PNAS
www.pnas.org

Supplementary Information for

Lactobacillus Bile Salt Hydrolase Substrate Specificity Governs Bacterial Fitness and Host Colonization

Matthew H. Foley[1], Sarah O'Flaherty[2], Garrison Allen[1], Alissa J. Rivera[1], Allison K. Stewart[3], Rodolphe Barrangou[2*], Casey M. Theriot[1*].

Rodolphe Barrangou & Casey M. Theriot
Email: rbarran@ncsu.edu & cmtherio@ncsu.edu

This PDF file includes:

Figures S1 to S11
Tables S1 to S3

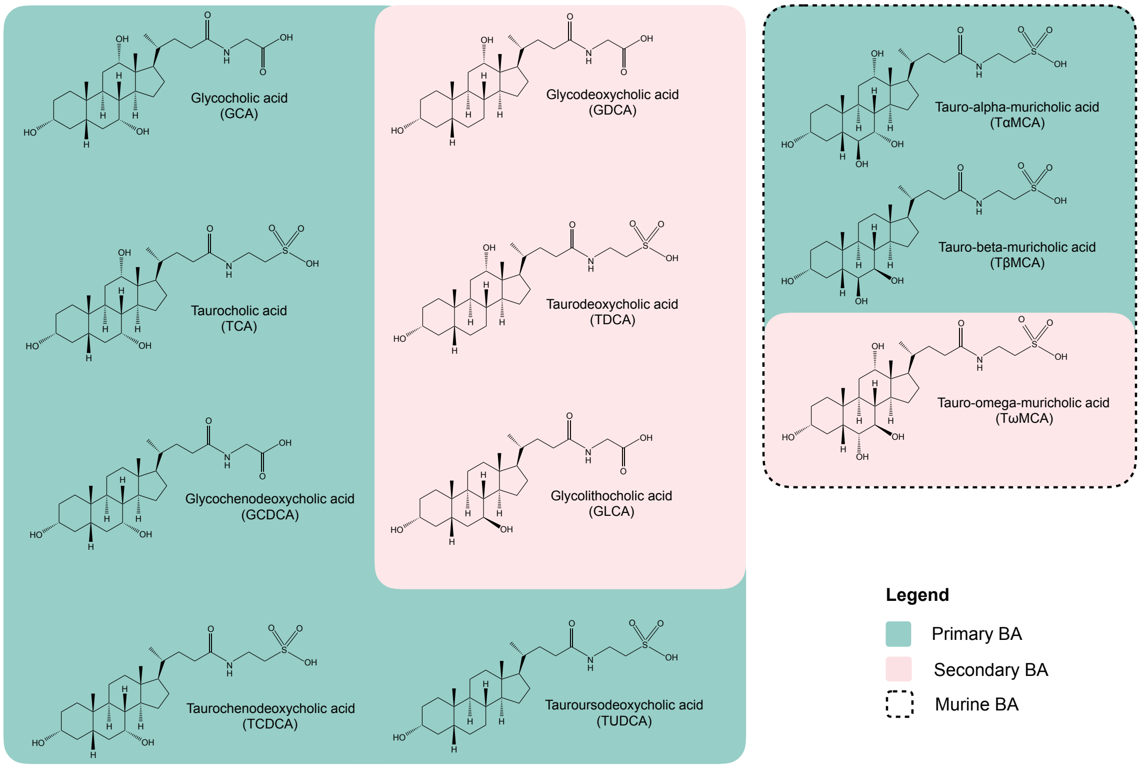

Fig. S1. Bile acid structures and abbreviations referenced in this study.

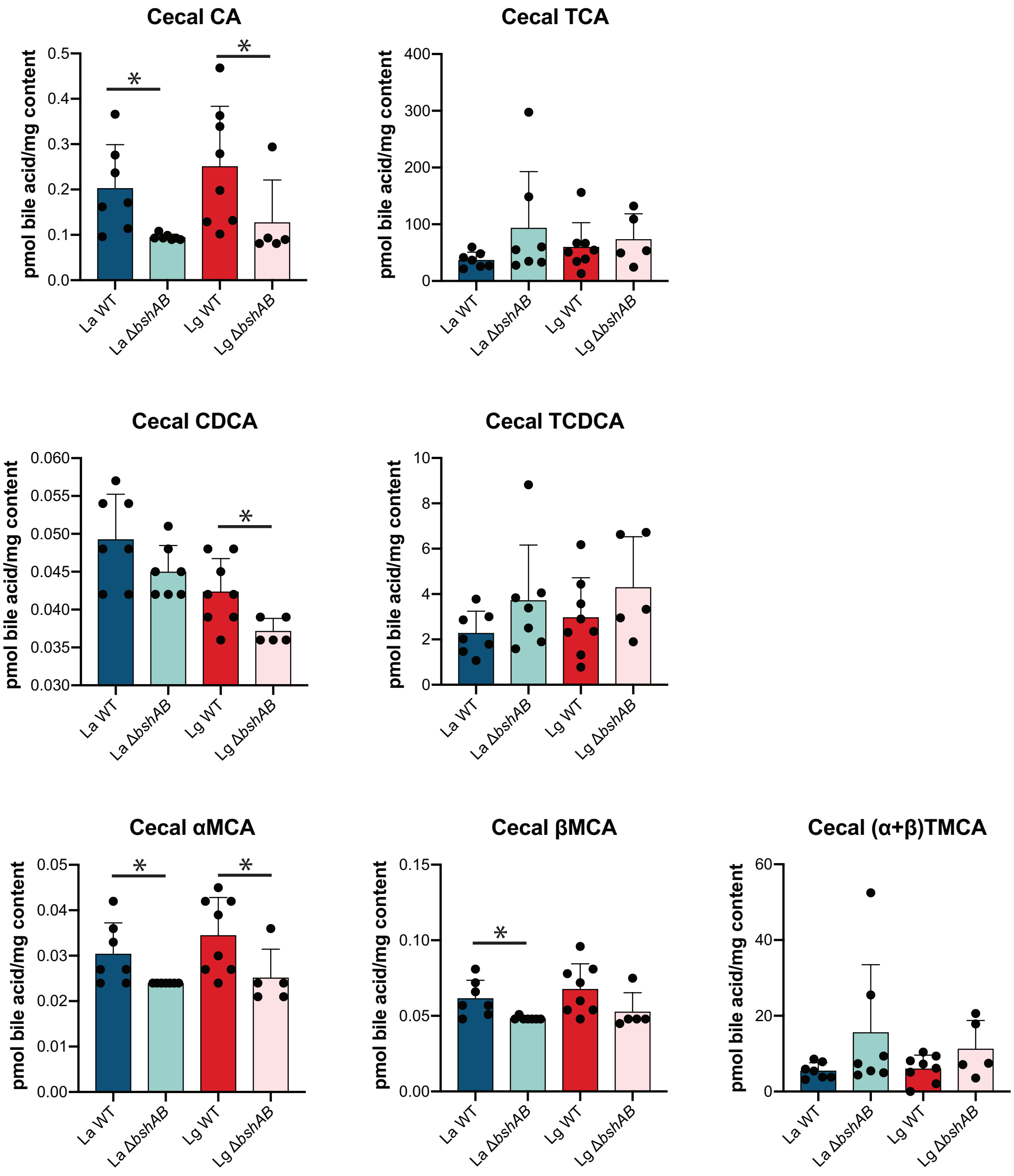

Fig. S2. Targeted bile acid metabolomics of *Lactobacillus*-associated mice cecum. Cecal BA concentrations from germ free C57BL/6J mice colonized with WT (n=7) or Δ*bshAB* (n=7) *L. acidophilus*, and WT (n=8) or Δ*bshAB* (n=5) *L. gasseri*. Asterisks represent significant (p < 0.05) differences between groups by Mann-Whitney test.

