

1 **The capacity to produce hydrogen sulfide (H_2S)**
2 **via cysteine degradation is ubiquitous in the**
3 **human gut microbiome**

4
5 **Domenick J. Braccia¹, Xiaofang Jiang², Mihai Pop^{1,3}, A. Brantley Hall^{1,4}**

6
7 ¹ Center for Bioinformatics and Computational Biology, University of Maryland, College Park,
8 Maryland, USA

9 ²National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA

10 ³Department of Computer Science, University of Maryland, College Park, Maryland, USA

11 ⁴ Department of Cell Biology and Molecular Genetics, University of Maryland, College Park,
12 Maryland, USA

13 **Abstract**

14
15 As one of the three mammalian gasotransmitters, hydrogen sulfide (H_2S) plays a major role in
16 maintaining physiological homeostasis. Endogenously produced H_2S plays numerous beneficial
17 roles including mediating vasodilation and conferring neuroprotection. Due to its high membrane
18 permeability, exogenously produced H_2S originating from the gut microbiota can also influence
19 human physiology and is implicated in reducing intestinal mucosal integrity and potentiating
20 genotoxicity and is therefore a potential target for therapeutic interventions. Gut microbial H_2S
21 production is often attributed to dissimilatory sulfate reducers such as *Desulfovibrio* and
22 *Bilophila* species. However, an alternative source for H_2S production, cysteine degradation, is
23 present in gut microbes, but the genes responsible for cysteine degradation have not been
24 systematically annotated in gut microbes. To better understand the potential for H_2S production
25 via cysteine degradation by the human gut microbiome, we performed a comprehensive search
26 for genes encoding cysteine-degrading genes in 4,644 bacterial genomes from the Unified
27 Human Gastrointestinal Genome (UHGG) catalogue. We identified 407 gut bacterial species as
28 putative cysteine degrading bacteria, 328 of which have not been previously implicated in H_2S
29 production. We identified the presence of at least one putative cysteine degrading bacteria in
30 metagenomic data of 100% of 6,644 healthy subjects and the expression of cysteine-degrading
31 genes in metatranscriptomics data of 100% of 59 samples. Additionally, putative cysteine-
32 degrading bacteria are more abundant than sulfate reducing bacteria ($p < 2.2e-16$). Overall, this
33 study improves our understanding of the capacity for H_2S production by the human gut
34 microbiome and may help to inform interventions to therapeutically modulate gut microbial H_2S
35 production.

36

37 **Introduction**

38
39 Hydrogen sulfide (H_2S) is a consequential molecule produced by the gut microbiota with
40 pleiotropic effects on human physiology. It is one of the three physiological gasotransmitters,
41 along with carbon monoxide and nitric oxide, and is produced endogenously in many tissues
42 including, but not limited to, the brain, heart and liver (1). Endogenous H_2S production occurs
43 via the enzymes cystathionine beta-synthase (CBS), cystathionine gamma-lyase (CSE) and 3-
44 mercaptopyruvate sulfur transferase (MST) (2). CBS, CSE and MST are tightly regulated
45 pyridoxal-5'-phosphate (PLP)-dependent enzymes and produce H_2S primarily from the
46 degradation of cysteine (3) (Figure 1B). H_2S produced by these enzymes plays a litany of
47 physiological roles including: suppression of oxidative stress in the brain, regulation of blood
48 pressure through vasodilation and protection of hepatic stellate cells from cirrhosis in the liver
49 (4). As a result, abnormally low endogenous levels of H_2S are hypothesized to be an underlying
50 cause of peripheral artery disease, and efforts have been made to measure serum levels of H_2S
51 quickly and non-invasively as a proxy for early detection of peripheral artery disease (5).

52
53 Microbes in the gastrointestinal tract also contribute to H_2S production in humans. A majority of
54 the microbially produced H_2S originates in the colon, where estimates of luminal concentrations
55 of hydrogen sulfide range from 0.3 mM to 3.4mM (6). The serum concentration of H_2S in healthy
56 individuals is difficult to measure but is estimated to range from 34.0 to 36.4 μM (7). H_2S readily
57 diffuses across the intestinal epithelium and can enter circulation influencing host physiology
58 (8). Excessive production of H_2S by gut microbes has been linked with decreased mucosal
59 integrity through reduction of mucosal disulfide bonds (9), inhibition of colonocyte butyrate
60 oxidation via cytochrome-c inhibition (10), and genotoxicity (8) (Figure 1C).

61

62 While the mammalian pathways of H₂S production have been well studied, the contribution of
63 gut-microbial H₂S production to circulating H₂S levels and the subsequent systemic effects on
64 human physiology are largely unknown. The first step towards a better understanding of the
65 effects of H₂S on human physiology is to identify which microbial species are responsible for
66 H₂S production. There are two major sources for H₂S production in the human gut microbiota,
67 dissimilatory sulfate reduction and the degradation of the sulfur-containing amino acids cysteine
68 and methionine (11). We must note that sulfate is first reduced to sulfite before H₂S is produced,
69 however, we refer to this process as sulfate reduction for the remainder of this work.

70

71 In the literature, H₂S production is often attributed to the well-characterized dissimilatory sulfate
72 reduction pathway (4). Common representatives of sulfate reducing bacteria (SRB) are found in
73 the class *Delta proteobacteria* with *Desulfovibrio* spp. and *Bilophila wadsworthia* being the most
74 abundant representatives in the human gut (10). Sulfate and sulfite are used by SRB as
75 terminal electron acceptors for anaerobic respiration (12). While SRB are prevalent in human
76 populations, their relative abundances are generally very low and are dependent on ecological
77 interactions with other hydrogenotrophs, such as methanogens and acetogens (10,13,14).

78

79 Unlike the comprehensively-characterized pathways for dissimilatory sulfate reduction, the
80 species of the gut microbiome responsible for H₂S production via degradation of sulfur-
81 containing amino acids have not been comprehensively characterized. Gut microbial
82 involvement in amino acid fermentation has garnered recent attention, as many physiologically
83 relevant downstream metabolites are produced by gut microbial degradation of amino acids (15)
84 (Figure 1A). Depending on dietary intake, a pool of sulfur-containing amino acids is available for
85 fermentation by gut microbiota (16). Recent studies have demonstrated that cysteine
86 supplementation leads to far more H₂S production than inorganic sulfate supplementation

87 underscoring the comparative importance of the cysteine-degradation pathway in total H₂S
88 production (12–14).

89

90 Methane (CH₄) is primarily produced by the methanogen *Methanobrevibacter smithii* (17) and is
91 one of the primary gases present in mammalian flatus. Sulfate-reducing bacteria and
92 methanogens have been historically considered mutually exclusive in microbial communities
93 due to the competition for hydrogen (10). However, experiments carried out on human flatus
94 have shown that both H₂S and CH₄ production occurs simultaneously in some individuals,
95 seemingly contradicting the notion that methanogens and sulfate-reducing bacteria cannot co-
96 exist (6).

97

98 It is important to delineate between H₂S produced via dissimilatory sulfate reduction and H₂S
99 produced via cysteine degradation because different approaches are necessary to modulate
100 these two sources of H₂S production. Because of the poor annotation of the genes which
101 produce H₂S via cysteine degradation across species of the gut microbiome, the relative
102 contributions of H₂S production are unclear. To address this gap, we designed a bioinformatic
103 approach to first identify putative cysteine-degrading bacteria in the human gut microbiome and
104 then compare the relative abundances of putative cysteine-degrading bacteria and sulfate-
105 reducing bacteria across metagenomic data from inflammatory bowel disease, colorectal cancer
106 and healthy cohorts.

107

108 **Results**

109
110 **Cysteine-degrading genes are widely distributed in and expressed by the**
111 **human gut microbiome**

112

113 To identify species capable of H₂S production via cysteine degradation in the human gut
114 microbiome, we curated a list of enzymes experimentally proven to produce H₂S, and searched
115 for these enzymes across 4,644 species in the Unified Human Gastrointestinal Genome
116 (UHGG) collection (18) (Figure 2, Supplementary Table 1). This collection comprises 4,644 non-
117 redundant genome sequences from species representatives generated by clustering 204,938
118 genome sequences from bacteria known to inhabit the human gut.

119
120 Of the representative UHGG species, 18.4% (855/4,644) contain one or more cysteine-
121 degrading genes (Figure 3A) whereas just 0.6% (27/4,644) contain the sulfate-reducing genes
122 *dsrAB*. Aside from known cysteine-degrading bacterial species compiled in the manual curation
123 step, an additional 406 previously cultured species were found to contain one or more cysteine-
124 degrading genes (Figure 2, names in bold). Furthermore, 10.8% (44/406) of these species have
125 evidence of H₂S production, 8.6% (35/406) showed no signs of H₂S production, and 80.8%
126 (328/406) have had no prior test for H₂S production (Supplementary Table 1). Additionally, 550
127 metagenome-assembled genomes (MAGs) were found to contain one or more cysteine-
128 degrading genes. No UHGG genomes contain both *dsrAB* and a cysteine-degrading gene, while
129 many genomes contain multiple cysteine-degrading genes (Figure 3A).

