Sulfoquinovose is a select nutrient of prominent bacteria and a source of hydrogen sulfide in the human gut

Buck Hanson
University of Vienna https://orcid.org/0000-0001-6947-7286
Kerim Kits
University of Vienna
Jessica Löffler
University of Vienna
Anna Burrichter
University of Konstanz
Alexander Fiedler
University of Konstanz https://orcid.org/0000-0002-2852-3538
Karin Denger
University of Konstanz
Benjamin Frommeyer
University of Konstanz https://orcid.org/0000-0002-8817-0930
Craig Herbold
University of Vienna
Thomas Rattei
University of Vienna https://orcid.org/0000-0002-0592-7791
Nicolai Karcher
University of Trento
Nicola Segata
University of Trento https://orcid.org/0000-0002-1583-5794
David Schleheck (✉ david.schleheck@uni-konstanz.de)
University of Konstanz
Alexander Loy (✉ alexander.loy@univie.ac.at)
University of Vienna https://orcid.org/0000-0001-8923-5882

Short Report

Keywords: sulfoquinovose, human gut, bacteria, fecal metatranscriptome

Posted Date: August 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-49676/v1

Version of Record: A version of this preprint was published at The ISME Journal on March 31st, 2021. See the published version at https://doi.org/10.1038/s41396-021-00968-0.

Abstract

Diet selectively shapes the human gut microbiota and fuels production of diverse metabolites that influence host health. Responses of the microbiota to diet are highly personalized, yet mechanistically not well understood because the metabolic capabilities of human gut microorganisms remain largely unknown. Here we show that sulfoquinovose (SQ), an omnipresent monosaccharide in green vegetables, is a selective substrate for few but ubiquitous bacteria in the human gut. In anoxic incubations of human feces and in defined co-culture, Eubacterium rectale and Bilophila wadsworthia both use previously unrecognized pathways to cooperatively catabolize SQ to hydrogen sulfide (H2S), a key intestinal metabolite with disparate effects on host health. We find SQ degradation capability encoded in almost half of E. rectale genomes but otherwise sparsely distributed among microbial species in the human intestine. Re-analysis of fecal metatranscriptome datasets of four human cohorts showed that SQ degradation (mostly from E. rectale and Faecalibacterium prausnitzii) and H2S production (mostly from B. wadsworthia) pathways were expressed abundantly across various health states, suggesting their active contribution to gut functioning. The discovery of green diet-derived SQ as an exclusive microbial nutrient and an additional source of H2S in the human gut highlights the role of individual dietary compounds and organosulfur metabolism on microbial activity and has implications for precision editing of the gut microbiota by dietary and prebiotic interventions.

Main Text

Dietary habits largely modulate the highly personalized composition and temporal dynamics of the human intestinal microbiota and influence disease risk10,11. Each individual daily ingests thousands of different dietary compounds12 of which some, such as starch or glucose, are substrates for many microorganisms while others, such as the seaweed polysaccharide porphyran, are select energy sources of metabolically privileged microorganisms13,14. Disentangling the effects of single dietary compounds on gut microbiota membership and activity is important for understanding diet-microbiota interaction mechanisms15 but complicated because the genetic and physiological capabilities of microorganisms are insufficiently established2.

Here, we explored microbial metabolism of sulfoquinovose (6-deoxy-6-sulfoglucose, SQ) in the human gut. SQ is the polar head group of sulfoquinovosyl diacylglycerol (SQDG), a ubiquitous sulfolipid in the photosynthetic membranes of all land plants, algae, dinoflagellates, and cyanobacteria. SQDG is one of the most abundant organic sulfur compounds in the biosphere and can represent greater than 25% of the total lipids in common dietary, leafy green vegetables such as spinach, lettuce, and green onion3,16. Studies on the catabolism of dietary SQDG by animals are scarce; in guinea pigs SQDG is deacylated via host-derived lipases to sulfoquinovosyl glycerol (SQG)17. It is not known whether host tissues are able to metabolize SQG or SQ. Some *Proteobacteria* can catabolize SQDG/SQG and SQ analogously to the Embden-Meyerhof-Parnas (EMP) and the Entner-Doudoroff (ED) glycolytic pathways, hence, either via the sulfo-EMP (e.g. commensal and pathogenic *Escherichia coli*) or the sulfo-ED pathway (e.g. *Pseudomonas putida*) 18–20. Bacterial SQ degradation products are the C3-organosulfonates 2,3-dihydroxypropane-1-

sulfonate (DHPS) or 3-sulfolactate, which can serve as sources of sulfite for specialized sulfite- respiring and H2S-producing *Desulfovibrio* species[21].

To investigate SQ degradation by human gut microbiota, we constructed triplicate anoxic microcosms with fecal slurries mixed from eight vegetarians and incubated them with SQ (10 mM) or control substrates (**Extended Data Fig. 1a**). SQ was completely consumed within 20 h concomitant with a transient accumulation of DHPS (up to 5 mM) between 20 and 52 h (**Fig. 1a**). Concurrent with DHPS consumption, H2S production peaked at 96 h. Additionally, formate accumulated transiently with a peak at 73 h (about 3 mM), while acetate gradually increased over 144 h (to 7.5 mM) (**Fig. 1a**), compared to unamended control microcosms (**Fig. 1b**). Primary SQ and secondary DHPS degradation was accompanied by select changes in community composition, with nine species-level 16S rRNA gene operational taxonomic units (OTUs) increasing significantly ($P < .01$) in relative abundance compared to unamended microcosms (**Extended Data Fig. 2, Supplementary Information Tables 1, 2**). Temporal abundance changes of the two most strongly increasing OTUs, identified as *Eubacterium rectale* (*Agathobacter rectalis22*) and *Bilophila wadsworthia* (**Fig. 1c**, **Extended Data Fig. 2a**), were confirmed by fluorescence probe- based microscopy cell counting (**Fig. 1d**) and corresponded with the consumption of SQ and DHPS, respectively (**Fig. 1a, c**). Stable isotope probing of individual cells in SQ- and heavy water (D2O)-amended microcosms demonstrated high cellular activity (up to 29% D-labeling) of *E. rectale* (**Fig. 1e**), suggesting that SQ was utilized for growth23. Amplicon sequencing of *dsrB*, encoding the beta-subunit of a key enzyme for H2S production, dissimilatory sulfite reductase, revealed a significant increase in a *Bilophila* OTU and two *Desulfovibrio* OTUs ($P < .01$; **Extended Data Fig. 1i, Supplementary Tables 3, 4**) in concert with DHPS and H2S dynamics (**Fig. 1a**). Consistently, phylogenomic analysis of 93 metagenome-assembled genomes (MAGs) from the SQ-amended microcosm (**Extended Data Fig. 1a, Supplementary Table 5**) identified one MAG as *E. rectale* (SQ_MAG_41, average nucleotide identity (ANI) of >97% to 42 *E. rectale* strains; **Supplementary Tables 5, 6**) and a second MAG as *B. wadsworthia* (SQ_MAG_14, ANI of >97% to three *B. wadsworthia* strains; **Supplementary Tables 5, 7**). These data suggest that SQ is cooperatively metabolized mainly by *E. rectale* and *B. wadsworthia* with net production of acetate and H2S. We confirmed this finding by anoxically co-culturing isolates of *E. rectale* (DSM 17629) and *B. wadsworthia* (strain 3.1.6) with SQ. *E. rectale* grew by fermenting SQ and produced DHPS, which was rapidly consumed after co-inoculation of *B. wadsworthia* leading to near-stoichiometric production of H2S (**Fig. 1f**). The co-culture reproduced the sulfur metabolite dynamics observed in our SQ-amended microcosms and affirms DHPS-cross-feeding between *E. rectale*, an abundant microbiota member in healthy individuals24, and *B. wadsworthia*, a pathobiont associated with gastrointestinal inflammation and cancer 25,26.

We next applied genome-resolved metatranscriptomics on the fecal microcosms and differential proteomics on pure-culture experiments to identify the genes/proteins involved in complete catabolism of SQ to H2S. *E. rectale* SQ_MAG_41 was the only MAG from the SQ- amended microcosms metagenome that contained a cluster with SQ metabolism genes and *yihQ/sftG* (**Fig. 2a)**, which is co-located with all previously characterized SQ-degradation gene clusters and encodes a sulfoquinovosidase (*alpha*-

glucosidase) that cleaves SQ from SQDG/SQG 20,27 (BF, AWF, Oehler SR, BTH, AL, Franchini P, Spiteller D, DS, submitted; *manuscript is included for reviewers*) (**Extended Data Fig. 3**). Transcription of the *E. rectale yihQ/sftG* as well as of 10 flanking genes was significantly upregulated upon SQ degradation in the microcosms ($P < 0.02$, **Fig. 2d, Supplementary Table 8**). We independently confirmed co-expression of these 11 co-localized genes by constructing a gene expression network for *E. rectale* using publicly available metatranscriptomes from 1,090 human fecal samples 6–8 (**Extended Data Fig. 4**). Besides *yihQ/sftG*, this cluster includes genes of the 6-deoxy-6-sulfofructose transaldolase (SFT) pathway for SQ degradation that was described in the aerobic *Bacillus aryabhattai* isolate SOS1 and confirmed to be active during SQ fermentation by *E. rectale* strain DSM 17629 (BF, AWF, Oehler SR, BTH, AL, Franchini P, Spiteller D, DS, submitted). Here, a SQ isomerase (SftI) converts SQ to 6-deoxy-6-sulfofructose. A novel transaldolase (SftT) then catalyzes interconversion of 6-deoxy-6-sulfofructose with glyceraldehyde-3-phosphate to 3- sulfolactaldehyde and fructose-6-phosphate. Fructose-6-phosphate is funneled into glycolysis for fermentation and growth. 3-sulfolactaldehyde is reduced to DHPS by an NADH-dependent reductase (SftR) as an additional fermentation step21 (BF, AWF, Oehler SR, BTH, AL, Franchini P, Spiteller D, DS, submitted), and the DHPS is exported via SftE.

We subsequently queried 73,793 high-quality MAGs, assembled from 9,428 human gut metagenomes28, for the four core catalytic enzyme genes (*yihQ/sftG*, *sftI*, *sftT*, *sftR*) and found a surprisingly narrow distribution of the SFT pathway across human gut species. All four genes were present in only 23 (0.5%) of 4930 species-level genome bins28 and mostly at low frequency (**Fig. 2a**). A notable exception is the *E. rectale* genome bin with 44% of its 2301 MAGs encoding the SFT pathway (**Fig. 2a**), suggesting that *E. rectale* is the most prevalent SQ-metabolizing organism in the human gut.

DHPS is a previously unknown substrate for *B. wadsworthia* to generate H2S (**Extended Data Fig. 5a**) and we thus sought to identify the genes responsible for DHPS catabolism, such as genes for enzymes converting DHPS to 3-sulfolactate, the substrate for the desulfonating enzyme SuyAB21. Screening the genomes of human gut-associated *Desulfovibrionaceae* for known genes involved in the metabolism of DHPS and other organosulfonates for respiration of sulfite (**Extended Data Fig. 5b**)21,29 revealed that the currently described genes for linking DHPS catabolism to the 3-sulfolactate desulfonation pathways (the DHPS dehydrogenase genes *hpsO* and *dhpA*)21,30 are absent in all known *Bilophila* genomes, indicating that DHPS might not be metabolised via 3-sulfolactate in *B. wadsworthia*. Instead, we identified for the pure culture of *B. wadsworthia* 3.1.6 through differential proteomics that it strongly and specifically expresses an uncharacterized four-gene cluster during growth with DHPS, including genes for a new DHPS sulfite-lyase system (DslAB) (**Fig. 2b, c, Extended Data Fig. 5c, d, e**). DslA is a closely related but phylogenetically distinct paralog of isethionate (2-sulfoethanol) sulfite-lyase (IslA) in *B. wadsworthia* (**Extended Data Fig. 6**). IslA is a glycyl radical enzyme that catalyzes isethionate desulfonation, a key reaction in taurine degradation29,31. DslB is the putative activator of DslA, analogous to the IslA-activating, radical SAM enzyme IslB. Accordingly, *dslA* and *dslB* were the most differentially expressed *B. wadsworthia* genes in the SQ-degrading fecal microcosms (n-fold change of 83 and 93, respectively; $P = 3.8 \times 10^{-8}$ and 1.0×10^{-5}, respectively, **Fig. 2d, Supplementary Table 9**). Highly oxygen-sensitive glycyl radical enzymes are a diverse and abundant protein family of the human gut microbiota, yet largely

remain functionally uncharacterized32. While IslA cannot utilize DHPS as a substrate 29, desulfonation of DHPS into sulfite and hydroxyacetone by the predicted DslAB (**Extended Data Fig. 5e**) was confirmed in cell-free extracts of DHPS-grown *B. wadsworthia* cells, but only when assayed under strictly anoxic conditions (**Fig. 2b**), as for the highly oxygen-sensitive IslA reaction29. Furthermore, the pure culture excreted hydroxyacetone during DHPS degradation (**Extended Data Fig. 5a**). We thus propose that *B. wadsworthia* uses DslA, a glycyl radical C-S bond-cleaving DHPS sulfite- lyase, to produce sulfite from DHPS intracellularly, which is then respired to H2S *via* the DsrAB- DsrC-dissimilatory sulfite reductase pathway (**Fig. 2c**)33,34. During the submission of our study, another study was published that corroborated and complemented our findings by the biochemical and structural characterization of the DHPS sulfite-lyase 35. Collectively, we show that *E. rectale* and *B. wadsworthia* each employ previously unrecognized metabolic pathways in the human gut and engage in interspecies DHPS transfer for joint degradation of plant diet-derived SQ to H2S.