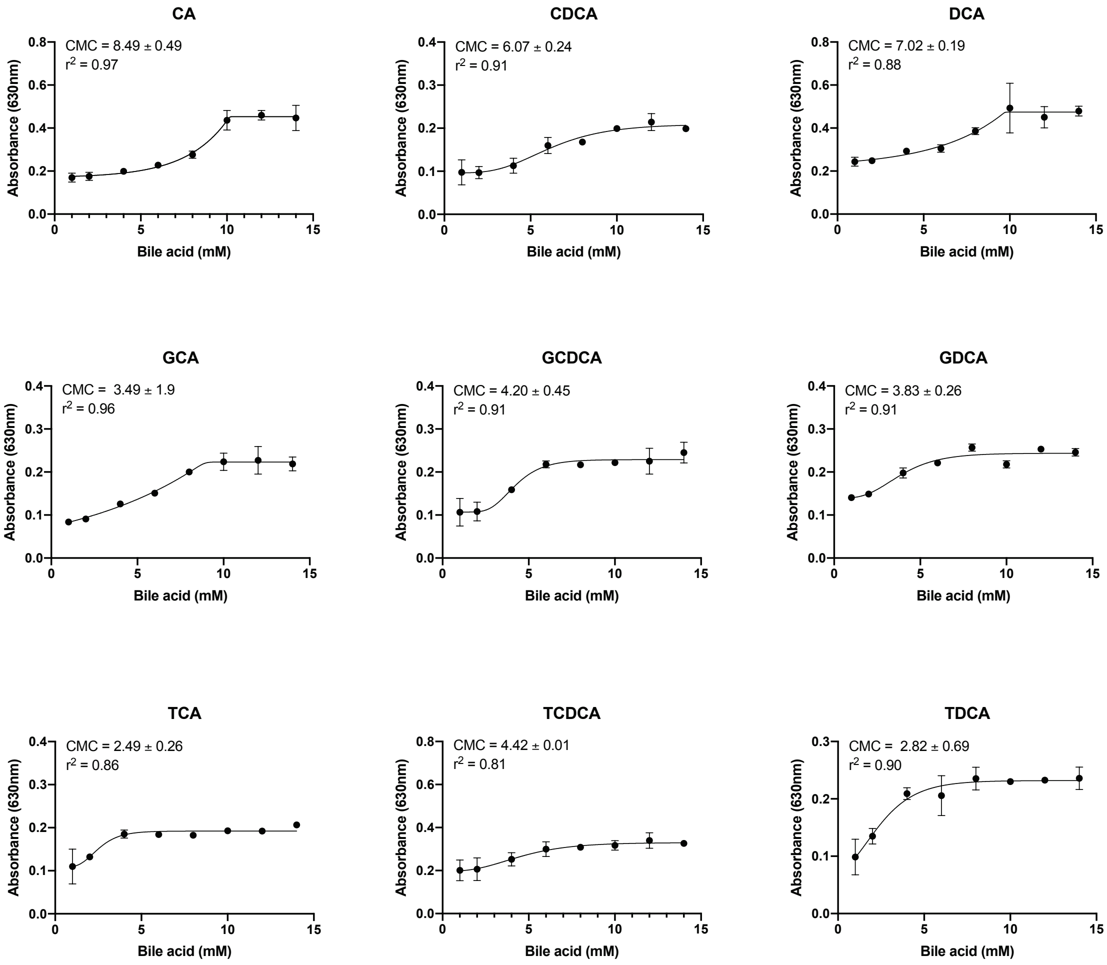

Fig. S3. Critical Micelle Concentrations (CMCs) of bile acids used in this study. CMCs were determined using Optimizer-BlueBALLS. Absorbance data from two independent experiments was plotted against BA concentrations for each molecule. A standard five-parameter logistic curve fit was performed using Graphpad Prism, and the CMC was determined by calculating the inflection point of each curve represented by the $Log_{10}EC50$. Inflection point values represent mean ± s.e.m n=2 replicates.

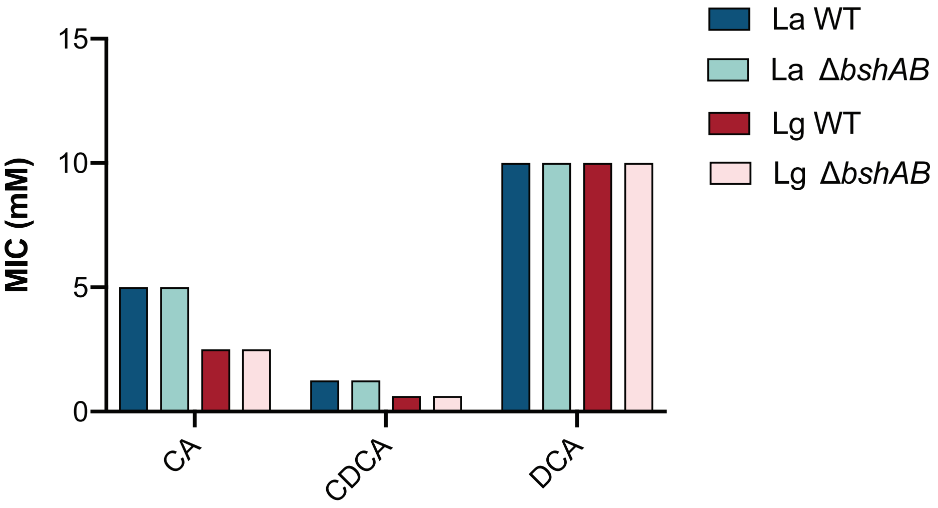

Fig. S4. WT and Δ*bshAB* deconjugated bile acid MICs. Bars represent mean MICs from n=3 independent experiments. MICs did not vary between experiments.

