Supplemental Material

The microbial *gbu* gene cluster links microbial L-carnitine catabolism to red meat diet-enhanced cardiovascular disease risk

Jennifer A. Buffa*[1,2], Kymberleigh A. Romano*[1,2], Matthew F. Copeland*[3], David B. Cody[3], Weifei Zhu[1,2], Rachel Galvez[3], Xiaoming Fu[1,2], Kathryn Ward[3], Marc Ferrell[1,2], Hong J. Dai[3], Sarah Skye[1,2,†], Ping Hu[3], Lin Li[1,2], Mirjana Parlov[3], Amy McMillan[1,2,‡], Xingtao Wei[3], Ina Nemet[1,2], Robert A. Koeth[1,2,4], Xinmin S. Li[1,2], Zeneng Wang[1,2], Naseer Sangwan[1,2], Adeline M. Hajjar[1,2], Mohammed Dwidar[1,2], Taylor L. Weeks[1], Nathalie Bergeron[5], Ronald M. Krauss[6], W.H. Wilson Tang[1,2,4], Federico E. Rey[7] Joseph A. DiDonato[1,2], Valentin Gogonea[8], G. Frank Gerberick[3], Jose Carlos Garcia-Garcia[3], and Stanley L. Hazen[1,2,4].

[1] Department of Cardiovascular & Metabolic Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH
[2] Center for Microbiome & Human Health, Cleveland Clinic, Cleveland, OH
[3] Life Sciences TPT, and Global Biosciences, The Procter & Gamble Company, Cincinnati, OH
[4] Department of Cardiovascular Medicine, Heart, Vascular and Thoracic Institute, Cleveland Clinic, Cleveland, OH
[5] Department of Biological and Pharmaceutical Sciences, College of Pharmacy, Touro University California, Vallejo, CA
[6] Departments of Pediatrics and Medicine, University of California, San Francisco, San Francisco, CA
[7] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706 USA
[8] Department of Chemistry, Cleveland State University, Cleveland, OH

* contributed equally

† current address: Abbott Structural Heart, Santa Clara, CA

‡ current address: Dose Biosystems Inc., Toronto, ON, Canada

Table of Contents for Supplemental Material

Supplemental Results………………………………………………………………………………………………...3

Supplemental Table 1: Baseline clinical characteristics…………………………………………..…………………5

Supplemental Table 2: CVD, CAD, PAD risk ratios……………….……………………………………………….....6

Supplemental Table 3: Functional annotation of *gbu* gene cluster members………………………………………7

Supplemental Table 4: Amino acids sequences used in Figure 4E……………………….………………………..…8

Supplemental Table 5: Gnotobiotic community member PCR primers...……………………………………………13

Supplemental Table 6: Plasmids used for heterologous gene expression in *E. coli*…...……………………………14

Supplemental Table 7: Composition of APPROACH baseline and intervention diets………………..…………….15

Supplemental Results

The schematic pathway presented in Extended Data Figure 3 suggests that γBB traverses the microbe's membrane through a carnitine/γBB antiporter and exchanges a CoA moiety, most 17 probably with an acetyl-CoA, after which it docks to the key catalytic polypeptide GbuA. This dual functioning enzyme in concerted reaction both generates TMA from γBB and creates a double bond resulting in crotonyl CoA. The combined γBB TMA lyase/dehydrogenase activity of GbuA seems to be redox neutral as it does not require an electron transport flavoprotein (ETF) as regular Acyl-CoA dehydrogenases do for their activity. The protein GbuD, identified by sequence homology as having a similar function to ubiquinone oxidoreductase (which has a 4Fe-4S cluster as a cofactor), seems to have an electron transport function, but is not required for TMA lyase activity (TMA generation) based on cloning and functional studies. The efflux of TMA from the microbe might go through the GbuE antiporter or through a yet unidentified (X) transporter. Alternatively, TMA has high diffusibility and may not require a dedicated transported to efflux the microbe. GbuF, identified by sequence homology as having the activity of a CoA-hydrolyze enzyme, likely further splits crotonyl-CoA into crotonate and CoA; crotonate ultimately might be released outside the microbe through an unidentified (e.g. Y) transporter.

The empirical results observed with combinatorial cloning and functional analyses of products generated / detected in media highlight the need for both encoded acyl-CoA transferase enzymes of the *gbu* gene cluster, GbuB and GbuC, to presumably activate the γBB substrate into the γ-butyrobetainyl-CoA intermediate (γBB CoA). To further ascertain the function of GbuA, a multiple sequence alignment of 18 biochemically characterized acyl-CoA dehydrogenase homologs to GbuA from *E. timonensis* revealed conservation of residues in GbuA implicated in flavin adenine dinucleotide binding (Extended Data Fig. 3), suggesting that GbuA may leverage this coenzyme to dehydrogenate the α-carbon to β-carbon bond of γBB-CoA to produce a crotonobetainyl-CoA intermediate en route to cleavage of the critical C-N bond to liberate TMA and crotonyl-CoA. While the exact mechanism of C-N bond cleavage by GbuA is presently unclear, it may mirror the FAD-mediated catalysis employed by glutaryl-CoA dehydrogenase to eliminate CO2 and generate crotonyl-CoA.1 Alternatively, GbuA may operate akin to the desulfinase Acd, another member of the acyl-CoA dehydrogenase family like GbuA, that while unable to perform dehydrogenation, catalyzes cleavage of the C-S bond from 3-Sulfinopropionyl coenzyme A to produce sulfite and propionyl-CoA.2 In this scenario GbuA may liberate the saturated 4-carbon product butyryl-CoA as part of C-N bond cleavage, instead of crotonyl-CoA. While, the acyl-CoA thioester hydrolase homolog GbuF was not necessary for production of TMA from γBB in recombinant *E. coli* (Fig. 4F), it is presumed to play a role in CoA re-cycling in E. timonensis. The activity of this thioesterase would result in the generation of crotonate (or possibly the related four-carbon by-product butyrate if GbuA operates akin to Acd). Promiscuous thioesterase activity within *E. coli*, encoded by genes such as *tesA*, *tesB*, *yciA*, may explain why GbuF was unnecessary in our recombinant gain-of-function studies. Finally, GbuE is presumably responsible for uptake of the γBB substrate. Whether it promotes the antiport of the resulting four carbon product and TMA, or alternative transporters (e.g. x, y) present in E. coli provide this function, are unknown. Interestingly, the natively encoded *E. coli* L-carnitine: γBB antiporter CaiT was unable to functionally substitute for recombinantly expressed GbuE, though it is possible it was not expressed under the conditions used in these combinatorial functional studies.

1. Rao, K. S., Albro, M., Dwyer, T. M. & Frerman, F. E. Kinetic Mechanism of Glutaryl-CoA Dehydrogenase. *Biochemistry.* **45**, 15853–15861 (2006).
2. Schurmann, M., Deters, A., Wubbeler, J. H. & Steinbuchel, A. A Novel 3-Sulfinopropionyl Coenzyme A (3SP-CoA) Desulfinase from Advenella mimigardefordensis Strain DPN7T Acting as a Key Enzyme during Catabolism of 3,3'-Dithiodipropionic Acid Is a Member of the Acyl-CoA Dehydrogenase Superfamily. *J.Bacteriol*. **195**, 1538–1551 (2013).

Supplemental Table 1. Baseline clinical characteristics of *n* = 2918 Genebank subjects used in analyses with γBB.

