

Supplemental Information for Capsular polysaccharide correlates with immune response to the human gut microbe, *Ruminococcus gnavus*

Matthew T. Henke,¹ Eric M. Brown,² Chelsi D. Cassilly,¹ Hera Vlamakis,² Ramnik J. Xavier,² and Jon Clardy^{1,*}

¹Harvard Medical School, Boston, MA 02115 ²Massachusetts General Hospital, Boston, MA; and Broad Institute of Harvard and MIT, Cambridge, MA

*corresponding author: jon_clardy@hms.harvard.edu

Table S1. Strains used in this study.

Strain	Origin
ATCC 29149	stool; healthy adult
ATCC 35913	stool; healthy adult
RJX1119	stool; child treated with antibiotics
RJX1120	biopsy, adult with IBD
RJX1121	biopsy, adult with IBD
RJX1122	biopsy, adult with IBD
RJX1123	biopsy, adult with IBD
RJX1124	biopsy, adult with IBD
RJX1125	biopsy, adult with IBD
RJX1126	biopsy, adult with IBD
RJX1127	biopsy, adult with IBD
RJX1128	stool; adult with IBD

ATCC 29149 and ATCC 35913 came from the feces of healthy adults and were purchased from ATCC. RJX1119 was isolated from an infant treated with antibiotics (24), RJX1120 through RJX1127 were isolated from biopsies from IBD patients, strain RJX1128 was isolated from feces of an IBD patient (10). Strains were grown in anaerobically at 37°C in defined medium (34).

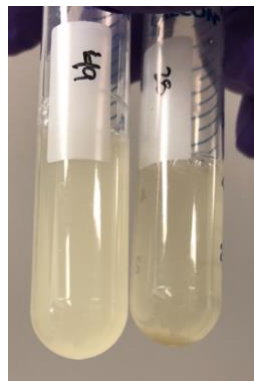


Figure S1. Differences in settling phenotype between some strains of *Ruminococcus gnavus*. ATCC 29149 (“49”) does not settle in culture, while RJX1128 (“28”) settles readily.