To explore how widely distributed and active SQDG/SQ metabolism is in the human gut, we analyzed the expression of the SFT pathway, and *dslAB*, *islAB*, and *dsrABC* in the 1,090 stool metatranscriptomes from the Health Professionals Follow-Up Study (365 samples from 96 healthy men)6 and an inflammatory bowel disease (IBD) study (189, 203, and 333 samples from 27 healthy individuals, 28 with ulcerative colitis, and 50 with Crohn's disease, respectively)8. Consistent with previous reports6,8,9, glycolysis (from glucose and glucose-6-phosphate) and starch degradation were in the top 5% of 445 expressed microbial pathways across all datasets (**Fig. 3a**). We found that the SFT pathway was within the top 33% (136th) of all 445 expressed pathways and transcribed prevalently among individuals across all cohorts (177 out of 201 individuals). Further underlining the importance of SQ as a microbial nutrient in the gut, mean SFT pathway transcription was at a similar level to pathways for the usage of fucose (125th) and N-acetylneuraminate (149th) that are abundant components of glycoproteins and glycolipids in the colonic epithelium (**Fig. 3a**). Contrary to diet-derived SQ, these host-derived sugars represent permanently available substrates for microorganisms in the gut15. Furthermore, we determined that the mean relative abundance of the SFT pathway was two orders of magnitude higher than the proteobacterial sulfo-EMP pathway (321st), which was actively expressed in only 28 samples (**Fig. 3a**). In many individual samples (51%), a single bacterial species was responsible for more than 50% of SFT pathway transcription. Eighteen putative SQ degraders of the families *Lachnospiraceae* and *Ruminococcaceae* (*Firmicutes*) expressed the SFT pathway (**Fig. 3b**). *E. rectale*, *Faecalibacterium prausnitzii*, *Clostridium aldenense*, *Roseburia* sp. AM16-25, and *Clostridium clostridioforme* contributed most to expression of the SFT pathway but *E. rectale* was the single most dominant species across both healthy individuals (**Fig. 3b**) and those with ulcerative colitis or Crohn's disease (276, 148, and 198 metatranscriptomes, respectively) (**Extended Data Fig. 7**). Genes for the DHPS sulfite- lyase (*dslA*) and its activating enzyme (*dslB*) were also expressed in the human stool (136 positive samples out of 1,090), with some samples showing co-expression of both the SFT pathway and the DHPS (*dslAB*, *dsrAB*, *dsrC*) pathway (**Fig. 3c, Extended Data Fig. 7**). However, the isethionate sulfite-lyase genes (*islAB*) that enable *B. wadsworthia* to utilize the sulfite from taurine and isethionate as an electron acceptor29 were transcribed more prevalently (736 samples) and at a higher relative abundance when compared to *dslAB* (**Fig. 3c)**. Through meat consumption and

microbial deconjugation of host-secreted taurocholic bile acids11,25,36, taurine is a continuously available substrate in the gut and thus could decouple the DHPS-mediated physiological interaction of *B. wadsworthia* with the primary SQ fermenters. Transcription of the DsrAB-DsrC pathway for H2S production was abundant and prevalent in the dataset (843 positive samples out of 1,090), and dominated by *B. wadsworthia* (**Fig. 3c, Extended Data Fig. 7**). Notably, we did not detect significant differences in expression of SQ degradation and H2S production pathways between cohorts of IBD patients and healthy individuals, which is likely because, as reported9, stool samples were not selectively taken from patients with active IBD and thus, contrary to previous studies37,38, the observed species composition in stool was not significantly different between IBD and control cohorts. Overall, our re-analysis of stool metatranscriptomes from the four cohorts highlighted *B. wadsworthia* as the most important sulfidogen and DHPS as an additional substrate for microbial H2S production in the human gut. *B. wadsworthia* is specialized in utilizing diverse organosulfonates that are energetically more favorable compared to sulfate for sulfite-respiring bacteria21,29,39.

By uncovering the identities and activities of genes for degradation of the plant-derived sulfonated monosaccharide SQ in the human gut microbiota, we have shown that SQ is an exclusive substrate for the abundant *E. rectale* and few other *Firmicutes*. The concept of exclusive nutrient access13 promises SQ dosage-dependent control over the abundances and activities of these bacteria that are generally associated with a positive impact on human health1,40. We have also demonstrated that interspecies transfer of the SQ degradation product DHPS is a previously unknown physiological link between a plant-based diet and H2S production by the intestinal pathobiont *B. wadsworthia*, which uses DHPS in addition to taurine (**Fig. 4**). H2S is an important intestinal metabolite that influences the activity of epithelial and microbial cells and has beneficial as well as detrimental effects on the colonic environment41,42. The complex host-endogenous and microbial processes regulating intestinal H2S homeostasis and how they are influenced by diet or health state are insufficiently defined42,43. Our work emphasizes the roles of microbial cooperation and sulfur metabolism in the human intestinal tract and uncovers new physiological and genetic features of prevalent microbiota members that will inform dietary and drug-based therapies42 for targeted modification of microbiota membership and H2S levels in intestinal disease.

Methods

Chemicals

SQ and 3-sulfolactate were synthesized by MCAT GmbH (Donaueschingen, Germany). DHPS was synthesized and validated by NMR and HPLC-MS as reported previously18. Taurine, isethionate and all other purchased, routine chemicals were of at least *puriss* p.a. grade if not otherwise stated.

Human fecal microcosms

Human fecal samples were collected from healthy individuals adhering to ovo-lacto-vegetarian diets and who had not received antibiotics in the prior 6 months. Sampling and microbiota analysis of human fecal

samples was approved by the University of Vienna Ethics Committee (reference #00161). Study participants provided informed consent and self-sampled using an adhesive paper- based feces catcher (FecesCatcher, Tag Hemi, Zeijen, NL) and a sterile 107 x 25 mm polypropylene tube with a screw-cap-attached sampling spoon (Sarstedt, Nümbrecht, DE). Samples were kept cool with ice during transport periods and stored at 4°C for up to a 48 h period. Feces from eight individuals were transferred to an anaerobic chamber (Coy, Grass Lake, MI) having a mixed-gas atmosphere of 5% H2, 10% CO2, 85% N2 (Air Liquide, Austria). Fecal materials were pooled and a final mass of 3.9 g was homogenized in 30 mM anaerobic bicarbonate buffer, pH 6.8. Fecal homogenate was distributed into 20 ml hungate tubes across five treatment conditions in triplicate including buffer alone, 10 mM glucose, 10 mM taurine with 10 mM formate, 10 mM sulfoquinovose, and 10 mM sulfoquinovose with 50% D2O (99.9 atom % D, Sigma-Aldrich) having a final volume of 10 ml with a fecal concentration of 19.5 mg/ml. Tubes were sealed with butyl-rubber stoppers and screw-on caps and incubated at 37°C. One ml of each incubation was subsampled at 0, 6, 20, 28, 44, 52, 73, 99, 114, and 140 h for metabolite and nucleic acid analyses. Twenty µl of each subsample was added to 100 µl of 2 g/l zinc acetate for fixation of H2S. Two-hundred µl were fixed with paraformaldehyde (PFA) for FISH-based analyses as described previously 47,48. The remaining sample volume was centrifuged at 20,000 X *g* at 4°C to pellet biomass and store at -80°C for nucleic acid extraction. For SQ and DHPS quantification, 500 µl of the biomass-free supernatant was transferred to a 2 ml glass vial containing 170 µl acetonitrile (Sigma-Aldrich, Austria) and crimp-sealed with butyl-rubber stoppers and stored at - 20°C. The remaining supernatant (~300 µl) was transferred to a new microcentrifuge tube for quantification of anionic metabolites (e.g. short-chain fatty acids).

Pure culture and defined co-culture experiments

E. rectale DSM 17629 (purchased from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) and *B. wadsworthia* strain 3.1.6 (kindly provided by Dr. Emma Allen-Vercoe, University of Guelph, Canada) were grown under conditions as described previously21,29 (BF, AWF, Oehler SR, BTH, AL, Franchini P, Spiteller D, DS, submitted). Briefly, a carbonate-buffered mineral-salts medium reduced with titanium(III)- nitriloacetate (Ti(III)-NTA) was used under a N2/CO2 gas phase (80:20) in butyl-rubber stopperedserum flasks (20-ml scale) or culture tubes (10-ml scale). All cultures were incubated at 37°C in the dark without shaking. For growth of *E. rectale* with SQ or glucose as substrates for fermentation, the medium was supplemented with yeast extract (0.1% *w/v*), hemin (0.5 µg/ml) and vitamin K1 (0.1 µg/ml). For growth of *B. wadsworthia* with DHPS, taurine, 3-sulfolactate, or isethionate as electron acceptors (10 mM each), the medium was supplemented with 1,4-naphthochinone (0.2 µg/ml) and with lactate (20 mM) (or formate) as an electron donor. At intervals, growth was monitored as optical density (OD580nm) of the resuspended cultures (aftermixing by inversion), of which samples were taken to determine sulfonate disappearance and formation of products (see below). Cells were harvested from *B. wadsworthia* 3.1.6 replicate growth experiments in the late exponential growth phase for differential proteomics when grown with the four different organosulfonates (**Extended Data Fig. 5c**) or for preparation of anoxic cell- free extracts for DHPS sulfite-lyase enzyme tests (**Fig. 2b**) (described below). For co-cultivation experiments, medium containing SQ as the sole substrate for fermentation and organosulfonate respiration plus all supplements (above) was used and either both

strains were co-inoculated at the start of the growth experiment or (**Fig. 1f**) *E. rectale* DSM 17629 was inoculated first (n=3) and after 81 hours *B. wadsworthia* 3.1.6 was co-inoculated.

Anoxic cell-free extracts and enzyme tests

Preparation of anoxic cell-free extracts of *B. wadsworthia* 3.1.6 and DHPS sulfite-lyase assays were done under conditions as described previously for the isethionate sulfite-lyase (IslA) reaction[29]. Briefly, cultures were harvested in the late exponential growth phase (OD580 appox. 0.3 – 0.4) by centrifugation of the serum bottles. Cells were resuspended in anoxic Tris-HCl buffer (50mM, pH 8.0) containing MgCl2 (5 mM) and disrupted by three passages through a cooled French pressure cell, which had been flushed with N2 gas. The enzyme was assayed discontinuously at room temperature by monitoring the formation of the two products sulfite and hydroxyacetone. The reaction mixture (1 ml) contained 20 mM DHPS, 1 mM *S*-adenosylmethionine chloride (SAM), 1 mM Ti(III)-NTA, and about 0.2 – 0.9 mg total protein in 50 mM Tris-HCl (pH 8.0) containing MgCl2 (5 mM). The buffer, including DHPS and SAM, was initially degassed under vacuum in glass cuvettes with rubber stoppers, and flushed with nitrogen gas, three times each. Then, Ti(III)-NTA was added, and the reaction was initiated by the addition of anoxic crude extract, each with a syringe and needle through the rubber stoppers. At appropriate time intervals, samples (100 µl) were taken by syringes for quantification of sulfite by a colorimetric assay and of hydroxyacetone by derivatization and HPLC-UV analysis (see below).

Differential proteomics

Total proteomic analyses of cell-free extracts of *B. wadsworthia* 3.1.6 were done as previously described29 at the Proteomics Centre of the University of Konstanz (https://www.biologie.uni-konstanz.de/proteomics-centre/). Briefly, each trypsin-digested and purified sample was analyzed twice on a Orbitrap Fusion with EASY-nLC 1200 (Thermo Fisher Scientific), and tandem mass spectra were searched against the protein database using Mascot (Matrix Science) and Proteome Discoverer v1.3 (Thermo Fisher Scientific) with “Trypsin” enzyme cleavage, static cysteine alkylation by chloroacetamide, and variable methionine oxidation.