	All patients (n=2918)
Age(years)	62.2±11.0
Male (%)	67.9
Former/Current smokers (%)	67.6
Diabetes mellitus (%)	34.2
Hypertension (%)	72.9
Hyperlipidemia (%)	86.5
Prior coronary artery disease (%)	77.9
CAD (%)	82.0
PAD (%)	23.7
CVD (%)	83.4
MACE 3 year(%)	13.9
Framingham ATP III Risk Score	9.0(6.0-11.0)
BMI (kg/m^2)	28.7(25.6-32.6)
LDL cholesterol (mg/dL)	95.0(77.0-114.0)
HDL cholesterol (mg/dL)	33.5(28.0-40.2)
Total cholesterol (mg/dL)	159.8(138.8-186.6)
Triglycerides (mg/dL)	121.0(85.0-173.0)
hsCRP (mg/dL)	2.38(1.01-5.74)
Creatinine (mg/dL)	0.86(0.76-1.0)
eCrCl (ml/min/1.73m^2)	87.6(73.5-97.3)
Gammabutyrobetaine (µM)	0.93(0.77-1.16)
TMAO (µM)	3.79(2.44-6.25)
Carnitine (µM)	38.0(31.7-45.0)
Baseline Medications	
ACE inhibitors (%)	50.5
Beta-Blocker (%)	65.6
Statin (%)	63.9
Aspirin (%)	76.2

Baseline characteristics, comorbidities, and medications for subjects used in Figure 1. Continuous data are presented as mean ± standard deviation or median (interquartile range). Categorical variables are presented as %. CAD = coronary artery disease; PAD = peripheral artery disease; CVD = cardiovascular disease; MACE = major adverse cardiovascular events (MI, stroke, or death); BMI = body mass index; LDL = low density lipoprotein; HDL = high density lipoprotein; hsCRP = high sensitivity C-reactive protein; eCrCl = estimated creatinine clearance; ACE = angiotensin-converting-enzyme.

Supplemental Table 2. CVD, CAD, PAD risk ratio.

CVD (no=484,yes=2434)	RR	Low	High	p value
Q1(0.01-0.07)	1	1	1	-
Q2(0.07-0.93)_unadjusted	1.04	0.99	1.09	0.09
Q2_Adj1	1.04	0.99	1.09	0.10
Q3(0.93-1.16)_unadjusted	1.10	1.05	1.15	2.59E-05
Q3_Adj1	1.10	1.05	1.15	3.45E-05
Q4(1.16-4.14)_unadjusted	1.12	1.07	1.18	5.54E-07
Q4_Adj1	1.11	1.06	1.17	1.19E-05

CAD (no=525,yes=2390)	RR	Low	High	p value
Q1(0.01-0.07)	1	1	1	-
Q2(0.07-0.93)_unadjusted	1.04	0.99	1.09	0.10
Q2_Adj1	1.04	0.99	1.09	0.13
Q3(0.93-1.16)_unadjusted	1.11	1.06	1.16	2.06E-05
Q3_Adj1	1.11	1.05	1.16	4.20E-05
Q4(1.16-4.14)_unadjusted	1.14	1.09	1.19	1.19E-07
Q4_Adj1	1.12	1.07	1.18	4.68E-06

PAD (no=2227,yes=691)	RR	Low	High	p value
Q1(0.01-0.07)	1	1	1	-
Q2(0.07-0.93)_unadjusted	0.92	0.75	1.11	0.38
Q2_Adj1	0.96	0.79	1.17	0.68
Q3(0.93-1.16)_unadjusted	1.02	0.84	1.23	0.84
Q3_Adj1	1.11	0.92	1.35	0.29
Q4(1.16-4.14)_unadjusted	1.44	1.21	1.71	5.15E-05
Q4_Adj1	1.37	1.14	1.65	0.0007

Risk ratios (RR) for CVD, CAD and PAD, as well as corresponding 95% confidence intervals (CI), were estimated using both univariable (unadjusted) and multivariable (adjusted) dispersion parameters for a generalized linear models (GLM) in R version 3.63 with family=quasipoisson selected.[1,2,3] Adjustments (Adj1) were made for individual traditional cardiac risk factors in the Framingham risk score (includes age, LDL cholesterol, HDL cholesterol, blood pressure (SBP and DBP), diabetes, and smoking) and high-sensitivity C-reactive protein (CRP), creatinine and BMI level.

1. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004; 159(7):702-706. 3.

2. Zou G. Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. Stat Methods Med Res 2013 Dec;22(6):661-70. doi: 10.1177/0962280211427759

3. https://www.r-bloggers.com/2018/08/relative-risk-ratios-and-odds-ratios-by-ellis2013nz/

Supplemental Table 3. Functional annotation of the protein sequences from the γBB utilization cluster using the EMBL-EBI bioinformatics tool Hmmer3 phmmer[a] with the UniprotKB protein database.

***gbu* protein**	**Annotation Score**	**Homology based description**	**UniprotKB entry**
GbuA	927.6	Butyryl-CoA dehydrogenase (cofactor: FAD)	A0A1H5EDR0_RHOJO
GbuB	363.2	CoA-transferase	A0A386PL67_CLOSE
GbuC	340.5	CoA-transferase	A0A386PL67_CLOSE
GbuD	152.1	NADH:ubiquinone oxidoreductase chain I-like protein (cofactor: 4Fe-4S cluster)	I9KTI1_9THEO
GbuE	517.7	L-carnitine/gamma-butyrobetaine antiporter	A0A1C6ENF5_9CLOT
GbuF	104.7	Acyl-CoA thioester hydrolase	A0A1M5CIM1_9FIRM

Protein amino acid sequences for members of the *gbu* gene cluster were loaded into Hmmer3 phmmer (version 3.2.1 https://www.ebi.ac.uk/Tools/pfa/hmmer3_phmmer/), which provides a score and predicted functional annotation. Using this in combination with functional studies from heterologous gene expression studies we have developed the overall schematic detailed in Fig. 4C and Supplemental Fig. 3A.

Supplemental Table 4: Amino acid sequences for proteins encoded within the *gbu* operon from the organisms presented in Figure 4E.