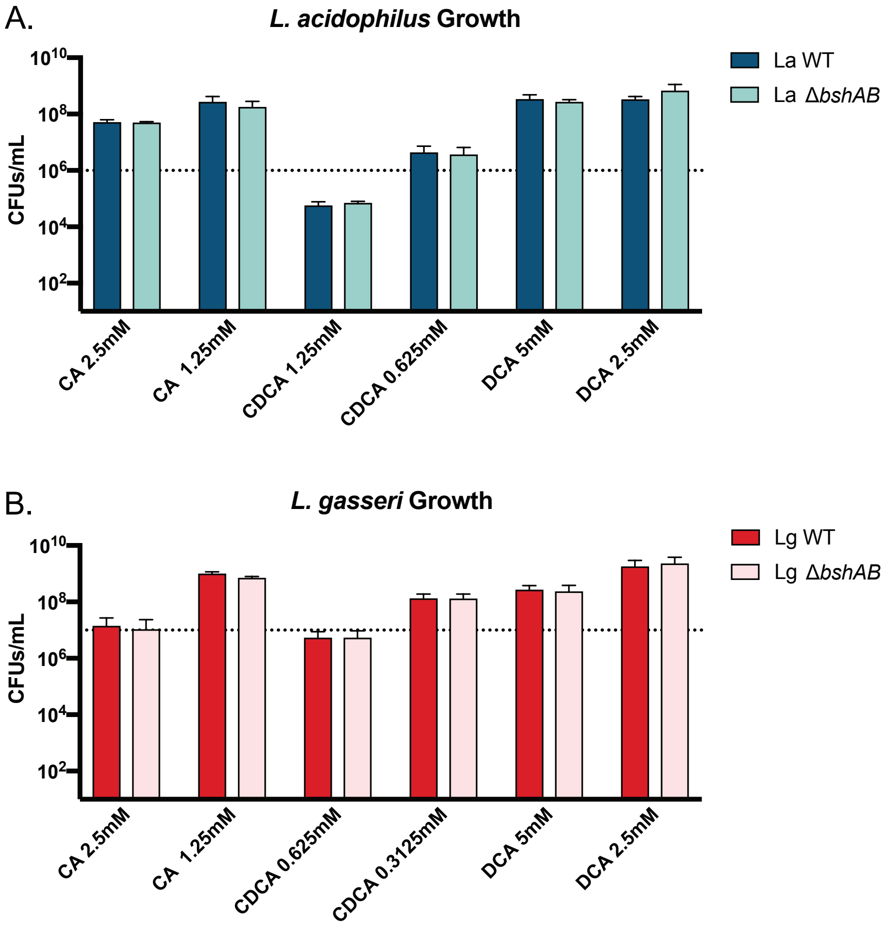

Fig. S5. BSH activity does not affect growth in the presence of deconjugated bile acids. (A) *L. acidophilus* and (B) *L. gasseri* WT and Δ*bshAB* strains grown with CA, CDCA, and DCA. Bars represent mean CFUs and error bars represent s.d. from n=4 independent experiments. Dashed lined denotes the approximate starting CFUs/mL at 0h.

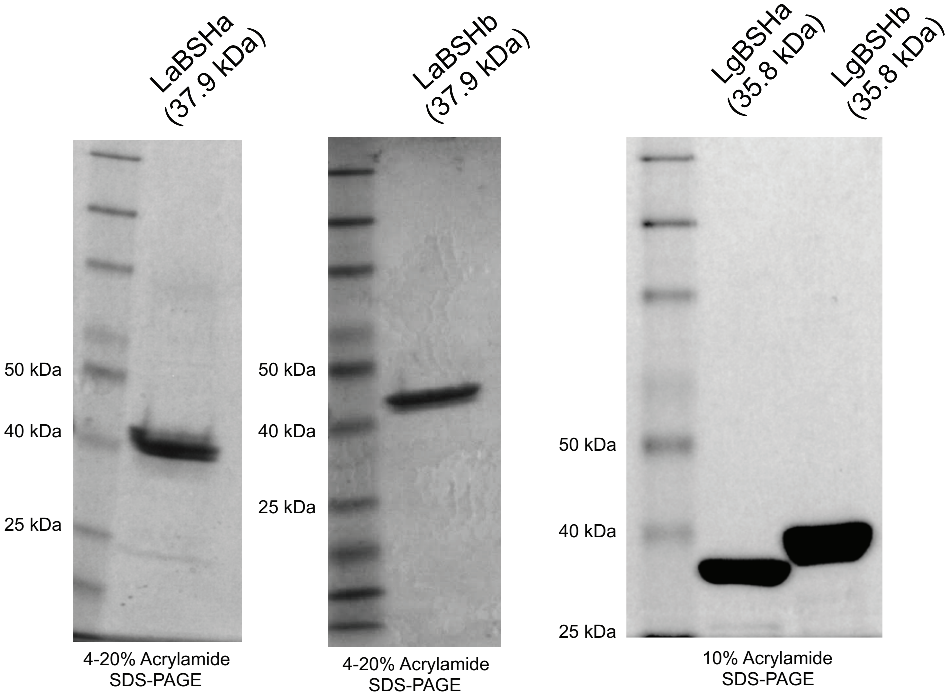

Fig. S6. SDS-PAGE of purified *Lactobacillus* BSHs. Molecular weights for each protein are labeled at the top of each gel.

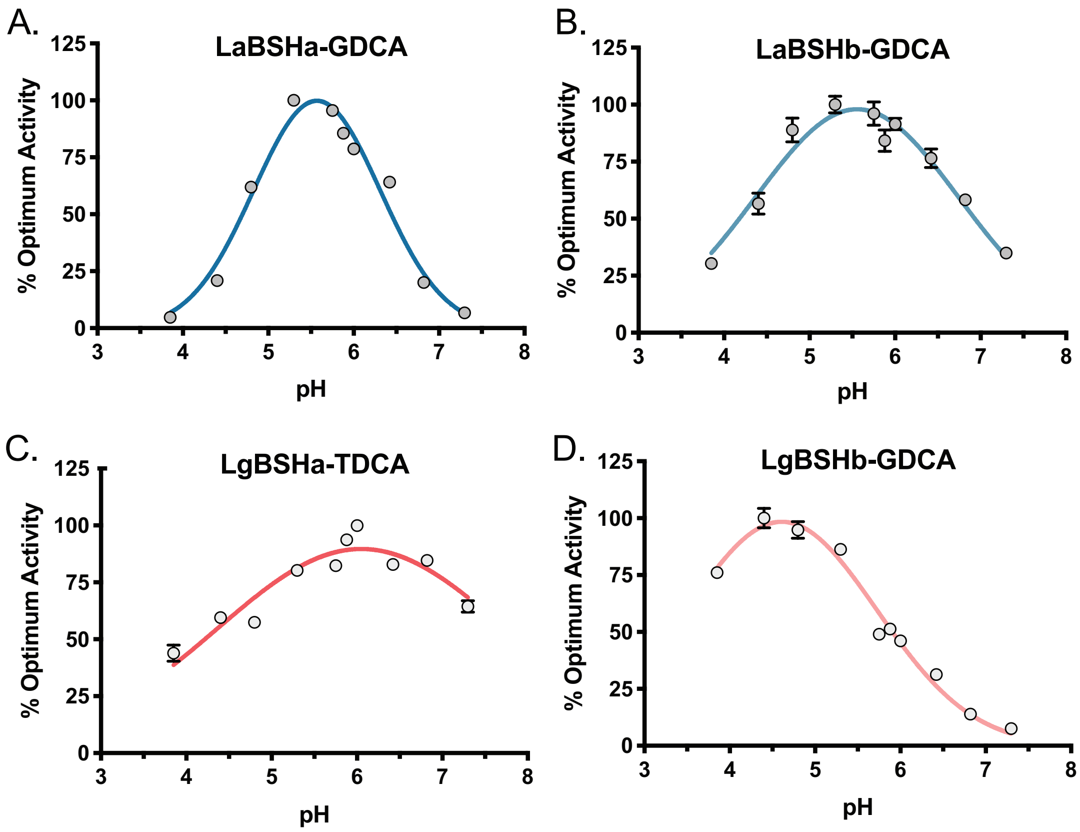

Fig. S7. pH optima of *Lactobacillus* BSH. Buffer conditions ranging from pH 3.5–7.5 were assayed with each BSH using the substrate with highest activity. Citrate buffer was used to buffer pHs >6 and phosphate was used to buffer pHs <6. Error bars represent s.d.

