

1 **Demonstrating the Beneficial Effect of Low Protein Diet in**  
2 **Primary Sclerosing Cholangitis through a Randomized**  
3 **Clinical Trial and Multi-omics Data Analysis**

4 Xiaole Yin<sup>1,\*</sup>, Gila Sasson<sup>2,\*</sup>, Zheng Sun<sup>1</sup>, Shanlin Ke<sup>1</sup>, Demsina Babazadeh<sup>3</sup>, Shaikh Danish  
5 Mahmood<sup>2</sup>, Macie Andrews<sup>2</sup>, Shelley Hurwitz<sup>2</sup>, Tinashe Chikowore<sup>1</sup>, Maia Paul<sup>2</sup>, Nadine Javier<sup>2</sup>,  
6 Malav Dave<sup>2</sup>, Alexandra Austin<sup>2</sup>, Linda Gray<sup>4</sup>, Francene Steinberg<sup>5</sup>, Elaine Souza<sup>4</sup>, Christopher  
7 Bowlus<sup>4</sup>, Yang-Yu Liu<sup>1,6,#</sup>, Joshua Korzenik<sup>2,#</sup>

8 *<sup>1</sup>Channing Division of Network Medicine, Department of Medicine, Brigham and Women's*  
9 *Hospital, Harvard Medical School, Boston, MA 02115, USA*

10 *<sup>2</sup>Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital,*  
11 *Harvard Medical School, Boston, MA 02115, USA*

12 *<sup>3</sup>Center for Clinical Investigation, Brigham and Women's Hospital, Harvard Medical School,*  
13 *Boston, MA 02115, USA*

14 *<sup>4</sup>Division of Gastroenterology and Hepatology, University of California Davis School of Medicine,*  
15 *Sacramento, CA 95817, USA*

16 *<sup>5</sup>Department of Nutrition, University of California Davis, College of Agricultural and*  
17 *Environmental Sciences, Davis CA 95616, USA*

18 *<sup>6</sup>Center for Artificial Intelligence and Modeling, The Carl R. Woese Institute for Genomic Biology,*  
19 *University of Illinois at Urbana-Champaign, Champaign, IL 61801, USA*

20 *\* These authors contributed equally to this work.*

21 *# Correspondence should be addressed to J.K. ([JKORZENIK@bwh.harvard.edu](mailto:JKORZENIK@bwh.harvard.edu)) and Y.-Y.L.*  
22 *([yyl@channing.harvard.edu](mailto:yyl@channing.harvard.edu))*

23

24 **ABSTRACT**

25 Primary sclerosing cholangitis (PSC), a progressive cholestatic hepatobiliary disease characterized  
26 by inflammation and fibrosis of the bile ducts, has a pathophysiology that is not understood. No  
27 effective therapies exist. The only treatment option for PSC is liver transplant. We undertook a  
28 pilot randomized trial of diet to investigate the pathophysiology of the disease, the role of diet and  
29 to advance potential therapy. We enrolled 20 patients with PSC and randomly assigned them to a  
30 Low Protein/low sulfur Diet (LPD, n=10) or the Specific Carbohydrate Diet (SCD, n=10) for 8  
31 weeks. Results showed that low protein intake benefits PSC patients, whereas higher protein levels  
32 exacerbate the condition. We further identified gut bacterial markers useful for distinguishing LPD  
33 responders (mostly PSC with concomitant ulcerative colitis) from non-responders. Additionally,  
34 by integrating multi-omics data, we propose that this diet modifies the intestinal sulfur cycle  
35 reducing hydrogen sulfide (H<sub>2</sub>S) production. Our findings provide an understanding of the  
36 beneficial effect of LPD as well as insights into a possible key driver of inflammation in PSC.  
37

38 **KEYWORDS**

39 Primary Sclerosing Cholangitis; Inflammatory Bowel Disease; Metabolomics; Low protein diet;  
40 Specific carbohydrate diet; Sulfur metabolic pathway

41

## 42 INTRODUCTION

43 Primary sclerosing cholangitis (PSC), a chronic cholestatic liver disease characterized by  
44 inflammation and fibrosis of the biliary tree, can lead to debilitating consequences including end-  
45 stage liver disease and malignancies. Currently, no therapy has been found to be effective for PSC,  
46 other than liver transplant, which remains the mainstay of therapy. However, the disease can recur  
47 after transplant. PSC is intimately associated with inflammatory bowel disease (IBD), particularly  
48 ulcerative colitis (UC). Approximately 70-80% of patients with PSC also have underlying IBD [1,  
49 2], but only around 3- 5% of UC patients will develop PSC [3]. Thus, the combination of the two  
50 creates a distinct disease phenotype, which is a prototypical disease of the gut-liver axis [1, 4].

51  
52 The pathophysiology of PSC remains poorly understood. The paradigm of IBD is relied on as a  
53 model to understand PSC, presuming an interaction between the environment, the microbiome and  
54 immune system with a genetic predisposition [5, 6]. Other hypotheses suggest a toxic bile acid or  
55 a cross reacting T cell response between the colonocytes and cholangiocytes [5, 7]. The immune  
56 modifying approaches that can be effective in IBD do not benefit PSC. A distinct dysbiotic fecal  
57 microbiome has been detailed in PSC, which overlaps with that of IBD [1, 4, 8]. While numerous  
58 studies have investigated changes in the gut microbiome in the context of IBD [9-12], fewer have  
59 focused on PSC [13-18].

60  
61 Our approach to PSC in this study is based on the hypothesis that sulfur-containing amino acids  
62 such as methionine may be substrates for sulfate-reducing bacteria to generate hydrogen sulfide  
63 (H<sub>2</sub>S). H<sub>2</sub>S production is likely coupled with increased bacterial degradation of mucin, which is  
64 rich in cysteine and methionine [19, 20]. Further, H<sub>2</sub>S can also impair colonocyte uptake and beta-  
65 oxidation of butyrate [21]. At higher levels, H<sub>2</sub>S can be toxic to mitochondria, generating radical  
66 oxygen species and activating the NLRP10 inflammasome, generating a broader inflammatory  
67 reaction [22]. In this way, reducing H<sub>2</sub>S by manipulating dietary sulfur content may limit cellular  
68 impairment and inflammation. Practically speaking, this would involve the reduction of sulfur-  
69 containing amino acids, which are primarily derived from animal and plant-based proteins.