GbuA

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_067536470.1	393	100	100	MDFKLSKEHRALQVKAKEFTEQVLFPYEMECEENNGITPETHKYITQQVMDWGFNATNHSKEHGGQGMTLFEQALCSEQFGMSTGAIWDAIPQPSFPMKFGTQEQIEEYLIPSNQCKRRDAYAITEADAGSDPTQCQTTAVKCDGGYKINGEKWYVTVGNVADFLLVHTHIDGDPEKATVFFVEKDAPGVSVKRTPEFTHHFAFKHPEFTFTDVIVDESKILKGIGEGFNMTKDWFVEARLGIAARCVGGAERVLNVANDWAAERIQGGRPIREYQVIEHMIADMTMEIMAAKSLLYRVCWEISDPDMDRKLKHARASAIKLWCSEMVNHVTDKGVQILGGRGYMRENPVERLWRDQRVDRIWEGTSEIQRNVIGGQIKKRGVELYTGWDYEK*
E. timonensis 71.3 (P&G isolate)	NA	393	100	100	MDFKLSKEHRALQVKAKEFTEQVLFPYEMECEENNGITPETHKYITQQVMDWGFNATNHSKEHGGQGMTLFEQALCSEQFGMSTGAIWDAIPQPSFPMKFGTQEQIEEYLIPSNQCKRRDAYAITEADAGSDPTQCQTTAVKCDGGYKINGEKWYVTVGNVADFLLVHTHIDGDPEKATVFFVEKDAPGVSVKRTPEFTHHFAFKHPEFTFTDVIVDESKILKGIGEGFNMTKDWFVEARLGIAARCVGGAERVLNVANDWAAERIQGGRPIREYQVIEHMIADMTMEIMAAKSLLYRVCWEISDPDMDRKLKHARASAIKLWCSEMVNHVTDKGVQILGGRGYMRENPVERLWRDQRVDRIWEGTSEIQRNVIGGQIKKRGVELYTGWDYEK*
A. desmolans 43058	WP_031476031.1	392	90.1	93.9	MDFKLSKEHRALQEKAREFTEQVLFPYEMECEENNGISPETHKYITEQVMKWGFNATNHSKEHGGAGLTLFEQALVSEQLGMSTGAIWDAVPQPSFPMKFGTQEQIDEYLIPSCQCKRRDAYAITEADAGSDPTMCQTTAVKCDGGYKINGEKWYVTVGNIADFLLVHTHLDGDPNKATVFFVEKDAPGVSVKRTPEFTHHFAFKHPEFTFEDVVVDESKILKGIGEGFNMTKDWFVEARLGIAARCVGGAERVLKEANDFAATRVQGDRVIRNYQVIEHMLVDMTMDIMAAKSLLYRVCWEISGDMDRKLKHARASAIKLWCSEMVNRVTDKGVQILGGRGYMRENAIERLWRDQRVDRIWEGTSEIQRNVIGGMIKKRGTELYTDWAYEE*
E. minutum 700079	ASS41451.1	392	86.8	92.4	MDFTLSKKDRALSEKAREFTEQHLIPYEIECEENNGLSPESHKAITEQVMAWGFNATNHSKEHGGAGFTLFEQMLASEQFGKVTGAIWDAVPQPSFPMKFGTKEQIEEYLIPANQCKRRDAYAITEADAGSDPTECQTTAVKCDGGYKINGEKWYVTVGNIADFLLVHTHLDGDPAKATVFFVEKDTPGVSIKRTPEFTHHFAFKHPEFTFEDVVVDESKILKGIGEGFTMTKDWFVEARLGIAARCVGGAERVLEVANEWAANRVQGGQIIRDYQVIEHMLVDMTMEIMAAKSLLYRVCWEISGDMDRKLKHARASAIKLWCSEMVNRVTDMGVQILGGRGYMRENPVERLWRDQRVDRIWEGTSEIQRNVIGGMIKKRGTDLYTEWAYEE*
C. bacterium UC5.1-1D1	WP_054327869.1	392	88.8	93.6	MDFKLSKEHRALQEKAREFTEQVLFPYEMEVEENNGLSPESHKYINGQVMEWGFNATNHSKEHGGTGYTLFEQALISEQLGMSTGAIWDAVPQPSFPMKFGTKEQIEEYLIPTCQCKRRDAYAITEADAGSDPTMCQTTAVKCDGGYKINGEKWYVTVGNIADYLLVHTHIDGDPNKATVFFVEKDTPGVSIKRTPEFTHHFAFKHPEFLFEDVVVDESKILKGIGEGFNMTKDWFVEARLGIAARCVGGAERVLKEANDFAANRVQGDRVIRDYQVIEHMLVDMTMDIMAAKSLLYRVCWEISGDMDRKLKHARASAIKLWCSEMVNRVTDKGVQILGGRGYMRENAVERLWRDQRVDRIWEGTSEIQRNVIGGQIKKRGIELYTGWAYEE*
Uncultured *Eubacterium* sp.	SCJ98641.1	393	99	99.5	MDFKLSKEHRALQVKAKEFTEQVLFPYEMECEENNGITPETHKYITQQVMDWGFNATNHSKEHGGQGMTLFEQALCSEQFGMSTGAIWDAIPQPSFPMKFGTQEQIEEYLIPSNQCKRRDAYAITEADAGSDPTQCQTTAVKCDGGYKINGEKWYVTVGNVADFLLVHAHIDGDPDKATVFFVEKDAPGVSVKRTPEFTHHFAFKHPEFTFTDVIVDESKILKGIGEGFNMTKDWFVEARLGIAARCVGGAERVLNVANDWAAERVMGGRPIREYQVIEHMIADMTMEIMAAKSLLYRVCWEISDPDMDRKLKHARASAIKLWCSEMVNHVTDKGVQILGGRGYMRENPVERLWRDQRVDRIWEGTSEIQRNVIGGQIKKRGVELYTGWDYEK*
Epulopiscium Nuni2H_MBin003	OOB78507.1	390	77.9	86.5	MDFSLSKEHLELQDKAKRFTQEVLFPYEMECEENDGISPETHAIINQKVLDWEFNATNHKKEHGGQGMTLFEQMLVSEQLGQSTGAIWDAVPQPSFPMRFGTQAQIDEYLYPACQNKRRDAYAITEADAGSDPTQCLTTAVKVDGGWKINGEKWYVTVGDVADFILVHAHVDGDVNKAAVFFVEKNAPGVSVKRTPLYTHHFAFKHPEFIFEDVIVGDDKMLGEIGDGFNMTKDWFVEARLGIAARCVGGATRILEVANEFANNRVQFGKPIIEFQAIEFMIADMAMDILAAKSMLYRYCWEISDPDADRKVRHAKAAAMKLWCSEMVGRVADKGVQILGGRGYMRENPVERLARDTRVDRIWEGTSEIQRVIIGGQIKKRGTELYTGW*