Metabolite analyses

H2S was quantified colorimetrically[49]. Absorbances at 670 nm were measured on a transparent 96- well plate (Greiner Bio-One, Austria) using an Infinite 200 PRO spectrophotometric microplate reader (TECAN Group Ltd., Männedorf, Switzerland). The samples were quantified with external standards prepared in parallel using sodium sulfide (Sigma-Aldrich, St. Louis, Missouri, USA). Biomass-free supernatants were subjected to capillary electrophoresis for separation and quantification of short-chain fatty acids on a P/ACE–MDQ apparatus (Beckman Coulter, Krefeld, Germany) equipped with a UV-detector with a wavelength filter at 230 nm. The capillary was a fused silica column (TSP075375; 75 µm ID, Polymicro Technologies) 60 cm long (50 cm to the detector) and 75 µm in diameter. Samples were processed using a CEofix Anions 5 Kit (Beckmann Coulter, Krefeld, Germany) according to the manufacturer's instructions. Analytes were separated in reverse polarity mode at 30 kV (ramp: 0.5 kV/s) for 10 min. All samples were

diluted 1:10 with a working solution consisting of 0.01 M NaOH, 0.5 mM CaCl2 and 0.1 mM caproate (internal standard). Analytes were quantified using an external standard mixture of SQ, Na2SO4, formate, succinate, acetate, lactate, propionate, butyrate, and valerate. SQ, DHPS, 3-sulfolactate, taurine and isethionate were quantified in culture experiments by HPLC against authentic standards, as described previously18, using a hydrophilic interaction ZIC-HILIC (Merck, Darmstadt, Germany) column and an evaporative light scattering detector. Hydroxyacetone, short-chain fatty acids, and alcohols in culture experiments were analyzed by a HPLC method described previously21, using an Aminex column (BioRad) and a refractive index detector. Hydroxyacetone was quantified also by derivatization with dinitrophenylhydrazine and HPLC-UV, using a method described for acetaldehyde29. Sulfite in enzyme reactions was quantified using a colorimetric assay (Fuchsin assay) or by derivatization with *N*-(9-acridinyl)maleimide and HPLC-UV29.

Fluorescence cell counting

FISH probes targeting *E. rectale* (EREC996-Cy5, EREC1252-Cy5) and *B. wadsworthia* (BWA829-Cy3) were designed using ARB50 and evaluated using SILVA TestProbe 3.0 against the SSU REF database (Release 132; December 13, 2017) (**Supplementary Table 10**)51. Optimal hybridization conditions were evaluated for fluorescently labeled probes with formamide concentrations ranging 0-70% using PFA-fixed cells of pure cultures of *E. rectale* DSM 17629 and *B. wadsworthia* DSM 11045 (purchased from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany). For microcosm cell counts, aliquots of PFA-fixed cells (5 or 20 µl) were filtered onto black polycarbonate filters (0.2 µm pore size, GTBP Isopore Membrane Filters, Millipore, Cork Ireland) using a 25 mm glass vacuum filter holder (Sartorius Stedim Biotech, Goettingen, Germany). One-sixth filter sections were cut and stained with FISH probes using standard procedures47. Samples were stained with DAPI (1 µg/mL; Sigma– Aldrich) for 5 min and washed with deionized water. Filter sections were placed on a glass microscope slide and overlain with one drop of Citifluor AF1 (Electron Microscopy Services, Hatfield, PA) and a coverslip. Cells were imaged using a Zeiss Axioplan 2 equipped with a Plan- Neofluar 100 x oil objective with a 1.3 numerical aperture (Zeiss, Germany). Cells from twelve fields of view were manually counted in a grid measuring 0.1225 mm2.

Single-cell stable-isotope probing by FISH-Raman microspectroscopy

PFA-fixed cells (5-10 µl) from fecal microcosms incubated with 10 mM SQ and 50% D2O were diluted into 100 µl of 1X PBS and sonicated at 35 kHz for 2 min at 20°C (Sonorex, Brandelin Electronic, Berlin, Germany) followed by centrifugation for 20 min at 20°C and 20,913 X *g*. Liquid-based FISH was performed as previously described23 using FLUOS-labeled probes targeting *E. rectale* (**Supplementary Table 10**) and most *Bacteria* (EUB338mix)52. Stained cell pellets were resuspended in 5-10 µl sterile filtered water. One µl of a 1:10 dilution of each sample was spotted onto an aluminum-coated glass slide (Al136; EMF Corporation) and allowed to air dry, followed by a 2 s wash in cold sterile filtered water. After air drying, fluorescence was used to identify cells for Raman analyses. Single cell Raman spectra were acquired using a LabRAM HR800 confocal Raman microscope (Horiba Jobin-Yvon) equipped with a 532

nm neodymium- yttrium aluminium garnet laser and a 300 grooves/mm diffraction grating. D-labeling (%CD) was quantified from integrated C-D and C-H peak areas (2,040-2,300 and 2,800-3,100 cm-1, respectively) as previously described23. Cells having %CD values greater than the average plus 3X the standard deviation of unlabeled cells, measured at the 0 h time point (threshold 3.54% CD), were considered D-labeled and metabolically active. Significant enrichment of %CD-labeling in FISH-positive *E. rectale* cells was evaluated using a Wilcoxon rank-sum test comparing FISH- positive cell counts for each time point with the 0 h time point.

Extraction of nucleic acids

DNA for amplicon and metagenome sequencing and RNA for metatranscriptome sequencing were extracted following standard procedures53 that included bead beating for 30 s on dry ice using a Lysing Matrix E tube (MP Biomedicals). Purified nucleic acid pellets were air dried at ambient room conditions and resuspended in 30 µl DNase/RNase-free H2O (Carl Roth GmbH, Germany) and stored at -80°C.

16S rRNA gene and *dsrB* amplicon sequencing

Established two-step PCR barcoding protocols were used for amplicon sequencing of the 16S rRNA gene (with primers H_515F_mod-5'-(head) GCTATGCGCGAGCTGC- GTGYCAGCMGCCGCGGTAA and H_806R_mod-5'-(head) GCTATGCGCGAGCTGC-GGACTACNVGGGTWTCTAAT that target most *Bacteria* and *Archaea*)54 and the dissimilatory sulfite reductase gene *dsrB55*. Blank nucleic acid extractions and negative (water only) PCR reactions were included as controls. Barcoded amplicon libraries were pooled at equivalent copy numbers (20x109) for 300 bp paired-end sequencing on an Illumina MiSeq sequencer (Microsynth AG, Balgach, Switzerland). Sequencing results were analyzed according to the procedures outlined previously54,55 and clustered into operational taxonomic units (OTU) according to sequence similarity cut-off of 97% using Uparse56, which also screens out chimeras. 16S rRNA amplicons were classified using RDPclassifier57 as implemented in Mothur v1.42.358 using Release 132 of the SILVA database. 16S rRNA gene amplicons that showed significant increases over time in the SQ-amended microcosms were classified to the species-level by phylogenetic inference by constructing a maximum likelihood tree (MEGA7 v7.0.18)59 using a collated dataset of amplicons and their top 10 subjects from querying against the NCBI 16S rRNA reference gene database. Classification of *dsrB* sequences was performed with phylogenetic placement using a curated DsrB reference sequence database and corresponding consensus tree55,60. 16S rRNA gene and *dsrB* libraries were analysed using the software package Phyloseq61 for R Studio v1.0.143 (rstudio.com). Statistical tests between treatment time points were performed with DESeq2 using the variance stabilizing transformation (vst)62.

Metagenomics

Samples from three time points (28, 73, and 114 h) from the SQ-amended microcosms were chosen for metagenome sequencing. DNA samples were diluted to 0.1 ng/µl in 130 µl and sheared using a Covaris M220 Focused-ultrasonicator Instrument (Covaris, Woburn Massachusetts, USA) to a target length of 350 bp. Library preparation was conducted according to the NEBNext Ultra II DNA Library Prep Kit for Illumina

(New England BioLabs Inc., Ipswich, Massachusetts) protocol. Indexing primers were ligated conforming to NEBNext Multiplex Oligos for Illumina manual (Index Primers Set 1, New England BioLabs Inc., Ipswich, Massachusetts). Library fragment sizes and purity were screened using a Bioanalyzer 2100 and the High Sensitivity DNA kit (Agilent Technologies, Santa Clara, CA). Samples were pooled for 150 bp paired-end sequencing on an Illumina HiSeq 3000/4000 at the Biomedical Sequencing Facility (BSF) of the Research Center for Molecular Medicine (CeMM), Vienna, Austria. Metagenomic datasets consisted of 87, 93, and 158 M reads (28, 73, and 114 h time points, respectively). For assembly and binning of metagenome-assembled genomes (MAGs), adapters and low quality base-calls were trimmed from Illumina reads using BBduk (https://jgi.doe.gov/data-and-tools/bbtools/ - right -trimmed, q=15, minlength = 149 bp). Trimmed reads were error-corrected using BayesHammer63 and assembled using metaSPAdes (SPAdes v3.13.1)64. MAGs were binned using MetaBAT v2.12.2 using composition and abundance under all presets on each dataset65. MAGs were assessed for quality using checkM v1.0.1266 and de-replicated using drep (v1.4.3)67. The highly fragmented and incomplete original *Bilophila* sp. MAG_14 MAG was missing *dslA*, however an unbinned 3913 bp contig (99.3% identity to *B. wadsworthia* 3.1.6) containing the *dslA* sequence had a similar coverage to the MAG (19.7 vs. 20.4 fold coverage). This contig was tested for consistency with the original MAG through iterative reassembly using MAGspinner v0.10 (github.com/hexaquo/MAGspinner). MAGspinner removes contigs with abnormal coverage and/or tetranucleotide signatures during the reassembly process. The inclusion of the contig containing the *dslA* sequence in the final bin (30 rounds of reassembly) was interpreted as evidence that the contig belongs to *Bilophila* sp. MAG_14. Representative MAGs were assigned taxonomy through alignment and phylogenetic placement of concatenated marker genes into the Genome Taxonomy Database reference tree (GTDB-Tk)22. 16S rRNA sequences were extracted from each MAG with nhmmer68 using rfam69 models for bacterial and archaeal small subunit rRNAs (RFAM: RF00177, RF01959). Minimum overlap between sequence and model was 300 nucleotides. Sequences were classified using the RDP classifier57 as implemented in Mothur. The number of tRNAs and associated amino acids were determined using tRNAscan-SE70. Pairwise, whole- genome, average nucleotide identity (ANI) between genomes of *Eubacterium rectale* and between genomes of 16 representative *dsrAB*-containing bacteria from the gut, including *Bilophila*, were calculated using FastANI (v1.2) using >95% ANI and <83% as the intra- and inter-species cutoffs71, respectively.

Metatranscriptomics

Metatranscriptome libraries (n = 9) were prepared from nucleic acid extracts from three timepoints (6, 20, and 52 h) of the triplicate SQ-amended microcosms. DNase digestion of nucleic acid extracts was performed up to three times (ezDNase; Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Removal of DNA was confirmed using a standard 16S rRNA gene PCR followed by electrophoretic analysis. Total RNA quality was evaluated using a Bioanalyzer 2100 and Agilent RNA 6000 Pico kit (Agilent Technologies, Santa Clara, CA). rRNA was removed from total RNA using the Ribo-Zero Gold rRNA Removal Kit (Epidemiology) (Illumina Inc., San Diego, CA) according to the manufacturer’s instructions. Purified mRNA pools were quantified (Quant-iT RiboGreen RNA Assay Kit, ThermoFisher Scientific) and 1-100 ng was used in each library preparation for multiplexed Illumina

RNASeq analysis (NEBNext Ultra RNA Library Prep Kit for Illumina with NEBNext Multiplex Oligos for Illumina primer set 1, New England BioLabs Inc., Ipswich, MA). Indexed libraries were purified from adapter sequences (Agencourt AMPure XP Beads, Beckman Coulter), quantified (Quant-iT PicoGreen RNA Assay Kit, ThermoFisher Scientific), evaluated for size and purity (Bioanalyzer 2100 and the High Sensitivity DNA kit, Agilent Technologies, Santa Clara, CA), and pooled at equimolar contributions. Sequencing (50 bp, single-end) was performed using Illumina HiSeq 3000/4000 (BSF, Vienna, Austria). The reads were demultiplexed and bam files were converted to fastq using SAMtools (v1.9)72 followed with trimming using Trimmomatic (v0.39)73 with the following settings: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:40. The trimmed reads were then mapped to the genomes of *Eubacterium rectale* ATCC 33656 (GCF_000020605.1_ASM2060v1) and *Bilophila wadsworthia* 3.1.6 (GCF_000185705.2_Bilo_wads_3_1_6_V2) using bbmap (v37.61; sourceforge.net/projects/bbmap/) with a minid of 97%. Feature counting was done using featureCounts within the Subread package (v1.6.3)74; reads that mapped equally well to multiple locations were assigned as fractions of the uniquely mapping reads for the same region (-M – fraction). After removal of reads that mapped to any rRNA, the raw read counts for all libraries were then converted to copies per million counts (cpm) and normalized using a trimmed mean of M-values (TMM) between each pair of samples using calcNormFactors (edgeR, v3.9)75. Differential expression analysis was done by pair-wise comparison of the samples from 20 h incubation to the samples from the 6 h incubation using the exactTest function (edgeR, v3.9). Genes were considered differentially expressed if the adjusted p-value was below 0.05 ($P < .05$).