70  
71 In contrast, the comparator diet examined in this study, the Specific Carbohydrate Diet (SCD), is  
72 a popular diet among IBD patients that is grain-free and low in sugar and lactose, but does not  
73 limit protein content. This diet has been studied to a limited degree in IBD with variable  
74 effectiveness [23, 24]. The SCD excludes the intake of complex carbohydrates, disaccharide-  
75 containing foods, grain, corn, potato, and dairy while permitting carbohydrates obtained from  
76 fruits, honey, some vegetables, and fermented yogurt. The SCD is based on the principle that  
77 specifically selected carbohydrates in SCD require minimal digestion, so they are subsequently  
78 well absorbed, avoiding further gut microbial metabolism.

79 Diet composition may influence gut microbial activity in PSC, both through supporting or  
80 enhancing the growth of particular bacterial communities and modifying metabolic profiles. These  
81 adaptations may consequently influence disease progression through the gut-liver axis [25-29]. Yet,  
82 the interplay between these parameters is complex and poorly understood [4] especially with  
83 regard to its effect on disease course [30, 31]. Despite the strong patient interest, clinical trials  
84 targeting dietary interventions specifically for PSC are notably lacking. To date, only one study  
85 has examined a gluten-free diet in individuals with PSC and associated colitis, and it failed to  
86 demonstrate clinical or biomedical improvement [32].

87  
88 In this study, we conducted a pilot randomized controlled trial to assess the clinical efficacy of a  
89 low-protein diet (LPD) compared to the SCD in PSC (**Fig. 1a & 1b**), and leveraged high-  
90 throughput technologies (shotgun metagenomic sequencing and global untargeted metabolomic  
91 profiling) and multi-omics data analyses to gain mechanistic insights into response to dietary  
92 intervention (**Fig. 1c**).

## 93 94 **RESULTS**

### 95 **Patient characteristics**

96 We screened ~100 individuals with PSC, and 25 met eligibility criteria and 5 eligible individuals  
97 chose not to participate prior to randomization. Ten participants were randomly assigned to the  
98 LPD and ten to the SCD (**Fig. 1a, Fig. 1b & Table 1**). Participant characteristics were balanced  
99 between the groups. Two participants withdrew due to cholangitis, and their data until withdrawal  
100 were incorporated into the analysis.

101 Participants in the LPD group demonstrated lower protein consumption in 3-day food recordings  
102 at all study time points during dietary intervention (week 2 to week 10) compared to SCD group  
103 participants (Wilcoxon rank-sum test,  $4.57e-05 < p\text{-value} < 0.010$ ), supporting the compliance with  
104 the dietary interventions (**Fig. 2a**).

### 105 106 **LPD benefits PSC patients, but SCD does not**

107 Comparing the ALP levels at W4 and W8 with those at W0, we found that LPD led to improvement  
108 (i.e., decrease in ALP) in eight of the ten individuals (80%), with six of them (75%) returning to  
109 their baseline ALP level when dietary interventions were removed at W10 (**Fig. 2b**). In contrast,  
110 only two of the ten participants (20%) in the SCD group responded to SCD (i.e., with lower ALP  
111 levels than their baseline levels). Moreover, for most of the participants in the SCD group, their  
112 ALP values did not appear to be affected after discontinuing the dietary intervention, suggesting  
113 that SCD may not be a potentially effective treatment for PSC. For those eight participants who

114 responded to LPD, hereafter referred to as LPD-responders, their ALP values were reduced by  
115  $18.5\% \pm 10.4\%$  at W4 (i.e., after two weeks of treatment), and by  $22.6\% \pm 15.6\%$  at W8 (i.e., after  
116 another four weeks of treatment).

117 When accounting for an ALP decrease of at least 10%, LPD showed efficacy in 60% of patients,  
118 compared to a 20% efficacy rate in SCD. With a reduction of at least 20%, LPD demonstrated  
119 efficacy in 30% of patients, while SCD showed a 10% rate. Considering a decrease of at least 30%,  
120 LPD maintained its efficacy in 30% of patients, with no cases of efficacy in SCD. Finally, with an  
121 ALP decrease of at least 50%, LPD is effective in 10% of patients, with SCD also showing no  
122 cases of effectiveness. Changes of other parameters, including Aspartate Aminotransferase (AST)  
123 and Alanine Aminotransferase (ALT) values, during the intervention were shown in **Fig. S1**.

124 Direct evidence obtained from Lasso Regression analysis revealed a significant positive  
125 correlation between the dietary methionine levels and the ALP levels (**Fig. 2c**). This finding  
126 indicates that lower methionine intake is associated with lower ALP levels compared to baseline  
127 values. However, it is crucial to note that this relationship is correlative rather than causative.

128 In addition to methionine, levels of magnesium, total sugar, iron, total folate, other amino acids,  
129 vitamin K and vitamin D were also positively associated with the ALP levels (**Fig. 2c**). By contrast,  
130 the levels of histidine, calcium, tryptophan, soluble dietary fiber, beta-carotene, starch, and vitamin  
131 B-12 were negatively associated with the ALP levels. At the same time, we observed that  
132 carbohydrate consumptions (nutrient density) exhibited significant differences between LPD and  
133 SCD, with LPD being higher than SCD at all time points during dietary intervention (week 2 to  
134 week 10) compared to baseline (Wilcoxon rank-sum test,  $0.0014 < p\text{-value} < 2.17e\text{-}05$ ) (**Fig. S2**).

135

### 136 **Three disease categories respond differently to dietary intervention**

137 We performed a stratified analysis to determine if comorbid IBD influences ALP response to  
138 dietary interventions. Though the numbers of individuals in each subgroup were small, marked  
139 differences were evident for certain disease categories. For instance, for PSC-UC patients (n= 10  
140 in total), LPD (n=6) was significantly more effective in lowering ALP than SCD (n=4), (Wilcoxon  
141 rank-sum test, p-value = 0.038). Moreover, all the PSC-UC patients in the LPD arm responded to  
142 LPD assessed by decreasing ALP levels, see **Fig. 2d**. For PSC-CD patients (n= 3 in total), there  
143 was only one patient in the LPD arm (n=1), and this patient responded to LPD. Interestingly, for  
144 PSC patients without IBD (PSC-alone) (n=7 in total), one patient in the LPD arm (n=3) responded,  
145 while no response in SCD (n=4). Moreover, the two interventions didn't demonstrate any  
146 significant difference in reducing ALP levels (**Fig. 2f**).