GbuB

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_067536474.1	392	100	100	MLPLEGIRILDLTTLSGYCGMELADYGAEVIKVETPSGGDPLRELAPLKNGASPHHVFRDRGKKSVTLNLMHPEGKELFKKLVATADAVLENFTPGTMEKLGLGYEDLSAIKPSLVYGRVSAYGSTGPETDIPQSDLIAQAKSGVMHFTGFPENPPTRIGFSISEHYAASFLASAVCVAIYYARSTGEGQMVETSLCGSAIAVSEDKVITWGADGEDPMRTGNAHPLINPYDILKCKNGYVAMGISSDAQWTKFCKAFNTPEWDEDEKYCSNLVRGYHYFGDLRNKIEDLFSNYTMQEIAAICDENLIPGTMCSTTREALEEPQLQVRNMIVSLEDEKIGKLEMPGKPVKFVGEEEEPLKAAPNLGEHNAEVYGALAIDSVELAHLQQQGVI*
E. timonensis 71.3 (P&G isolate)	NA	392	99.7	99.7	MLPLEGIRILDLTTLSGYCGMELADYGAEVIKVETPSGGDTLRELAPLKNGASPHHVFRDRGKKSVTLNLMHPEGKELFKKLVATADAVLENFTPGTMEKLGLGYEDLSAIKPSLVYGRVSAYGSTGPETDIPQSDLIAQAKSGVMHFTGFPENPPTRIGFSISEHYAASFLASAVCVAIYYARSTGEGQMVETSLCGSAIAVSEDKVITWGADGEDPMRTGNAHPLINPYDILKCKNGYVAMGISSDAQWTKFCKAFNTPEWDEDEKYCSNLVRGYHYFGDLRNKIEDLFSNYTMQEIAAICDENLIPGTMCSTTREALEEPQLQVRNMIVSLEDEKIGKLEMPGKPVKFVGEEEEPLKAAPNLGEHNAEVYGALAIDSVELAHLQQQGVI*
A. desmolans 43058	WP_031476033.1	392	77.9	88.5	MFPLEGIKVLDLMTLSGYCGMELADYGAEVIKVESPEGDPLRSLAPLKNGASVHHAFRDRGKKSITLNLQHPEGKEIFKKLVATADVVLENFAPNTMEELGLGYDVLSAIKPSLVYGRISGYGSTGEGANMPQLDLIAQAKSAVMHFTGFPENEPTRIGFAISERYAASFLSSAICLALYHARETGEGQMVETTLCGSAIAISEDKVITYGAEHEDPMRTGNAHPLINPYDILKCKDGYVAMGISSDAQWEKFCNAFNCPKEWTEDEKYCSNLVRGYHYFGDLRVKLEDLFSNYSMQEIADICDQALIPGTMCSTTKEALQQPQLQVRNMIVSVKDAELGELEMPGRPVKFCNQDEEPLASAPTVGAQNEALYTALGLDRAAMDRLQQEGVI*
E. minutum 700079	ASS41452.1	392	73.7	85.7	MEPLKGIKVLDLTINSSYCGMELADYGAEVIKVEEPSGGDPLRGLAPLKDGASPHHIYRDRGKMSITLDLQKPEGQDILRKLVKESDAIIENFVPGTLEKLGLGYEDLSKINPTIVYGRISAYGSTGKGAETPQSDLVAQAKCGVMHLTGFPENPPTRIGFNISEHYAASFLASAICIAIYRAKKYGEGQMVEATLAGSAIAISEDKVITYGADNEDPMRTGNAHPLINPYDILKCKDGYVAMGISSDAQWTKFCKAFNTPEWDEDEKYCSNLVRGYHYFGDLRIKIEELFSNYTMQEITDICDEHLIPGTMCSTTKEALNEQQLKERNMIVDLKTKELGTIRMPGRPIKFSGESESAIKDAPALGEQTEKILTDLGYDSEKIDQLKDNGII*
C. bacterium UC5.1-1D1	WP_082397279.1	235	59.2	73.2	MGFAISEHYAASFLSSAVCLAVYHARETGEGQVVETTLCGSAVAITEDKVITYGAEHEDPMRTGNAHPLINPYDILKCKDGYVAMGISSDAQWEKFCNAFNCPKEWTEDEKYCSNLVRGYHYFGDLRVKLEDLFSGYTMQEIADICDQALIPGTMCSTTKEALQQPQLRVRNMIVSVNDTDLGELEMPGKPVKFCGEEEEPLSAAPEIGAQNEQVYTSLGIDAAAMENLHREGVI*
	WP_082397277.1	125			MTLDLRHPEGKEIFKKLVATADVVLENFPPKTMEELGLGYDVLSVIKPSLVYGRISAYGSTGDEADTPQLDLIAQAKSGVMHFTGFPENHPPALALRYRSTTRLASFPLLCAWRFIMPEKPARGR*
	WP_082397275.1	58			MFPLEGIRVLDLMTLSGYCGMELADYGAEVIKVESPSGDPLRTLAPLKKWCQCSSCLS*
	Compiled	390			MFPLEGIRVLDLMTLSGYCGMELADYGAEVIKVESPSGDPLRTLAPLKKWCQCSSCLSPRKEEHNARSAASGREGDFQKAGGYGRCCSGKFPAQNNGRAWARICVISHQTISCIRTHFRIRFHRRRSGYPAVRSDCAGEERRDAFYWFSGKPPTRIGFAISEHYAASFLSSAVCLAVYHARETGEGQVVETTLCGSAVAITEDKVITYGAEHEDPMRTGNAHPLINPYDILKCKDGYVAMGISSDAQWEKFCNAFNCPKEWTEDEKYCSNLVRGYHYFGDLRVKLEDLFSGYTMQEIADICDQALIPGTMCSTTKEALQQPQLRVRNMIVSVNDTDLGELEMPGKPVKFCGEEEEPLSAAPEIGAQNEQVYTSLGI
Uncultured *Eubacterium* sp.	SCJ98638.1	392	99	99.7	MLPLEGIRILDLTTLSGYCGMELADYGAEVIKVETPSGGDPLRELAPLKNGASPHHVFRDRGKKSVTLNLMHPEGKELFKKLVATADAVLENFTPGTMEKLGLGYEDLSAIKPSLVYGRVSAYGSTGPETDIPQSDLIAQAKSGVMHFTGFPENPPTRIGFSISEHYAASFLASAVCVAIYYARSTGEGQMVETSLCGSAIAVSEDKVITWGADGEDPMRTGNAHPLINPYDILKCKNGYVAMGISSDAQWTKFCKAFNTPEWDEDEKYCSNLVRGYHYFGDLRNKIEDLFSNYTMQEIAAICDENLIPGTMCSTTREALEEPQLQVRNMIVSLEDEKIGKLEMPGKPVKFVGEEEEPLQAAPNLGEHNAEIYGALAIDSVELTHLQQEGVI*
Epulopiscium Nuni2H_MBin003	OOB78505.1	388	62.8	76.5	MLPLTGLRILDLTTLSGYCGMEFADHGAEVIKIETPNGGDALRKFSSFKGEKSRHHQFRDRGKKSITLDIYNIEGQKLFKELVAKSDAVLENFPAGTLEKLGLGYEDMEKINPAIVYGRISALGSTGNAVNIPQSEIIAQAKTGAMHVTGFPEEEPTRIGFPIASRYSAGFLTSGVILAICLAKQSGEGQLVEVTLCGSLATITEDKVITYGVTDEDPMRTGNAHPLVNPYDIIKCKNGYVALGIGSDFQWIKLCKVLGKEDWSSDEKYCTNDVRGLHYFGDLRDKIEEFFQDYTMQELAEKCDSVLIPGTMCSTIKEALSEEQLHVRNMIVEVDGLKMSNRPVKFVGGTEPDFVKAPELGANNAEIFGLIDVDDAKLAKLRADGVI*