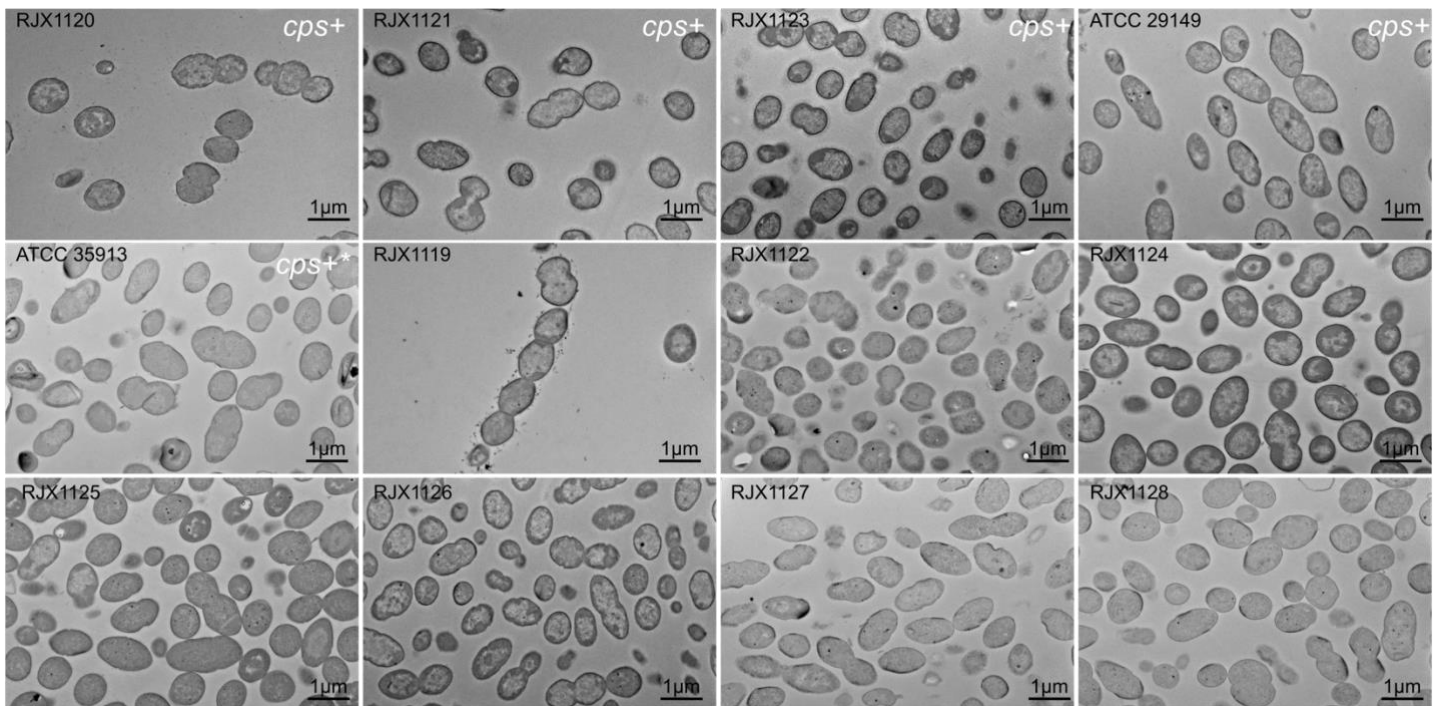


Figure S2. Electron micrographs of the 12 *R. gnavus* strains included in this study after 24 hours of growth. Some strains possess a capsule polysaccharide gene cluster (*cps*⁺) and show increased capsular material on their cell surfaces, except for ATCC 35913, which has several nonsense mutations in the *cps* gene cluster. While other strains lack the *cps* gene cluster, and have lesser or nonexistent capsules. Uniquely, strain RJX1119 lacks the *cps* gene cluster, but forms chains and appears to be encapsulated.

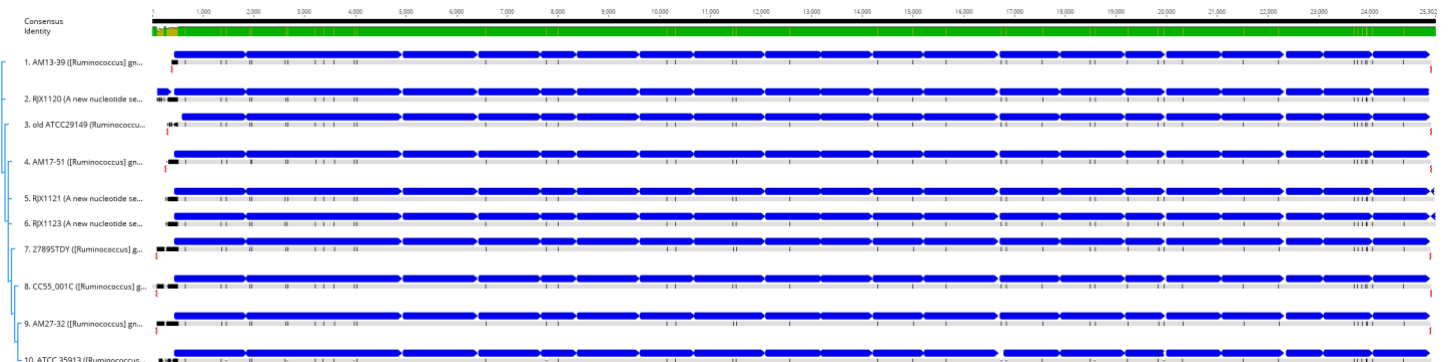


Figure S3. The capsule gene cluster (RUMGNA_02411 to RUMGNA_02392) is only found in some strains, and is conserved at >99% sequence identity. This indicates it was likely recently shared among these strains. Note that several genes in the ATCC 35913 *cps* cluster are disrupted by nonsense mutations, likely rendering this cluster nonfunctional.

Table S2. Gene annotations for the capsular polysaccharide biosynthetic gene cluster (*cps*) found in several strains of *Ruminococcus gnavus*, shown in Figure 1C.



ATCC 29149 (AAYG) locus	Consensus description	Gene size (bp)	Protein Size (aa)
RUMGNA_02411	polyprenyl glycosylphosphotransferase	1,272	424
RUMGNA_02410	bifunctional fucokinase/L-fucose-1-P-guanylyltransferase	3,060	1,020
RUMGNA_02409	LytR family transcriptional regulator	1,464	488
RUMGNA_02408	DegT/DnrJ/EryC1/StrS aminotransferase family protein	1,230	410
RUMGNA_02407	glycosyl transferase	696	232
RUMGNA_02406	glycosyltransferase family 1 or 4 protein	1,233	411
RUMGNA_02405	glycosyltransferase family 2 protein	1,068	356
RUMGNA_02404	O-antigen ligase/polymerase family protein	1,386	462
RUMGNA_02403	glycosyltransferase family 1 protein	1,119	373
RUMGNA_02402	glycosyltransferase family 2 protein	1,011	337
RUMGNA_02401	glycosyltransferase family 2 protein	1,008	336
RUMGNA_02400	flippase	1,455	485
RUMGNA_02399	hypothetical protein	1,203	401
RUMGNA_02398	hypothetical protein	1,263	421
RUMGNA_02397	WecB/TagA/CpsF family glycosyltransferase	771	257
RUMGNA_02396	glycosyltransferase family 1 protein	1,098	366
RUMGNA_02395	NAD(P)-binding/amine oxido reductase protein	1,218	406
RUMGNA_02394	hypothetical protein	753	251
RUMGNA_02393	glycosyltransferase, group 2 family protein	954	318
RUMGNA_02392	DUF3048 domain-containing protein	1,104	368

Table S3. Locus name comparison of capsular polysaccharide biosynthesis gene cluster (*cps*) among the *Ruminococcus gnavus* strains. Genes that have non-sense mutations in ATCC 35913 genome are labeled in red.