147

148

149

## 150 Gut microbial features distinguish LPD responders from non-responders

151 Given the efficacy of LPD, we aimed to determine distinguishing microbial biomarkers between  
152 LPD responders and non-responders using shotgun metagenomic sequencing data. We found that  
153 the alpha diversity (measured by the Shannon index) of stool samples in the PSC-UC group was  
154 significantly lower than that in the PSC-CD and PSC-alone groups (**Fig. S3a & Fig. S3b**,  
155 Wilcoxon rank-sum test, p-value = 0.016 and p-value = 0.00062, respectively). Samples in each  
156 category were clustered together and distinct from the other two groups (PERMANOVA p-  
157 value=0.001, F=3.53) (**Fig. S2c**). When individuals were simply classified into responders  
158 (individuals showing reduced ALP) and non-responders (individuals showing reduced ALP),  
159 variations in alpha diversities and differences in beta diversity can be found in **Fig 3a, Fig. S4 &**  
160 **Fig. S5**. The clusters of LPD were separated from LPD non-responders in the PCoA plot based on  
161 the Bray-Curtis dissimilarity, suggesting that gut microbiota is key in differentiating response  
162 outcomes (**Fig. 3a**). Similar results were revealed by UMAP clustering (**Fig. S4b**).

163 To account for three disease categories as covariates and to eliminate random effects of inter-  
164 individual variation due to temporal changes, we employed MaAsLin2 to perform a multivariate  
165 linear regression analysis on the taxonomic profiles of the stool samples (see **Methods, Formula**  
166 **1**). We aimed to find microbial features that significantly differentiate LPD responders from LPD  
167 non-responders (with cutoff p-value < 0.05 and q-value < 0.2). We identified 21 (or 15) microbial  
168 species as significantly positively (or negatively) associated with responders, respectively (**Fig.**  
169 **3b**). Among these, *Ruminococcus gnavus*, *Blautia caecimuris*, *Blautia producta*, *Enterocloster*  
170 *bolteae*, and *Flavonifractor plautii*, were more abundant in responders across samples of all three  
171 disease categories. Conversely, *Clostridium bacteria*, *Roseburia bacteria*, *Senegalimassilia*  
172 *anaerobia*, and others were found in much lower abundance in responders.

173 In the heatmap of microbial profiles, samples are clustered based on the similarity of their bacterial  
174 structures (**Fig. 3b**). Interestingly, we found that the selected bacterial markers not only  
175 differentiate responders from non-responders but also distinguish between disease categories. For  
176 instance, PSC-UC samples are clustered together, as are PSC-CD samples, while non-responders,  
177 who were all PSC-alone, also form a distinct group. Samples from the same individuals are  
178 clustered together on both the heatmap (**Fig. 3b**) and the PCoA plot (**Fig. 3a**), indicating a  
179 personalized gut microbiome, i.e., intra-individual differences are smaller than inter-individual  
180 ones. Therefore, disease categories can also be considered as a general predictor of LPD response.

181

## 182 Gut metabolome is affected by multiple variables

183 To better understand metabolomic parameters of LPD response, we conducted the global  
184 untargeted metabolomic analysis of all stool samples using the ultra-performance liquid  
185 chromatography-tandem mass spectrometry (UPLC-MS) (see **Methods**). A total of 1,428  
186 metabolites were quantified from at least one sample in this study, and a subset of 1,057



187 metabolites with known chemical names and corresponding identities in at least one of the  
188 following three databases: the Human Metabolome Database (HMDB) [33], the Kyoto  
189 Encyclopedia of Genes and Genomes (KEGG) database [34], and the PubChem database [35].  
190 Metabolite compositions exhibited greater temporal variations within individuals than microbial  
191 profiles, as indicated by ordination analysis (**Fig. 3a** versus **Fig. 3c**, **Fig. S4a** versus **Fig. S4c**),  
192 revealing that after dietary interventions, samples from the same individuals at different time points  
193 were interspersed with samples from others, leading to a mixed clustering pattern. Using the  
194 previously described multivariate linear regression model, we identified 29 metabolomic markers  
195 associated with clinical outcomes ( $p$ -value  $< 0.05$  and  $q$ -value  $< 0.2$ ) (see **Methods, Formula 1**).  
196 Of these, 27 metabolites were positively associated with LPD responders, while 2 metabolites were  
197 negatively associated (**Fig. 3d**). The intensity of metabolite markers in LPD responders versus  
198 LPD non-responders was less distinctive than that observed with microbial markers (**Fig. 3d**).  
199 Additionally, in contrast to microbial markers, clustering based on metabolite markers composition  
200 did not show a clear correlation with disease category (**Fig. S6**). Moreover, the observable intra-  
201 individual variation in these markers after dietary intervention indicates baseline metabolic profile  
202 cannot serve as a reliable predictor of response outcome to diet intervention, unlike microbial  
203 markers.

204

### 205 **Sulfur-related metabolites predict the LPD-induced changes**

206 Diet-induced temporal changes in metabolite profiles were further analyzed to determine the  
207 primary mechanism underlying the effect of LPD. We employed multivariate linear regression  
208 models, which are different from those described earlier, to distinguish metabolite markers under  
209 LPD intervention (W4 and W8) from baseline (W0), incorporating the individual as a random  
210 effect and disease category, age, gender, and BMI as fixed effects (**see Methods, Formula 2**). We  
211 identified 39 metabolite markers that significantly changed throughout LPD compared to baseline  
212 ( $p$ -value  $< 0.05$  and  $q$ -value  $< 0.25$ ), 21 of which have known chemical annotations. Of these, 6  
213 (or 15) metabolites' abundances increased (or decreased) with LPD. Notably, the majority of  
214 depleted metabolites (8/15) were either amino acids or fatty acids. Among the 6 enriched  
215 metabolites, four were classified as carbohydrates ( $n=1$ ), medium-chain hydroxy acids ( $n=1$ ),  
216 tricarboxylic acids ( $n=1$ ), or quinoline carboxylic acids ( $n=1$ ) (**Fig. 4a**).