GbuC

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_067536479.1	400	NA	NA	MKKRPLEGIKILDFSQVLSAPFCGMLLADMGAEIIKVERPGNGDISREYGPYKNDISLYFCQYNRGKKGVAVDMRSEEGKNVVMDLVSKVDIVIENFKSGTLEKLGIGYEEMLKVNPGLIYGSIAGFGTYGPLSHLPCMDIIAAARSGLVSTTGQGADAPIKPGFSLCDTWAGLQLFRGLTMALLHKQKTGKGLRVDIAMLDCAFYMCEAPVLEHSLTGEFTPRTGNHHPWYAPCGEFVTNDGNAVIAVTREEEWTQLCETLGLSALEQDPRFNNNESRVRNREALIAELEKVTSAMGRYEIEKKLCEVGVPASAVQTLAEFAGNPQTKELKVITEINQPGVGDYTVVNTPIRFSKTPVDPNGVAAGFPGANSKDILNELGYSQEKIDELIDSGAIYQSV*
E. timonensis 71.3 (P&G isolate)	NA	400	100	100	MKKRPLEGIKILDFSQVLSAPFCGMLLADMGAEIIKVERPGNGDISREYGPYKNDISLYFCQYNRGKKGVAVDMRSEEGKNVVMDLVSKVDIVIENFKSGTLEKLGIGYEEMLKVNPGLIYGSIAGFGTYGPLSHLPCMDIIAAARSGLVSTTGQGADAPIKPGFSLCDTWAGLQLFRGLTMALLHKQKTGKGLRVDIAMLDCAFYMCEAPVLEHSLTGEFTPRTGNHHPWYAPCGEFVTNDGNAVIAVTREEEWTQLCETLGLSALEQDPRFNNNESRVRNREALIAELEKVTSAMGRYEIEKKLCEVGVPASAVQTLAEFAGNPQTKELKVITEINQPGVGDYTVVNTPIRFSKTPVDPNGVAAGFPGANSKDILNELGYSQEKIDELIDSGAIYQSV*
A. desmolans 43058	WP_031476035.1	400	77.5	87	MKKRPLEGIKVLDFSQVLSAPFCGMMLADMGAEVIKVERPGAGDISREYGPYVNDISLYFCQYNRGKKGIAIDMRSEEGKKVVLDLVAQVDVVIENFKAGTLEKLGIGYDKMLEVNPGLIYGSISGFGTYGPLSHLPCMDIIAAARSGLVAQSGEAGSAPIKPGFSMCDTWAGLQLMRGLSMALLHKQKTGKGLRVDIAMLDCAFYMCEASVLEHSLTGEFTEKSGNHYVWYAPYGEFKAKDGNVVIAVSTEEEWESLCRTLGLDALAADPRFATNELRVQNRDALIAELDKTTGAMGRYDLENKLCAAGVPAAAVQTLREFDNNPQTKSLNVITSIHQQGVGDYVVTNTPILFSKTPVDPNASAAGHPGEHSVEILQELGYGQDRIDSLLAEGAVHQAG*
E. minutum 700079	ASS41456.1	262	66.7	79.9	MDIIAAARSGLVGTSGGENEAPIKPGFSLCDTWAGLQLLRGLSMGLLQKEMTGKGCRVDIAMLDCAFYMCEYPVLEYSITGEFSKRTGNHIENYAPYGEFKTKDGNVVIAASSEEQWAEVLKILNLEELAKDKRFSDNEQRVKNREALIAAIEEGTSKRGRYELETALCAAGVPAAAVQSLAEFDNHPETKELKVIIDVNQPEEGDYIAVNTPIIFSKTPVDPNHPAAAHPGENSVEIMKEIGYSDELINTLLENGSISQES*
	ASS41453.1	141			MMKNRPLEGIKVLDFSQVLAAPFCGMMLADMGAEVIKVERPGLGDISREYGPYKNDISLYFCQFNRGKKDIAVDMRSEGGKKVIMDLVAGADIVIENFKYGTLDKMGLGFDDMIKVNPQLIYDGVKFLWSFLLTSTPKGIF*
	Compiled	402			MMKNRPLEGIKVLDFSQVLAAPFCGMMLADMGAEVIKVERPGLGDISREYGPYKNDISLYFCQFNRGKKDIAVDMRSEGGKKVIMDLVAGADIVIENFKYGTLDKMGLGFDDMIKVNPQLIYDGVKFLWSFLLTSTPKGIFDIIAAARSGLVGTSGGENEAPIKPGFSLCDTWAGLQLLRGLSMGLLQKEMTGKGCRVDIAMLDCAFYMCEYPVLEYSITGEFSKRTGNHIENYAPYGEFKTKDGNVVIAASSEEQWAEVLKILNLEELAKDKRFSDNEQRVKNREALIAAIEEGTSKRGRYELETALCAAGVPAAAVQSLAEFDNHPETKELKVIIDVNQPEEGDYIAVNTPIIFSKTPVDPNHPAAAHPGENSVEIMKEIG
C. bacterium UC5.1-1D1	WP_082397281.1	400	76.3	87.5	MKKGPLKGIKVLDFSQVLSAPFCGMMLADMGADVIKVERPGNGDISREYGPYVNDISLYFCQYNRGKKGIAIDMRSEEGKKVVLDLVAQVDIVIENFKAGTLEKLGLGYDKMVEVNPSLIYGSISGFGTYGPLSHLPCMDIIAAARSGLVAQSGQAGDAPIKPGFSMCDTWAGLQMMRGLSMALLYKQKTGKGLRVDIAMLDCAFYMCEAPVLEHSATGKFTEKSGNHYVWYAPYGEFAAKDGNVVIAVSSEGEWKTLCYTLGVDDLAADPRFCNNELRVQNRDALIVELEKATGAMGRYDVENKLHEAGVPAAAVQSLQEFDNNPQTKALNVITQINQLGVGEYTVTNTPILFSETPVDPDASTAGFPGANSIEILNSLGYDQERIDKLLAEGAIHQAG*
Uncultured *Eubacterium* sp.	SCJ98636.1	400	96	97.5	MKKRPLEGIKILDFSQVLSAPFCGMLLADMGAEIIKVERPGNGDISREYGPYKNDISLYFCQYNRGKKGVAVDMRSEEGKGVVMDLVSKVDIVIENFKSGTLEKLGIGYEEMLKVNPGLIYGSIAGFGTYGPLSHLPCMDIIAAARSGLVSTTGQGADAPIKPGFSLCDTWAGLQLFRGLTMALLHKQKTGKGLRVDIAMLDCAFYMCEAPVLEHSLTGEFTPRTGNHHPWYAPCGEFATKDGNAVIAVTCEEEWTQLCETLGLSALAQEPRFNNNESRVQNREALIAELEKVTLTLGRYELEKNLCGVGVPASAVQTLAEFAGNPQTKELKVITEINQPGVGDYTVVNTPIRFSKTPVDPNGVAASFPGANSKDILSELGYSQEKIDQLIDSGAIYQSV*
Epulopiscium Nuni2H_MBin003	OOB78504.1	407	67.2	80	MHTKSMKKRPLEGIKVLDFSQVLAAPFCGMLLADLGAEVIKIERPGDGDISRGFGPYVNGVSLYFCQYNRGKKGIAIDLGSPEGKKVAMDLVAEVDIVIENYKSGTLEKLGIGYDEMIKVNPSLIYGSIAGFGTYGPLSHLPCMDIIAAARSGLVSQSGVEGFAPIKPGFSFCDTWSGLNLLRALSMALYHKQQTGEGTRIDLAMLDIAFYMGDDALLDYAQTGEFTPRTGNSDPWFSPNGEYTTNDGNVVIATTTEEEWQALCDALKLDALLHDERFATNELRNINRDVLRTEIEVATKAMGRYEVEKTLIEAGVPASAVQTLAEFIHSDEAKELELISIIKDQPGVGDYMATSTPIHFSKTPINHQAHAAQKPGVKTVEVLEELGYSKERIDELINCGAVHQEG*

GbuD

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_074038086.1	92	100	100	MKLISPDTTKCVKCYACTEVCPMKVIGISASGFPEPKSYAYRLCINCGYCVDICVFDALHHKVRKQRSSDPGAAIRRYEEMKKKRKSEQEIK*
E. timonensis 71.3 (P&G isolate)	NA	70	76.1	76.1	MKVIGISASGFPEPKSYAYRLCINCGYCVDICVFDALHHKVRKQRSSDPGAAIRRYEEMKKKRKSEQEIK*
A. desmolans 43058	WP_078097032.1	103	42.3	56.3	MLFVQANELISPDKEKCVRCDACVDICPMHVIHVDEDGYPQAVGKAFLTCINCGYCVDVCAPGALRHRVRKRSVNPTAALHAALKRLRAARERDDKTKRGESK*
E. minutum 700079	ASS41457.1	95	41.1	57.9	MKELFYPDKEKCIRCRWCAEICPMGVIDLTKDGYPTPVEEAFNLCINCGYCVDVCESAAFYHAVRKRNTTSKAAKKRLQRKREVNKLLGKPYKRL*
C. bacterium UC5.1-1D1	WP_082397283.1	108	41.3	55.6	MDGFCFEGDFVVEAVELISPDKEKCVKCNACVDICPMRVIGVDEEGYPHAVGKAFLTCINCGYCVDVCAPGALKHKVRKRSANSAAALRAAVKRLNRSREYRKKRGEP*
Uncultured *Eubacterium* sp.	SCJ98634.1	70	66.3	71.7	MKVIGISASGFPEPESFAYRLCINCGYCVDICVFDALHHRVRKQRSSDSRAAIRRYEEMKKKRKKEQKNQ*
Epulopiscium Nuni2H_MBin003	OOB78503.1	92	46.7	64.1	MSLIHIEEDKCAKCGRCIVFCPMGLLENKPDGFPQPTKEAYQRCVNCGYCVDICIFDALFHKIRKRNANSKDAALKRYETLLKRGVKKNAK*