130
131 **Widespread potential for H₂S production via cysteine degradation in the**
132 **human gut microbiome**

133
134 The prevalence and relative abundance of putative cysteine-degrading bacteria and sulfate-
135 reducing bacteria was calculated for 10,700 metagenomic samples from healthy, inflammatory
136 bowel disease, colorectal cancer and adenoma cohorts (19–22). Among the 6,644 healthy
137 subjects, there is a markedly higher ($W = 44,133,484$; $p < 2.2\text{e-}16$, two-sided Wilcoxon rank
138 sum test) relative abundance of putative cysteine-degrading bacteria compared to sulfate-
139 reducing bacteria suggesting that cysteine degradation contributes considerably to H₂S

140 production for the average healthy person (Figure 3B). Cysteine degrading genes are also
141 widespread in healthy populations with 100% of the 6,644 healthy subjects containing at least
142 one putative cysteine-degrading bacteria in their gut microbiome.

143

144 To confirm the transcription of cysteine-degrading genes and sulfate-reducing genes, we
145 analyzed metatranscriptomic data from a dietary intervention study by Lawrence et al. (23). The
146 study relied on stool sample collections before and after plant-based and animal-based diet
147 interventions to evaluate the effects of diet on microbial gene regulation and community
148 composition. Our analysis revealed that 100% of samples from this study (59/59) showed non-
149 trivial expression (RPKM ≥ 1) of one or more cysteine-degrading genes. While 61% of samples
150 (36/59) contained expression of both dissimilatory sulfate reduction and cysteine degradation,
151 36% of samples (21/59) had cysteine degradation genes as the sole source of H₂S production
152 (Figure S4). Methionine gamma-lyase is the most actively transcribed H₂S producing gene, and
153 3-mercaptopropyruvate sulfurtransferase is generally the least transcribed (Figure S2). The primary
154 dissimilatory sulfate reductase genes, *dsrAB*, appear constitutively expressed across all three
155 diets. Comparatively, expression of cysteine-degrading genes appears to be more sporadic
156 across diet conditions and slightly lower in both plant and animal-based diets compared to
157 baseline. These results suggest that for more than one third of individuals cysteine degradation
158 may be the dominant pathway for H₂S production.

159

160 **Core genes from dissimilatory sulfate reduction and methanogenesis are**
161 **co-expressed**

162

163 Though *in vitro* assays have indicated that methanogens and sulfate-reducing bacteria compete
164 for hydrogen and may thus mutually exclude one another (10), core genes involved in
165 dissimilatory sulfate reduction (Figure S2) and methanogenesis (Figure S3) are simultaneously
166 expressed in 8% (5/59) of samples obtained from healthy individuals (Figure S4). Additionally,

167 one or more cysteine-degrading genes were expressed simultaneously with methanogenesis
168 genes in 10% (6/59) of samples.

169

170 **Increased relative abundance of H₂S producing bacteria in the colorectal**
171 **cancer gut microbiome**

172

173 Next, we assessed the relative abundance of putative cysteine-degrading bacteria and sulfate-
174 reducing bacteria in individuals with the two most common clinical manifestations of
175 inflammatory bowel disease (IBD), Crohn's disease and ulcerative colitis, colorectal cancer
176 (CRC) and healthy controls. Putative cysteine-degrading bacteria are significantly more
177 abundant than sulfate-reducing bacteria across IBD and CRC populations from metagenomic
178 samples derived from curatedMetagenomicData (19), the Human Microbiome Project 2 (HMP2)
179 (20), PRISM (21) and Lewis *et al.* (22) (all p < 2.2x10⁻¹⁶) (Figure 4A-D).

180

181 Both putative cysteine-degrading bacteria and sulfate-reducing bacteria are significantly more
182 abundant in CRC than in the respective control groups (Figure 4A). The strength of the
183 association is similar for putative cysteine-degrading bacteria (W = 114,615; p < 3.4x10⁻⁵) and
184 sulfate-reducing bacteria (W = 118,888, p < 1.5x10⁻⁷). In adenomas sulfate-reducing bacteria
185 were found to be moderately differentially abundant (W = 44,330; p = 0.048) while putative
186 cysteine-degrading bacteria were not found to be significantly differentially abundant (W =
187 38,494; p = 0.48).

188

189 The relative abundance of putative cysteine-degrading bacteria is moderately higher than
190 controls for adults with Crohn's disease in the HMP2 cohort, but not in the PRISM cohort.
191 (HMP2: W = 114,116, p = 0.05; PRISM: W = 2,736, p = 0.27) and relatively similar for infants
192 with Crohn's disease (W = 862, p = 0.08). Likewise, putative cysteine-degrading bacteria are

193 seen at similar relative abundance in adults with ulcerative colitis compared to healthy controls
194 (HMP2: W = 70,404, p = 0.11; PRISM: W = 2,407, p = 0.20).

195
196 The relative abundance of sulfate-reducing bacteria tends to be lower in individuals with IBD
197 compared to healthy controls. In adults with Crohn's disease, there is a significantly lower
198 relative abundance of sulfate-reducing bacteria compared to healthy controls (HMP2: W =
199 87,939, p = 1.0x10⁻⁵; PRISM: W = 1,435, p = 2.5x10⁻⁵). Infants with Crohn's disease do not have
200 a significant difference in relative abundance of sulfate-reducing bacteria (W = 964, p = 0.29)
201 compared to healthy controls. Adults with ulcerative colitis have markedly lower relative
202 abundance of sulfate-reducing bacteria compared to healthy controls (HMP2: W = 45,423, p =
203 4.5x10⁻¹³; PRISM: W = 1,382, p = 6.0x10⁻⁴).

204
205 **Discussion**
206
207 Due to its role as a mammalian gasotransmitter, H₂S plays important roles in maintaining
208 physiological homeostasis. However, H₂S may also cause deleterious effects in a concentration-
209 dependent manner. Therefore, it is of great importance to understand the sources of exogenous
210 H₂S production in the gut in order to tease out the links between H₂S and human physiology.
211 The source of gut microbial H₂S production is often attributed to dissimilatory sulfate reduction,
212 with far less attention given to H₂S production via the degradation of the sulfur-containing amino
213 acid cysteine. In fact, there has not been a microbiome-wide annotation of the potential for H₂S
214 production via cysteine degradation. The systematic annotation we performed in this study
215 expands our understanding of which species can produce H₂S in the gut, many of which have
216 not been previously reported to have the capability for H₂S production (Supplementary Table 1).
217 Our analysis of shotgun sequenced metagenomic data from 6,644 metagenomic samples
218 revealed that putative cysteine-degrading bacteria are ubiquitous inhabitants of the human gut

219 microbiome and have significant higher relative abundance than sulfate-reducing bacteria.
220 Furthermore, analysis of metatranscriptomic data demonstrates that cysteine-degrading genes
221 are in fact expressed in the gut. These results suggest that cysteine degradation is likely a
222 major source of microbial H₂S production and may be the sole source of microbially produced
223 H₂S in some individuals. Therefore, cysteine degradation is an important aspect to consider
224 when designing studies to assess the effects of H₂S on human health or modulate gut microbial
225 H₂S production.

226

227 We also explored the relative abundance of putative cysteine-degrading bacteria in IBD and
228 CRC to understand whether these bacteria could contribute to, or promote disease progression.
229 We found that putative cysteine-degrading bacteria are significantly more abundant in CRC
230 samples than in healthy controls. While sulfate-reducing bacteria are also increased in CRC
231 compared to healthy controls, putative cysteine-degrading bacteria are far more abundant. This
232 finding corroborates previous studies linking H₂S and the progression of CRC (24) and
233 highlights the need to identify the dominant source of H₂S in the CRC gut. Putative cysteine-
234 degrading bacteria were not differentially abundant between samples from IBD patients and
235 healthy individuals but are more abundant than sulfate-reducing bacteria. Importantly, it still
236 remains to be elucidated whether or not this difference in relative abundance directly translates
237 to higher production of H₂S via cysteine degradation in comparison with sulfate reduction.