217 To identify metabolic pathways that are significantly altered, we performed metabolite set  
218 enrichment analysis (MSEA) [36, 37] using all the 21 metabolite features that significantly  
219 changed throughout treatment and with known chemical names. We found significant changes in  
220 purine metabolism (MetaboAnalyst v5.0,  $p$ -value = 0.0083), sulfur metabolism ( $p$ -value = 0.0083),  
221 and biotin metabolism ( $p$ -value = 0.021) with LPD. Notably, sulfur metabolism had the highest  
222 enrichment ratio (enrich = 6.26), suggesting it may be a primary mechanism behind the effect of  
223 LPD (**Fig. 4b**).

224 To extract metabolites potentially converted from or to sulfur-containing chemicals, out of the  
225 1,043 detected metabolites, we identified 40 sulfur-related metabolites referenced in the SMILES  
226 database [38]. Four key metabolites were found to have significantly changed (**Fig. 4c-f & Fig.**  
227 **4k**). Methionine (Wilcoxon signed-rank test, p-value=0.039 at W4) and its oxide methionine  
228 sulfone (Wilcoxon signed-rank test, p-value=0.015 at W8) were significantly depleted compared  
229 to baseline. This suggests that reduced protein intake led to a continual decrease in methionine and  
230 methionine sulfone in the gut. Taurine significantly enriched by W8 (Wilcoxon signed-rank test,  
231 p-value=0.047) compared to baseline. Cysteine level varied over time, increasing at W4 but  
232 decreasing at W8. Notably, cysteine intensity changes showed a significant negative association  
233 with ALP changes, indicating that cysteine accumulation correlates with lower ALP values  
234 (Spearman correlation:  $R = -0.42$ , p-value= 0.0172) (**Fig. 4j**). Furthermore, in association with  
235 taurine accumulation, significant changes in luminal bile acids were observed. Notably, we found  
236 the levels of cholic acid was significantly elevated with LPD, while taurocholic acid were  
237 significantly decreased with LPD (**Fig. 4k & Fig. S7**, pathways referred to KEGG database [45])  
238 A complete list of significantly changed sulfur-containing metabolites (n=45) and bile acids (n=19)  
239 is in **Table S1**.

240

#### 241 **Changes in functional metagenomic features associated with LPD intervention**

242 To explain the microbial contributions to the metabolite variations and to understand the temporal  
243 changes of functional consequences of the microbial community leading to response, we profiled  
244 gene families in all stool samples. Based on the shotgun metagenomic sequencing data, we  
245 analyzed the functional gene profiles using HUMAnN3.6 and identified 2,427 enzymes (KOs)  
246 from different bacterial species. Although the functional pathways were relatively stable across  
247 individuals and across times (**Fig. S8**), due to microbial functional redundancy, we then selected  
248 genes that significantly changed throughout dietary intervention in LPD responders by fitting a  
249 multivariate linear model to each enzyme, adjusting for the random effect of each individual and  
250 other covariates (**see Methods, Formula 2**). In total, 99 (or 11) genes significantly increased (or  
251 decreased) at W4 and W8 compared to baseline (W0) (target p-value < 0.05).

252 We focused on genes involved with sulfur metabolism to decipher the relationship between sulfur-  
253 related metabolite changes and microbial functional changes. We referred to a comprehensive list  
254 of 74 genes involved in microbial sulfur metabolism in humans [39, 40]. Of these, 27 were detected  
255 in at least one sample in this study, with 6 (or 5) of these genes significantly decreasing (or  
256 increasing) during intervention in LPD responders (Wilcoxon rank-sum test, p-value < 0.05), and  
257 4 genes increasing with marginal significance ( $0.05 < \text{p-value} < 0.06$ ) (**Fig. 4k & Table S2**).  
258 Interestingly, LPD led to a reduced amount of *cysJ*, a critical enzyme for producing hydrogen  
259 sulfide ( $\text{H}_2\text{S}$ ) from sulfite, and a reduced amount of *dycD*, key for  $\text{H}_2\text{S}$  production from cysteine  
260 (**Fig. 4k**) [39, 40]. Conversely, the *cysK* gene, involved in metabolizing  $\text{H}_2\text{S}$  in the biosynthesis of  
261 cysteine, increased throughout LPD (**Fig. 4k**) [39, 40]. Similarly, there was an increase in the *bsh*



262 gene that metabolizes taurine and aids in the deconjugation of tauro-conjugated bile acids excreted  
263 to the gut from the biliary tree (**Fig. 4k**) [41]. These findings highlight the relationship between  
264 microbial functional profile with metabolite changes, supporting the potential mechanistic role of  
265 altered sulfur metabolism in attenuating luminal inflammation and enhancing ALP reduction in  
266 PSC.

267

## 268 **Bacterial species associations with functional genes**

269 Given the significant changes in microbial functional profiles, we explored the bacterial species  
270 connected to these enzymes. We constructed a large-scale association network among functional  
271 genes, metabolites, and bacterial species. To identify covariations strictly linked to diet change, we  
272 first residualized each feature in either measurement type using the same multivariate model  
273 employed to determine differential changes in responders during LPD intervention (see **Methods**,  
274 **Formula 2**). This residualization process, using longitudinal measurements, minimizes inter-  
275 individual variation and highlights within-person associations over time, considering diagnostic  
276 categories as covariates. The resulting network contained 74 edges linking genes and metabolites,  
277 906 edges between genes and species, and 57 edges linking gene pairs, where at least one  
278 connected gene was sulfur-related and the paired correlation was significant (Spearman correlation  
279  $FDR < 0.05$ ) (**Fig. 5a & Table S3**). The network encompassed 331 nodes spanning features from  
280 three measurement types. *tauD*, a gene converting taurine to sulfite (further convertible to  $H_2S$ )  
281 (**Fig. 4k**), was central in the network with the most connections ( $n = 174$ ) to different bacterial  
282 species. However, for most of these species (90.2%), the abundance change of a single species was  
283 not significant (multivariate model  $p$ -value  $< 0.05$ ). *Eggerthella lenta*, significantly positively  
284 associated with *tauD* (Spearman correlation  $FDR < 0.05$ ) (**Table S3**), showed a marked decrease  
285 in abundance during LPD intervention (Wilcoxon rank-sum test  $p$ -value =  $8.7e-3$ ; multivariate  
286 model:  $p$ -value = 0.016) (**Fig. 5d, Fig. S9 & Table S4**). This species is critical in diet-induced  
287 changes, because it is also linked to *dmsC* (K00185), involved in the conversion from methionine  
288 to sulfite (**Fig. 4k**), estimating via the taxonomic stratified functional profile by HUMAnNv3.6,  
289 and carries 10 other genes associated with sulfur metabolism, including *iscS*, and *ahcY*, which are  
290 genes significantly reduced with LPD (Wilcoxon rank-sum test,  $p$ -value  $< 0.05$ ) as well (**Table**  
291 **S2**). Genes *cysJ* and *dcyD* carried the second and fourth most connections, contributed by 92 and  
292 74 different species, respectively (**Fig. 5a & Table S3**). These two genes were significantly  
293 interrelated (Spearman correlation coefficient = 0.95,  $R = 5.06e-11$ ) and associated with other  
294 reduced genes including *tauD*, *dmsC*, *aspC*, and *iscC*. However, these genes were contributed by  
295 more than a single bacterium because no significant changes were found in any single species in  
296 *cysJ* or *dcyD* based on stratified abundance. *bsh*, a gene that metabolizes taurine and aids in the  
297 deconjugation of tauro-conjugated bile acids, did not exhibit positive correlations in any species  
298 in the complex association network. However, the taxonomic stratified analysis showed a  
299 significant increase in the abundance of *bsh* in *Roseburia intestinalis* in LPD responders (Wilcoxon  
300 rank-sum test,  $p$ -value =  $1.7e-3$ ), a decrease in the SCD group and no change in LPD non-