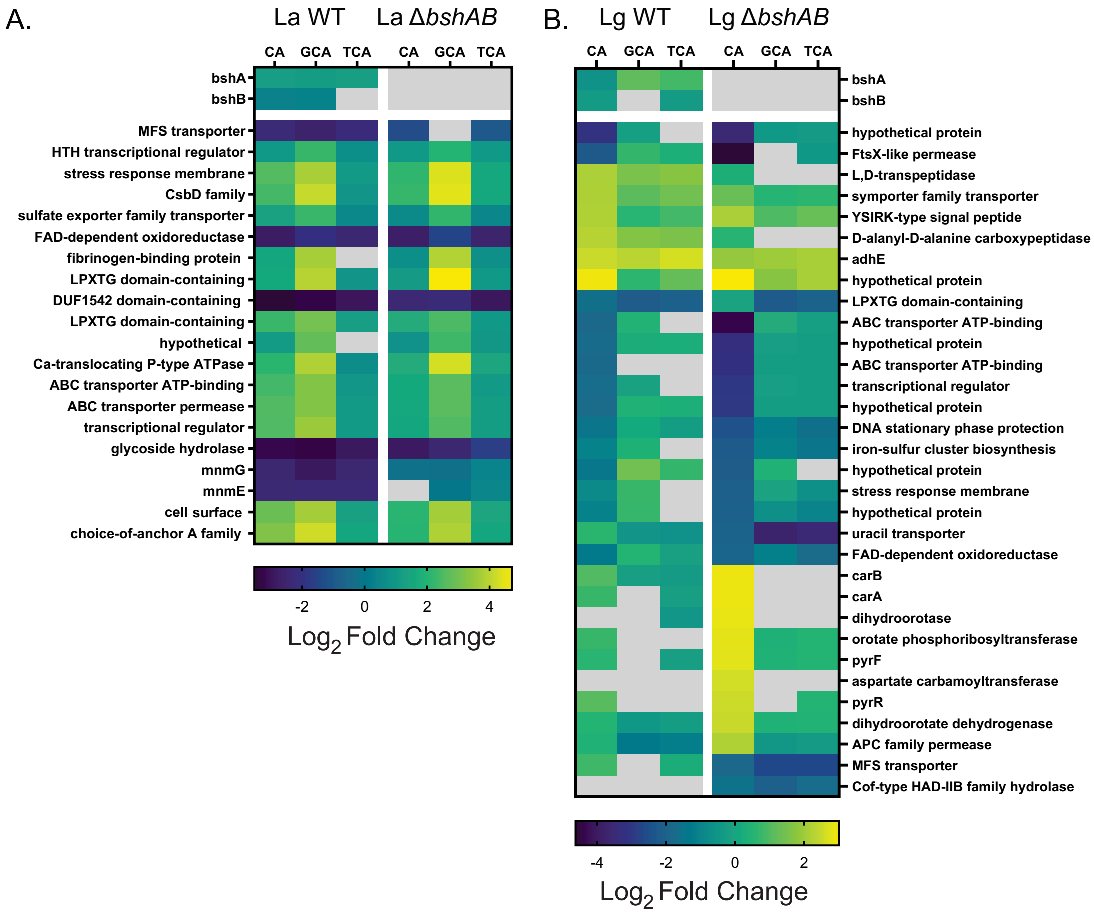

Fig. S8. Heatmap of differential gene expression during growth with various bile acids (A,B) Differential expression for each gene in response to CA, GCA, and TCA. The scale represents Log2 ratios and exact values are shown in Table S1. In addition, the Log$_2$ ratios for each of the *bsh* genes is shown at the top of the heat map. The gray color cells indicate no significant differential expression.

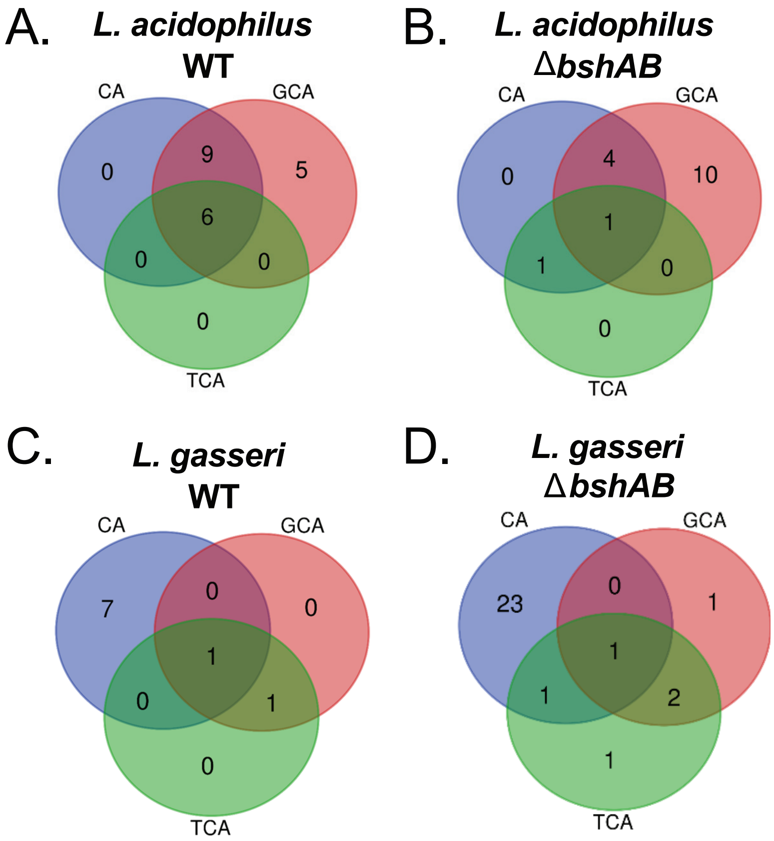

Fig. S9. Venn diagram summary of differential gene expression. Venn diagrams showing the number of genes and the common genes differentially expressed under growth in each condition (CA, GCA or TCA compared to MRS medium for (A) *L. acidophilus* NCK1909 (WT) and *L. acidophilus* NCK2523 (Δ*bshAB)*, (B) *L. gasseri* NCK2253 (WT) and *L. gasseri* NCK2680 (Δ*bshAB*). For the Venn diagram genes have a Log2 ratio ≤ -2 or ≥ 2 and a p value <0.05 and gene details are found in Table S1.