238

239 Prior studies suggested that methanogens and sulfate-reducing bacteria are mutually exclusive,
240 potentially due to their competition for hydrogen. These experiments did not consider cysteine
241 degradation as a potential source of H₂S. However, subsequent studies have reported the
242 presence of both CH₄ and H₂S in the human flatus (6), seemingly contradicting this notion of
243 mutual exclusivity of CH₄ and H₂S producing bacteria. To resolve this discrepancy, we
244 examined the transcriptional co-occurrence of methanogens, sulfate-reducing bacteria, and

245 cysteine-degrading bacteria in the human gut and found the co-occurrence of all three
246 pathways. This discrepancy between *in vitro* experiments and *in vivo* observations could be
247 explained by the complex biogeography of the gut in which methanogens and sulfate-reducing
248 bacteria occupy distinct niches or from H₂S production via cysteine degradation.

249

250 There are many reactions in which H₂S is formed as an intermediate, such as assimilatory
251 sulfate reduction, however, these reactions do not result in significant production of H₂S and are
252 thus not relevant to total H₂S production by the gut microbiome. Therefore, we limited our
253 search for H₂S producing bacteria to pathways in which H₂S was the endpoint, or byproduct,
254 and not just an intermediate of the pathway. Our search identified the genes for dissimilatory
255 sulfate reduction in *Eggerthella* and *Gordinobacter* species. We have included these species as
256 sulfate-reducing bacteria though there is little evidence to suggest that these species are true
257 sulfate reducers (25,26). Further wet-lab validation of these claims is necessary to confirm
258 *Eggerethella* spp. and *Gordinobacter* spp. as non-sulfate-reducing bacteria. We also note that
259 our search for H₂S producing genes included only the 4,644 representative genomes in UHGG.
260 The full UHGG collection contains 204,938 non-redundant genomes with core and accessory
261 gene information that may contain other putative H₂S-producing sub-species that we did not
262 analyze. Another potential shortcoming of this analysis is the overrepresentation of western
263 countries in the data pool used. An expanded set of samples would be required to claim that
264 putative-cysteine degrading bacteria are globally prevalent in the human gut microbiome.
265 Finally, we note that sulfate-reducing bacteria may be mucosally associated and present at low
266 relative abundances which could mean that stool metagenomics may underestimate the true
267 abundance of sulfate-reducing bacteria in the human gut.

268

269 In conclusion, we show that the relative abundance of putative cysteine-degrading bacteria is
270 significantly higher than sulfate-reducing bacteria across healthy individuals as well as

271 individuals with colorectal cancer and inflammatory bowel disease. These results bolster
272 previous studies suggesting the importance of dietary cysteine in gut microbial H₂S production.
273 We also provide a comprehensive overview of putative cysteine-degrading bacteria complete
274 with experimental evidence, or lack thereof, for H₂S production in numerous experimental
275 contexts. The systematic annotation of putative H₂S-producing species performed in this study
276 can serve as a resource for future studies examining the links between H₂S and disease and
277 could help these studies to tease out the concentration-dependent effects of H₂S on human
278 health. Overall, this work informs future approaches to modulate gut microbial H₂S production
279 via dietary interventions and may lead to an improved understanding of the complex interplay
280 between H₂S and human health and disease.

281

282 **Methods**

283

284 **Curation of cysteine-degrading and sulfate-reducing genes**

285

286 A search for potential hydrogen sulfide producing bacteria was conducted using *a priori*
287 knowledge of dissimilatory sulfate reduction and sulfur-containing amino acid degradation by gut
288 microbes. First, amino acid sequences of enzymes responsible for H₂S production (11,27–31)
289 were downloaded using the UniProtKB (32) web browser (www.uniprot.org). Links to UniProt
290 entries used in the search space are listed in Supplementary Table 1 under the “search_space”
291 sheet.

292

293 **Search for putative H₂S producing bacteria in the human gut**

294

295 Note that all BLAST tools used in this work are from the `blast+` command line package,
296 version 2.8.1 (33). A protein BLAST database was created using the protein sequences as input
297 to `makeblastdb` with option `-dbtype prot`. These protein sequences were then searched
298 against 4,644 genome sequences from UHGG (18) using `blastp` (34). Hits were filtered based
299 on two fairly conservative thresholds: E-value < 1x10⁻¹¹⁰ and amino acid identity > 50%. Then,
300 depending on the nature of the sequence matches, the bacterial genomes receiving hits to H₂S
301 producing genes were labeled as either putative cysteine-degrading bacteria or sulfate-reducing
302 bacteria. For instance, any hit to one or more of the cysteine-degrading genes was enough to
303 consider the species to be a putative cysteine-degrading bacterium. However, in order to be
304 considered capable of dissimilatory sulfate reduction, the genome must have received a hit for
305 both *dsrA* and *dsrB* as they are subunits of the final functional protein. Putative cysteine-
306 degrading bacteria across UHGG were then visualized by uploading a taxonomic tree in newick
307 tree format to the iTOL (35) web interface (Figure 2). Pie charts in Figure 2 were generated by

308 parsing the `blastp` output GFF files and uploading to the iTOL web interface. Figure 3A
309 showing the overlap of gene hits to the UHGG collection was generated using UpSetR (36).

310
311 **Calculating relative abundances with Kraken2**
312
313 The raw sequencing reads for the metagenomic samples used in this study were downloaded
314 and extracted with NCBI's SRA toolkit v2.10.9 (37). Quality control and adapter trimming of the
315 fastq sequence files were done with the Trim Galore wrapper v0.6.6 (38). To remove potential
316 human contaminants, quality-trimmed reads were screened against the human genome (hg19)
317 with Bowtie2 v2.4.2 (39). Taxonomy profiling of the metagenomic cleaned reads were generated
318 using Kraken2(2.0.8-beta) (40) to map against the pre-built database of the Unified Human
319 Gastrointestinal Genome (UHGG) catalog (18).

320
321 **Analysis of putative sulfate-reducing bacteria**
322
323 Amino acid sequences encoding dissimilatory sulfate reductase genes *dsrA* and *dsrB* were
324 downloaded from UniProt (accession links in Supplementary Table 1). `blastp` was used to
325 query 4,644 genomes from UHGG for additional species potentially performing dissimilatory
326 sulfite reduction.

327
328 The search returned 27 valid hits ($\geq 50\%$ amino acid identity and E-value $\leq 1e-110$ for both *dsrA*
329 and *dsrB*) to bacteria under the phyla *Proteobacteria*, *Firmicutes*, and *Actinobacteria*
330 (Supplementary Table 1, sheet labeled “sulfate_reduction_hits”). Hits to bacteria within the
331 *Proteobacteria* phylum were expected, as the subphylum *Deltaproteobacteria* contains well-
332 known sulfate-reducing bacteria. Hits to the *Firmicutes* species *Desulfitobacterium hafniense*
333 were also expected since this taxon has been shown to reduce sulfite compounds to sulfide
334 (41,42). Per Muller et al. (25), the presence of *dsrAB* in *Gordonibacter pamelaeae* (phylum
335 *Actionobacteria*) is likely due to a lateral gene transfer event from the genus *Desulfitobacterium*

336 as evidenced by the incongruence between phylogenies built using 16S rRNA gene sequencing
337 and dsrAB gene sequencing. We replicated this phenomenon by constructing a phylogenetic
338 tree from 27 sequences that match dsrAB within UHGG genomes (Figure S2). To construct the
339 tree, a multiple sequence alignment of the 27 sequences using mafft version 7.307 with
340 options --maxiterate 1000 and --localpair was fed to FastTree version 2.1.9 with
341 options -nt. The tree was then visualized using the iTOL web interface.

342

343 Aside from the expected cases, we decided to include hits to *Firmicutes* and *Actinobacteria*
344 species without experimental evidence of H₂S production via dissimilatory sulfate reduction. Our
345 rationale for including these species is to stay consistent with our inclusion of cysteine-
346 degrading bacteria lacking experimental evidence of H₂S production from cysteine degradation.
347 We did not add the “putative” descriptor to this group of bacteria because, unlike the putative
348 cysteine-degrading bacteria we identified, the vast majority of the species that turned up in our
349 results are experimentally validated sulfate-reducing bacteria.

350

351 **Transcriptomic analysis of H₂S producing genes and CH₄ producing genes**

352

353 We sought to confirm the active expression of H₂S producing genes and CH₄ producing genes
354 alongside the existing genomic evidence presented using data from David et al. 2014 (23). In
355 this study, 10 participants had RNA sequencing performed on their stool samples before and
356 after a plant-based and animal-based diet intervention. Subjects waited 6 days before switching
357 to the next diet and getting a baseline reading. Confirming the expression of H₂S producing
358 genes involved the following steps: 1. Metadata for samples was downloaded from the SRA run
359 selector <https://trace.ncbi.nlm.nih.gov>. 2. Raw sequencing data was downloaded using
360 fastq-dump from the SRA toolkit version 2.10.9 (37). 3. Manually curated H₂S producing
361 genes were given as input to diamond makedb (43). 4. Raw RNA-seq data were then aligned

362 against the manually curated protein database using the `diamond blastx` command with
363 options `-b42.0 -c1` for better performance. The raw counts of reads mapped per gene were
364 normalized to RPKM values for downstream analysis. The threshold for considering an H₂S
365 gene “expressed” was RPKM >=1. A sample was said to be “methane producing” if >= 90% of
366 the genes involved in the methanogenesis pathway recruited one or more read mapping. The
367 results were then parsed with a custom shell script and visualized in Figure S3 using the R
368 package `ggplot` (44).