ATCC 29149 (AAYG) locus	Consensus description	ATCC 29149 (PUEL0)	ATCC 29149 (CP0)	RJX1120	RJX1121	RJX1123	ATCC 35913	2789STD Y	AM13-39	AM17-51	AM27-32	CC55_001C	Lachnospiraceae bacterium 2_1_58FAA
RUMGNA_02411	polyprenyl glycosylphospho transferase	C5Y99_R S00105	FXV78_R S12430	CDL27_ 10975	CDL19_R S08915	CDL25_R S01380	RGNV35913_ RS12015	ARA00_R S00965	DW153_ 04070	DW207_R S00740	DW722_R S05785	HMPREF1201_RS03490	HMPREF0991_RS01530
RUMGNA_02410	<i>bifunctional fucokinase/L-fucose-1-P-guanylyltransferase</i>	C5Y99_R S00110	FXV78_R S12435	CDL27_ 10980	CDL19_R S08920	CDL25_R S01375	RGNV35913_ RS12010	ARA00_R S00970	DW153_ 04075	DW207_R S00745	DW722_R S05780	HMPREF1201_RS03485	HMPREF0991_RS01525
RUMGNA_02409	LytR family transcriptional regulator DegT/DnrJ/EryC1/StrS	C5Y99_R S00115	FXV78_R S12440	CDL27_ 10985	CDL19_R S08925	CDL25_R S01370	RGNV35913_ RS12005	ARA00_R S00975	DW153_ 04080	DW207_R S00750	DW722_R S05775	HMPREF1201_RS03480	HMPREF0991_RS01520
RUMGNA_02408	aminotransferase family protein	C5Y99_R S00120	FXV78_R S12445	CDL27_ 10990	CDL19_R S08930	CDL25_R S01365	RGNV35913_ RS12000	ARA00_R S00980	DW153_ 04085	DW207_R S00755	DW722_R S05770	HMPREF1201_RS03475	HMPREF0991_RS01515
RUMGNA_02407	glycosyl transferase	C5Y99_R S00125	FXV78_R S12450	CDL27_ 10995	CDL19_R S08935	CDL25_R S01360	RGNV35913_ RS11995	ARA00_R S00985	DW153_ 04090	DW207_R S00760	DW722_R S05765	HMPREF1201_RS03470	HMPREF0991_RS01510
RUMGNA_02406	glycosyltransferase family 1 or 4 protein	C5Y99_R S00130	FXV78_R S12455	CDL27_ 11000	CDL19_R S08940	CDL25_R S01355	RGNV35913_ RS11990	ARA00_R S00990	DW153_ 04095	DW207_R S00765	DW722_R S05760	HMPREF1201_RS03465	HMPREF0991_RS01505
RUMGNA_02405	glycosyltransferase family 2 protein	C5Y99_R S00135	FXV78_R S12460	CDL27_ 11005	CDL19_R S08945	CDL25_R S01350	RGNV35913_ RS11985	ARA00_R S00995	DW153_ 04100	DW207_R S00770	DW722_R S05755	HMPREF1201_RS03460	HMPREF0991_RS01500
RUMGNA_02404	O-antigen ligase/polymerase family protein	C5Y99_R S00140	FXV78_R S12465	CDL27_ 11010	CDL19_R S08950	CDL25_R S01345	RGNV35913_ RS11980	ARA00_R S01000	DW153_ 04105	DW207_R S00775	DW722_R S05750	HMPREF1201_RS03455	HMPREF0991_RS01495
RUMGNA_02403	glycosyltransferase family 1 protein	C5Y99_R S00145	FXV78_R S12470	CDL27_ 11015	CDL19_R S08955	CDL25_R S01340	RGNV35913_ RS11975	ARA00_R S01005	DW153_ 04110	DW207_R S00780	DW722_R S05745	HMPREF1201_RS03450	HMPREF0991_RS01490
RUMGNA_02402	glycosyltransferase family 2 protein	C5Y99_R S00150	FXV78_R S12475	CDL27_ 11020	CDL19_R S08960	CDL25_R S01335	RGNV35913_ RS11970	ARA00_R S01010	DW153_ 04115	DW207_R S00785	DW722_R S05740	HMPREF1201_RS03445	HMPREF0991_RS01485
RUMGNA_02401	glycosyltransferase family 2 protein	C5Y99_R S00155	FXV78_R S12480	CDL27_ 11025	CDL19_R S08965	CDL25_R S01330	RGNV35913_ RS11965	ARA00_R S01015	DW153_ 04120	DW207_R S00790	DW722_R S05735	HMPREF1201_RS03440	HMPREF0991_RS01480
RUMGNA_02400	flippase	C5Y99_R S00160	FXV78_R S12485	CDL27_ 11030	CDL19_R S08970	CDL25_R S01325	RGNV35913_ RS11960	ARA00_R S01020	DW153_ 04125	DW207_R S00795	DW722_R S05730	HMPREF1201_RS03435	HMPREF0991_RS01475; HMPREF0991_RS01470; HMPREF0991_RS01465; and
RUMGNA_02399	hypothetical protein	C5Y99_R S00165	FXV78_R S12490	CDL27_ 11035	CDL19_R S08975	CDL25_R S01320	RGNV35913_ RS11955	ARA00_R S01025	DW153_ 04130	DW207_R S00800	DW722_R S05725	HMPREF1201_RS03430	HMPREF0991_RS01460
RUMGNA_02398	hypothetical protein	C5Y99_R S00170	FXV78_R S12495	CDL27_ 11040	CDL19_R S08980	CDL25_R S01315	RGNV35913_ RS11950	ARA00_R S01030	DW153_ 04135	DW207_R S00805	DW722_R S05720	HMPREF1201_RS03425	HMPREF0991_RS01455