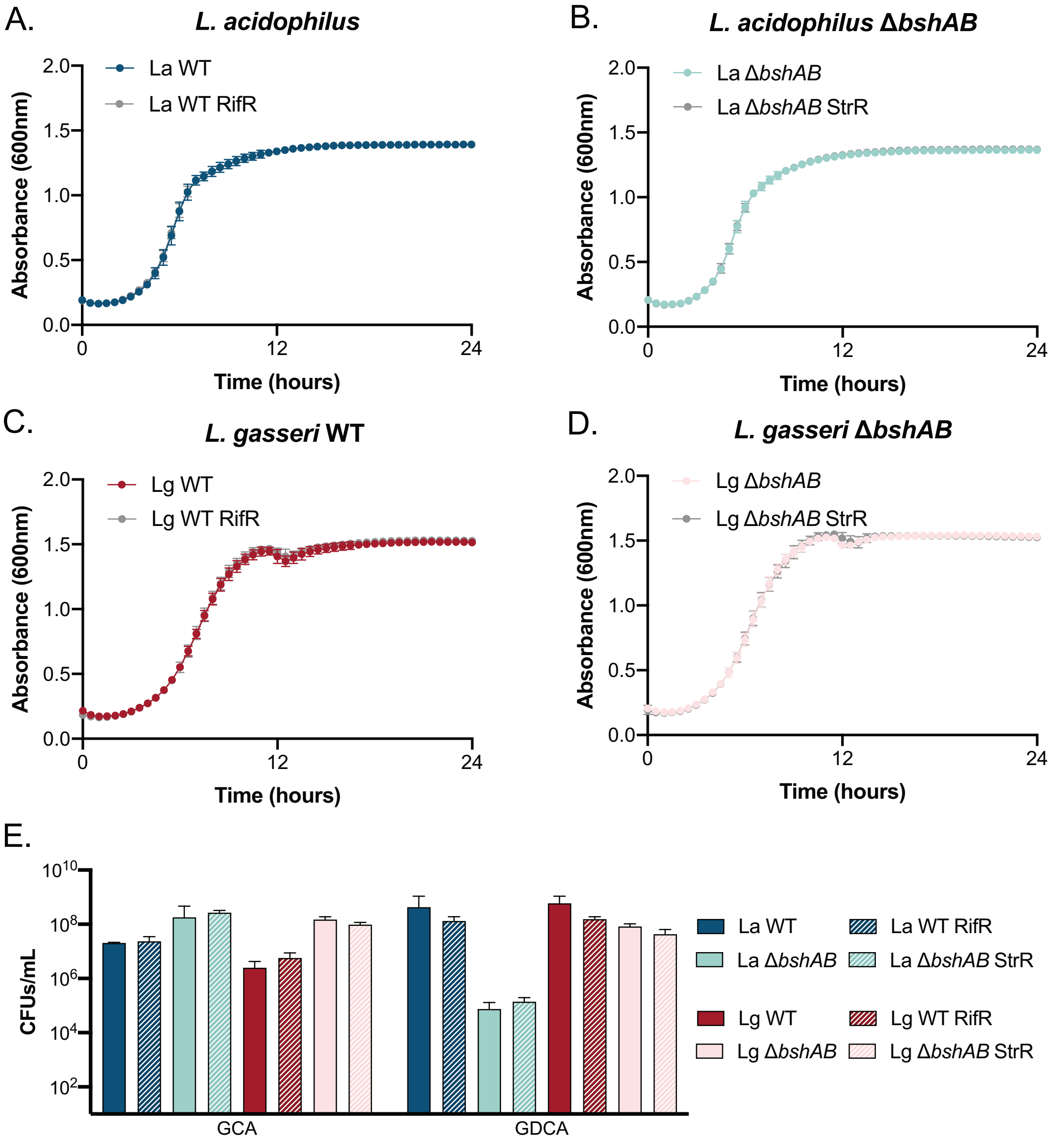

Fig. S10. Rifampicin and streptomycin-resistant strains are phenotypically similar to their isogenic parent strains. Spontaneous rifampicin resistance mutants (RifR) of (A) WT *L. acidophilus* or (C) WT *L. gasseri* and streptomycin resistance mutants (StrR) of (B) Δ*bshAB L. acidophilus* or (D) Δ*bshAB L. gasseri* were selected for by plating ~10^{10} bacteria on MRS agar containing 100 µg/mL rifampicin and 500 µg/mL streptomycin, respectively. Growths were carried out in MRS media in an anaerobic plate reader in replicate n=5. (E). *L. acidophilus* and *L. gasseri* spontaneous antibiotic-resistant mutants grown for 24 h in 5 mM GCA or 2.5 mM GCDCA. Error bars represent s.d. from n=3 independent experiments.

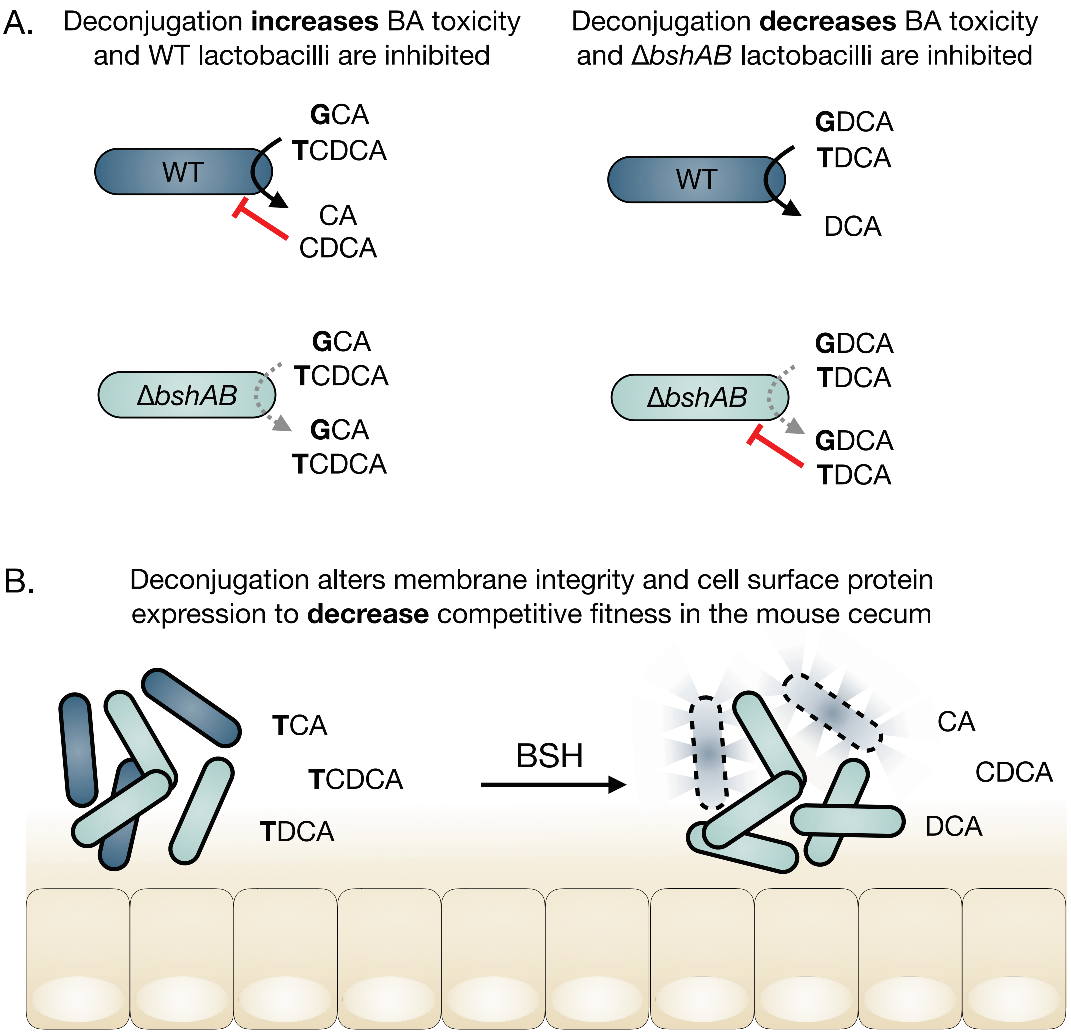

Fig. S11. Model of BSH-governed *Lactobacillus* fitness and colonization. *L. acidophilus* is depicted here as a representative *Lactobacillus* species. (A) On the left, our *in vitro* experiments demonstrate how WT *L. acidophilus* (portrayed as dark blue) can generate deconjugated BAs that are more toxic than their conjugated forms, resulting in growth inhibition. Δ*bshAB* (portrayed as light blue) cannot deconjugate these BAs thereby limiting exposure to toxic deconjugated BAs. On the right, we also demonstrate how WT *L. acidophilus* can generate deconjugated BAs that are less toxic than their conjugated forms. Δ*bshAB* cannot deconjugate these BAs thereby resulting in growth inhibition. The toxicity of a given conjugated/deconjugated BA is species-specific and deconjugation is dependent on the substrate preferences of the BSH. (B) *in/ex vivo* competitions demonstrate that Δ*bshAB is* more competitive than WT when exposed to the murine cecal BA pool. Deconjugated BAs in this niche may be more toxic to *Lactobacillus,* explaining why BSH activity is detrimental. Additionally, deconjugated BA exposure may alter the membrane integrity and transcriptome of *Lactobacillus in vivo* differently based on the expression of a BSH, and we hypothesize that these changes can also govern competitive fitness.