369

Acknowledgements

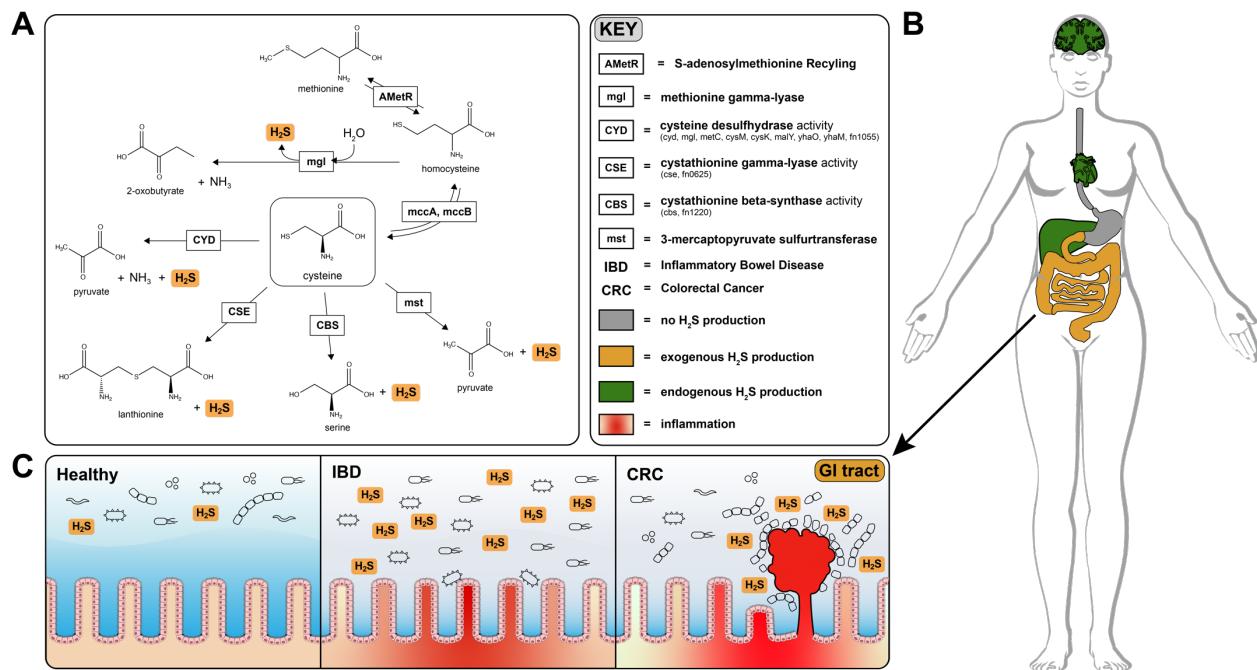
370
371 D.J.B. is supported by National Science Foundation-funded Research Traineeship (COMBINE),
372 X.J. is supported by the Intramural Research Program of the NIH, National Library of Medicine.
373 M.P. is supported by NIH grant R01-AI-100947. B.H. is supported by startup funding from the
374 University of Maryland. The authors declare no conflict of interests.

375
376

Figures

377

378



379

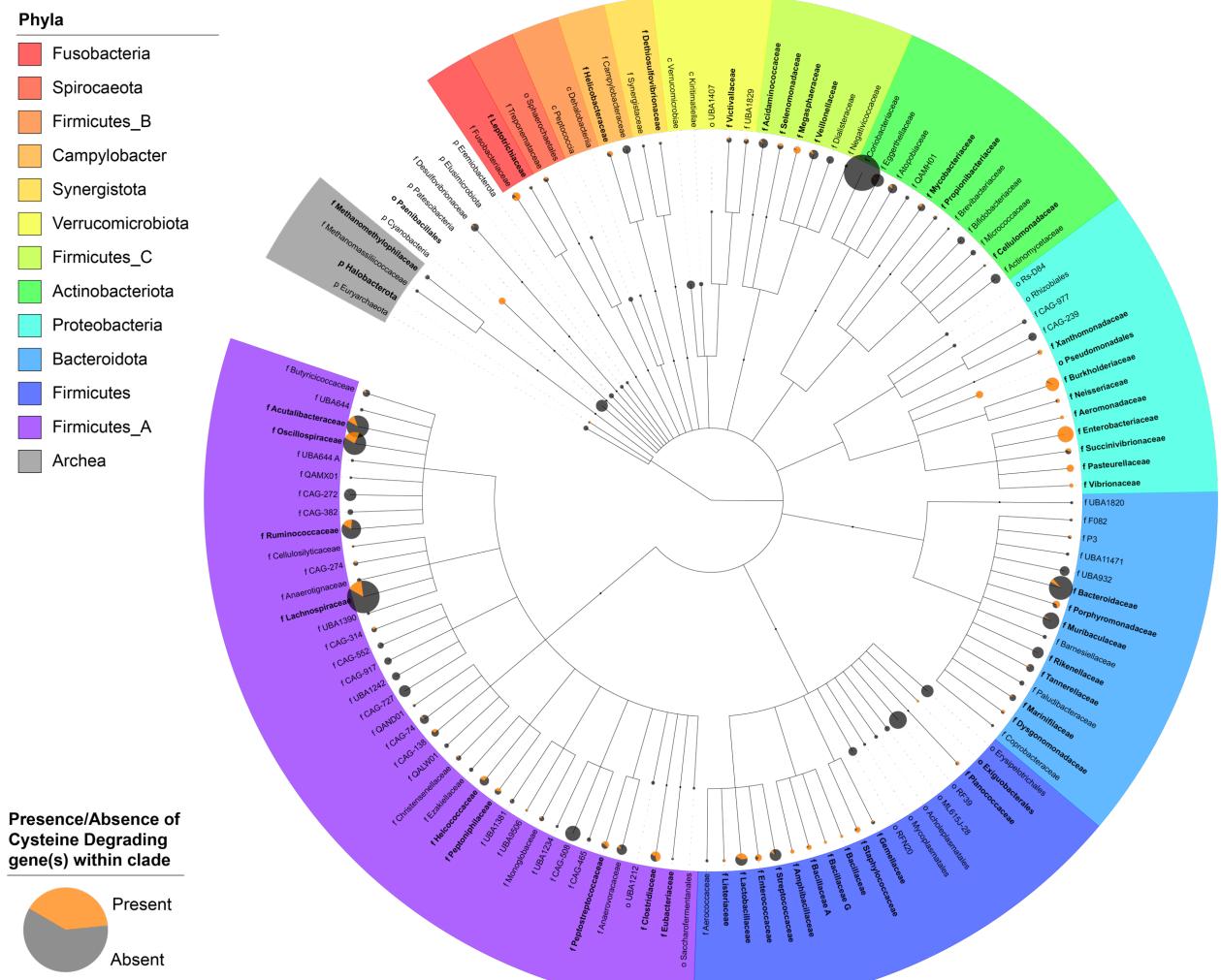
380

381 **Figure 1. H₂S production via cysteine degradation in the human gut microbiome.**