	WecB/TagA/Cps F family												
RUMGNA _02397	glycosyltransferase	C5Y99_R S00175	FXV78_R S12500	CDL27_ 11045	CDL19_R S08985	CDL25_R S01310	RGNV35913_ RS11945	ARA00_R S01035	DW153_ 04140	DW207_R S00810	DW722_R S05715	HMPREF1201 _RS03420	HMPREF0991 _RS01450
RUMGNA _02396	glycosyltransferase family 1 protein	C5Y99_R S00180	FXV78_R S12505	CDL27_ 11050	CDL19_R S08990	CDL25_R S01305	RGNV35913_ RS11940	ARA00_R S01040	DW153_ 04145	DW207_R S00815	DW722_R S05710	HMPREF1201 _RS03415	HMPREF0991 _RS01445
RUMGNA _02395	NAD(P)-binding/amine oxidoreductase protein	C5Y99_R S00185	FXV78_R S12510	CDL27_ 11055	CDL19_R S08995	CDL25_R S01300	RGNV35913_ RS11935	ARA00_R S01045	DW153_ 04150	DW207_R S00820	DW722_R S05705	HMPREF1201 _RS03410	HMPREF0991 _RS01440
RUMGNA _02394	hypothetical protein	C5Y99_R S00190	FXV78_R S12515	CDL27_ 11060	CDL19_R S09000	CDL25_R S01295	RGNV35913_ RS11930	ARA00_R S01050	DW153_ 04155	DW207_R S00825	DW722_R S05700	HMPREF1201 _RS03405	HMPREF0991 _RS01435
RUMGNA _02393	glycosyltransferase, group 2 family protein	C5Y99_R S00195	FXV78_R S12520	CDL27_ 11065	CDL19_R S09005	CDL25_R S01290	RGNV35913_ RS11925	ARA00_R S01055	DW153_ 04160	DW207_R S00830	DW722_R S05695	HMPREF1201 _RS03400	HMPREF0991 _RS01430
RUMGNA _02392	DUF3048 domain-containing protein	C5Y99_R S00200	FXV78_R S12525	CDL27_ 11070	CDL19_R S09010	CDL25_R S01285	RGNV35913_ RS11920	ARA00_R S01060	DW153_ 04165	DW207_R S00835	DW722_R S05690	HMPREF1201 _RS03395	HMPREF0991 _RS01425