GbuE

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_067536484.1	541	100	100	MENLNKKDRGNIEIGLIIVSVAVLAVFIIFMVLNPETTIRGIGGFFNKMISGLGPVFEVVAFVTFLIGLYLGLGKYGKIRLGDCKPEYSTFSYLAMMLLASLASAALYWSFTEWAFYYEAPGLGMEPHSTEALESSLGYQFFHWGMVNQAMYTVMGVAIAYGVYVRKIKSFQTSAICCSMLGEKVKGKSAIGKVIDFLVIFGVLGALSSSLGLAVPLAAGGLKTLFGIEATPVVQIGIIVVIALVYTFTSYLGTEKGMKVISNAAAGLCIVFLLYVLCAGPTTFILKNIVNSLGHMISKLPRMSLFTDPIENTGFAESWTIYFNAFYLNYVAMMGIFIAKVSKGRTIREVAVATIFGMTAGGWFIFGINGSFSIKTFLDGSADVVGLVNSGVGDAAIYSILKVLPGGATLLPFLILLLIVGFVAPSMDSASLALAETVTKRGTPKMALRIFFCVLLAIIPMAIILSGAGFDPIKQVAIIISAPFMVILVGVEIGLFKWLKHDSRSGLHAKNIALQEKEDHEAFLKEEAELQGEEEETSETA*
E. timonensis 71.3 (P&G isolate)	NA	541	100	100	MENLNKKDRGNIEIGLIIVSVAVLAVFIIFMVLNPETTIRGIGGFFNKMISGLGPVFEVVAFVTFLIGLYLGLGKYGKIRLGDCKPEYSTFSYLAMMLLASLASAALYWSFTEWAFYYEAPGLGMEPHSTEALESSLGYQFFHWGMVNQAMYTVMGVAIAYGVYVRKIKSFQTSAICCSMLGEKVKGKSAIGKVIDFLVIFGVLGALSSSLGLAVPLAAGGLKTLFGIEATPVVQIGIIVVIALVYTFTSYLGTEKGMKVISNAAAGLCIVFLLYVLCAGPTTFILKNIVNSLGHMISKLPRMSLFTDPIENTGFAESWTIYFNAFYLNYVAMMGIFIAKVSKGRTIREVAVATIFGMTAGGWFIFGINGSFSIKTFLDGSADVVGLVNSGVGDAAIYSILKVLPGGATLLPFLILLLIVGFVAPSMDSASLALAETVTKRGTPKMALRIFFCVLLAIIPMAIILSGAGFDPIKQVAIIISAPFMVILVGVEIGLFKWLKHDSRSGLHAKNIALQEKEDHEAFLKEEAELQGEEEETSETA*
A. desmolans 43058	WP_031476039.1	550	70.9	82.2	MENNTKQKKNSDVEVGLIVVSVAILVAFIILMIFNPDGTLNAITKFFNMMITGLGPFFEIIAFVSFIVGLYLLFGKYGKVRLGNCKPEYSNFSYFSMMLLASLASAALYWSFTEWASYYQAPGLGMQPGSVEAIESSLGYQFFHWGIVNQSMYTIMGVAIAYGVYVRKVPSFQTSAVCCAMLGEKAKGKSVIGKVIDFLVIFGVLGALSSSLGLAVPLFTGGLNALFGIPVNAVSQVGTIVFIAIVYSITSFLGTKKGMQTISNIASVTCVIFLLFVLFAGPTTFILKNIVNSFGHMIEKLPRMALFTDPVNNTGFVESWTVYFVAFYLNYVAMMGIFIAKVSKGRTIREVAMYTMFIMTGGGILIFGINGSFSIFQHLHGLVDVIGLVTSGVGDVAIYSILKTLPLGSTVLPFLILVLVVGFVAPSMDSASLALAETVTKRGTPKMALRMFFCILLAVIPLSIILVGAEFTAIKNIAIIISAPFLIILAGTILGLLKWLRHDDRIGLHAKIVAQQEAEMRAEALAEDAGEVPVPEAEGELKAVMAEQSE*
E. minutum 700079	ASS41458.1	533	63.4	79.7	MDKSKKNNTEIGLMITALVVLGVFLVFMILNPKKTIGGIGGFFQGMIGTLGPLFLLLAFGSLIVSLILMFSKYGKVRLGDSKPEYSTFSWMAMMMLASLAAAALYWSLTEWAFYYETPGIGIKPRTDEALESSLSFQFYNWGMVNQALYALAGVAIGYSVYVRKVRSFQTSAVCNAMIGKRLGDNGKTILGKVIDFCVILGILGALSSSLGLAVPLASGGLKVLFGMEATPIVQIGIIVVIAAVYIFTSYLGTNKGMKVLSNAAAVLCVAFLIYVLCVGPTQFIFKNTVNSLGKCFHNLIPIALFTDPIKKTGFPEGWTIYFVAFYLNYVAMMGIFIAKISKGRTIREVAAATVIGMSTCTWIFFGIDGSFSMHAFITGKVDVVSLANSGAGQSAIYSILDILPFGGTLIPLVMLLLIVGFVAPSMDSAALALSETVTKSGEPKMRVRLFWCVALAVIPLAVTITGSGFDAIKQVAIIVSVPFIVIIAGMALGLFKWLKEDSLNGVHDRNLEIQREEMLQDRGMDATTENATE*
C. bacterium UC5.1-1D1	WP_082397285.1	555	71.9	81.6	MRRRNRKMGNDSKKNKNADVEVGLIIVSVAILVAFIICMVVNAEGTLNTIMNFFNVMIDALGPFFLLVAFISFLVSLYLGFGKYGKIRLGNCKPEYSNFSYFAMMLLASLASAALYWSFVEWASYYQAPGLGMEPQSLEALESSLAYQFFHWGFVNQSLYTIAGVAIAYGVYVRKLPSFQTSAVCCAMLGKKVKGKSIIGKVIDFLVIFGILGALSSSLGLGVPLATSGLNTLWGIEATPAVQIGVIVFIALVYTFTSFLGTKKGMQTISNIASVVCIFFLLYVLFMGPTTFIFKNIVNSLGYLIQDLPRMALFTDPIANTGFAESWTIYFVAFYLNYVAMMGIFIAKVSKGRTIREIAIYTIFIMSFGGMCIFGIDGSFSIYAHTQGMVDVTGLVASGIGDAAIYEILKALPLGSTLLPFLILVLIVGFVAPSMDSASLALAETVTKRGTPKMALRIFFCVLLAIIPMSVLLVGASFDSIKYIAIIISAPFLLILAGMMLGLLKWLKEDERSGLHAEIIAIKEAEDREGTIEENANVRGALQEASECEIIEKKV*
Uncultured *Eubacterium* sp.	SCJ98632.1	541	98.2	99.3	MENLNKKDRGNIEIGLIVVSVAVLAVFIIFMVLNPETTISGIGGFFNKMISGLGPVFEVVAFVTFLIGLYLGLGKYGKIRLGDCKPEYSTFSYLAMMLLASLASAALYWSFTEWAFYYEAPGLGMEPHSTEALESSLGYQFFHWGMVNQAMYTVMGVAIAYGVYVRKIKSFQTSAICCSMLGEKVKGKSVIGKVIDFLVIFGVLGALSSSLGLAVPLAAGGLKTLFGIEATPVVQIGIIVVIALVYTFTSYLGTEKGMKVISNAAAGLCIIFLLYVLCAGPTTFILRNIVNSLGHMISKLPRMSLFTDPIENTGFAESWTIYFNAFYLNYVAMMGIFIAKVSKGRTIREVAVATIFGMTAGGWFIFGINGSFSIKTFLDGSADVVGLVNSGVGDAAIYKILEVLPGGASLLPFLILLLIVGFVAPSMDSASLALAETVTKRGTPKMALRIFFCVLLAIIPMAIILSGAGFDPIKQVAIIISAPFMVILVGVEIGLFKWLKHDSRSGLHAQNIALQEKEDHEAFLKEEAELQGEEEETSEPA*
Epulopiscium Nuni2H_MBin003	OOB78506.1	527	66.4	79.7	MSNKNRDTSNVEVGLMITSIVILIAFIILMMTNAEGTISVISSFFWSSMSTLGPFFQLFTFGTFIAALYLAFGKYGKVRLGDCKPEYSTPSYVSMMLLASLASSALYWSFTEWAYYYASPGLGMEPYSVLALENSLSYQFYHWGIVNQAMYTAVGVSIAYGVYVKKVKSFQTSAVCSAMMGDKVKGKSAIGKVIDFAVIFGILGGLSSSLGLGVPLAAGALERSLGIPQTAIVQVGIIVFIACVYAFTSYLGTKKGMKVLTDITTVISICVLLFILLAGPTTFILKNIVSSLGHMIDNTVRMTLFTDPVENTGFAEGWTIYFQAFYLNYVAMMGIFIAKVSKGRTIKEVALATLFGISAGSWIFFGINGSFSISTFLNGQIDVVGLVNGGAGESVIFDILATLPLGAFMPFIILILIVGFLAPSLDSASLALAETVTKEGTPKMAVRMVWCVLLAVIPMSIILTGTSFDAIKYISIIISIPFVVIVTGMQIGLFKWLAQDSRNGVHKANIELQEKEIAAELAQKNN*