L. acidophilus wild type

CA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS07885	DUF1542 domain-containing protein CDS	-3.53	0
LBA_RS09100	glycoside hydrolase family 65 protein CDS	-3.28	1.00E-250
LBA_RS06090	FAD-dependent oxidoreductase CDS	-2.83	2.50E-256
LBA_RS09595	mnmG CDS	-2.52	0
LBA_RS09600	mnmE CDS	-2.44	0
LBA_RS00225	SLC45 family MFS transporter CDS	-2.38	8.27E-58
LBA_RS08260	calcium-translocating P-type ATPase, PMCA-type CDS	2.03	9.28E-292
LBA_RS08110	LPXTG cell wall anchor domain-containing protein CDS	2.22	8.65E-182
LBA_RS08960	ABC transporter ATP-binding protein CDS	2.30	0
LBA_RS04455	CsbD family protein CDS	2.34	0
LBA_RS08970	TetR/AcrR family transcriptional regulator CDS	2.44	1.73E-239
LBA_RS08965	ABC transporter permease CDS	2.48	0
LBA_RS04450	GlsB/YeaQ/YmgE family stress response membrane protein CDS	2.54	0
LBA_RS09685	cell surface protein CDS	2.80	1.30E-52
LBA_RS09690	choice-of-anchor A family protein CDS	3.12	3.60E-68

Venn Diagram

Names	total	locus_tag
CA GCA TCA	6	LBA_RS09100
		LBA_RS09595
		LBA_RS06090
		LBA_RS09600
		LBA_RS00225
		LBA_RS07885
CA GCA	9	LBA_RS08970
		LBA_RS04455
		LBA_RS08960
		LBA_RS09685
		LBA_RS08260
		LBA_RS08110
		LBA_RS08965
		LBA_RS09690
		LBA_RS04450
GCA	5	LBA_RS08125
		LBA_RS07330
		LBA_RS05625
		LBA_RS07335
		LBA_RS04415

GCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS09100	glycoside hydrolase family 65 protein CDS	-3.41	7.44E-249
LBA_RS07885	DUF1542 domain-containing protein CDS	-3.37	0
LBA_RS09595	mnmG CDS	-2.92	0
LBA_RS00225	SLC45 family MFS transporter CDS	-2.69	6.87E-71
LBA_RS09600	mnmE CDS	-2.41	1.67E-263
LBA_RS06090	FAD-dependent oxidoreductase CDS	-2.05	9.75E-147
LBA_RS04415	helix-turn-helix transcriptional regulator CDS	2.20	5.02E-21
LBA_RS05625	putative sulfate exporter family transporter CDS	2.21	1.35E-144
LBA_RS08125	hypothetical protein CDS	2.71	5.23E-25
LBA_RS08110	LPXTG cell wall anchor domain-containing protein CDS	2.97	0
LBA_RS08960	ABC transporter ATP-binding protein CDS	3.14	0
LBA_RS08965	ABC transporter permease CDS	3.16	0
LBA_RS08970	TetR/AcrR family transcriptional regulator CDS	3.48	0
LBA_RS09685	cell surface protein CDS	3.61	8.10E-104
LBA_RS07330	fibrinogen-binding protein CDS	3.63	2.02E-48
LBA_RS04450	GlsB/YeaQ/YmgE family stress response membrane protein CDS	3.67	0
LBA_RS08260	calcium-translocating P-type ATPase, PMCA-type CDS	3.76	0
LBA_RS07335	LPXTG cell wall anchor domain-containing protein CDS	3.85	2.16E-237
LBA_RS04455	CsbD family protein CDS	4.05	0
LBA_RS09690	choice-of-anchor A family protein CDS	4.15	6.90E-146

TCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS07885	DUF1542 domain-containing protein CDS	-3.06	0
LBA_RS09100	glycoside hydrolase family 65 protein CDS	-2.93	4.35E-58
LBA_RS06090	FAD-dependent oxidoreductase CDS	-2.52	5.51E-238
LBA_RS09595	mnmG CDS	-2.51	0
LBA_RS09600	mnmE CDS	-2.41	0
LBA_RS00225	SLC45 family MFS transporter CDS	-2.31	4.72E-61

L. acidophilus ΔbshAB

CA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS09100	glycoside hydrolase family 65 protein CDS	-2.91	1.08E-182
LBA_RS06090	FAD-dependent oxidoreductase CDS	-2.78	6.00E-273
LBA_RS07885	DUF1542 domain-containing protein CDS	-2.47	0
LBA_RS09685	cell surface protein CDS	2.01	1.73E-28
LBA_RS04450	GlsB/YeaQ/YmgE family stress response membrane protein CDS	2.06	4.75E-270
LBA_RS04455	CsbD family protein CDS	2.17	2.41E-256

Venn Diagram

Names	total	locus_tag
CA GCA TCA	1	LBA_RS07885
CA GCA	4	LBA_RS09100
		LBA_RS04455
		LBA_RS04450
		LBA_RS09685
CA TCA	1	LBA_RS06090
GCA	10	LBA_RS08110
		LBA_RS08125
		LBA_RS08260
		LBA_RS07330
		LBA_RS05625
		LBA_RS08960
		LBA_RS07335
		LBA_RS08965

GCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS09100	glycoside hydrolase family 65 protein CDS	-2.51	2.28E-126
LBA_RS07885	DUF1542 domain-containing protein CDS	-2.34	0
LBA_RS05625	putative sulfate exporter family transporter CDS	2.10	6.52E-153

LBA_RS08125	hypothetical protein CDS	2.28	3.52E-22			LBA_RS09690
LBA_RS08110	LPXTG cell wall anchor domain-containing protein CDS	2.41	8.44E-237			LBA_RS08970
LBA_RS08970	TetR/AcrR family transcriptional regulator CDS	2.48	7.14E-264			
LBA_RS08960	ABC transporter ATP-binding protein CDS	2.59	0			
LBA_RS08965	ABC transporter permease CDS	2.60	0			
LBA_RS09685	cell surface protein CDS	3.60	2.19E-126			
LBA_RS09690	choice-of-anchor A family protein CDS	3.74	1.37E-155			
LBA_RS07330	fibrinogen-binding protein CDS	3.80	6.39E-83			
LBA_RS08260	calcium-translocating P-type ATPase, PMCA-type CDS	4.21	0			
LBA_RS04450	GlsB/YeaQ/YmgE family stress response membrane protein CDS	4.30	0			
LBA_RS04455	CsbD family protein CDS	4.37	0			
LBA_RS07335	LPXTG cell wall anchor domain-containing protein CDS	4.71	0			

TCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LBA_RS07885	DUF1542 domain-containing protein CDS	-2.99	0
LBA_RS06090	FAD-dependent oxidoreductase CDS	-2.62	1.77E-241