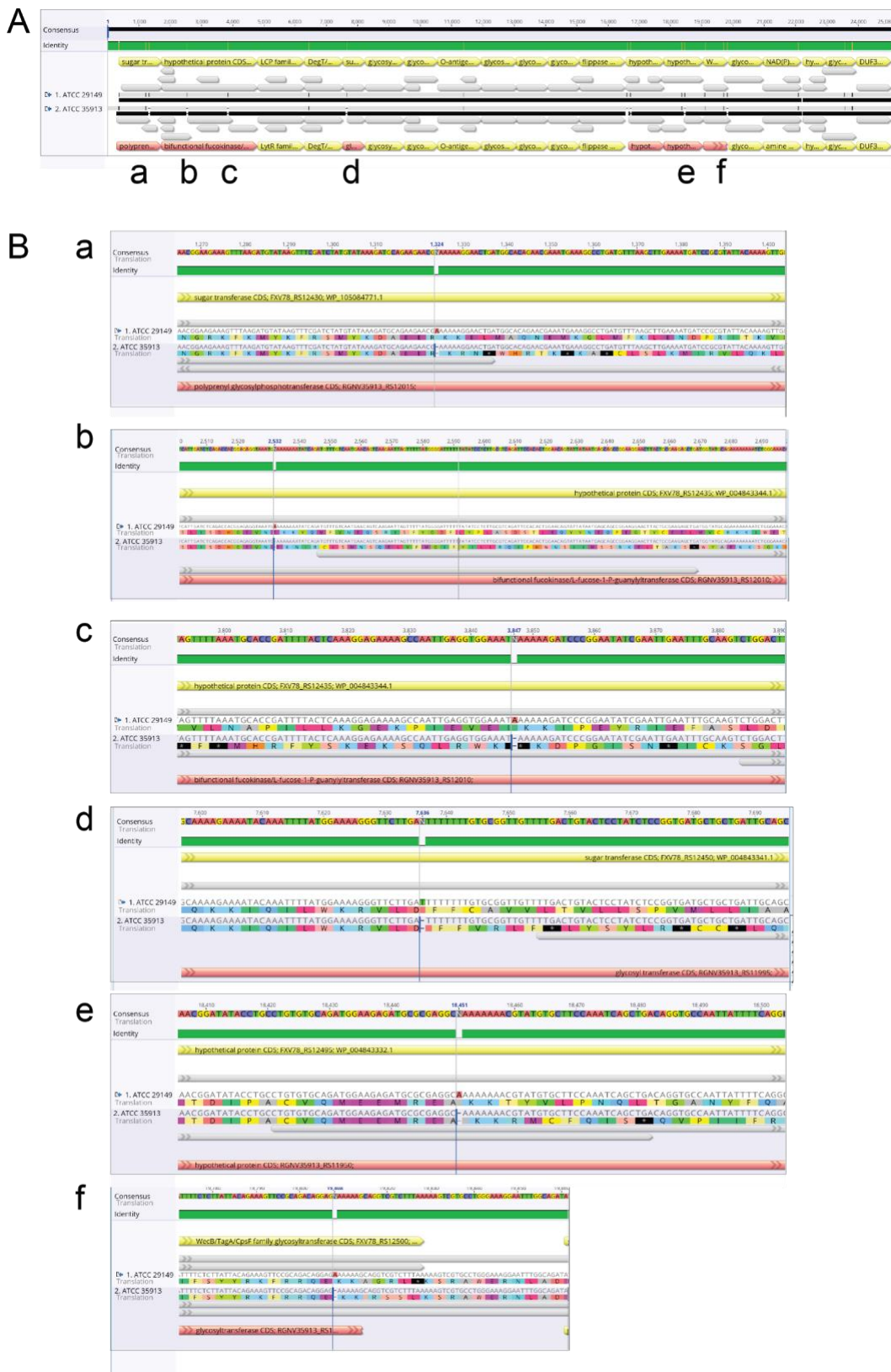


Figure S4. ATCC 35913 harbors a nonfunctional *cps* gene cluster. (A) The entire *cps* gene cluster compared between ATCC 29149 (top) and ATCC 35913 (bottom) where several genes (B) have nonsense mutations likely rendering these genes nonfunctional. Small letters indicate corresponding genes between A and B.

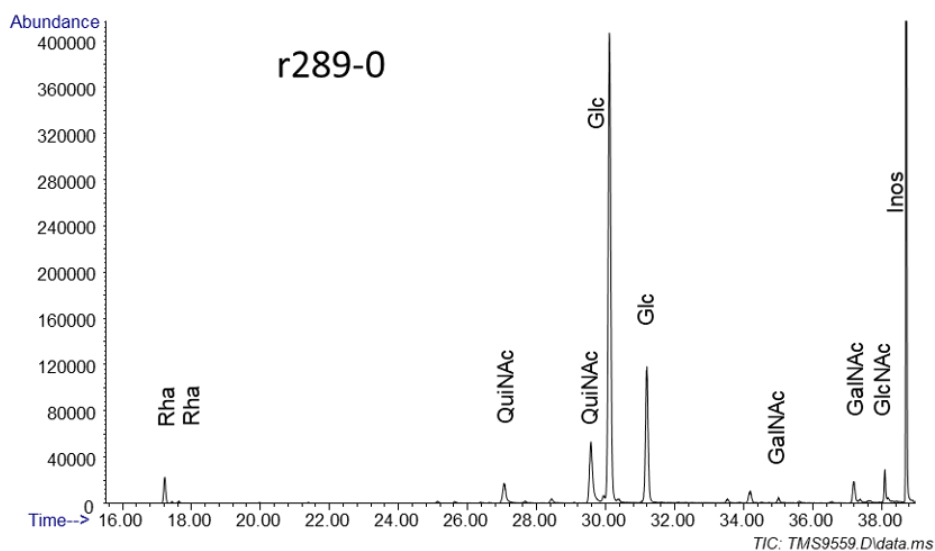
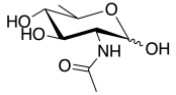
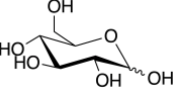
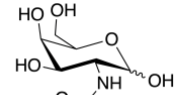


Figure S5. GCMS chromatogram of TMS-derivatives from hydrolysis of purified capsular polysaccharide shows it to be composed of glucose, N-acetyl-quinovosamine, N-acetyl-galactosamine. Trace amount of N-acetyl-glucosamine and rhamnose are also detected.