GbuF

Organism	Accession	# AA	% Identity (relative to *E. timonensis* SN18)	% Similarity (relative to *E. timonensis* SN18)	Amino Acid sequence
E. timonensis SN18	WP_067536487.1	137	100	100	MRSDCIFFCPIRVRYNEVDQQGIVYNGHYVIYTDIAFEEFFRSKGYPYKTLVEEHDSEVCHKKSTYEFFASAYEGDLIEVGMRVLKIGNRSMTIGFEIYREGEDDVIVTCETVYVGYDTKNRASKPISDLMRQLLGA*
E. timonensis 71.3 (P&G isolate)	NA	137	100	100	MRSDCIFFCPIRVRYNEVDQQGIVYNGHYVIYTDIAFEEFFRSKGYPYKTLVEEHDSEVCHKKSTYEFFASAYEGDLIEVGMRVLKIGNRSMTIGFEIYREGEDDVIVTCETVYVGYDTKNRASKPISDLMRQLLGA*
A. desmolans 43058	WP_034972294.1	137	66.4	86.1	MRSDCFFFCPVRVRFGEVDRQGIVYNGNYVVYTDLAFEEFLRSRGYSYRELAEEHESEVCHRKATYDFCSSAFEGDDLEVGISNIHVGNKSFTITFEIYRKDEDELLVRCESVYVGYDVENRASRPITDLMRSLLTP*
E. minutum 700079	ASS41459.1	173	45.1	59	MNRSSSVCESIMISNDSHRTVAGFFKTWNGKNEGITMRDECIFFCTVRVRYGEVDQQGIVYNGNYAVYGDFAFCEFLRNRGCSYNDLSKTLDFEVCHKASSFEFFSSAYAEDLVEIGVKIVNIGTRSFTIGYEFRREGEDELLVYSEVVYVGYDVENRASRELTPAFVNLLTA*
C. bacterium UC5.1-1D1	WP_054327874.1	138	69.6	83.3	MRRDCFYFCPPFRVRYGEVDQQGIVYNGNYVAYTDLAFEEFLRSKGYSYRELAEKHDSEVCHKKSTYEFVSSAYAGDMLEVGISKIKIGKKSFTIQFEIYHQGEDDLILLCESVYVGYDVENRTSRPITDLMRALLSS*
Uncultured *Eubacterium* sp.	SCJ98628.1	137	94.2	100	MRSDCIFFCPIRVRYNEVDQQGIVYNGHYVIYTDIAFEEFFRSKGYPYKTLVEEHDSEVCHKKSTYEFYSSAYEGDLLEVGMRILKIGNKSMTIGFEIYREGEDELIITCETVYVGYDTKNRASKPISDLMRQLLGA*
Epulopiscium Nuni2H_MBin003	OOB78502.1	137	58.4	78.1	MRNDFDFFHQIRVRFGEVDQQGIVYNGNYSIYTDVAFEEYFRAKGHPYKYLAEGESEICHRKYTATYSSSAVYDDLLDVGVRLKEIGTRSFTLIFEIYRAGENELLFDAQSVYVGYDSTNRCSQPISPLMRKLLTE*

For each ORF within the gbu gene cluster (GbuA, GbuB, GbuC, GbuD, GbuE, GbuF), primary amino acid sequences and NCBI Accession numbers are presented for each organism purported to contain a gbu cluster. The percent identity and percent similarity of each sequence relative to the corresponding E. timonensis SN18 ORF were determined using MatGAT (Matrix Global Alignment Tool) . Amino acid sequences from E. timonensis 71.3 were utilized to create codon-optimized ORFs for heterologous gain of-function assays in E. coli BL21 Star (DE3) in Figure 4d and Supplemental Figure 4

Supplemental Table 5. Gnotobioitc community member PCR primers.

Primer Target	Primer Sequence (5'-3')	Product Length	Annealing Temp	Study
B. caccae Fwd	GTGCAGAGTGGCATACTGAAAG	93	56	Goodman et al. 2009
B. caccae Rev	AGTGTAGGCTTTCGTACCGTTC			Goodman et al. 2009
B. ovatus Fwd	TTCCCGGAGAGTAAGCTAACAG	89	56	Goodman et al. 2009
B. ovatus Rev	GGAACAGTACAAGACCCTTTGG			Goodman et al. 2009
B. thetaiotaomicron Fwd	GAGGGTGTCGTATTTCCGAAGG	81	58	Goodman et al. 2009
B. thetaiotaomicron Rev	GTTCCCTGATCCAGTGTGTTGG			Goodman et al. 2009
C. aerofaciens Fwd	CTTTCAGCAGGGAAGAGTCAA	219	58	Kageyama, Sakamoto, and Benno 1999
C. aerofaciens Rev	TGTTCCACCCTCCCCTAC			This Study
E. rectale Fwd	GCAGGAACATCAATTACACTGAC	95	55	This Study
E. rectale Rev	CCAAACTCGTAATCATTTCCCGGC			This Study
E. timonensis Fwd	GGTACACAGGAACAGATCGAAG	487	56	This Study
E. timonensis Rev	CAGCCCAGTCGTTTGCTAC			This Study
P. penneri Fwd	TAACCACCGCCAACAGTACA	77	56	This Study
P. penneri Rev	AGTAACTATGGTGATGGCGAC			This Study

Primers used for confirmation of presence (or absence) of indicated community members following colonization of germ-free mice (see Methods).