Construction of a bacterial YihQ/SftG profile hidden markov-model (HMM)

Bacterial genomes and MAGs that were publicly available in Genbank as of July 27, 2017 (N=103411) were screened using hmmsearch (v3.1b2)76 for genes that contained the Glycosyl hydrolases family 31 pfam (PF01055.25). The resulting dataset (N=127608) was further screened using a YihQ/SftG-specific motif that we constructed using the signature sulfoquinovosidase residues reported by Speciale *et al*.27 corresponding to positions 290-508 in YihQ (Genbank accession AAB03011.1) from *E.coli*: W-X(2)-D-W-X-G-X(6)-G-X(4)-W-X-W-X(1,300)-D-X-G-G-[YWF] (expressed as a PROSITE pattern). Subsequences that matched this motif (N=585, mean length = 225 nt, length standard deviation = 2 nt) were extracted from the full dataset using a simple grep command in a Unix environment and aligned using MAFFT v7.39777. Positions corresponding to the variable-length linker corresponding to positions 308-505 in AAB03011.1 (corresponding to -X(1,300)- in the prosite pattern) were manually removed from the alignment. Remaining subsequences were clustered at 70% identity using Usearch78. Centroid sequences (N=15, length=26) were aligned using MAFFT v7.397. The resulting alignment was used to construct an HMM using hmmbuild (v3.1b2)76. The performance of the hmm for identifying the clade of sulfoquinovosidase proposed by Speciale *et al*.27 was examined phylogenetically. The set of genes that contained the Glycosyl hydrolases family 31 pfam (PF01055.25, N=127608) were length-filtered (min. 50 amino acids) and clustered at 90% identity with Usearch. Centroid sequences (N=2591) were aligned using MAFFT v7.397 and a phylogenetic tree was constructed in FastTree279. E-values < 1 were diagnostic for the proposed clade of sulfoquinovosidase. Of the 2591 Glycosyl hydrolases family 31 test

sequences, 103 belonged to the YihQ/SftG clade and 2488 did not. There were 4 false negatives and 1 false positive, resulting in a balanced accuracy of 0.98.

Genome analyses and comparative genomics

We searched all contigs and MAGs from the SQ-amended microcosm for known SQ metabolism genes using blastp and the YihQ/SftG profile HMM.

To identify publically available genomes that encode the SFT pathway, we searched the NCBI genome database using blastp for organisms that contain all four of the essential genes for the pathway - the sulfoquinovosidase (YihQ/SftG), SQ isomerase (SftI), 6-deoxy-6-sulfofructose transaldolase (SftT), and the SLA reductase (SftR). Each of these genes from the *E. rectale* ATCC 33656 genome was searched against all NCBI genomes using blastp with -max_target_seqs set to 20,000, an e-value cutoff of 1e-30, and a minimum 70% query coverage. All hits for each gene were then collated, sorted, and cross referenced to identify organisms which had hits for all four genes. Each genome was then checked for co-localization and synteny of the genes of interest to identify those that potentially encoded the complete pathway. This analysis identified 83 bacterial genomes that encode all four core enzymes of the SFT fermentation pathway (**Supplementary Table 11**).

Additionally, the presence of the SFT gene cluster in 73,798 high-quality MAGs from several thousands of publically available human gut metagenomes28, was searched using a two- step approach. Firstly, all translated open reading frames, as inferred with Prokka v1.14.080/Prodigal v2.6.381, were queried with our YihQ/SftG profile HMM (**Supplementary file YihQ-SftG.hmm**). Among 1,976 hits, all had an e-value of less than 0.1, and 84% had an e-value of less than 10-5. We did not apply an e-value filter in this step since we expected that the proximity check (see below) removes false positives. Secondly, blast-based searches including the three additional SFT pathway core protein sequences (SftI, SftT, and SftR) were performed against all translated open reading frames of genomes with an YihQ/SftG HMM hit. Physical proximity of the four genes in the genome was checked to validate SQ-gene clusters. We retained blast hits if they were above 75% of the mean reference gene length and had a mean identity of at least 70%. For each query genome and gene, we then kept only that blast hit with the largest e-value. After these steps, all blast hits had an e-value < 10-114. We then considered only those genomes where all blast hits were found on the same contig and ensured that all of them were within a 10-ORF window. 1118 MAGs passed these filters (**Supplementary Table 12**). The reference genome set used for taxonomic annotation of MAG clusters consisted of 80,853 genomes obtained from GenBank (March 2018) corresponding to 17,607 microbial species for which at least one proteome was available in UniProt. 137 isolate genomes82 were added to this set for a total of 80,990 isolate genomes with confident taxonomic annotation. Concatenated marker gene alignments using representative MAGs from each of the 23 species-level MAG clusters containing the SQ operon were generated using checkM v1.0.1266 and placed into a phylogenomic tree using FastTree279. Genes that enable taurine (*tpa*), DHPS (*dslAB, hpsO, hpsN, dphA*), 3-sulfolactaldehyde (*slaB*), 3- sulfolactate (*suyAB, slsC, comC*), sulfoacetaldehyde (*xsc, sarD*), and isethionate (*islAB*) catabolism and sulfite reduction (*dsrABC*) were initially identified in genomes of 16

representative *dsrAB*-containing human gut bacteria by blastp screening all protein coding genes from the genomes of interest against custom amino acid sequence databases established for each gene. The blastp results were filtered with an e-value cutoff of 1e-30 and a minimum 70% query coverage. Neighbourhoods of some genes of interest were checked manually for synteny: *tpa* with *sarD*, *adhE* with *islAB*, *dhpA* with *slaB*, *hpsOPN,* and *slsC* with *comC*.

Metatranscriptome analysis of publically available datasets

Reads from a total of 1,090 paired gut metagenomes/metatranscriptomes were downloaded from https://ibdmdb.org/ (725 paired samples from 105 individuals8) and the sequence read archive at https://www.ncbi.nlm.nih.gov/sra (365 samples from 96 individuals, BioProject accession PRJNA3542356. The downloaded reads, which were already screened for reads originating from the host, were then trimmed using Trimmomatic (v0.39)73 with the following settings:

LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:80. Taxonomic and functional profiling of the metagenomic and metatranscriptomic samples was done using HUMAnN2 v0.11.183 and MetaPhlAn2 v2.7.784. To assess the contribution of various SFT pathway encoding organisms to total transcription of the SFT pathway, we mapped all of the trimmed reads from the 1090 metatranscriptomes to all of the SFT pathway-encoding genomes we identified in our NCBI search (total of 83) with bbmap (v37.61, sourceforge.net/projects/bbmap/) using a minid of 97%. Additionally, we used the same pipeline and settings to map the trimmed reads from the 1090 metatranscriptomes to *dslAB*, *islAB*, and *dsrABC* genes identified in genomes of 16 representative *dsrAB*-containing human gut bacteria (see previous section). Feature counting was done using featureCounts within the Subread package (v1.6.3)74; reads that mapped equally well to multiple locations were assigned as fractions of the uniquely mapping reads for the same region (-M – fraction). Read counts were then converted to reads per kilobase per million reads (RPKM) or reads per kilobase (RPK). To estimate the relative abundance of the SFT pathway in the 1,090 metatranscriptomes, we analyzed all of the metatranscriptome samples using HUMANn2 with default settings and then manually added the RPK (read per kilobase) values from each organism found to be expressing the SFT pathway as well as the sum pathway RPK values to the pathway abundance output from HUMANn2 prior to normalizing all of the RPK values to relative abundance using the humann2_renorm_table script provided in HUMAnN2. Proportional contribution of each organism to total expression of the SFT pathway was calculated as a ratio between the RPKM value for the organism of interest divided by the total RPKM value for the pathway in that sample. We compared mean expression (normalized RPKM values) of the SFT pathway and the DsrAB-DsrC pathway between the different cohorts in the metatranscriptomes (non-IBD vs UC, non-IBD vs CD, and UC vs CD) using a one-factor ANOVA test followed by Tukey's test; there were no statistically discernible differences in the expression means of both pathways between the different cohorts.

E. rectale transcription network analysis

Downloaded and trimmed reads (processed as described above) from the 1090 gut metatranscriptomes were mapped to the *E. rectale* ATCC 33656 genome using bbmap (v37.61,

sourceforge.net/projects/bbmap/) with at a minid of 97%. Feature counting was done using featureCounts within the Subread package (v1.6.3)74. Samples with greater than □ of all *E. rectale* ATCC 33656 genes with at least 1 read mapped (n = 244) were selected for network construction.

Raw read counts for these 244 samples were then converted to copies per million counts (cpm) and normalized using a trimmed mean of M-values (TMM) between each pair of samples using calcNormFactors (edgeR, v3.9)75. The normalized matrix was then used as input into SPIEC-EASI (v1.0.5)85 to construct a sparse and compositionally robust gene expression network using the following settings: method='mb', lambda.min.ratio=0.1, nlambda=10, rep.num=60. Combinations of different lambda min ratios (1e-3 to 1e-1) and nlambda (10 to 100) were used to find a stability threshold near the target (0.05); a lambda min ratio of 0.1 and nlambda of 10 yielded an optimal stability threshold of 0.047. Significant correlations between gene pairs (positive or negative interaction strength of >0.099) with any functional prediction (n=1349 unique genes, 2326 gene- pairs) were then imported into Cytoscape (v3.7.1)86 to visualize the network. Significant clusters were identified using ClusterONE87 with the following settings: min. size=4, min.density=auto, edge.weights=unweighted. The list of significantly upregulated *E. rectale* ATCC 33656 genes from the SQ-amended fecal microcosms (see metatranscriptomics above) was then cross referenced with this network to identify co-expressed gene clusters that respond to SQ amendment.

Phylogenetic and phylogenomic tree construction

Phylogenomic analysis of the 93 MAGs from the fecal incubations was performed with GToTree (v1.4.2)88 using 74 bacterial single-copy genes and a minimum 70% hit fraction for a MAG to be included in the analysis. With this restriction, 66 MAGs were retained for further analysis. FastTree279 was then used to construct approximate maximum likelihood trees using the JTT+CAT model.

Comparative phylogenetic analysis of the glycoside hydrolase family 31 (which includes YihQ/SftG sulfoquinovosidases) was done by downloading the glycoside hydrolase family 31 seed alignment from the pfam database (780 amino acid sequences), aligning the sequences using MAFFT (v7.427)77 (FFT-NS-i-x2 strategy), and then adding 219 putative novel YihQ sequences (70 characterized glycoside hydrolase family 31 proteins, 70 amino acid hits from the Uniprot database using the yihQ HMM model at an e-value of 1e-6, and 79 unique YihQ homologs from the 83 genomes encoding the SFT pathway) and an additional 8696 glycoside hydrolase 31 sequences using the --add and --keeplength options. The 8696 glycoside hydrolase family 31 sequences correspond to 9126 amino acid sequences from uniprot reference proteomes that matched the glycoside hydrolase HMM with an e-value of 1e-40 or less and then clustered at 90% identity using usearch. TrimAl (v1.4.rev15) was used in strict mode to trim the alignment to a final length of 284 amino acid positions. An approximate maximum likelihood tree was then constructed from the alignment using the JTT+CAT model.

For comparative phylogenetic analysis of pyruvate formate-lyase like enzymes and to distinguish isethionate sulfite-lyases (IslA) from DHPS sulfite-lyases (DslA), we retrieved amino acid sequences of known choline trimethylamine-lyases, 4-hydroxyphenylacetate decarboxylases, hydroxyproline

dehydratases, aryl- alkyl-succinate synthases, glycerol dehydratases, and pyruvate formate lyases from the KEGG ligand enzyme nomenclature database (abbreviation E.C.) and clustered them at 90% identity using usearch (v11.0.667)78. All of the sequences were aligned with MAFFT (v7.427)77 using the FFT-NS-i-x2 strategy. 23 glycyl radical enzyme sequences, including known isethionate sulfite-lyases, from genomes of members of the family *Desulfovibrionaceae* originating from the human gut were added to the 1204 reference amino acid sequence alignment with MAFFT using the --add and --keeplength options. TrimAl (v1.4.rev15)89 was used in strict mode to trim the alignment to a final length of 569 positions. An approximate maximum likelihood tree was then constructed from the alignment with FastTree2 using the JTT+CAT model.