Table S4. Quantification of area under peaks from GCMS chromatogram above for each of the major glycosyl residue detected, along with their structures.

Structure	Glycosyl residue	Mass (mg)	Mol % ¹	Ratio
	Rhamnose (Rha)	1.3	2.2	
	N-Acetyl Quinovosamine (QuiNAc)	15.8	23.6	2
	Glucose (Glc)	39.8	59.4	6
	N-Acetyl Galactosamine (GalNAc)	9.0	11.0	1
	N-Acetyl Glucosamine (GlcNAc)	3.1	3.8	
	SUM	69.0	100.0	

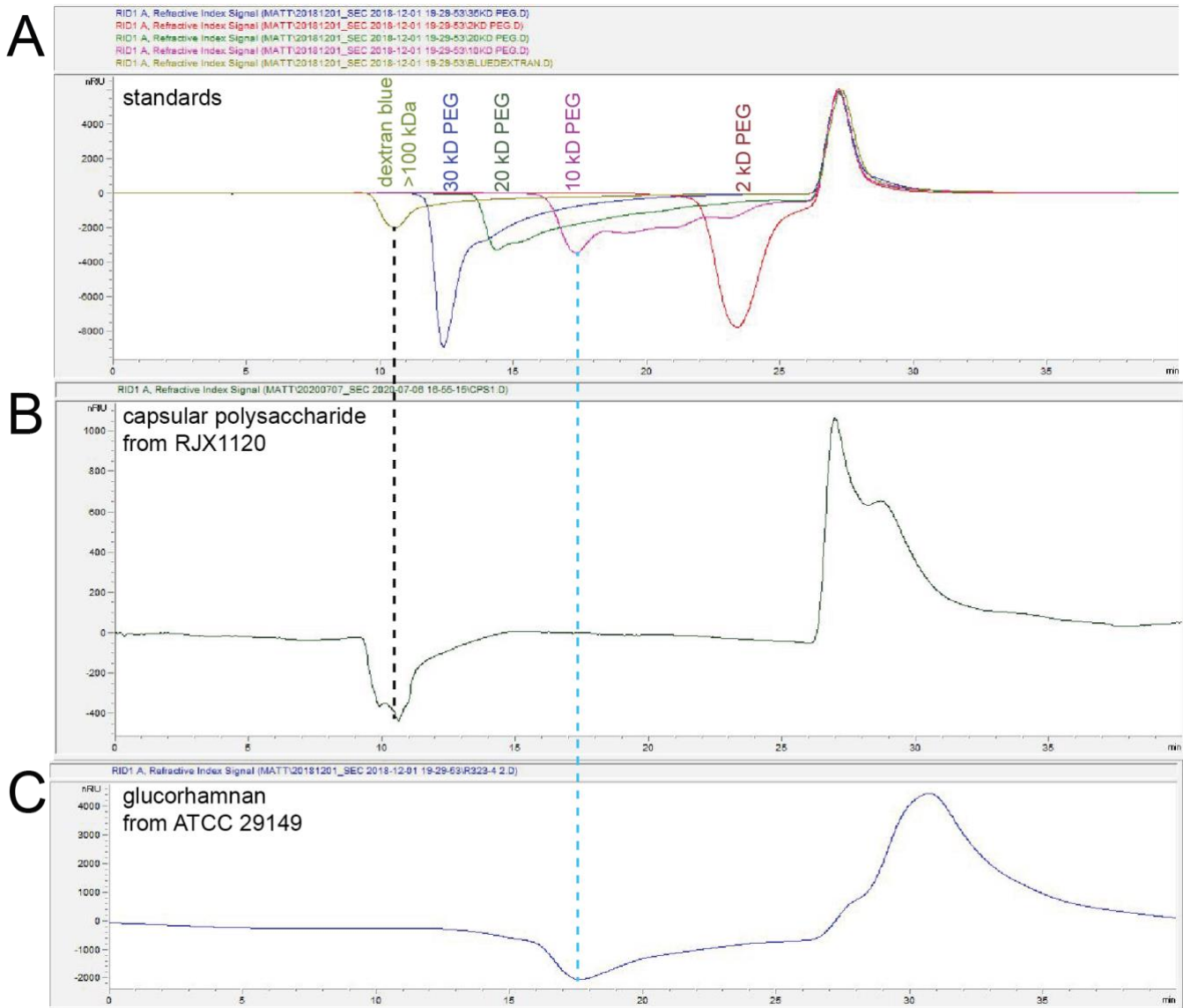


Figure S6. Analytical size exclusion chromatography of (A) molecular weight size standards, (B) purified capsular polysaccharide, and (C) purified glucorhamnan, indicates that the capsular polysaccharide is >100 kDa in size.