301 responders (**Fig. 5b & Table S5**). Notably, *Roseburia intestinalis* is a key bacterium in sulfur  
302 metabolism, containing 10 relevant genes that significantly changed, including *cysK*, *metK*, *dcm*,  
303 *metB*, *metY*, *luxS*, and others (**Fig. 5c & Table S5**). Not only gene abundances, the bacterial  
304 abundance of *Roseburia intestinalis* showed a marked increase in LPD responders (Wilcoxon rank-  
305 sum test: p-value = 3.8e-3; multivariate model: p-value = 4.4e-4) (**Fig. 5d, Fig. S8 & Table S4**).

306

## 307 DISCUSSION

308 The relationship between diet-microbiome-host interactions and their influence on the course of  
309 PSC is complex and poorly understood. To our knowledge, this is the first randomized controlled  
310 trial comparing the effects of two diets to assess a biochemical improvement in PSC through their  
311 modulatory effects on the gut microbiome. We observed that protein restriction was associated  
312 with improvement in ALP, while a diet limited in carbohydrates but not protein yielded no such  
313 benefit.

314 Moreover, the integration of multi-omics measurements, including state-of-the-art whole  
315 metagenomic sequencing and untargeted metabolite measurements, enabled us to reveal  
316 significant diet-induced changes in sulfur metabolism, purine metabolism, and biotin metabolism.  
317 Previous studies have suggested that H<sub>2</sub>S can be generated from multiple sulfur-containing amino  
318 acids, or from sulfate-reducing bacteria [19, 42]. In turn, H<sub>2</sub>S overproduction inhibits the  
319 cytochrome c oxidase activity in the human mitochondrial respiratory chain, thus disrupting the  
320 mitochondrial energy metabolism and promoting mucosal inflammation [43]. Although a  
321 potentially detrimental role of H<sub>2</sub>S has been linked to intestinal dysbiosis, much less is known  
322 about its relationship to liver disease progression. Our study provides evidence of significant diet-  
323 induced changes in sulfur-containing metabolites and related enzymes. Notably, decreased protein  
324 intake led to reduced levels of methionine and its oxide, associated with lower microbial enzymatic  
325 activity, particularly in enzymes involved in H<sub>2</sub>S production. We also observed a decrease in H<sub>2</sub>S-  
326 producing bacteria, including *Eggerthella lenta*, significantly associated with enzymes like *tauD*,  
327 *dmsC*, *iscS*, and *ahcY*. Thus, our data confirm, from various perspectives, that a low-protein diet  
328 regulates H<sub>2</sub>S production by correcting the imbalance in sulfur metabolism.

329 Alterations in sulfur metabolism as the primary mechanism for the beneficial effects of low protein  
330 are also reflected in the accumulation of taurine and cysteine. This accumulation results from  
331 increased activity of the *bsh* gene, associated with taurine accumulation, and the *malY* and *metB*  
332 genes, linked to cysteine production, when the conversion of taurine and cysteine to H<sub>2</sub>S was  
333 significantly reduced, as discussed previously. An unhealthy gut microbiome that lacks the  
334 required deconjugation of host primary bile acids into secondary bile acids, impairs liver and gut  
335 function, in addition to overall health [44, 45]. Importantly, our data showed that taurine  
336 accumulation and cholate increase correspond with the increase of deconjugation genes, such as  
337 *bsh*, in response to taurocholate—a taurine-conjugated bile acid, indicating amelioration of the  
338 imbalanced gut microbiome in PSC patients with restricted protein intake. Intriguingly, the

339 bacteria found to increase after LPD, primarily carriers of *bsh*, *cysK*, *maly*, and *metB* genes, include  
340 *Roseburia intestinalis*. The increase in *Roseburia intestinalis*, known for its beneficial role in  
341 producing beneficial short-chain fatty acids which possess anti-inflammatory properties [46, 47],  
342 suggests a rebalancing of the bacterial structure in PSC patients, where deconjugation is otherwise  
343 lacking compared with healthy individuals. Therefore, our data suggest that LPD might enhance  
344 the gut's deconjugation of primary bile acids, offering potential protective mechanisms against  
345 inflammation and helping to restore a healthier microbial balance/composition.

346 While the role of H<sub>2</sub>S has not been previously investigated in PSC, more has been done to  
347 investigate its role in UC with a small but largely supportive literature [48, 49]. An increase in  
348 sulfate-reducing bacteria has been associated with UC and flares in particular [50, 51]. With regard  
349 to diet, flares have been found to be associated with a higher protein diet [52]. Higher protein diets  
350 have been demonstrated to increase H<sub>2</sub>S production by intestinal bacteria increasing sulfate-  
351 reducing bacteria [53]. Several interventional studies of a low sulfur diet have shown to be of  
352 benefit in UC though more detailed investigations into the metagenomic and metabolomic  
353 elements and pathologic pathways have not been part of those studies [54]. Still, the intimate  
354 connection between UC and PSC would suggest these studies might be relevant and H<sub>2</sub>S may offer  
355 a potential physiologic link between these diseases.