Declarations

Acknowledgments

We acknowledge support from the Biomedical Sequencing Facility (https://www.biomedical-sequencing.org) for metagenome and metatranscriptome sequencing, from the Konstanz Proteomics Centre (https://www.biologie.uni-konstanz.de/proteomics-centre/) and from Patricia Wolf for optimizing hybridization conditions for the BWA829 FISH probe during the 2016 International FISH Course at the Division of Microbial Ecology, University of Vienna, Austria. We are thankful to Frank Maixner for initiation of collaboration and to the DOME gut group members in Vienna as well as to our colleagues at University of Konstanz for fruitful discussions and support. This work was financially supported by the Austrian Science Fund (FWF; project grants I2320-B22 and DOC 69-B) and the Deutsche Forschungsgemeinschaft (DFG; grants SCHL1936/3-4; KoRS-CB).

Author contributions

BTH, DK, DS, and AL conceived the study. BTH, DK, JL, AGB, BF, AF, KD, CWH, and NK performed experiments and analyzed data. TR and NS provided bioinformatic support. DS and AL helped with experimental design and data interpretation, and DS helped with writing. BTH, DK, and AL wrote the article. All authors discussed the results and revised the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

All sequence data generated in this project is available at NCBI under BioProject PRJNA593787.

Additional Information

Supplementary Information is available for this paper.

Material requests and correspondence should be addressed to Alexander Loy, alexander.loy@univie.ac.at and David Schleheck, david.schleheck@uni-konstanz.de.

References

1. Zeevi, D. *et al.* Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079–1094 (2015).
2. Heintz-Buschart, A. & Wilmes, P. Human Gut Microbiome: Function Matters. *Trends Microbiol.* **26**, 563–574 (2018).
3. Goddard-Borger, E. D. & Williams, S. J. Sulfoquinovose in the biosphere: occurrence, metabolism and functions. *J* **474**, 827–849 (2017).
4. Barton, L. L., Ritz, N. L., Fauque, G. D. & Lin, H. C. Sulfur Cycling and the Intestinal Microbiome. *Dis. Sci.* **62**, 2241–2257 (2017).
5. Wang, R. Physiological implications of hydrogen sulfide: a whiff exploration that blossomed. *Rev.* **92**, 791–896 (2012).
6. Abu-Ali, G. S. *et al.* Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat Microbiol* **3**, 356–366 (2018).
7. Mehta, R. S. *et al.* Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol* **3**, 347–355 (2018).
8. Schirmer, M. *et al.* Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nat Microbiol* **3**, 337–346 (2018).
9. Lloyd-Price, J. *et al.* Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).
10. Thaiss, C. A. *et al.* Persistent microbiome alterations modulate the rate of post-dieting weight regain. *Nature* **540**, 544–551 (2016).
11. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
12. Barabási, A.-L., Menichetti, G. & Loscalzo, J. The unmapped chemical complexity of our diet. *Nature Food* **1**, 33–37 (2019).
13. Shepherd, S., DeLoache, W. C., Pruss, K. M., Whitaker, W. R. & Sonnenburg, J. L. An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434–438 (2018).
14. Hehemann, J.-H. *et al.* Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908–912 (2010).
15. Bell, A. *et al.* Elucidation of a sialic acid metabolism pathway in mucus-foraging Ruminococcus gnavus unravels mechanisms of bacterial adaptation to the gut. *Nat Microbiol* **4**, 2393–2404 (2019).

16. Benning, C. Biosynthesis and function of the sulfolipid sulfoquinovosyl *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 53–75 (1998).

17. Gupta, S. D. & Sastry, P. S. Metabolism of the plant sulfolipid– sulfoquinovosyldiacylglycerol: degradation in animal tissues. *Biochem. Biophys.* **259**, 510–519 (1987).

18. Denger, K. *et al.* Sulphoglycolysis in Escherichia coli K-12 closes a gap in the biogeochemical sulphur cycle. *Nature* **507**, 114–117 (2014).

19. Felux, A.-K., Spiteller, D., Klebensberger, J. & Schleheck, D. Entner-Doudoroff pathway for sulfoquinovose degradation in Pseudomonas putida SQ1. *Natl. Acad. Sci. U. S. A.* **112**, E4298–305 (2015).

20. Abayakoon, P. *et al.* Structural and Biochemical Insights into the Function and Evolution of Sulfoquinovosidases. *ACS Cent Sci* **4**, 1266–1273 (2018).

21. Burrichter, *et al.* Anaerobic Degradation of the Plant Sugar Sulfoquinovose Concomitant With H2S Production: Escherichia coli K-12 and Desulfovibrio sp. Strain DF1 as Co- culture Model. *Front. Microbiol.* **9**, 2792 (2018).

22. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Biotechnol.* **36**, 996–1004 (2018).

23. Berry, D. *et al.* Tracking heavy water (D2O) incorporation for identifying and sorting active microbial cells. *Natl. Acad. Sci. U. S. A.* **112**, E194–203 (2015).

24. Kraal, L., Abubucker, S., Kota, K., Fischbach, M. A. & Mitreva, M. The prevalence of species and strains in the human microbiome: a resource for experimental efforts. *PLoS One* **9**, e97279 (2014).

25. Ridlon, J. M., Wolf, P. G. & Gaskins, H. R. Taurocholic acid metabolism by gut microbes and colon cancer. *Gut Microbes* **7**, 201–215 (2016).

26. Devkota, S. *et al.* Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice. *Nature* **487**, 104–108 (2012).

27. Speciale, G., Jin, Y., Davies, G. J., Williams, S. J. & Goddard-Borger, E. D. YihQ is a sulfoquinovosidase that cleaves sulfoquinovosyl diacylglyceride sulfolipids. *Chem. Biol.* **12**, 215–217 (2016).

28. Pasolli, E. *et al.* Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* **176**, 649–662.e20 (2019).

29. Peck, S. C. *et al.* A glycyl radical enzyme enables hydrogen sulfide production by the human intestinal bacterium. *Natl. Acad. Sci. U. S. A.* **116**, 3171–3176 (2019).

30. Mayer, J. *et al.* 2,3-Dihydroxypropane-1-sulfonate degraded by Cupriavidus pinatubonensis JMP134: purification of dihydroxypropanesulfonate 3-dehydrogenase. *Microbiology* **156**, 1556–1564 (2010).

31. Xing, M. *et al.* Radical-mediated C-S bond cleavage in C2 sulfonate degradation by anaerobic bacteria. *Nature Communications* 10 (2019).

32. Levin, B. J. & Balskus, E. P. Discovering radical-dependent enzymes in the human gut microbiota. *Opin. Chem. Biol.* **47**, 86–93 (2018).
33. Laue, H., Friedrich, M., Ruff, J. & Cook, A. M. Dissimilatory sulfite reductase (desulfoviridin) of the taurine-degrading, non-sulfate-reducing bacterium Bilophila wadsworthia RZATAU contains a fused DsrB-DsrD subunit. *Bacteriol.* **183**, 1727–1733 (2001).
34. Santos, A. A. *et al.* A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science* **350**, 1541–1545 (2015).
35. Liu, J. *et al.* Two radical-dependent mechanisms for anaerobic degradation of the globally abundant organosulfur compound dihydroxypropanesulfonate. *Natl. Acad. Sci. U. S. A.* (2020) doi:10.1073/pnas.2003434117.
36. Huxtable, R. J., Franconi, F. & Giotti, A. *The Biology of Taurine: Methods and Mechanisms*. (Springer Science & Business Media, 1987).
37. Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).
38. Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, R79 (2012).
39. Rey, F. E. *et al.* Metabolic niche of a prominent sulfate-reducing human gut *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13582–13587 (2013).
40. Tian, Y., Wang, K. & Ji, G. Faecalibacterium Prausnitzii Prevents Tumorigenesis in a Model of Colitis-Associated Colorectal Cancer. *Gastroenterology* 152 S354–S355 (2017).
41. Ijssennagger, N. *et al.* Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Natl. Acad. Sci. U. S. A.* **112**, 10038–10043 (2015).
42. Wallace, J. L., Motta, J.-P. & Buret, A. G. Hydrogen sulfide: an agent of stability at the microbiome-mucosa interface. *J. Physiol. Gastrointest. Liver Physiol.* **314**, G143–G149 (2018).
43. Carbonero, F., Benefiel, A. C., Alizadeh-Ghamsari, A. H. & Gaskins, H. R. Microbial pathways in colonic sulfur metabolism and links with health and disease. *Physiol.* **3**, 448 (2012).
44. Pomare, E. W., Branch, W. J. & Cummings, J. H. Carbohydrate fermentation in the human colon and its relation to acetate concentrations in venous blood. *Clin. Invest.* **75**, 1448– 1454 (1985).
45. Zambell, K. L., Fitch, M. D. & Fleming, S. E. Acetate and butyrate are the major substrates for de novo lipogenesis in rat colonic epithelial cells. *Nutr.* **133**, 3509–3515 (2003).
46. Fu, M. *et al.* Hydrogen sulfide (H2S) metabolism in mitochondria and its regulatory role in energy production. *Natl. Acad. Sci. U. S. A.* **109**, 2943–2948 (2012).
47. Daims, H., Stoecker, K. & Wagner, M. Fluorescence in situ hybridization for the detection of prokaryotes. *Molecular microbial ecology* (2004).
48. Berry, D. *et al.* Phylotype-level 16S rRNA analysis reveals new bacterial indicators of health state in acute murine colitis. *ISME J.* **6**, 2091–2106 (2012).

49. Cline, J. D. Spectrophotometric determination of hydrogen sulfide in natural *Limnol. Oceanogr.* **14**, 454–458 (1969).
50. Ludwig, W. *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
51. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–6 (2013).
52. Daims, H., Brühl, A., Amann, R., Schleifer, K. H. & Wagner, M. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* **22**, 434–444 (1999).
53. Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Environ. Microbiol.* **66**, 5488– 5491 (2000).
54. Herbold, C. W. *et al.* A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. *Microbiol.* **6**, 731 (2015).
55. Pelikan, C. *et al.* Diversity analysis of sulfite- and sulfate-reducing microorganisms by multiplex dsrA and dsrB amplicon sequencing using new primers and mock community- optimized bioinformatics. *Microbiol.* **18**, 2994–3009 (2016).
56. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon *Nat. Methods* **10**, 996–998 (2013).
57. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Environ. Microbiol.* **73**, 5261–5267 (2007).
58. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community- supported software for describing and comparing microbial communities. *Environ. Microbiol.* **75**, 7537–7541 (2009).
59. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
60. Müller, A. L., Kjeldsen, K. U., Rattei, T., Pester, M. & Loy, A. Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J.* **9**, 1152–1165 (2015).
61. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
62. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
63. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14 Suppl 1**, S7 (2013).
64. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).

65. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
66. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
67. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
68. Wheeler, T. J. & Eddy, S. R. nhmmer: DNA homology search with profile *Bioinformatics* **29**, 2487–2489 (2013).
69. Kalvari, I. *et al.* Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* **46**, D335–D342 (2018).
70. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
71. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Commun.* **9**, 5114 (2018).
72. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078– 2079 (2009).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
75. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
76. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
77. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Biol. Evol.* **30**, 772–780 (2013).
78. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
79. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
80. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
81. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
82. Browne, H. P. *et al.* Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543–546 (2016).

83. Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Methods* **15**, 962–968 (2018).
84. Truong, D. T. *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Methods* **12**, 902–903 (2015).
85. Kurtz, Z. D. *et al.* Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput. Biol.* **11**, e1004226 (2015).
86. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
87. Nepusz, T., Yu, H. & Paccanaro, A. Detecting overlapping protein complexes in protein- protein interaction networks. *Methods* **9**, 471–472 (2012).
88. Lee, M. D. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* **35**, 4162–4164 (2019).
89. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
90. Denger, K. & Cook, A. M. Racemase activity effected by two dehydrogenases in sulfolactate degradation by Chromohalobacter salexigens: purification of (S)-sulfolactate dehydrogenase. *Microbiology* **156**, 967–974 (2010).