Supplemental Table 6 Heterologous gene expression plasmids.

	Codon optimized *E. timonensis gbu* genes for expression in *E. coli* BL21 Star (DE3)									
Plasmid name	**gbuA**	**gbuB**	**gbuC**	**gbuD**	**gbuE**	**gbuF**	**Parent plasmid**	**Selection marker**	**Origin of replication**	**Reference**
pACYC_acdA6_caiT	x				x		pACYCDuet-1	Cam^{R}	P15A	This study
pACYC_caiT					x		pACYCDuet-1	Cam^{R}	P15A	This study
pACYC_caiT_mnyN				x	x		pACYCDuet-1	Cam^{R}	P15A	This study
pACYC_mnyN				x			pACYCDuet-1	Cam^{R}	P15A	This study
pACYCDuet-1							pACYCDuet-1	Cam^{R}	P15A	Novagen
pCDF_acdA6_caiT	x				x		pCDFDuet-1	Sm^{R}	CDF	This study
pCDFD_acdA6	x						pCDFDuet-1	Sm^{R}	CDF	This study
pCDFD_acdA6_ybgC	x					x	pCDFDuet-1	Sm^{R}	CDF	This study
pCDFDuet-1							pCDFDuet-1	Sm^{R}	CDF	Novagen
pCOLA_frc3		x					pCOLADuet-1	Kan^{R}	ColA	This study
pCOLA_frc3_frc4		x	x				pCOLADuet-1	Kan^{R}	ColA	This study
pCOLA_frc4			x				pCOLADuet-1	Kan^{R}	ColA	This study
pCOLA_frc4_frc3		x	x				pCOLADuet-1	Kan^{R}	ColA	This study
pCOLADuet-1							pCOLADuet-1	Kan^{R}	ColA	Novagen
pET11a							pET11a	Amp^{R}	pMB1	Novagen
pET11a_frc3		x					pET11a	Amp^{R}	pMB1	This study
pET11a_ybgC						x	pET11a	Amp^{R}	pMB1	This study
pET15b							pET15b	Amp^{R}	pMB1	Novagen
pET15b_acdA6	x						pET15b	Amp^{R}	pMB1	This study
pET15b_caiT					x		pET15b	Amp^{R}	pMB1	This study
pET15b_frc4			x				pET15b	Amp^{R}	pMB1	This study

The plasmids constructed for use in gain-of-function assays in the non-native *E. coli* host BL21 Star (DE3) (see Methods). Table captures the plasmid name, the *gbu* gene cluster gene(s) from *E. timonensis* encoded on each plasmid as well as the antibiotic selection marker and origin of replication. Plasmid encoded open-reading frames are codon-optimized versions of the primary amino acid sequences described in Supplemental Table 4.

Supplemental Table 7. Composition of APPROACH baseline and intervention diets

	Baseline diet	High saturated fat			Low saturated fat		
		Red meat	White meat	Non-meat	Red meat	White meat	Non-meat
Carbohydrate (%E)	49	41	42	41	39	46	41
Protein (%E)	14	24	24	24	26	23	25
Red meat[a]	mixed	12	0	0	12	0	0
White meat[a]		0	12	0	0	12	0
Vegetable protein[a]		4	4	16	4	4	16
Dairy and eggs[a]		9	9	9	9	9	9
Fat (%E)	37	35	34	35	35	31	34
Saturated fat	13[a]	13	14	14	8	7	7

Dietary methylamine content							
Choline [a] (mg)	410	573	498	447	654	575	*516*
Carnitine [b] (mg)	56	258	56	22	202	68	47
Choline+Carnitine (mg)	466	831	554	469	856	644	*562*

[a] Calculated values (Nutrition Data System for Research; University of Minnesota).

[b] Estimated values, based on published carnitine content of commonly consumed foods [1]

Briefly, the APPROACH study was an NIH funded dietary intervention study originally designed to examine the impact of both dietary protein source and saturated fat on lipoprotein composition changes.[2] Subjects enrolled in the APPROACH trial were all healthy adults (n=113 total, 44 males, 69 females). Subjects had to follow an omnivorous dietary pattern in the preceding months to be eligible for enrollment, and be of overall good health without major diseases or taking of chronic medications. Additional inclusion criteria included being 30-65 years old (median 45; IQR 34-56), non-smoking, agreeing to abstain from alcohol and dietary supplements during the study, and willing to consume all study foods as instructed. Exclusion Criteria include a history of heart disease, cerebrovascular disease, peripheral vascular disease, bleeding disorder, liver or renal disease, lung disease, diabetes, Human immunodeficiency virus (HIV), or cancer (other than skin cancer) in the last 5 years. Additional exclusion criteria also included body mass index (BMI) > 35 kg/m^2 or < 20 kg/m^2 (median 25.93; IQR 22.74-27.96), a history of unstable weight, abnormal thyroid stimulating hormone, blood pressure > 150/90, fasting glucose >126 mg/dl, fasting triglyceride levels >500 mg/dl, total- and LDL cholesterol >95th percentile for age and sex, being pregnant or breastfeeding, and taking hormones or any drugs known to affect lipid metabolism.

The APPROACH study had a factorial design enabling independent examination of (i) the impact of dietary protein source in the form of red meat, white meat, vs non-meat; and (ii) the impact of low vs high saturated fat. In analyses examining the impact of low vs high saturated fat, no differences in plasma or urine TMAO levels were observed.[3] Similarly, dietary saturated fat content showed no effect on plasma levels of any examined TMA and TMAO nutrient precursor (choline, γBB, carnitine).[3] In contrast, plasma TMAO levels rose markedly amongst subjects following 1 month of red meat as a primary source of dietary protein, especially compared to non-meat (predominantly vegetarian protein source, but also egg white).[3]

Diet compositions (also published in Wang *et al.* 2019 and Bergeron *et al.* 2019) based on compositional analysis of 2500 kcal 4-day rotating menus (energy % in the form of protein, carbohydrate, fat), and both saturated fat content in low vs high saturated fat arms, and dietary methylamine content (carnitine, total choline, choline + carnitine) for each dietary arm were calculated and are presented above. Of note, while total choline content between the intervention diets were similar, the carnitine content of the red meat diet was approximately 10-fold higher than the non-meat diet.

All diets were prepared in the metabolic kitchen of the Bionutrition Unit, University of California San Francisco-based Clinical and Translational Science Institute throughout the entire duration of the study. Participants picked up their meals weekly, and had weekly visits with the nutritionist. All subjects had a 2 week run-in phase where a general omnivorous run-in diet was provided, and then each of the intervention diets, in random order, was fed to subjects for a month each. At least a 2 week break occurred between each intervention diet (2 weeks was the default time period, but could be extended to avoid intervening holidays, weddings, and other issues to enhance compliance by subjects). All intervention diets were isocaloric, and had comparable calorie content derived from protein.

1. Demarquoy, J. *et al.* Radioisotopic determination of L-carnitine content in foods commonly eaten in Western countries. *Food Chemistry*. **86**, 137-142 (2004).

2. Bergeron, N. *et al.* Diets high in resistant starch increase plasma levels of trimethylamine- *N* -oxide, a gut microbiome metabolite associated with CVD risk. *Br. J. Nutr.* **116**, 2020–2029 (2016).

3. Wang, Z. *et al.* Impact of chronic dietary red meat, white meat, or non-meat protein on trimethylamine N-oxide metabolism and renal excretion in healthy men and women. *Eur. Heart J*. **40**, 583–594 (2019).