L. gasseri wild type

CA V MRS

				Venn Diagram Names	total	locus_tag
locus_tag	Name	DE Log2 Ratio	DE p-value	CA GCA TCA	1	LGAS_RS07675
LGAS_RS01970	hypothetical protein CDS	-3.17	2.85E-57	GCA TCA	1	LGAS_RS05250
LGAS_RS08330	FtsX-like permease family protein CDS	-2.31	5.00E-55	CA	7	LGAS_RS08145
LGAS_RS08285	L,D-transpeptidase family protein CDS	2.09	1.49E-139			LGAS_RS08165
LGAS_RS08165	cation:dicarboxylase symporter family transporter CDS	2.09	0			LGAS_RS08140
LGAS_RS08140	YSIRK-type signal peptide-containing protein CDS	2.14	3.41E-136			LGAS_RS08290
LGAS_RS08290	D-alanyl-D-alanine carboxypeptidase CDS	2.21	3.50E-66			LGAS_RS01970
LGAS_RS07675	adhE CDS	2.42	0			LGAS_RS08330
LGAS_RS08145	hypothetical protein CDS	2.88	1.05E-185			LGAS_RS08285

GCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LGAS_RS05250	LPXTG cell wall anchor domain-containing protein CDS	-2.27	1.53E-165
LGAS_RS07675	adhE CDS	2.25	4.04E-231

TCA V MRS

locus_tag	Name	DE Log2 Ratio	DE p-value
LGAS_RS05250	LPXTG cell wall anchor domain-containing protein CDS	-2.10	4.50E-296
LGAS_RS07675	adhE CDS	2.56	0

L. gasseri ΔbshAB

CA V MRS

				Venn Diagram Names	total	locus_tag
locus_tag	Name	DE Log2 Ratio	DE p-value	CA GCA TCA	1	LGAS_RS02410
LGAS_RS08330	FtsX-like permease family protein CDS	-4.63	0	CA TCA	1	LGAS_RS08145
LGAS_RS08335	ABC transporter ATP-binding protein CDS	-4.45	0	GCA TCA	2	LGAS_RS05250
LGAS_RS01970	hypothetical protein CDS	-3.64	3.00E-98			LGAS_RS02575
LGAS_RS05435	hypothetical protein CDS	-3.36	0	CA	23	LGAS_RS08140
LGAS_RS05440	ABC transporter ATP-binding protein CDS	-3.25	0			LGAS_RS05340
LGAS_RS05445	GntR family transcriptional regulator CDS	-3.11	6.69E-224			LGAS_RS02730
LGAS_RS05430	hypothetical protein CDS	-3.09	3.82E-284			LGAS_RS04770
LGAS_RS02730	DNA starvation/stationary phase protection protein CDS	-2.49	1.33E-233			LGAS_RS00160
LGAS_RS00380	iron-sulfur cluster biosynthesis family protein CDS	-2.27	4.34E-219			LGAS_RS08315
LGAS_RS08315	hypothetical protein CDS	-2.26	1.13E-31			LGAS_RS05440
LGAS_RS00160	GlsB/YeaQ/YmgE family stress response membrane protein CDS	-2.13	1.41E-192			LGAS_RS05370
LGAS_RS02735	hypothetical protein CDS	-2.10	2.70E-98			LGAS_RS00380
LGAS_RS02410	uracil transporter CDS	-2.09	8.97E-174			LGAS_RS08330
LGAS_RS04770	FAD-dependent oxidoreductase CDS	-2.01	3.99E-133			LGAS_RS05360
LGAS_RS08145	hypothetical protein CDS	3.01	2.59E-120			LGAS_RS05435
LGAS_RS05345	carB CDS	2.83	1.04E-299			LGAS_RS05375
LGAS_RS05350	carA CDS	2.82	9.34E-161			LGAS_RS08335
LGAS_RS05355	dihydroorotase CDS	2.79	1.59E-136			LGAS_RS01970
LGAS_RS05380	orotate phosphoribosyltransferase CDS	2.72	9.69E-133			LGAS_RS05380
LGAS_RS05375	pyrF CDS	2.69	9.59E-123			LGAS_RS05350
LGAS_RS05360	aspartate carbamoyltransferase CDS	2.53	1.16E-80			LGAS_RS05430
LGAS_RS05340	pyrR CDS	2.46	4.63E-95			LGAS_RS02735
LGAS_RS05370	dihydroorotate dehydrogenase CDS	2.42	2.17E-120			LGAS_RS05345

LGAS_RS09250	APC family permease CDS	2.13	4.33E-204			LGAS_RS05445
LGAS_RS08140	YSIRK-type signal peptide-containing protein CDS	2.07	6.63E-106			LGAS_RS09250
						LGAS_RS05355
GCA V MRS				GCA	1	LGAS_RS02395
LGAS_RS02410	uracil transporter CDS	-3.80	0	TCA	1	LGAS_RS07675
LGAS_RS02575	MFS transporter CDS	-2.77	0			
LGAS_RS05250	LPXTG cell wall anchor domain-containing protein CDS	-2.29	0			
LGAS_RS02395	Cof-type HAD-IIB family hydrolase CDS	-2.13	2.85E-44			
TCA V MRS						
LGAS_RS02410	uracil transporter CDS	-3.61	0			
LGAS_RS02575	MFS transporter CDS	-2.76	0			
LGAS_RS05250	LPXTG cell wall anchor domain-containing protein CDS	-2.07	1.78E-251			
LGAS_RS07675	adhE CDS	2.05	3.80E-254			
LGAS_RS08145	hypothetical protein CDS	2.02	4.52E-53			

Table S1. Table of differentially regulated *Lactobacillus* genes

Strains:		
E. coli	**Details**	**Reference**
EC101	RepA+ JM101; repA from pWV01 integrated in chromosome; cloning host for pORI-based plasmids	Law et al 1995
EC1000	RepA+ MC1000, repA from pWV01 integrated in chromosome; cloning host for pORI-based plasmids	Leenhouts et al 1996
NCK1391	Host for pTRK669	Russell and Klaenhammer 2001
NCK1911	Host for pTRK935	Goh et al 2009
NCK2519	Host for pTRK1123 (EC1000)	this study
NCK2520	Host for pTRK1124 (EC1000)	this study
NCK2714	Host for pTRK1206 (EC101)	this study
NCK2715	Host for pTRK1207 (EC101)	this study
L. acidophilus		
NCK1909	Δ*upp* (control strain; background strain for *upp*-based counterselective gene replacement)	Goh et al 2009
NCK1910	NCK1909 harboring pTRK669; host for pORI-based counterselective integration plasmids	Goh et al 2009
NCK2521	NCK1909 with Δ*bshA*	this study
NCK2522	NCK1909 with Δ*bshB*	this study
NCK2523	NCK1909 with Δ*bshA* and Δ*bshB*	this study
L. gasseri		
NCK2253	Δ*upp* (control strain; background strain for *upp*-based counterselective gene replacement)	Selle et al 2014
NCK2254	NCK2253 harboring pTRK669; host for pORI-based counterselective integration plasmids	Selle et al 2014
NCK2678	NCK2253 with Δ*bshA*	this study
NCK2679	NCK2253 with Δ*bshB*	this study
NCK2680	NCK2253 with Δ*bshA* and Δ*bshB*	this study
Plasmids:		
pTRK669	Ori (pWV01), RepA+	Russell and Klaenhammer 2001
pTRK935	Counterselective integration vector with a upp expression cassette	Goh et al 2009
pTRK1123	pTRK935-based plasmid with flanking regions to facilitate deletion of *bshA* from NCK1909	this study
pTRK1124	pTRK935-based plasmid with flanking regions to facilitate deletion of *bshb* from NCK1909	this study
pTRK1206	pTRK935-based plasmid with flanking regions to facilitate deletion of *bshA* from NCK2253	this study
pTRK1207	pTRK935-based plasmid with flanking regions to facilitate deletion of *bshb* from NCK2253	this study
pETite-LaBSHa	pETite C-his plasmid containing LaBSHa	this study
pETite-LaBSHb	pETite C-his plasmid containing LaBSHb	this study
pETite-LgBSHa	pETite C-his plasmid containing LgBSHa	this study
pETite-LgBSHb	pETite C-his plasmid containing LgBSHb	