Supplemental Legends

Extended data figure legends

Extended Data Fig. 1 Sulfoquinovose is rapidly degraded in anoxic microcosms with human feces. a. Overview of fecal microcosm experimental plan. Five microcosm treatments were constructed in triplicates and subsampled at 10 time points. At all time points, metabolite quantification and 16S rRNA gene and *dsrB* amplicon sequencing were performed. Time points selected for metagenome sequencing are shown as green dots. Time points selected for metatranscriptome sequencing are shown as blue dots. **b**-**f**. Microbial metabolite dynamics in microcosms **b**) without amendment, **c**) with 10 mM SQ, **d**) with 10 mM glucose, **e**) with 10 mM taurine plus 10 mM formate. In **b**, residual microbial activity showed no endogenously produced DHPS or H2S. In **c**, added SQ is rapidly degraded to DHPS and acetate, while internally produced formate is consumed likely as an electron donor for sulfite respiration. In **d**, glucose addition promotes microbial production of mainly acetate, propionate, and butyrate. In **e**, formate is rapidly consumed as an electron donor for taurine-derived sulfite respiration with the production of more acetate. Sulfite accumulation after the depletion of formate might be due to lack of sufficient electron donor. An accumulation of hydroxyacetone as for the *B. wadsworthia* pure culture (**Extended Data Fig. 5a**) was not detectable in the microcosms, suggesting consumption of hydroxyacetone by other microorganisms. In **f**, comparison of H2S amounts in the five microcosm treatments indicated strong production of H2S in the presence of taurine (serving as a positive control for organosulfonate respiration) and of SQ (both SQ alone and in the presence of 50% D2O). **g, h.** Relative abundance of *E.*

rectale (OTU 7, panel **g**) and *B. wadsworthia* (OTU 15, panel **h**) 16S rRNA genes over time in the 5 microcosm treatments. **i**. Three *dsrB* OTUs showed significant increases over time in SQ-amended microcosms (closed circles) compared to no amendment controls (open circles). *dsrB* OTUs were characterized at the genus level as *Desulfovibrio* (OTU 1, black lines; OTU 8, blue lines) and *Bilophila* (OTU 5, red lines). * indicate time points with significant differences in relative abundances between the no amendment and SQ- amended microcosms ($P < .01$). **j**. Quantification of %CD-labeling using Raman microspectrometry of individual microbial cells in SQ-amended microcosms containing 50% heavy water (D2O). The total number of cells measured at each time point is shown with the percent of those cells with a %CD above a threshold cutoff of 3.54 %CD (black bar). In panels **b-i**, points represent averages of triplicate measures. Error bars represent one standard deviation. SQ, sulfoquinovose; DHPS, 2,3-dihydroxypropane-1-sulfonate; OTU, operational taxonomic unit; Glu, glucose treatment; NA, no amendment control.

Extended Data Fig. 2 Addition of sulfoquinovose to human feces microcosms triggers only minor shifts in microbiota composition. a. Relative abundance of the nine OTUs (out of 791 OTUs total) that showed significant increases ($P < .01$) in SQ-amended microcosms (triangles) compared to no-amendment controls (circles). Points represent averages of triplicate measures. Error bars represent one standard deviation. Identity of the OTUs was revealed by phylogenetic placement into a reference tree with close relatives. Percentages in brackets indicate the 16S rRNA gene similarity to the closest reference species. **b, c**. Principal component analysis plotting the shifts in microcosm beta-diversity (Bray-Curtis dissimilarity) over time. Points represent 16S rRNA gene amplicon libraries generated from biologically replicated microcosms. In **b**, all treatments are shown. In **c**, only the no-amendment and sulfoquinovose treatments are shown to emphasize community shifts driven by SQ. SQ, sulfoquinovose; OTU, operational taxonomic unit.

Extended Data Fig. 3 Comparative phylogenetic analysis of glycoside hydrolase family 31 enzymes, including YihQ/SftG sulfoquinovosidases. The glycoside hydrolase family 31 phylogeny was reconstructed using FastTree279 and the scale bar indicates the number of amino acid substitutions per position. The boxed inset depicts the phylogeny of putative and biochemically characterized (depicted in red) YihQ/SftG sulfoquinovosidases. Organisms that were identified to encode all four core enzymes of the SFT fermentation pathway are highlighted in green. Experimental characterizations of function from the CAZy database are depicted as red circles on the branch tips.

Extended Data Fig. 4 A genome-wide *E. rectale* transcription network reveals co-expression of a sulfoquinovosidase gene-containing gene cluster in human fecal samples. Expression of *yihQ/sftG* and 10 co-localized genes from *E. rectale* ATCC 33656 are positively correlated to one another across 1,090 human gut metatranscriptomes6–8. Circles (nodes) represent individual *E. rectale* genes and arrows depict significant interactions (edges); grey arrows represent significant positive (> 0.099) correlations and red arrows represent significant (< -0.099) negative correlations between genes. Nodes highlighted in green are genes whose expressions are significantly upregulated in the presence of 10 mM SQ in human fecal microcosms. Clusters are arranged in order of cluster size, with the largest cluster at the top of the

network. The only cluster composed solely of genes upregulated in SQ-amended fecal microcosms is the SFT fermentation gene cassette. The inset depicts the identities of the individual genes in this cluster, with the operon from the genome sequence below the inset. The gene denoted with an asterisk (EUBREC_RS11655) encodes a putative transcriptional regulator and does not appear in the network due to low interaction strength with any node (correlation 0.000 < * < 0.099).

Extended Data Fig. 5 Identification of a DHPS-desulfonating enzyme (DslA) in *B. wads- worthia* 3.1.6 and its distribution among selected sulfur compound-metabolising human gut bacteria. a. Growth of *B. wadsworthia* 3.1.6 (OD580 nm; blue line) with DHPS as an electron acceptor (black line) and lactate as an electron donor (orange line), showing production of acetate (gray line), H2S (yellow line), and of hydroxyacetone (red line). Production of hydroxyacetone from the desulfonation of DHPS implies a new type of reaction (attributed to DslA) analogous to the desulfonation of isethionate29. **b.** Presence/absence of organosulfur metabolism genes in the genomes of selected *dsrABC*-encoding microbes from the human gut and the *Bilophila* sp. MAG from the microcosm (depicted in boldface). **c**. Differential proteomics of *B. wadsworthia* 3.1.6 grown with lactate and either DHPS, SL, isethionate or taurine as electron acceptors (each n=1) illustrates substrate-specific induction of different (predicted) pathway enzymes. For all identified inducible genes, RefSeq locus tag numbers are shown (prefix HMPREF0179_RS) (panels **c**, **d**). In addition, DslAB in *Desulfovibrio desulfuricans* strain DSM642 (panel **b**) was also strongly induced during DHPS utilization, as confirmed by differential proteomics in comparison to sulfate, and strain DSM642 also excreted hydroxyacetone during growth with DHPS (data not shown). **d**. Overview of the identified genes located in separate gene clusters in *B. wadsworthia* 3.1.6 that are strongly induced during growth with either DHPS (blue), SL (red), or isethionate (dark green), and with taurine (light green, for taurine metabolism *via* the isethionate-desulfonation pathway29). In addition to the inducible TRAP-transport-system receptors identified by proteomics (**c**), the gene clusters for DHPS and SL also encode TRAP-transport system membrane permeases that were not identified by total proteomics (indicated in white); they most likely mediate transport of DHPS and SL, respectively. The two SL dehydrogenase genes identified in the SL-utilization gene cluster (14460, 14455 in **c**, **d**) are homologous to SlcC and ComC of *Chromohalobacter salexigens*, which catalyze an interconversion of (S)-3-sulfolactate to (R)-3-sulfolactate via sulfopyruvate, respectively90 (not shown in **d,** shown in **b**). **e**. Illustration of the enzyme reaction for the newly identified DHPS-inducible GRE (blue), shown desulfonating DHPS to hydroxyacetone and sulfite (termed DHPS sulfite-lyase, DslAB) in a reaction analogous to that of the first identified desulfonating GRE, isethionate sulfite-lyase (green) (IslAB)29, and the desulfonation reaction of 3-sulfolactate sulfo-lyase (SuyAB) (red), a thiamindiphosphat [TDP] dependent enzyme21. The desulfonation reaction of DHPS to sulfite and hydroxyacetone was detected in strictly anoxic cell- free extracts of DHPS-grown cells (**Fig. 2b**) but not in the presence of oxygen (**Fig. 2b**). Note that isethionate sulfite-lyase IslAB of *B. wadsworthia* 3.1.6 cannot catalyze a desulfonation reaction with DHPS29. 2-HES, 2-hydroxyethane-1-sulfonic acid (isethionate); DHPS, 2,3- dihydroxypropane-1-sulfonate; HSA, 2-oxo-3-hydroxy-propane-1-sulfonate; SLA, 3- sulfolactaldehyde; SL, 3-sulfolactate; SA, sulfoacetaldehyde; SP, sulfopyruvate; GRE, glycyl radical enzyme.

Extended Data Fig. 6 Genes for the DHPS sulfite-lyase (*dslA*) and isethionate sulfite-lyase (*islA*) of *B. wadsworthia* are paralogous. Phylogenetic analysis of pyruvate formate lyases and glycyl radical enzymes from the KEGG ligand enzyme nomenclature database and *Desulfovibrionaceae* family genomes shows that isethionate sulfite-lyases (encoded by *islA*) and DHPS sulfite-lyases (encoded by *dslA*) are closely related but phylogenetically distinct. Amino acid identities between IslA and DslA ranged from 54% to 64%; and were greater than 76% and 81% within DslA and IslA groups, respectively. Experimental characterization of function from the Uniprot database are depicted as red circles on the branch tips. DHPS, 2,3- dihydroxypropane-1-sulfonate.

Extended Data Fig. 7 *Eubacterium rectale* and *Bilophila wadsworthia* are the dominant contributors to expression of the SFT fermentation and H2S production pathways in human stool metatranscriptomes, respectively. Percentage and normalized read contribution (in RPKM) of various *Firmicutes* and gut sulfidogens to expression of the SFT pathway, *dslA* dependent DHPS catabolism (*dslAB*), taurine/isethionate catabolism (*islAB*), and sulfite reduction to sulfide (*dsrABC*) in publically available human stool metatranscriptomes from a non-IBD cohort (n=130), ulcerative colitis cohort (n=203), Crohn's disease cohort (n=333), and HPFS study (n=365)6–9. Each bar corresponds to 1 sample. Contributions to expression (%) are fractional with 1.0 corresponding to 100%. Abbreviations: IBD, inflammatory bowel disease; HPFS, health professionals follow-up study; RPKM, reads per kilobase per million reads; SQ, sulfoquinovose; DHPS, 2,3-dihydroxypropane-1-sulfonate.

Supplementary information table legends

Supplementary Information Table 1. 16S rRNA gene amplicon OTU table. Summary of 16S rRNA gene amplicon read counts for each microcosm treatment (No amendment, 10 mM Glucose, 10 mM taurine + 10 mM formate, 10 mM sulfoquinovose, and 10 mM sulfoquinovose + 50% D2O), replicates, and time points. OTUs were clustered at 97% and taxonomic assignment is shown (domain_id, phylum_id, class_id, order_id, family_id, and genus_id) with the associated confidence scores (domain_conf, phylum_conf, class_conf, order_conf, family_conf, and genus_conf). Contaminant OTUs were determined from replicate DNA extraction blank controls and no-template PCR controls. Removed contaminant OTUs (149 OTUs) are presented at the bottom of this sheet.

Supplementary Information Table 2. Results of 16S rRNA gene DESeq2 analysis. Results from significance testing using DESeq2 comparing OTU counts between 16S rRNA gene libraries generated from fecal microcosms receiving no amendment vs. 10 mM sulfoquinovose across paired time points. OTU_name; Time_point_(hours), time point of incubation in which DNA was harvested from; default output from DESeq2 (baseMean; log2FoldChange; lfcSE, log2FoldChange standard error; stat; pvalue; padj, adjusted pvalue); followed by taxonomic assignment (domain_id, phylum_id, class_id, order_id, family_id, and genus_id) and confidence scores(domain_conf, phylum_conf, class_conf, order_conf, family_conf, and genus_conf)54.

Supplementary Information Table 3. *dsrB* amplicon OTU table. Summary of *dsrB* amplicon read counts for each microcosm treatment (No amendment, 10 mM Glucose, 10 mM taurine + 10 mM formate, 10 mM sulfoquinovose, and 10 mM sulfoquinovose + 50% D2O), replicates, and time points. OTUs were clustered at 97% and taxonomic assignment is shown (Gene_type, Domain, Phylum, Class, Order, Family, and Genus) according to Pelikan et al., 2015 and Müller et al. 201555,60. No contaminant OTUs were detected as determined from replicate DNA extraction blank controls and no-template PCR controls.

Supplementary Information Table 4. Results of *dsrB* DESeq2 analyses. Results from significance testing using DESeq2 comparing OTU counts between dsrB gene libraries generated from fecal microcosms receiving no amendment vs. 10 mM sulfoquinovose across paired time points. OTU_name; Time_point_(hours), time point of incubation in which DNA was harvested from; default output from DESeq2 (baseMean; log2FoldChange; lfcSE, log2FoldChange standard error; stat; pvalue; padj, adjusted pvalue); followed by taxonomic assignment (Gene_type, Domain, Phylum, Class, Family, Genus)55,60.