356 In addition to our findings regarding sulfur metabolism, we were able to identify microbial  
357 signatures that can be used to predict response to LPD. Besides, we observed that the majority of  
358 responders had concomitant UC, and that the predictive microbial signatures were enriched in all  
359 baseline samples of PSC-UC participants. Collectively, these findings serve to enhance patient  
360 selection for this intervention and warrant further exploration in larger cohorts.

361 Although this study provides significant insights into better understanding PSC pathophysiology,  
362 there are several limitations, including a small cohort size and lack of a traditional control arm.  
363 Future inclusion of healthy controls or IBD-only controls could offer a more comprehensive  
364 understanding of mechanisms involved in disease progression. Additionally, while the 8-week-  
365 long intervention and 4-week self-guided follow-up provided interesting findings, exploring the  
366 longer-term effects and potential changes after prolonged intervention remains a key area for future  
367 research.

368 Overall, the well-documented dietary adherence and records of different subgroups of individuals  
369 with PSC, in combination with high-throughput technologies and the multi-omics approach,  
370 enabled us to determine the clinical, microbial and metabolic signatures associated with  
371 responsiveness to dietary interventions. This study enhanced our understanding of PSC  
372 pathogenesis and paved the way for large-scale dietary intervention studies in PSC.

373

374

375

## 376 **METHODS**

377 **Study setting and design.** This was a decentralized randomized controlled trial across the United  
378 States. Participant enrollment occurred between August 28, 2020 and June 9, 2021. Recruitment  
379 occurred in-person and by video remotely, and participants were able to conduct the study remotely.  
380 Recruitment was facilitated via physician referral, patient advocacy organizations (PSC Partners  
381 Seeking a Cure and Consortium for Autoimmune Liver Disease), and self-referral.

382 Eligible participants were randomized in 1:1 fashion to the LPD and SCD at a screening visit (week  
383 0, W0, referred to as “baseline”) during which medical records were reviewed, eligibility was  
384 assessed and bloodwork, stool, baseline Food Frequency Questionnaire (FFQ) and Patient-  
385 reported Outcome Measures (PROMs) were collected. Thereafter, participants were given two  
386 weeks for education on appropriate guidelines for the assigned dietary intervention as well as  
387 procurement of food. The dietary intervention began at week 2 (W2) and lasted for 8 consecutive  
388 weeks until week 10 (W10) under dietitian supervision, with an option to continue self-directed  
389 for an additional 4 weeks (W14). There was a total of 7 video visits with a research dietitian.  
390 Additional dietary counselling was also offered between screening and baseline visits to provide  
391 further instruction and support. Three-day food diaries, bloodwork, stool samples and PROMs  
392 were collected at 7 points throughout the study (**Fig. 1a**).

393 **Eligibility criteria.** Adults between 18 and 70 years of age with large-duct PSC diagnosed by  
394 typical cholangiogram findings with no evidence of a secondary cause of sclerosing cholangitis  
395 and with serum ALP >1.5 times the upper limit of normal (ULN) were potentially eligible for  
396 enrollment. Participants with concomitant UC or CD were eligible if the Simple Clinical Colitis  
397 Activity Index or Harvey-Bradshaw Index, respectively, were <5. Additional inclusion criteria  
398 included platelet count  $\geq 150,000/\text{mm}^3$ , serum albumin  $\geq 3.3$  g/dL, serum creatinine < ULN, and  
399 stable dose or no use of ursodeoxycholic acid (UDCA) for at least 3 months prior to enrollment  
400 for those taking and not taking UDCA respectively. Proficiency in English and ability to complete  
401 PROMs independently was also a requirement. Exclusion criteria included pregnancy or lactation,  
402 ALT above 10 times the ULN, total bilirubin at least twice the ULN, INR >1.2, decompensated  
403 cirrhosis, small duct PSC, other etiologies of liver disease, positive AMA, history of liver  
404 transplantation, history of hepatocellular carcinoma or cholangiocarcinoma, ascending cholangitis  
405 within 90 days of enrollment, antibiotic use within 6 weeks prior to enrollment or planned during  
406 the study period, current vegetarian or adherence to the SCD, nut allergy given that nut flour is a  
407 dietary staple of many SCD recipes and nut allergy could compromise diet adherence, celiac  
408 disease, history of malignancy within 5 years with the exception of adequately treated cervical  
409 carcinoma in situ and basal or squamous cell carcinoma, inability to complete a dietary log, or  
410 concurrent participation in another therapeutic clinical trial. Medical records of all individuals  
411 were reviewed at the screening visit to determine eligibility.

412  
413 **Dietary interventions.** The LPD was developed according to 2015-2020 USDA Dietary  
414 Guidelines of a vegan diet which state the diet should be rich in grains, legumes, nuts and other

415 plant-based proteins, with a focus on increased consumption of fruits, vegetables and healthy fats  
416 [55]. In addition to these guidelines, the diet restricted high-sulfate items, defined as  
417 containing >100mg sulfur per 100g of food item. A typical vegan diet contains approximately  
418 2.3g/day of sulfur-containing amino acids, falling within the estimated range of appropriate sulfur  
419 content of 2.1-3.0 g/day [56]. The SCD followed the dietary guidelines detailed in the book  
420 *Breaking the Vicious Cycle* by Elaine Gottschall [57]. The detailed LPD and SCD dietary guidance  
421 is provided in **Table S6**. Participants were provided with education materials, recipes and a food  
422 procurement stipend for their assigned diet at the screening visit. Total daily energy intake was  
423 calculated by the Mifflin St. Jeor equation using an activity factor of 1.3 reflecting light activity  
424 and exercise level.

425 **Assessment of diet composition and adherence.** Habitual diet was recorded using a validated  
426 FFQ at the screening visit (week 0) that queried dietary habits over the preceding year. During the  
427 intervention phase, 3-day food diaries were recorded at five time points (weeks 2, 4, 6, 8, 10) (**Fig.**  
428 **1a**). One of these days included a weekend. Participants recorded diet in real-time via  
429 photodocumentation of each meal, snack and beverage using a smartphone application that was  
430 customized for this study. The application transmitted time-stamped photographs taken during the  
431 recording period, along with a reference-sized study ruler to assess relative sizes of food items.  
432 Caloric intake and micronutrient composition of each meal was calculated and analyzed using  
433 Food Processor, a program with the ability to quantify protein intake, as well as cysteine,  
434 methionine and taurine content as a proxy for sulfur.