Table S2. Strains and plasmids used in this study

Primer name	**Primer sequence (5' to 3')**
L. acidophilus	
La_BshA_SOEA_F	GATC**GGATCC**GATGGAACGATTCCCACTTCC
La_BshA_SOEA_R	GGGACTGAATATAATTGATGTACACAT
La_BshA_SOEB_F	GTACATCAATTATATTCAGTCCCCATCAAAACTAAAACAATTC
La_BshA_SOEB_R	GATC**GAGCTC**CATCTGAAGGATAATGTGC
La_BshB_SOEA_F	GATC**GGATCC**GAAGGATCTGATGCTGATGCAG
La_BshB_SOEA_R	TGATGTACACATTTACTGCACGC
La_BshB_SOEB_F	CGTGCAGTAAATGTGTACATCAGAATATATGAATTAATTTTATATAC
La_BshB_SOEB_R	GATC**GAGCTC**GGATGATACGAATTAGTGAGATAAGTACC
La_Scr_BshA_F	CAAGTCTTACAATCATGACAAAGG
La_Scr_BshA_R	GGAACCACTTGCTAAGGCAC
La_Scr_BshB_F	GCGCATGCTATTGCAAATAGTG
La_Scr_BshB_R	CTGATAGGCGAACACATCCGTG
L. gasseri	
Lg_BshA_SOEA_F	GATC**GGATCC**CGTCGTTTCTGATGCTGAATTAGCTG
Lg_BshA_SOEA_R	CTTAAACCAGTACACATTTTAAAATCCTCC
Lg_BshA_SOEB_F	GGATTTTAAAATGTGTACTGGTTTAAGGTTAGACGATGAAACGATGAACGCTGAC
Lg_BshA_SOEB_R	GATC**GAGCTC**GCACTAGGAATTTCAGTAGTGGC
Lg_BshB_SOEA_F	GATC**GGATCC**AGACTACCTACTTGACGCAG
Lg_BshB_SOEA_R	TGATGTACACATATTAGATTTCTCCTC
Lg_BshB_SOEB_F	GAAATCTAATATGTGTACATCAAGTTTTATAAATTAAAATTAGCTAGAAATTTGTATC
Lg_BshB_SOEB_R	GATC**GAGCTC**CTGCTTCAAACTTTGAATGCG
Lg_Scr_BshA_F	CACTTGACCTCAATGGCACC
Lg_Scr_BshA_R	CGCTTTCATGACATATATTC
Lg_Scr_BshB_F	CATCTTGATTGACGAAGGAAC
Lg_Scr_BshB_R	CGACATTGGTTAATTGAAAGTTATC
Plasmid	
upp_ScF	GGGCTGGCTTAACTATGC
upp_ScR	CTTCCGGCTCGTATGTTG

* Restriction sites are shown in bold

Rosetta *E. coli*	
LaBSHa F	GAAGGAGATATACATATGTGTACATCAATTATATTCAGTCCCAAAGATCATTAC
LaBSHa R	GTGATGGTGGTGATGATGGTTTTGATGGTTAAATTTAGTTTTATCAAGCATATCATAAGTG
LaBSHb F	GAAGGAGATATACATATGTGTACCAGCATCTGTTATAATCCAAATGACC
LaBSHb R	GTGATGGTGGTGATGATG GTTCATATACTCAATGTCTTGTTTCTTAAAAAGATCATAGCA
LgBSHa F	GAAGGAGATATACATATGTGTACTGGTTTAAGATTTACAGATGATCAAGG
LgBSHa R	GTGATGGTGGTGATGATGATAGGTGATTAGCTTGTCAGCGTTCATCG
LgBSHb F	GAAGGAGATATACATATGTGTACATCAATTTTATATAGTCCCAAAGACCAC
LgBSHb R	GTGATGGTGGTGATGATGATTTATAAAACTAATATCCTGTTTCTTGAACAAGTCAAACACG
Synthesized Genes	
Codon optimized sequence of LaBSHb	ATG TGT ACC AGC ATC TGT TAT AAT CCA AAT GAC CAT TAT TTC GGT CGT AAT CTT GAT TAT GAA ATT GCC TAC GGG CAA AAA GTT GTA ATT GTA CCG CGT AAC TAT GAG TTT AAG TAT CGT GAG ATG CCT AGT CAG AAA ATG CAC TAC GCC TTT ATT GGT GTC AGT GTT GTC AAC GAT GAC TAC CCC CTG CTG TGT GAC GCC ATT AAC GAG AAG GGC CTT GGA ATT GCG GGA TTG AAC TTT CAG GGT CCC AAT CAT TAC TTC CCG AAG ATC GAA GGG AAA AAA AAT ATT GCG TCG TTC GAG CTG ATG CCT TAC TTA CTG TCA AAT TGT GAA AAT ACA GAC GAC GTT AAA GAA ATC CTT GAC AAC GCC AAC ATC TTG AAC ATC AGT TTC AGC GCT AAC TAC CCC GCC GCT GAT CTG CAT TGG ATT CTT TCT GAT AAA GCC GGC AAA AGC ATC GTA GTC GAA TCC ACG AAT AGC GGG CTG CAT ATT TAC GAC AAC CCA GTT AAT GTC TTG ACG AAC AAC CCT GAG TTC CCT GAC CAG TTG ATT AAG TTG TCA GAT TAT GCC GAC GTG ACT CCA CAC AAC CCC AAG AAT ACT CTG GTT CCA AAT GTT GAC CTT AAT TTA TAC AGC CGT GGT CTT GGC ACA CAT CAC TTA CCC GGT GGG ATG GAT TCA TCC TCC CGT TTC GTG AAA GTC GCG TTC GTT CTG GCC CAC ACA CCT CAA GGG AAG AAC GAA GTT GAG AAT GTT ACA AAT TAT TTT CAC ATT CTG CAT AGT GTT GAG CAA CCC GAC GGC TTG GAC GAG GTT GAG GAC AAT CGT TAT GAA TAT ACG ATG TAC ACT GAT TGC ATG AAC CTT GAT AAA GGC ATT CTT TAT TTC ACA ACT TAT GAC AAT AAT CGT ATC AAT GCA GTT GAT ATG CAT AAA GCT GAT CTT GAC TCG GAG GAC CTG ATT TGC TAT GAT CTT TTT AAG AAA CAA GAC ATT GAG TAT ATG AAC

Table S3. Oligonucleotides used in this study

Table S2 References

1. Law J*, et al.* (1995) A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol* 177(24):7011-7108.

2. Leenhouts K*, et al.* (1996) A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol Gen Genet* 253(1-2):217-224.

3. Russell WM, Klaenhammer TR*,* (2009) Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl Environ Microbiol* 67(9):4361-4364.

4. Goh YJ*, et al.* (2009) Development and application of a upp-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75(10):3093-3105.

5. Selle K*, et al.* (2014) Development of an integration mutagenesis system in *Lactobacillus gasseri*. *Gut Microbes* 5(3):326-323.