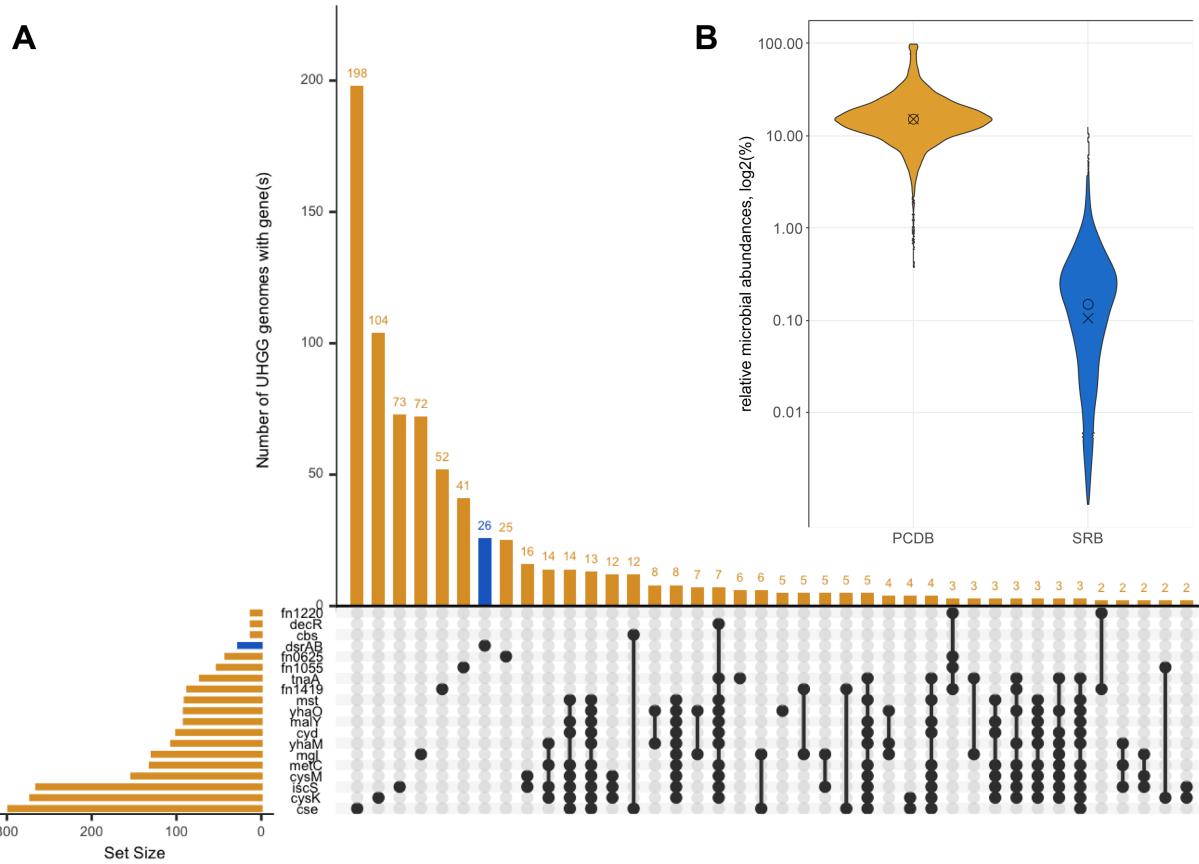
382 (A) Pathways of H₂S production via cysteine degradation in the human gut microbiome.

383 Pathways with labels ending in “activity” refer to a set of genes that convert cysteine to the
384 same products. Genes involved in: CYD = (cyd, mgl, metC, cysM, cysK, malY, yhaO, yhaM,
385 fn1055); CSE = (cse, fn0625); CBS = (CBS, fn1220). AMetR refers to the process of AdoMet
386 recycling present in *Bacillus subtilis* involving genes mtnN, luxS and various methylases (45).

387 (B) Visualization of H₂S production across human tissues. H₂S is produced endogenously in the
388 brain, liver and heart via cysteine degradation and is tightly regulated to avoid toxic effects of
389 H₂S overproduction. Refer to the KEY for descriptions of organ color coding. (C) Physiological
390 effects of H₂S on the gut. Emphasis is placed on the difference between healthy versus IBD and
391 CRC. In the IBD gut, H₂S is thought to contribute to the degradation of the protective mucosal
392 barrier which could cause or exacerbate inflammation. In CRC, various *Fusobacterium* species
393 are closely associated with colonic tumors and are known H₂S producers.

396 **Figure 2. Distribution of putative cysteine-degrading bacteria across the United Human**

397 **Gastrointestinal Genome (UHGG) collection.** A taxonomic tree showing the distribution of
 398 putative cysteine-degrading bacteria across the 4,644 genomes of the representative UHGG
 399 collection. Leaves of the tree are shown to the family level and only genera with ≥ 3
 400 subspecies are included. Phyla are labeled by color and pie charts at the leaf nodes correspond
 401 to presence or absence of cysteine-degrading genes whose expression results in H_2S
 402 production. The relative size of the pie chart represents the number of species in the family
 403 shown.

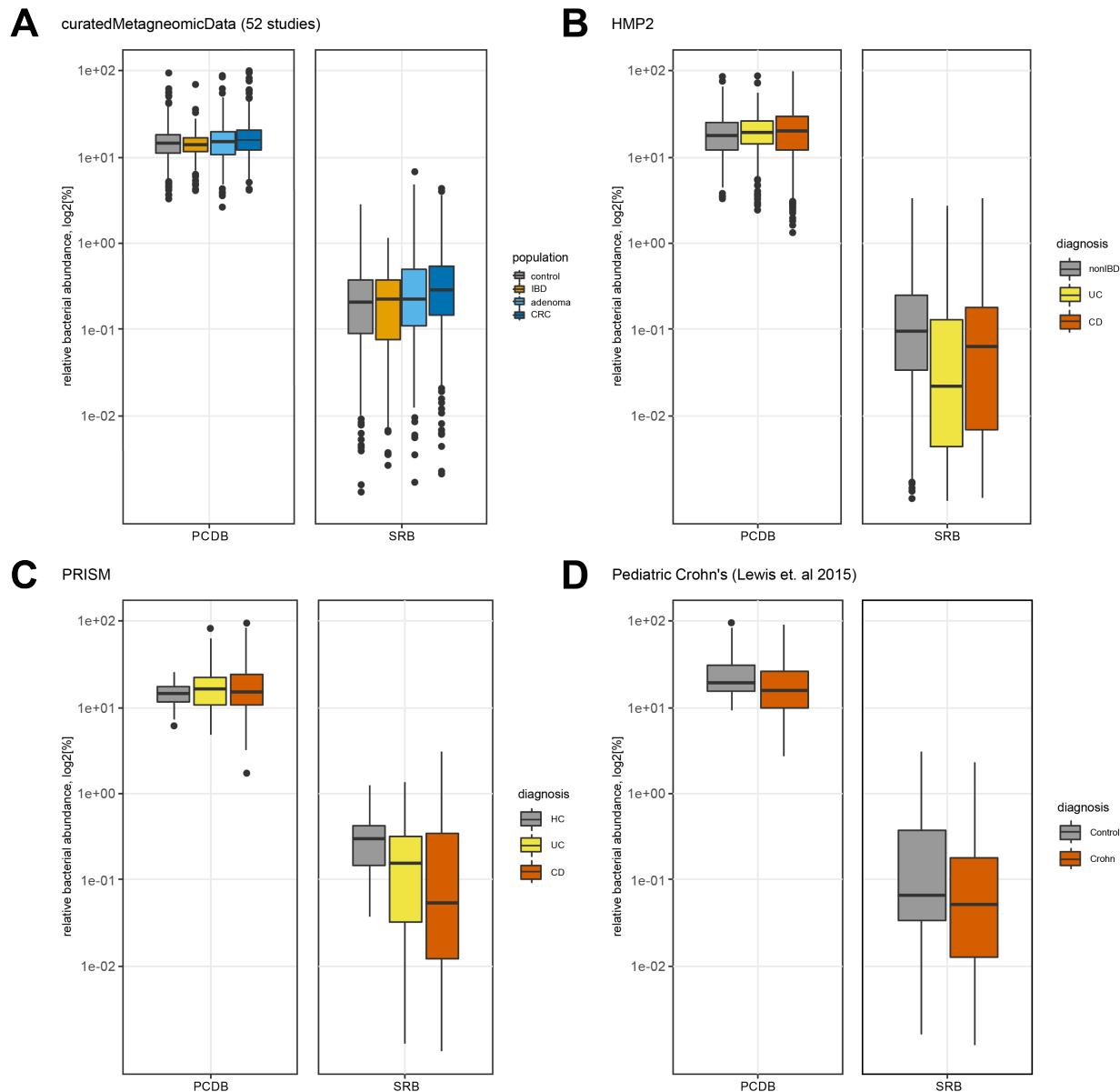


405
406
407 Figure 3. Comparing putative cysteine-degrading bacteria (PCDB) to sulfate reducing
408 **bacteria (SRB).** Color key: orange = cysteine-degrading genes and bacteria containing such
409 genes; blue = sulfate reducing genes and bacteria containing such genes. (A) Overlap of gene
410 hits in the UHGG collection. The y-axis shows the number of genomes receiving one or more
411 hits to H₂S producing genes CYD activity genes = (cyd, metC, cysM, cysK, malY, yhaO, yhaM,
412 fn1055); CSE activity genes = (cse, fn0625); CBS activity genes = (cbs, fn1220); sulfate
413 reducing genes = (dsrAB). The x-axis shows genes which co-occur in UHGG genomes. For
414 example, the gene cse appeared in 198 genomes individually and appeared alongside the gene
415 cbs in 12 genomes. The dissimilatory sulfate reducing gene dsrAB occurred in only 27
416 genomes, and did not co-occur with any cysteine-degrading genes searched. All species
417 receiving at least one hit to a cysteine-degrading gene are considered putative cysteine-
418 degrading bacteria. (B) Relative abundance of putative cysteine-degrading bacteria and sulfate

419 reducing bacteria among 6,644 healthy controls provided by curatedMetagenomicData (19)

420 ($p < 2.2 \times 10^{-16}$, two-sided Wilcoxon rank sum test).