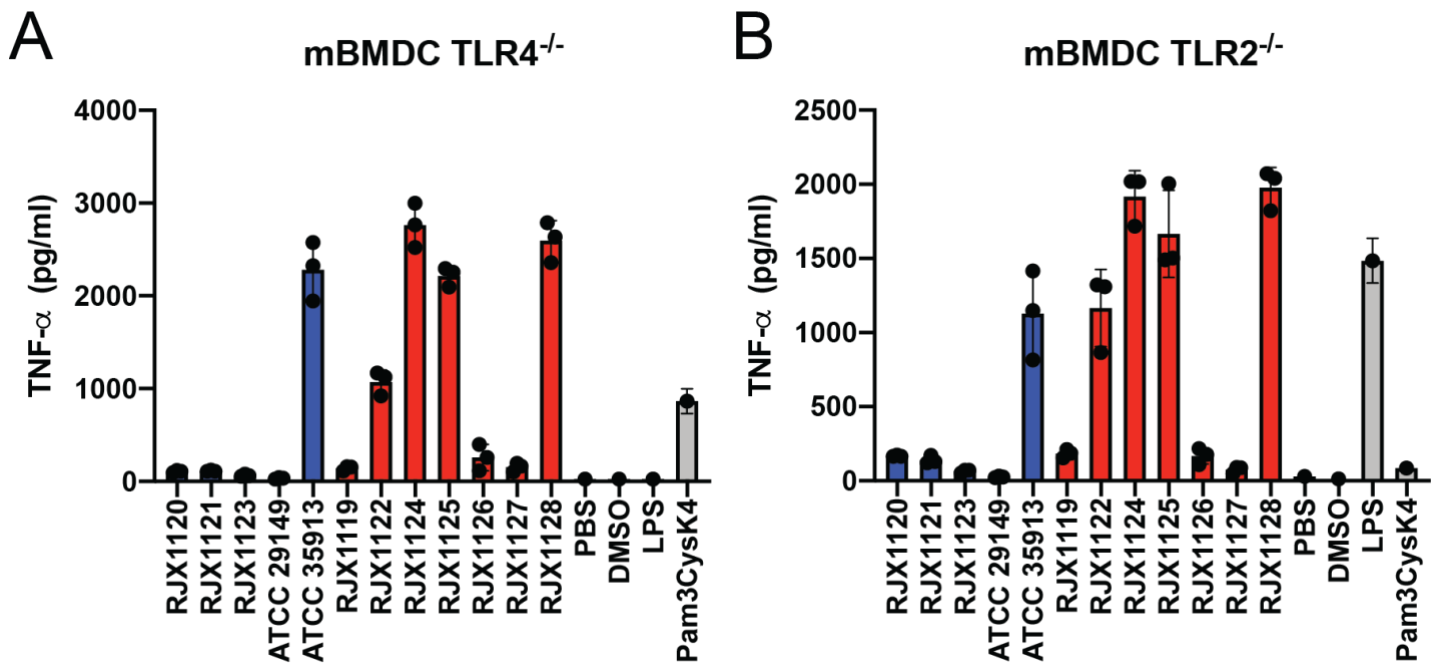


Figure S7. When mBMDCs lack either toll-like receptor (TLR) 4 (A) or TLR2 (B), they still secrete TNF α upon stimulation with unencapsulated strains of *R. gnavus*, while not responding robustly to strains that harbor a capsule. Strains harboring the *cps* locus are in blue, strains lacking are in red, controls in grey. ATCC35913 has a none functional *cps* cluster. LPS is a TLR4 agonist and Pam3CysK4 is a TLR2 agonist. *R. gnavus* strains were grown for 48 hrs, washed with PBS and then used to stimulate mBMDCs.