Supplementary Information Table 5. Metagenome assembled genomes from microcosms. Summary of metagenome assembled genomes (MAGs) dereplicated across three metagenomic datasets prepared from SQ-amended microcosms (time points 28, 73, and 144 h). MAG ID (column 1) sorted by MAG completeness. Bolded entries indicate MAGs identified as *Eubacterium rectale* and *Bilophila wadsworthia* according to average nucleotide identities (ANI) with genomes of publicly available type strains. Genome taxonomy based on concatenated genes sets (column 2) according to the Genome Taxonomy Database22. MAG quality was assessed using CheckM66 with columns 3-13 showing resultant summary data. Strain heterogeneity (%, column 14) was determined using dRep67. Presence of 16S SSU rRNA sequences in each MAG (column 15) was determined with nhmmer68 using rfam69 models for bacterial and archaeal SSU rRNAs (RFAM: RF00177, RF01959). Minimum overlap between sequence and model was 300 nucleotides. Sequences were classified using the RDPclassifer57 as implemented in Mothur. The number of tRNAs identified and the number of amino acids encoded by identified tRNAs (columns 16 and 17, respectively) were determined using tRNAscan-SE70.

Supplementary Information Table 6. *E. rectale* genome average nucleotide identity matrix.

Supplementary Information Table 7. *B. wadsworthia* genome average nucleotide identity matrix.

Supplementary Information Table 8. Significance values for *E. rectale* transcripts expressed in microcosms. Locus tags, gene annotations, normalized read counts, and EdgeR exact test results for read mapping to *Eubacterium rectale* ATCC 33656 in the triplicate SQ-amended microcosms at 6 hours, 20 hours, and 52 hours. P-values highlighted in red denote significantly differentially expressed genes ($P < 0.05$) at 20 hours compared to 6 hours. Genes highlighted in red denote genes that make up the SFT pathway gene cassette.

Supplementary Information Table 9. Significance values for *B. wadsworthia* transcripts expressed in microcosms. Locus tags, gene annotations, normalized read counts, and EdgeR exact test results for read mapping to *Bilophila wadsworthia* HMP 3.1.6 in the SQ-amended microcosms at 6 hours, 20 hours, and

52 hours. P-values highlighted in red denote significantly differentially expressed genes (P < 0.05) at 20 hours compared to 6 hours.

Supplementary Information Table 10. SQ gene-containing MAGs from the human gut.

Supplementary Information Table 11. FISH probes used in this study. Newly designed 16S rRNA-targeted oligonucleotide probes for fluorescence *in situ* hybridization. Probes and their respective unlabeled competitors were used in equimolar concentrations during hybridization.

Supplementary Information Table 12. SQ-containing MAGs from a MAG/genome database of human gut microorganisms. The MAG/genome database is available at http://segatalab.cibio.unitn.it/data/Pasolli_et_al.html and http://opendata.lifebit.ai/table/?project=SGB.

Figures

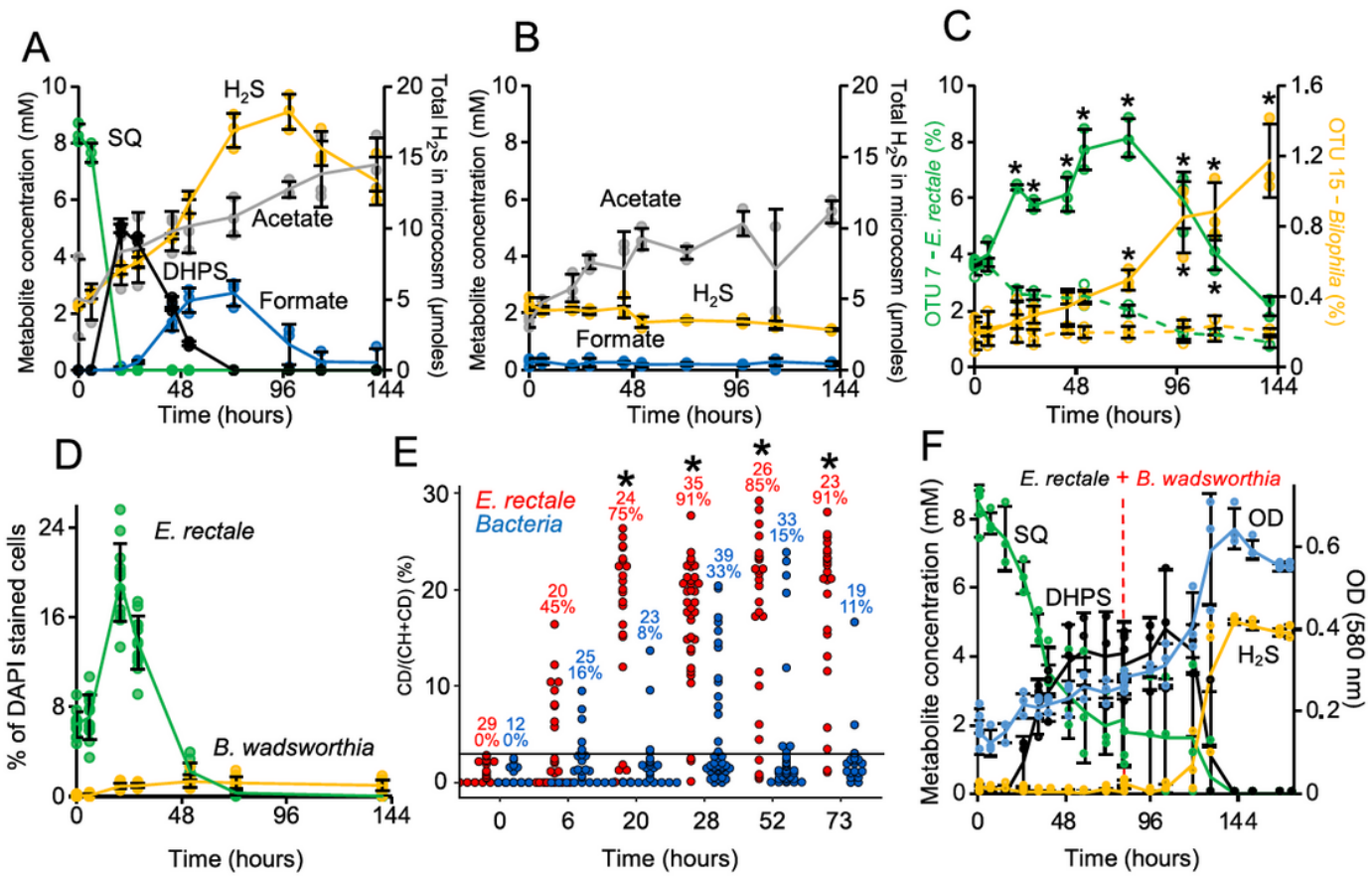

Figure 1

Sulfoquinovose is anaerobically degraded to H2S by DHPS-crossfeeding between Eubacterium rectale and Bilophila wadsworthia. a, Degradation of SQ (green) in anoxic human fecal microcosms resulted in transient production of DHPS (black) and formate (blue) and accumulation of H2S (yellow) and acetate

(grey). Formate is a possible electron donor for respiration of DHPS-derived sulfite. b, Unamended control microcosms did not produce DHPS, formate, or H2S. c, Relative abundance of E. rectale (green) and B. wadsworthia (yellow) 16S rRNA gene OTUs in microcosms with SQ (solid lines) and without amendment (dashed lines). d, FISH quantification of E. rectale and B. wadsworthia in SQ microcosms. Error bars represent one standard deviation of averages from 12 microscopic fields. e, FISH-Raman analysis of single cell activity (%CD labeling) of E. rectale (red) and Bacteria (blue) in SQ-amended fecal microcosms containing 50% D2O. Numbers indicate the number of cells analyzed and percentages of active cells (above a 3.54 %CD background threshold; black bar) and asterisks indicate that FISH- positive E. rectale cells are significantly enriched in CD vs. 0 h ($P < .01$). f, Two-step degradation of SQ via DHPS to H2S by a co-culture of E. rectale DSM 17629 (grown in pure culture for 81 h) and B. wadsworthia 3.1.6 (co-inoculated at 81 h, dotted line). In a, b, and c, lines represent averages of triplicate measures with error bars representing one standard deviation. Asterisks show significant differences ($P < .01$) in OTU relative abundance between matched time points of SQ and unamended microcosms. SQ, sulfoquinovose; DHPS, 2,3-dihydroxypropane-1-sulfonate; OTU, operational taxonomic unit; FISH, fluorescence in situ hybridization; DAPI, 4',6-diamidino- 2-phenylindole.

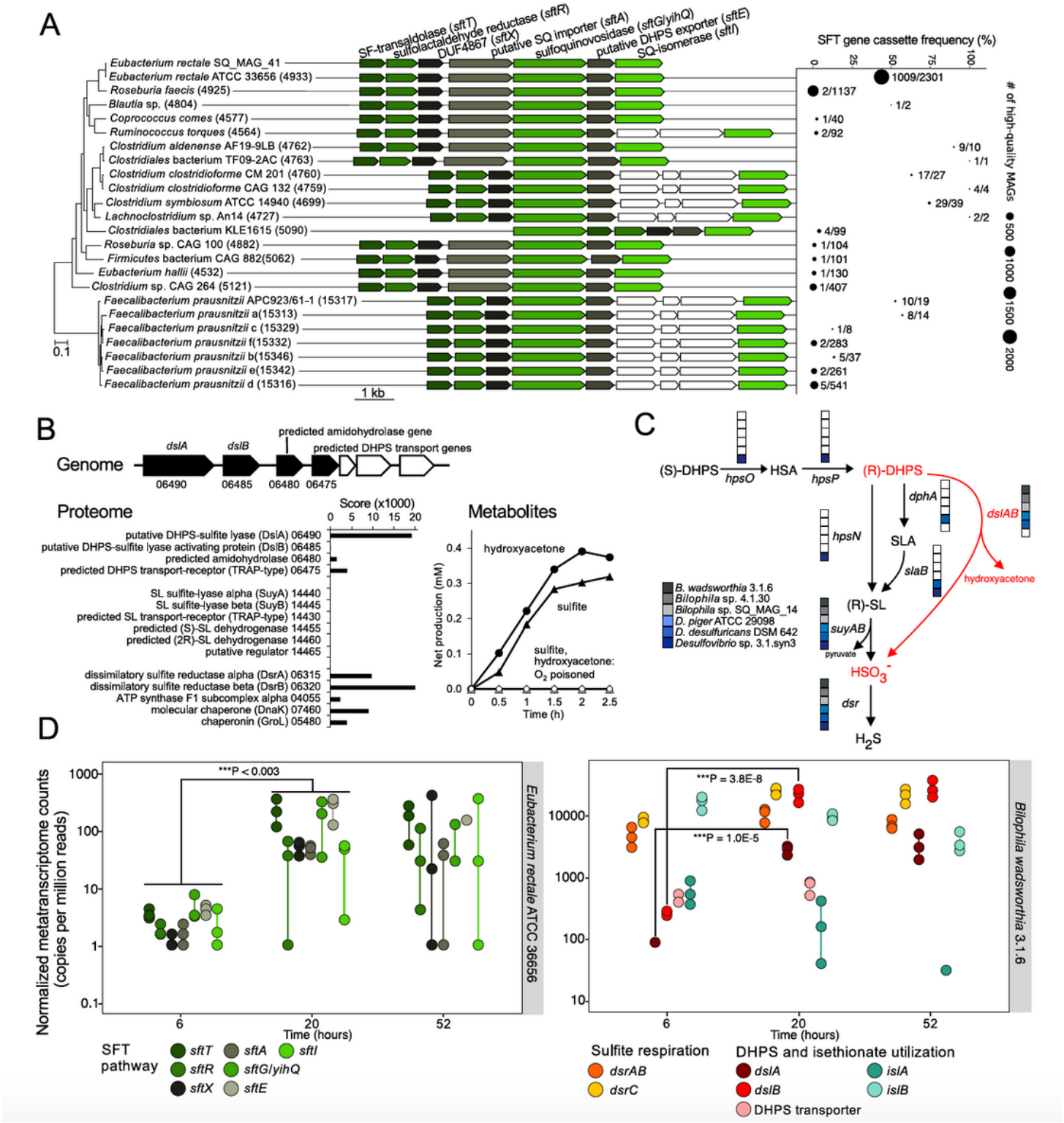

Figure 2

Distribution, composition, and expression of genes involved in sulfoquinovose degradation by Eubacterium rectale and DHPS degradation by Bilophila wadsworthia. a, Structure of the SQ-utilization gene cluster for SQ utilization via the SFT pathway in human gut bacteria and the E. rectale SQ_MAG_41 recovered from the microcosms. Genes depicted in white have no predicted function. The four core genes of the SFT pathway (in green shades) are present in 23 species-level MAG clusters (of 4930 total clusters;

each >95% MAG ANI) that are Lachnospiraceae and Ruminococcaceae family members. MAG clusters are ordered based on the phylogeny of representative genomes. Numbers in parentheses denote MAG species-cluster IDs28. SFT gene cassette frequency across each species-level MAG cluster with point labels indicating the number of MAGs containing the SQ-gene cassette with respect to the total number of MAGs in each species-level cluster. b, Structure of the DHPS utilization gene cluster in B. wads- worthia 3.1.6 and its inducibly expressed proteins (bar graph) during growth with DHPS as electron acceptor (numbers refer to RefSeq locus tag numbers; prefix HMPREF0179_RS) (Extended Data Fig. 5c shows a comparison to growth with taurine, isethionate and 3-sulfolactate as electron acceptors). Metabolite analysis (line graph) of cell-free extracts of DHPS-grown cells indicated cleavage of DHPS into sulfite and hydroxyacetone, if the reaction was performed under strictly anoxic conditions29; representative results (n=5). c, Presence/absence of DHPS-utilization pathway genes in B. wadsworthia genomes and other selected dsrABC-encoding human gut Desulfovibrionaceae. d, Expression of pathways for SQ utilization by E. rectale (transcripts mapped to E. rectale ATCC 36656) (left) and DHPS utilization and sulfite respiration by B. wadsworthia (transcripts mapped to B. wadsworthia 3.1.6) (right) in the metatranscriptomes of triplicate (n=3) fecal microcosms at 6, 20, and 52 h after amendment with 10 mM SQ. Vertical lines connect non-overlapping replicate data points and asterisks indicate significant difference in expression level. DUF, domain of unknown function; SF, 6-deoxy-6-sulfofructose; SQ, sulfoquinovose; DHPS, 2,3-dihydroxypropane-1-sulfonate; HSA, 2-oxo-3-hydroxy-propane-1- sulfonate; SLA, 3-sulfolactaldehyde; SL, 3-sulfolactate.