435 In order to optimize compliance, the research team reviewed dietary records in real-time and  
436 discussed barriers to compliance with participants when appropriate. Dietary analysis was also  
437 performed at the conclusion of each diet recording period to further assess compliance. Feedback  
438 was provided to participants via the application if there were fewer than 2 entries per day or if a  
439 24-hour period elapsed without any recording. The study coordinator followed up by email and/or  
440 telephone if no confirmation was received from the participant within 24 hours of the  
441 communication.

442 **Outcome measures.** The outcome measures were response in alkaline phosphatase (ALP),  
443 Alanine Aminotransferase (ALT) and Aspartate Aminotransferase (AST).

444 **Whole Metagenomic Sequencing.** Stool samples were collected at home frozen within 15  
445 minutes of collection and maintained at -80 °C before sending to Diversigen for DNA extraction,  
446 whole-genome shotgun library preparation, and Illumina sequencing. The metagenomics  
447 sequencing targeted approximately 5 Gb of sequences per sample, utilizing 151 base pair paired-  
448 end reads.

449 **Read-level quality control and metagenomic profiling.** Raw sequencing reads underwent  
450 quality control (QC) using KneadData version 0.12.0  
451 (<https://huttenhower.sph.harvard.edu/kneaddata/>). Briefly, the first step of this process utilizes  
452 Trimmomatic version 0.39 for forward/reverse adapter removal, low-quality reads trimming and



453 tandem repeats removal using the default parameters. Contaminants from the host (human)  
454 genomes were subsequently identified and removed by mapping against the reference database  
455 (hg37\_and\_human\_contamination) using Bowtie2 version 2.5.1.

456 After QC, each dataset contained over 2.8 Gb of clean, paired-end reads. Clean reads were then  
457 taxonomically profiled using MetaPhlan4 version 4.0.6 [58], which includes an updated database  
458 significantly larger than the previous version 3.1. Species-level relative abundances were  
459 considered all this study. Species that failed to exceed 0.1% average relative abundance or were  
460 detected in less than 4 samples were excluded. Functional profiling was performed using  
461 HUMAnN3 version 3.6 under default parameters according to the UniRef90 definition to get the  
462 relative abundance of gene families in the unit of reads per kilobase (RPK) [59]. And MetaCyc  
463 provides pathway definitions to group gene families to get pathway abundance and coverage. The  
464 abundances of gene families were further transformed into the unit of copies per million (CPM),  
465 and we regrouped the gene families to KEGG Orthogroups (KOs), to filter out 74 gene families  
466 that were involved in the microbial sulfur transformation.

467 **Global untargeted metabolomics.** Stool samples were sent to Metabolon for global untargeted  
468 metabolomics following the standard procedure. Briefly, samples were pretreated to remove  
469 proteins and recover chemically diverse metabolites. The resulting extract was divided into five  
470 fractions: two for analysis by two separate reverse phases (RP)/UPLC-MS/MS methods with  
471 positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with  
472 negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI,  
473 and one sample was reserved for backup. All methods utilized a Waters ACQUITY ultra-  
474 performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high  
475 resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II)  
476 source and Orbitrap mass analyzer operated at 35,000 mass resolution. Raw data were extracted,  
477 peak-identified and QC processed using the contractor's hardware and software. Compounds were  
478 identified by comparison to library entries of purified standards or recurrent unknown entities. The  
479 intensity of each compound was quantified using area-under-the-curve.

480 **Metabolite-level quality control and pretreatment.** Metabolite intensities were normalized as z-  
481 scores:  $\frac{x_i - \mu}{\sigma}$ , where,  $x_i$  represents the original metabolite intensity,  $\mu$  is the mean value of that  
482 metabolite intensity across all samples, and  $\sigma$  is the standard deviation of that metabolite intensity  
483 across all samples. Subsequently, missing values were imputed using the minimum normalized  
484 value for each compound. The normalized and imputed values were used in the association and  
485 multivariate linear regression analyses throughout the context. With the exception of the estimation  
486 of temporal changes of metabolite intensity in each individual, the original intensities of specified  
487 metabolites in W0 were subtracted from their original intensity in W4 (or W8, W14).

488 **Statistical analyses.** Prior to downstream analysis, metagenomic and metabolomic were combined.  
489 Samples with complete profiles of both types were used in the downstream analysis, totalling 64  
490 samples. Three samples with only metabolomics datasets were excluded. From a total of 108



491 nutrients, 59 nutrient features were selected based on domain knowledge. These features, measured  
492 in various units, were log-transformed before we adjusted them for total energy intake to get the  
493 nutrient densities. This was followed by Lasso regression utilizing the 'glmnet' library in RStudio  
494 to select significant dietary features (predictors) that are non-zero [60]. The response variable was  
495 the corresponding ALP value. Subsequently, we incorporated the selected features and included  
496 patient ages as an additional variable to construct a linear model. The coefficients and p-values  
497 were extracted for statistical significance. For each patient, the Change of ALP value is estimated  
498 by subtracting the baseline ALP (W0) from the ALP at W4 (or W8, W14), then dividing by the  
499 baseline ALP (W0). Then for each disease category (PSC-UC/PSC-CD/PSC-alone), the Potential  
500 to Response was further calculated as the number of samples with reduced ALP value divided by  
501 the total number of samples. Alpha diversity of gut microbiota was estimated using the Shannon  
502 index and microbial richness (detected species counts). Principal coordinates analysis was  
503 performed using the bray-curtis distance of the relative abundance of species in all samples.  
504 Metabolite set enrichment analysis (MSEA) was conducted quantitatively through the online  
505 platform of MetaboAnalyst 5.0 [36].