421

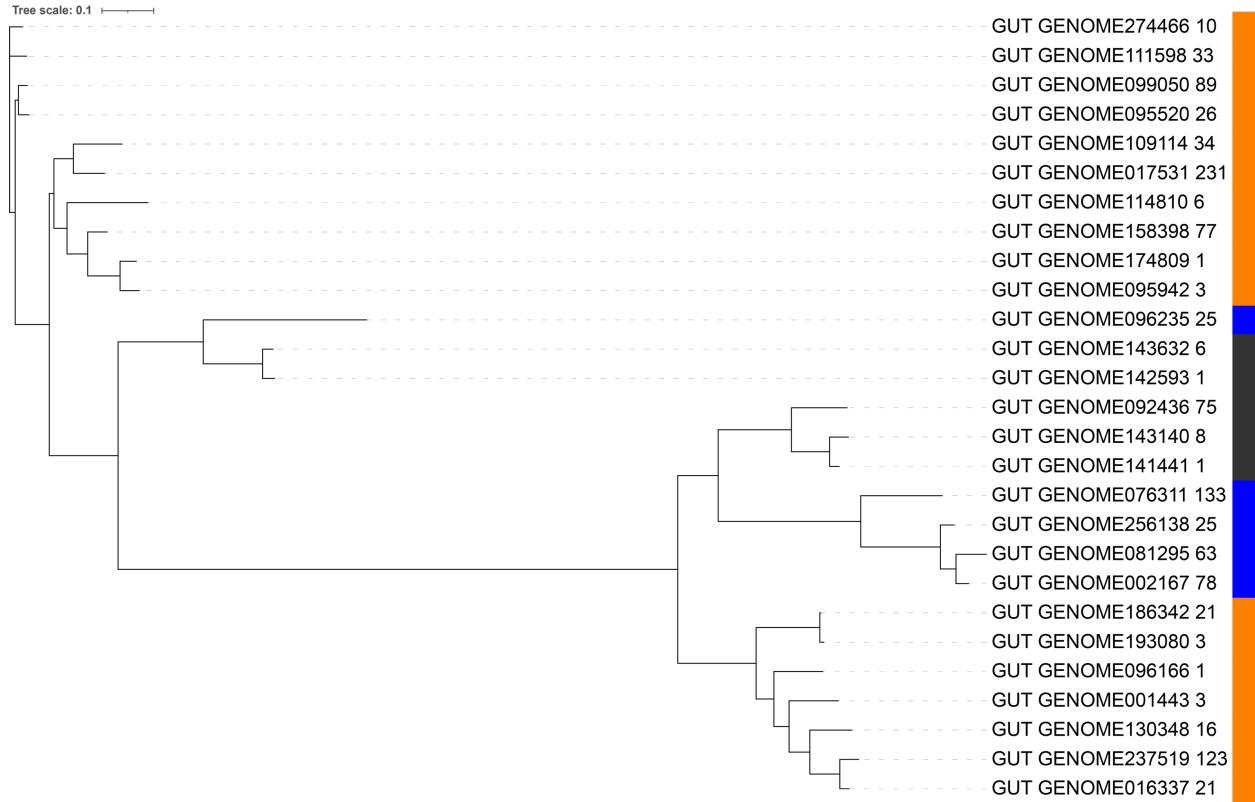


422
423
424 **Figure 4. Putative cysteine-degrading bacteria (PCDB) are more prevalent than sulfate-**
425 **reducing bacteria (SRB) among individuals with IBD, CRC, adenoma and healthy**
426 **controls.** Log2-transformed relative abundances of putative cysteine-degrading bacteria and
427 sulfate-reducing bacteria across healthy and diseased populations. Relative abundances were
428 calculated using Kraken2 (40) (see methods section). (A) Data obtained from
429 curatedMetagenomicData (19). Number of samples per disease category: control = 560, CRC =
430 352, adenoma = 143, IBD = 148. (B) Data obtained from HMP2 (20). Number of samples per

431 disease category: nonIBD = 359, ulcerative colitis (UC) = 367, Crohn's disease (CD) = 591. (C)
432 Data obtained from PRISM (21). Number of samples per disease category: control = 56, UC =
433 76, CD = 88. (D) Data obtained from Lewis et al. 2015 (22). Number of samples per disease
434 category: control = 26, CD = 86.
435

436 **Supplementary Material**

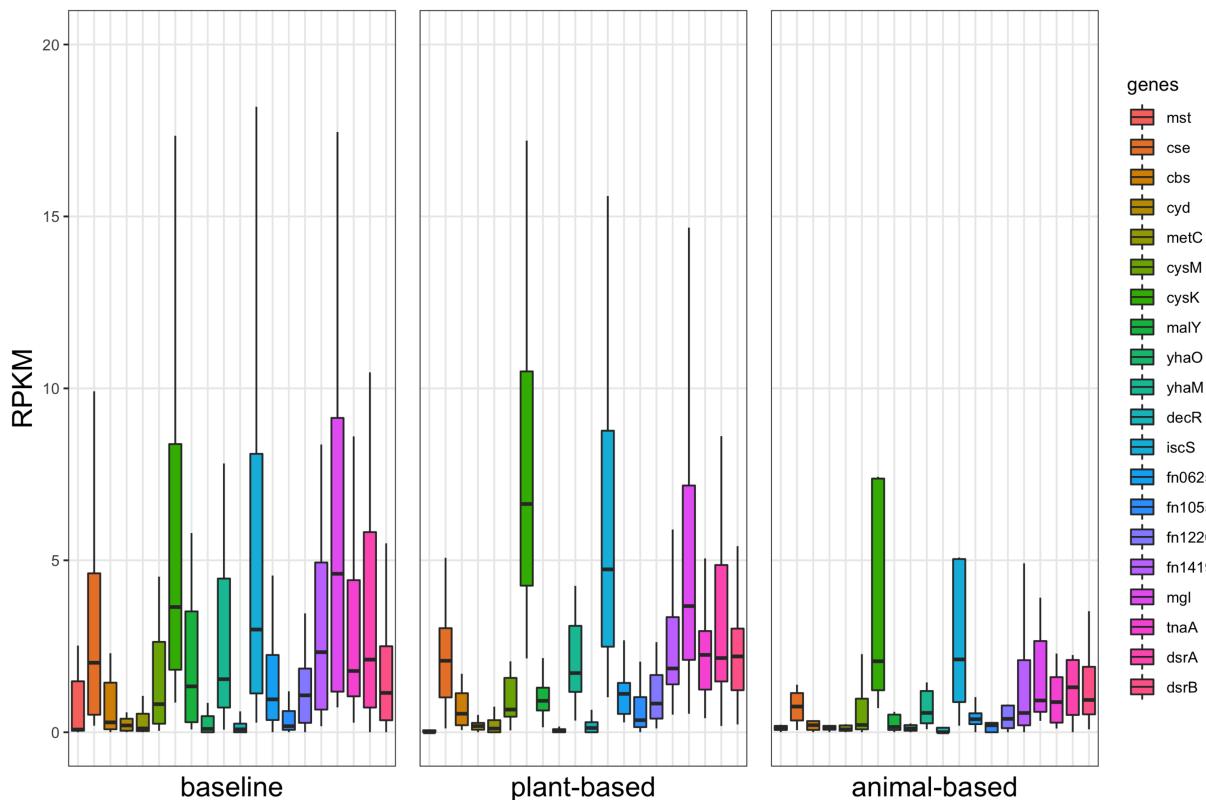
437
438



439
440
441 **Supplementary Figure 1. Phylogenetic analysis of dsrA hits from UHGG.** Colors on the
442 right signal the phylum that each copy of the gene originated from: orange =
443 *Desulfobacterota_A*, blue = *Firmicutes_B*, black = *Actinobacteriota*. Note: a phylogenetic tree
444 was generated for *dsrB* as well, and it returned an identical tree, therefore, we only included the
445 tree generated from analysis of *dsrA*.

