:14. https://doi.org/10.1186/s12864-016-2680-8.

56. Schofield WB, Zimmermann-Kogadeeva M, Zimmermann M, Barry NA, Goodman AL. 2018. The stringent response determines the ability of a commensal bacterium to survive starvation and to persist in the gut. Cell Host Microbe 24:120–132. https://doi.org/10.1016/j.chom.2018.06.002.

57. Bendak K, Loughlin FE, Cheung V, O'Connell MR, Crossley M, Mackay JP. 2012. A rapid method for assessing the RNA-binding potential of a protein. Nucleic Acids Res 40:e105. https://doi.org/10.1093/nar/gks285.

58. Teplova M, Farazi TA, Tuschl T, Patel DJ. 2016. Structural basis underlying CAC RNA recognition by the RRM domain of dimeric RNA-binding protein RBPMS. Q Rev Biophys 49:e1. https://doi.org/10.1017/S0033583515000207.

59. Afroz T, Cienikova Z, Cléry A, Allain FHT. 2015. One, two, three, four! How multiple RRMs read the genome sequence. Methods Enzymol 558:235–278. https://doi.org/10.1016/bs.mie.2015.01.015.

60. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208. https://doi.org/10.1093/nar/gkp335.

61. Azam MS, Vanderpool CK. 2018. Translational regulation by bacterial small RNAs via an unusual Hfq-dependent mechanism. Nucleic Acids Res 46:2585–2599. https://doi.org/10.1093/nar/gkx1286.

62. Stein EM, Kwiatkowska J, Basczok MM, Gravel CM, Berry KE, Olejniczak M. 2020. Determinants of RNA recognition by the FinO domain of the Escherichia coli ProQ protein. Nucleic Acids Res 48:7502–7519. https://doi.org/10.1093/nar/gkaa497.

63. Nithin C, Mukherjee S, Bahadur RP. 2019. A structure-based model for the prediction of protein-RNA binding affinity. RNA 25:1628–1645. https://doi.org/10.1261/rna.071779.119.

64. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550. https://doi.org/10.1073/pnas.0506580102.

65. Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE. 2001. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature 414:555–558. https://doi.org/10.1038/35107092.

66. Gloster TM, Turkenburg JP, Potts JR, Henrissat B, Davies GJ. 2008. Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora. Chem Biol 15:1058–1067. https://doi.org/10.1016/j.chembiol.2008.09.005.

67. Liu H, Shiver AL, Price MN, Carlson HK, Trotter VV, Chen Y, Escalante V, Ray J, Hern KE, Petzold CJ, Turnbaugh PJ, Huang KC, Arkin AP, Deutschbauer AM. 2021. Functional genetics of human gut commensal Bacteroides thetaiotaomicron reveals metabolic requirements for growth across environments. Cell Rep 34:108789. https://doi.org/10.1016/j.celrep.2021.108789.

68. Faith JJ, McNulty NP, Rey FE, Gordon JI. 2011. Predicting a human gut microbiota's response to diet in gnotobiotic mice. Science 333:101–104. https://doi.org/10.1126/science.1206025.

69. Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J, Buhler JD, Gordon JI. 2005. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science 307:1955–1959. https://doi.org/10.1126/science.1109051.

70. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1:6ra14. https://doi.org/10.1126/scitranslmed.3000322.

71. Görke B, Stülke J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol 6:613–624. https://doi.org/10.1038/nrmicro1932.

72. Hylemon PB, Phibbs PV. 1974. Evidence against the presence of cyclic AMP and related enzymes in selected strains of Bacteroides fragilis. Biochem Biophys Res Commun 60:88–95. https://doi.org/10.1016/0006-291x(74)90176-4.

73. Barabote RD, Saier MH. 2005. Comparative genomic analyses of the bacterial phosphotransferase system. Microbiol Mol Biol Rev 69:608–634. https://doi.org/10.1128/MMBR.69.4.608-634.2005.

74. Brigham CJ, Malamy MH. 2005. Characterization of the RokA and HexA broad-substrate-specificity hexokinases from Bacteroides fragilis and their role in hexose and N-acetylglucosamine utilization. J Bacteriol 187:890–901. https://doi.org/10.1128/JB.187.3.890-901.2005.