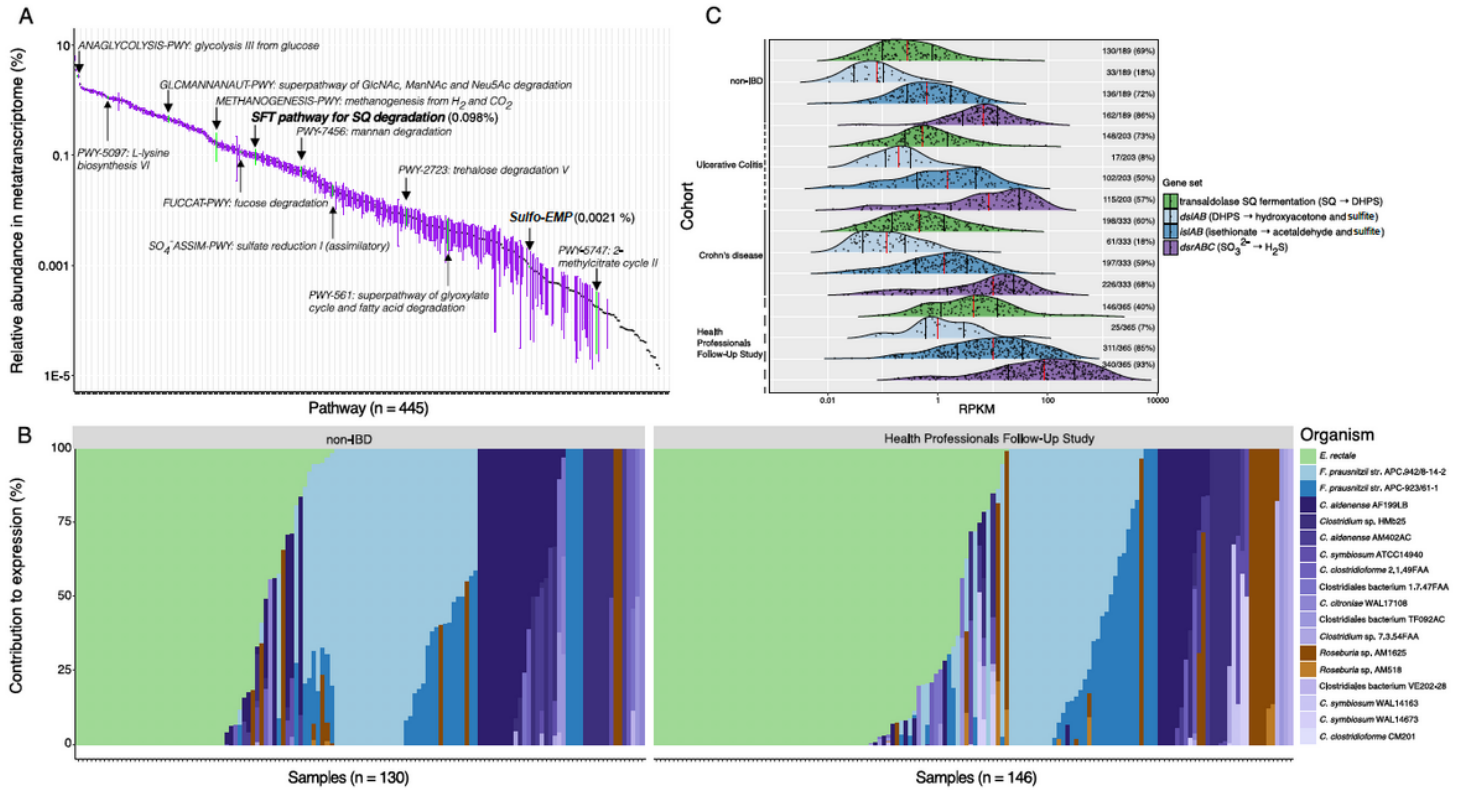

Figure 3

Expression of microbial pathways for SQ fermentation and H2S production are abundantly and frequently detected in human stool metatranscriptomes. a, Ranked relative abundance of 445 identified pathways in 1090 human stool metatranscriptomes from 201 individuals; each point is the mean relative abundance of a pathway and error bars correspond to the 95% confidence interval of the mean. b, Relative % contribution of various Firmicutes (depicted in various colors) to expression of the SFT pathway in a non-IBD cohort (n=130) and the HPFS cohort (n=146). Each bar corresponds to 1 sample. See methods for references. c, Expression of the SFT pathway, DHPS and isethionate utilization (dslAB and islAB, respectively), and sulfite respiration (dsrABC) in 1090 human stool metatranscriptomes from 201 individuals from 4 cohorts. The height and area of the ridgeline plots correspond to the distribution of expression. Each point corresponds to 1 sample, the red vertical lines depict the median, and the black lines correspond to the 25th and 75th percentiles. The number of samples with detectable expression within each group is indicated to the right of each plot. IBD, irritable bowel disease; HPFS, Health Professionals Follow-up Study; RPKM, reads per kilobase per million reads; SQ, sulfoquinovose; DHPS, 2,3-dihydroxypropane-1-sulfonate.

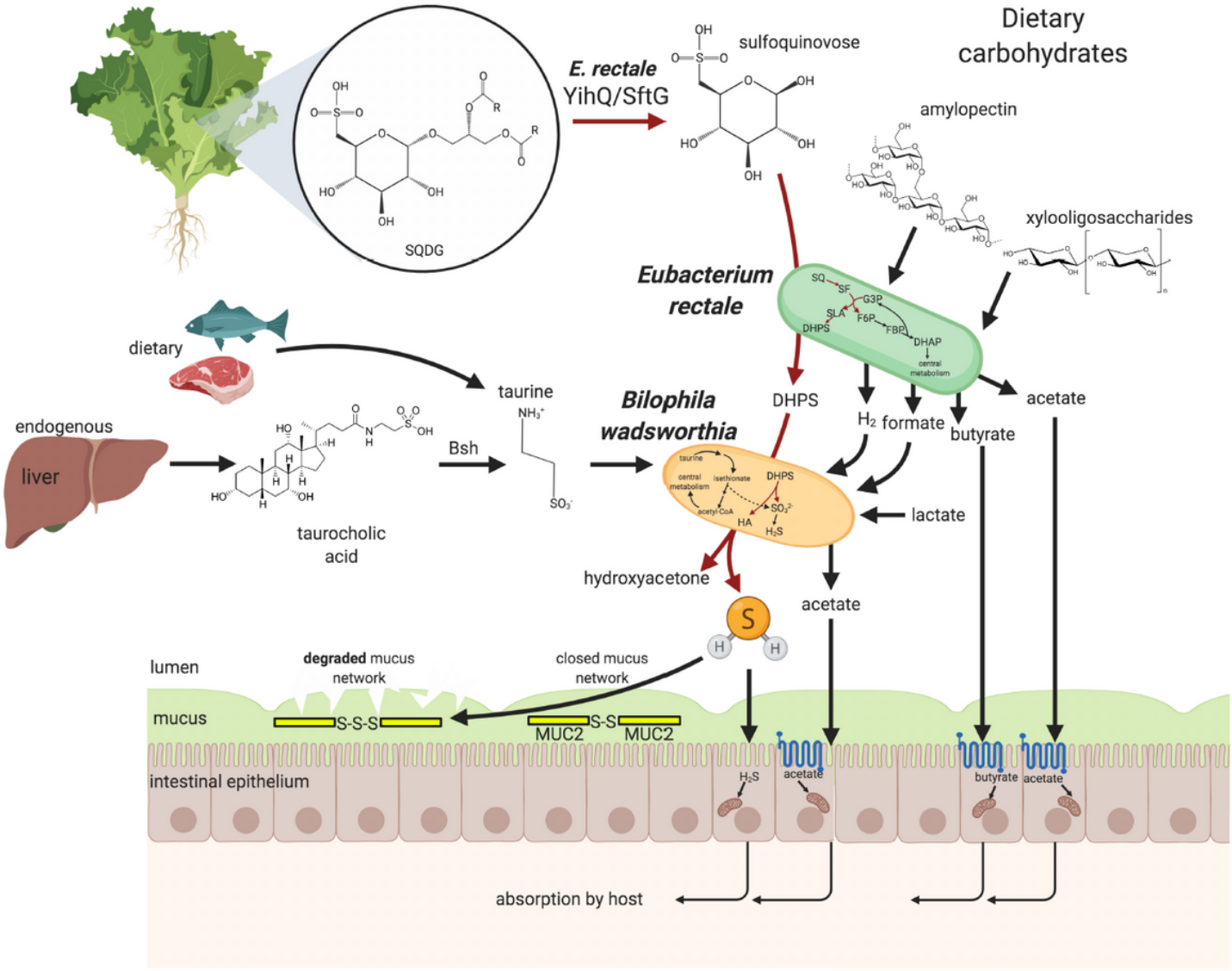

Figure 4

Sulfur energy metabolism and physiological interaction scheme of E. rectale and B. wadsworthia in the human gut. SQDG present in green vegetables enters the intestine as a common component of the human diet. Sulfoquinovosidase20, encoded by yihQ/sftG, cleaves the glycosidic linkage of the lipid tail to liberate SQ, which is then fermented by E. rectale to DHPS, acetate, and likely also to formate or H2 and CO2. Further plant-derived compounds, such as amylopectin, xylooligosaccharides and other carbohydrates, can also be catabolized and fermented by E. rectale. Other Firmicutes, such as members of the genera Faecalibacterium and Roseburia, can also function as primary SQ degraders in some individuals. B. wadsworthia cleaves the C-S bond in DHPS to produce sulfite for dissimilatory reduction to H2S gas coupled to oxidation of, e.g., formate, lactate, and hydrogen, while hydroxyacetone and also acetate are excreted. Taurine is present in dietary meat/fish and is liberated from host-secreted taurocholic acids by microbial bile salt hydrolases (encoded by bsh), and thus far was considered the

sole major source of sulfonate-sulfur for respiration and H2S production by Bilophila in the human intestine. H2S and short chain fatty acids are utilized by colonic epithelial cells as energy sources, act as signaling molecules, and are absorbed and further distributed by the host via the bloodstream44–46. Excessive concentrations of H2S gas can break the mucus layer in the colon by reduction of Muc2-mucin disulfide bonds to trisulfides41, thereby allowing bacteria to penetrate the mucus layer and interact with the host epithelial lining. Arrows colored in black represent previously described metabolic pathways, while those in dark red represent pathways revealed in this study. SQDG, sulfoquinovosyl diacylglycerol; SQ, sulfoquinovose (6-deoxy-6-sulfoglucose); SF, 6-deoxy-6- sulfofructose; G3P, glyceraldehyde-3-phosphate; F6P, fructose-6-phosphate; FBP, fructose-1,6- bisphosphate; DHAP, dihydroxyacetone phosphate; SLA, 3-sulfolactaldehyde; DHPS, 2,3- dihydroxypropane-1-sulfonate; HA, hydroxyacetone.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HansonKitsetalSupplementaryInformationTables.xlsx
- ExtendedFigures.pdf