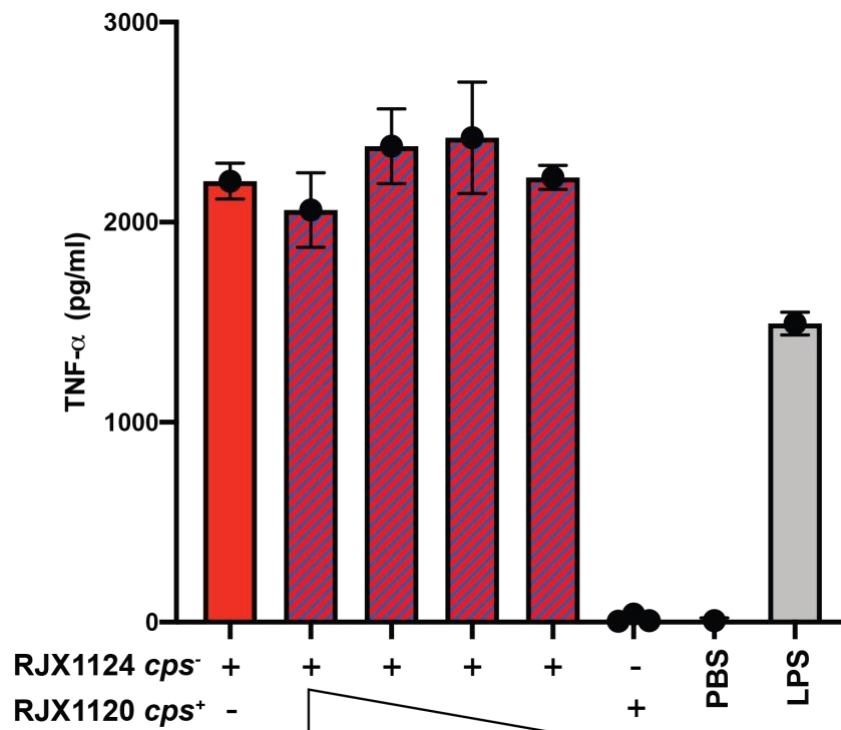


Figure S8. RJX1120 cells do not inhibit the ability of RJX1124 cells to stimulate mBMDCs to secrete TNF α , as detected by ELISA. *R. gnavus* cells grown for 48 hrs, washed with PBS, and then used to challenge mBMDCs.

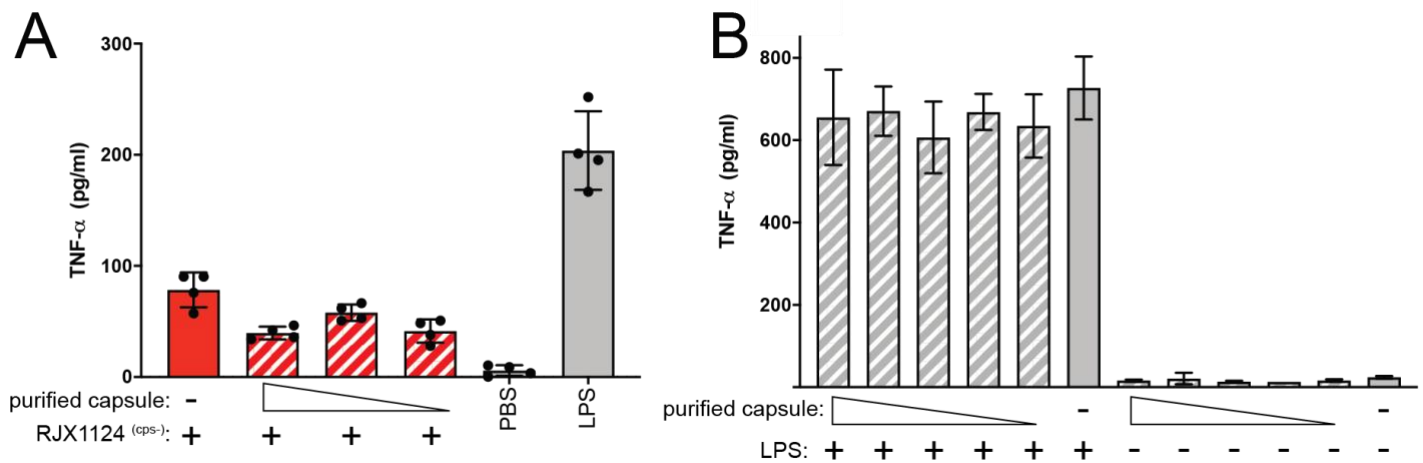


Figure S9. Purified capsular polysaccharide does not inhibit TNF α secretion when mBMDCs are stimulated with (A) an unencapsulated strain of *R. gnavus*, RJX1124, or (B) with LPS. The purified capsular polysaccharide also does not induce TNF α secretion. Dose of purified capsular polysaccharide are half-fold dilutions starting at 0.1 mg/mL, a typical concentration used for LPS or *R. gnavus* glucorhamnan in this assay.

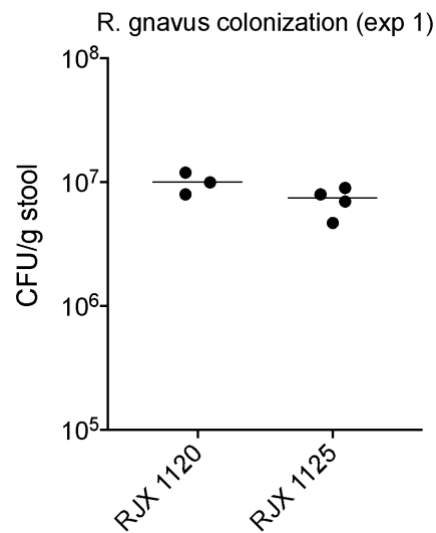


Figure S10. Both strains of *R. gnavus* colonized germ-free mice equally well as determined by cfus recovered from feces.