75. Holmqvist E, Berggren S, Rizvanovic A. 2020. RNA-binding activity and regulatory functions of the emerging sRNA-binding protein ProQ. Biochim Biophys Acta Gene Regul Mech 1863:194596. https://doi.org/10.1016/j.bbagrm.2020.194596.

76. Olejniczak M, Storz G. 2017. ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? Mol Microbiol 104:905–915. https://doi.org/10.1111/mmi.13679.

77. Updegrove TB, Zhang A, Storz G. 2016. Hfq: the flexible RNA matchmaker. Curr Opin Microbiol 30:133–138. https://doi.org/10.1016/j.mib.2016.02.003.

78. Coyne MJ, Reinap B, Lee MM, Comstock LE. 2005. Human symbionts use a host-like pathway for surface fucosylation. Science 307:1778–1781. https://doi.org/10.1126/science.1106469.

79. Porter NT, Canales P, Peterson DA, Martens EC. 2017. A subset of polysaccharide capsules in the human symbiont Bacteroides thetaiotaomicron promote increased competitive fitness in the mouse gut. Cell Host Microbe 22:494–506. https://doi.org/10.1016/j.chom.2017.08.020.

80. Bjursell MK, Martens EC, Gordon JI. 2006. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, Bacteroides thetaiotaomicron, to the suckling period. J Biol Chem 281:36269–36279. https://doi.org/10.1074/jbc.M606509200.

81. Peterson DA, McNulty NP, Guruge JL, Gordon JI. 2007. IgA response to symbiotic bacteria as a mediator of gut homeostasis. Cell Host Microbe 2:328–339. https://doi.org/10.1016/j.chom.2007.09.013.

82. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. 2009. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 6:279–289. https://doi.org/10.1016/j.chom.2009.08.003.

83. Holdeman L, Cato E, Moore W. 1977. Anaerobe laboratory manual, 4th ed. Anaerobe Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA.

84. Koropatkin NM, Martens EC, Gordon JI, Smith TJ. 2008. Starch catabolism by a prominent human gut symbiont Is directed by the recognition of amylose helices. Structure 16:1105–1115. https://doi.org/10.1016/j.str.2008.03.017.

85. García-Bayona L, Comstock LE. 2019. Streamlined genetic manipulation of diverse *Bacteroides* and *Parabacteroides* isolates from the human gut microbiota. mBio 10:e01762-19. https://doi.org/10.1128/mBio.01762-19.

86. Wang J, Shoemaker NB, Wang GR, Salyers AA. 2000. Characterization of a *Bacteroides* mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene. J Bacteriol 182:3559–3571. https://doi.org/10.1128/JB.182.12.3559-3571.2000.

87. Brown JR, Douady CJ, Italia MJ, Marshal WE, Stanhope MJ. 2001. Universal trees based on large combined protein sequence data sets. Nat Genet 28:281–285. https://doi.org/10.1038/90129.

88. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

89. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

90. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

91. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumby P, Genco CA, Vanderpool CK, Tjaden B. 2013. Computational analysis of bacterial RNA-Seq data. Nucleic Acids Res 41:e140. https://doi.org/10.1093/nar/gkt444.

92. Tjaden B. 2015. *De novo* assembly of bacterial transcriptomes from RNA-seq data. Genome Biol 16:1–10. https://doi.org/10.1186/s13059-014-0572-2.

93. SAS Institute, Inc. 1989. JMP®. SAS Institute, Inc., Cary, NC.

94. Ryder SP, Recht MI, Williamson JR. 2008. Quantitative analysis of protein-RNA interactions by gel mobility shift. Methods Mol Biol 488:99–115. https://doi.org/10.1007/978-1-60327-475-3_7.

95. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R. 2009. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics 25:2607–2608. https://doi.org/10.1093/bioinformatics/btp450.

96. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

97. Langmead B, Wilks C, Antonescu V, Charles R. 2019. Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics 35:421–432. https://doi.org/10.1093/bioinformatics/bty648.