446

Expression of H₂S producing genes



447

448

449 **Supplementary Figure 2. Transcriptomic confirmation for cysteine-degrading genes.** This

450 analysis confirms that the H₂S producing genes considered in this work are actively expressed

451 in healthy humans under a variety of dietary regimens. The y-axis displays log2 transformed

452 RPKM values for each gene (see methods for details on alignment and normalization

453 procedure). The x-axis separates reads counts by both gene and diet intervention. Baseline

454 represents samples taken before either plant- or animal-based diet. Some samples contained

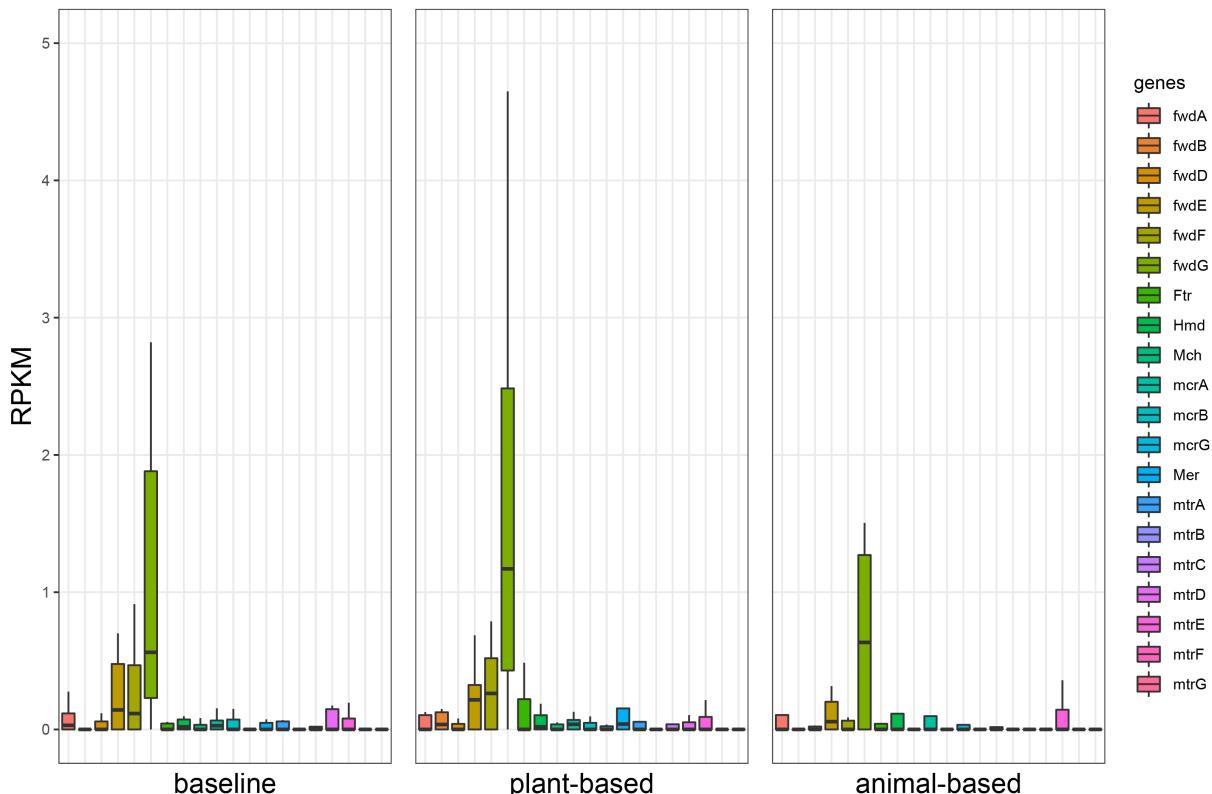
455 zero hits to one or more protein and can be seen along the bottom of the plots. Note that the

456 same individuals were fed the plant and animal based diets with a 6 day waiting period in

457 between the end of one diet and the taking of baseline samples for the next diet.

458

Expression of CH₄ producing genes



459

460

461 **Supplementary Figure 3. Gene expression of methane producing genes in the human**

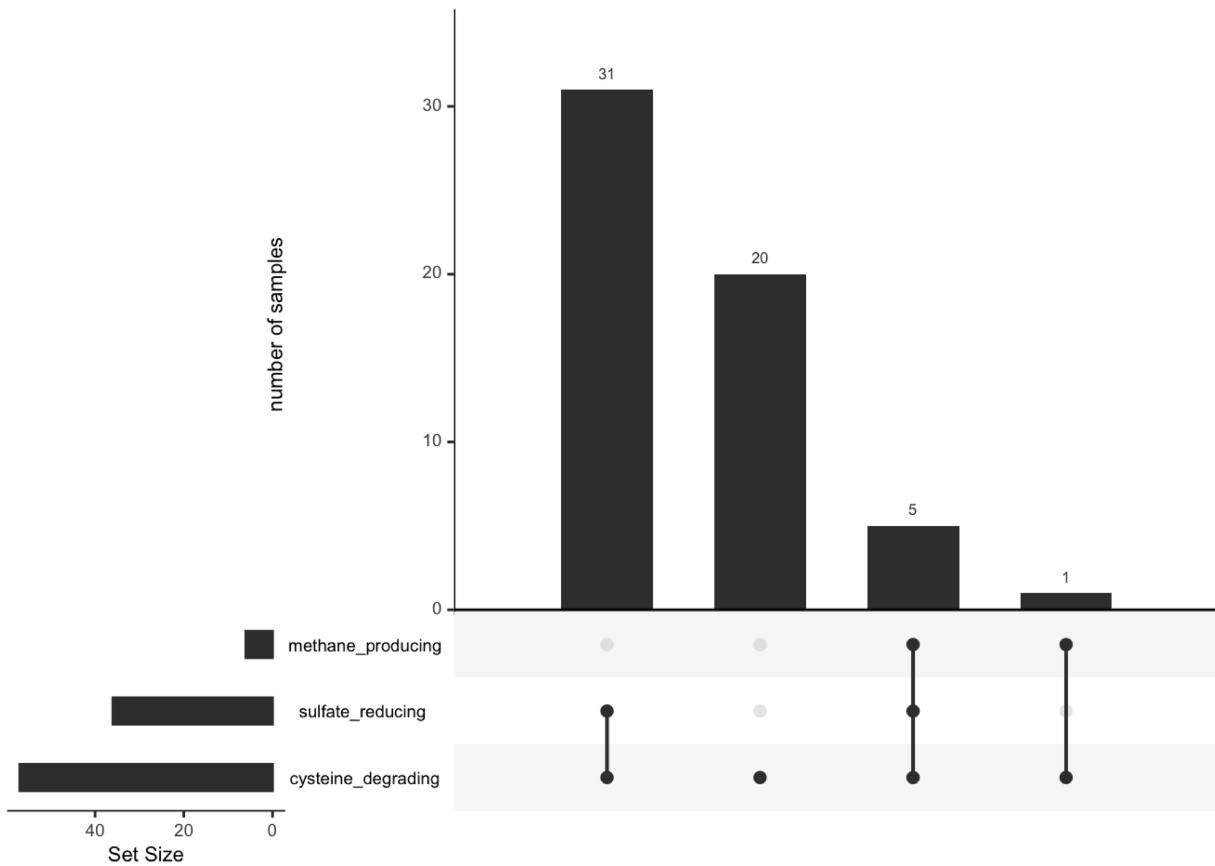
462 **gut.** Genes involved in the production of CH₄ by *Methanobrevibacter smithii* expressed in

463 healthy individuals. The three panels are from 10 participants over three legs of a diet

464 intervention study where metatranscriptomic reads were collected at baseline and diet

465 intervened time points. The y-axis shows RPKM adjusted read counts mapped to each gene

466 involved in the production of CH₄.



467
468
469 **Supplementary Figure 4. Presence of H₂S production and CH₄ production in the healthy**
470 **human gut microbiome.** For some individuals, there is simultaneous production of H₂S and
471 CH₄ in the gut microbiome. Out of 59 metagenomic stool samples from David et al. 2014, 57
472 expressed one or more cysteine-degrading genes, 36 expressed *dsrAB* and 6 expressed at least
473 90% of the genes required for methane production (see methods). The vertical bars represent
474 the number of samples containing one or more functions described. For example, 31 samples
475 showed expression of sulfate-reducing and cysteine-degrading genes and 5 samples contained
476 the same two functions plus methane production.
477

478 **References**

- 479
- 480 1. Wang R. Hydrogen Sulfide: The Third Gasotransmitter in Biology and Medicine. *Antioxid Redox Signal.* 2009 Oct 21;12(9):1061–4.
- 481
- 482 2. Kimura H. Hydrogen Sulfide: From Brain to Gut. *Antioxid Redox Signal.* 2009 Oct 5;12(9):1111–23.
- 483
- 484 3. Percudani R, Peracchi A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep.* 2003 Sep 1;4(9):850–4.
- 485
- 486 4. Wang R. Physiological Implications of Hydrogen Sulfide: A Whiff Exploration That Blossomed. *Physiol Rev.* 2012 Apr;92(2):791–896.
- 487
- 488 5. Shekarriz R, Friedrichsen DM, Brooks B, Silaski G, Rios L, Wiest E, et al. Sensor of transdermal biomarkers for blood perfusion monitoring. *Sens Bio-Sens Res.* 2020 Jun 1;28:100328.
- 489
- 490
- 491 6. Suarez F, Furne J, Springfield J, Levitt M. Insights into human colonic physiology obtained from the study of flatus composition. *Am J Physiol-Gastrointest Liver Physiol.* 1997 May 1;272(5):G1028–33.
- 492
- 493
- 494 7. Furne J, Saeed A, Levitt MD. Whole tissue hydrogen sulfide concentrations are orders of magnitude lower than presently accepted values. *Am J Physiol-Regul Integr Comp Physiol.* 2008 Nov;295(5):R1479–85.
- 495
- 496
- 497 8. Furne J, Springfield J, Koenig T, DeMaster E, Levitt MD. Oxidation of hydrogen sulfide and methanethiol to thiosulfate by rat tissues: a specialized function of the colonic mucosa. *Biochem Pharmacol.* 2001 Jul;62(2):255–9.
- 498
- 499
- 500 9. Blachier F, Andriamihaja M, Larraufie P, Ahn E, Lan A, Kim E. Production of hydrogen sulfide by the intestinal microbiota and epithelial cells and consequences for the colonic and rectal mucosa. *Am J Physiol-Gastrointest Liver Physiol.* 2021 Feb 1;320(2):G125–35.
- 501
- 502
- 503 10. Gibson G r., Macfarlane G t., Cummings J h. Occurrence of sulphate-reducing bacteria in

- 504 human faeces and the relationship of dissimilatory sulphate reduction to methanogenesis in
505 the large gut. *J Appl Bacteriol.* 1988 Aug 1;65(2):103–11.
- 506 11. Carbonero F, Benefiel AC, Alizadeh-Ghamsari AH, Gaskins HR. Microbial pathways in
507 colonic sulfur metabolism and links with health and disease. *Front Physiol* [Internet]. 2012
508 Nov 28 [cited 2020 Aug 13];3. Available from:
509 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3508456/>
- 510 12. Fecal hydrogen sulfide production in ulcerative colitis. *Am J Gastroenterol.* 1998 Jan
511 1;93(1):83–7.
- 512 13. Deng YF, Liao XD, Wang Y, Liang JB, Tufarelli V. Prebiotics Mitigate In Vitro Sulfur-
513 Containing Odour Generation in Caecal Content of Pigs. *Ital J Anim Sci.* 2015 Jan
514 1;14(1):3762.
- 515 14. Yao CK, Rotbart A, Ou JZ, Kalantar-Zadeh K, Muir JG, Gibson PR. Modulation of colonic
516 hydrogen sulfide production by diet and mesalazine utilizing a novel gas-profiling
517 technology. *Gut Microbes.* 2018 May 9;9(6):510–22.
- 518 15. Lin R, Liu W, Piao M, Zhu H. A review of the relationship between the gut microbiota and
519 amino acid metabolism. *Amino Acids.* 2017 Dec 1;49(12):2083–90.
- 520 16. Silvester KR, Cummings JH. Does digestibility of meat protein help explain large bowel
521 cancer risk? *Nutr Cancer.* 1995 Jan 1;24(3):279–88.
- 522 17. Miller TL, Wolin MJ, Macario EC de, Macario AJ. Isolation of *Methanobrevibacter smithii*
523 from human feces. *Appl Environ Microbiol.* 1982 Jan 1;43(1):227–32.
- 524 18. Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, et al. A unified catalog
525 of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol.* 2020 Jul
526 20;1–10.
- 527 19. Pasolli E, Schiffer L, Manghi P, Renson A, Obenchain V, Truong DT, et al. Accessible,
528 curated metagenomic data through ExperimentHub. *Nat Methods.* 2017 Nov;14(11):1023–
529 4.

- 530 20. Proctor L, Huttenhower C. The Integrative Human Microbiome Project. *Nature*. 2019
531 May;569(7758):641–8.
- 532 21. Franzosa EA, Sirota-Madi A, Avila-Pacheco J, Fornelos N, Haiser HJ, Reinker S, et al. Gut
533 microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol*.
534 2019 Feb;4(2):293–305.
- 535 22. Lewis JD, Chen EZ, Baldassano RN, Otley AR, Griffiths AM, Lee D, et al. Inflammation,
536 Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's
537 Disease. *Cell Host Microbe*. 2015 Oct 14;18(4):489–500.
- 538 23. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet
539 rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014 Jan
540 23;505(7484):559–63.
- 541 24. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al.
542 *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome*
543 *Res*. 2012 Feb 1;22(2):299–306.
- 544 25. Müller AL, Kjeldsen KU, Rattei T, Pester M, Loy A. Phylogenetic and environmental diversity
545 of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J*. 2015 May;9(5):1152–65.
- 546 26. Nguyen LH, Abu-Ali G, Mehta RS, Lloyd-Price J, Song M, Yan Y, et al. Su1940 - Dietary
547 Patterns, Sulfur Intake, and the Abundance of Sulfate-Reducing Bacteria. *Gastroenterology*.
548 2018 May;154(6):S-640.
- 549 27. Awano N, Wada M, Mori H, Nakamori S, Takagi H. Identification and Functional Analysis of
550 *Escherichia coli* Cysteine Desulphydrases. *Appl Environ Microbiol*. 2005 Jul 1;71(7):4149–
551 52.
- 552 28. Suwabe K, Yoshida Y, Nagano K, Yoshimura F. Identification of an L-methionine γ-lyase
553 involved in the production of hydrogen sulfide from L-cysteine in *Fusobacterium nucleatum*
554 subsp. *nucleatum* ATCC 25586. *Microbiology*,. 2011;157(10):2992–3000.
- 555 29. Shatalin K, Shatalina E, Mironov A, Nudler E. H2S: A Universal Defense Against Antibiotics

- 556 in Bacteria. *Science*. 2011 Nov 18;334(6058):986–90.
- 557 30. Nava GM, Carbonero F, Croix JA, Greenberg E, Gaskins HR. Abundance and diversity of
558 mucosa-associated hydrogenotrophic microbes in the healthy human colon. *ISME J*. 2012
559 Jan;6(1):57–70.
- 560 31. Shimada T, Tanaka K, Ishihama A. Transcription factor DecR (YbaO) controls detoxification
561 of L-cysteine in *Escherichia coli*. *Microbiology*. 2016 Sep 1;162(9):1698–707.
- 562 32. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the
563 Universal Protein knowledgebase. *Nucleic Acids Res*. 2004 Jan 1;32(suppl_1):D115–9.
- 564 33. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+:
565 architecture and applications. *BMC Bioinformatics*. 2009;10(1):421.
- 566 34. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J*
567 *Mol Biol*. 1990 Oct 5;215(3):403–10.
- 568 35. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.
569 *Nucleic Acids Res*. 2019 Jul 2;47(W1):W256–9.
- 570 36. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of
571 intersecting sets and their properties. Hancock J, editor. *Bioinformatics*. 2017 Sep
572 15;33(18):2938–40.
- 573 37. SRA-Tools - NCBI [Internet]. [cited 2021 Feb 10]. Available from: <http://ncbi.github.io/sra-tools/>
- 575 38. Babraham Bioinformatics - Trim Galore! [Internet]. [cited 2021 Feb 10]. Available from:
576 http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
- 577 39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012
578 Mar 4;9(4):357–9.
- 579 40. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome*
580 *Biol*. 2019 Dec;20(1):257.
- 581 41. CHRISTIANSEN N, AHRING BK. *Desulfitobacterium hafniense* sp. nov., an Anaerobic,

- 582 Reductively Dechlorinating Bacterium. Int J Syst Evol Microbiol. 1996;46(2):442–8.
- 583 42. Lie TJ, Godchaux W, Leadbetter ER. Sulfonates as Terminal Electron Acceptors for Growth
- 584 of Sulfite-Reducing Bacteria (*Desulfitobacterium* spp.) and Sulfate-Reducing Bacteria:
- 585 Effects of Inhibitors of Sulfidogenesis. Appl Environ Microbiol. 1999 Oct 1;65(10):4611–7.
- 586 43. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat
- 587 Methods. 2015 Jan;12(1):59–60.
- 588 44. Wickham H. ggplot2. WIREs Comput Stat. 2011;3(2):180–5.
- 589 45. Hullo M-F, Auger S, Soutourina O, Barzu O, Yvon M, Danchin A, et al. Conversion of
- 590 methionine to cysteine in *Bacillus subtilis* and its regulation. J Bacteriol. 2007
- 591 Jan;189(1):187–97.