506 **Linear regression models.** We used multivariate linear regression models to identify potential  
507 microbial markers distinguishing responders from non-responders after LPD. Abundances were  
508 fitted with the following species-specific linear mixed-effects model,

509 feature ~ **response outcomes** + diagnostic categories + gender + age +BMI + (1 | individual) +  $\epsilon$

510 **(Formula 1)**

511 In each species-specific multivariate model, the abundance of each species was modeled as the  
512 function of the binary response outcomes (responders/non-responders, with non-responder as  
513 reference) within each individual (as random effect), while adjusting for diagnostic categories  
514 (PSC-UC/PSC-CD/PSC-alone, with PSC-alone as reference), genders (male/female, with male as  
515 reference), age (continuous variable) and BMI (continuous variable). MaAsLin2 [61] in RStudio  
516 was the package used to fit the model, with p-values adjusted for multiple hypothesis testing and  
517 a target FDR of 0.2.

518 Separately, another series of multivariate linear regression models were fitted to find a differential  
519 abundance of features after LPD versus baseline using microbial profile, metabolic profile, and  
520 microbial functional profile, respectively. The model is as below,

521 feature ~ **treatment stages** + diagnostic categories + gender + age +BMI + (1 | individual) +  $\epsilon$

522 **(Formula 2)**

523 The difference between this model and the previous one is that the abundance of each feature is  
524 modeled as the function of the treatment stage (baseline/after LPD, with baseline as reference)  
525 instead of response results, while the rest fixed effect and random effect are not changed. The  
526 baseline corresponds to samples in W0, while after LPD corresponds to samples in W4 and W8.

527 Only those with a response of ALP were considered to fit these models to explain the dietary effects  
528 of the LPD.

529 **Clinicaltrials.gov ID:** NCT04678219

530  
531 **Author Contributions.** JK,YYL, GS, DB, MA, LG, FS, ES, and CB designed the project. GS led  
532 the clinical trial. XY led the data analysis and manuscript preparation. JK, YYL, GS, ZS, SK, DB,  
533 SDM, MA, SH, TC, MP, NJ, MD, AA, LG, FS, ES, and CB interpreted the results. XY prepared  
534 the manuscript. GS, JK, and YYL edited the manuscript. All authors reviewed and approved the  
535 manuscript. JK and YYL supervised the study.

536  
537 **Declaration of Interests.** No potential competing interest was reported by the authors.

538  
539 **Acknowledgements.** This work was supported by the Resnek Family Center for PSC Research.

540

541

## 542 REFERENCES

- 543 1. Hov, J.R. and T.H. Karlsen, *The microbiota and the gut-liver axis in primary sclerosing*  
544 *cholangitis*. Nat Rev Gastroenterol Hepatol, 2022.
- 545 2. Sayed, A., et al., *Predictors and Outcomes of Biologic/Immunosuppressive Therapy for*  
546 *PSC-Associated IBD*. Gastroenterology, 2022. 162(3): p. S76.
- 547 3. Mertz, A., et al., *Primary sclerosing cholangitis and inflammatory bowel disease*  
548 *comorbidity: an update of the evidence*. Ann Gastroenterol, 2019. 32(2): p. 124-133.
- 549 4. Hsu, C.L. and B. Schnabl, *The gut-liver axis and gut microbiota in health and liver disease*.  
550 Nat Rev Microbiol, 2023. 21(11): p. 719-733.
- 551 5. Jiang, X. and T.H. Karlsen, *Genetics of primary sclerosing cholangitis and*  
552 *pathophysiological implications*. Nature Reviews Gastroenterology & Hepatology, 2017.  
553 14(5): p. 279-295.
- 554 6. Hov, J.R. and T.H. Karlsen, *The microbiota and the gut-liver axis in primary sclerosing*  
555 *cholangitis*. Nature Reviews Gastroenterology & Hepatology, 2023. 20(3): p. 135-154.
- 556 7. Timur, L., et al., *Alterations of the bile microbiome in primary sclerosing cholangitis*. Gut,  
557 2020. 69(4): p. 665.
- 558 8. Vieira-Silva, S., et al., *Quantitative microbiome profiling disentangles inflammation- and*  
559 *bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses*. Nat  
560 Microbiol, 2019. 4(11): p. 1826-1831.
- 561 9. Fitzpatrick, J.A., et al., *Dietary management of adults with IBD - the emerging role of*  
562 *dietary therapy*. Nat Rev Gastroenterol Hepatol, 2022. 19(10): p. 652-669.
- 563 10. Lloyd-Price, J., et al., *Multi-omics of the gut microbial ecosystem in inflammatory bowel*  
564 *diseases*. Nature, 2019. 569(7758): p. 655-662.
- 565 11. Federici, S., et al., *Targeted suppression of human IBD-associated gut microbiota*  
566 *commensals by phage consortia for treatment of intestinal inflammation*. Cell, 2022.  
567 185(16): p. 2879-2898 e24.
- 568 12. Zhang, Y., et al., *Discovery of bioactive microbial gene products in inflammatory bowel*  
569 *disease*. Nature, 2022. 606(7915): p. 754-760.
- 570 13. Hov, J.R. and M.J.C.O.i.G. Kummen, *Intestinal microbiota in primary sclerosing*  
571 *cholangitis*. 2016. 33: p. 85-92.
- 572 14. Kentaro, I., et al., *Characterisation of the faecal microbiota in Japanese patients with*  
573 *paediatric-onset primary sclerosing cholangitis*. Gut, 2017. 66(7): p. 1344.
- 574 15. Nakamoto, N., et al., *Gut pathobionts underlie intestinal barrier dysfunction and liver T*  
575 *helper 17 cell immune response in primary sclerosing cholangitis*. Nature Microbiology,  
576 2019. 4(3): p. 492-503.
- 577 16. Kummen, M., et al., *Altered Gut Microbial Metabolism of Essential Nutrients in Primary*  
578 *Sclerosing Cholangitis*. Gastroenterology, 2021. 160(5): p. 1784-1798 e0.
- 579 17. Liu, Q., et al., *Altered faecal microbiome and metabolome in IgG4-related sclerosing*  
580 *cholangitis and primary sclerosing cholangitis*. Gut, 2022. 71(5): p. 899-909.

- 581 18. Ichikawa, M., et al., *Bacteriophage therapy against pathological Klebsiella pneumoniae*  
582 *ameliorates the course of primary sclerosing cholangitis*. Nat Commun, 2023. 14(1): p.  
583 3261.
- 584 19. Wolfson, S.J., et al., *Bacterial hydrogen sulfide drives cryptic redox chemistry in gut*  
585 *microbial communities*. Nature Metabolism, 2022. 4(10): p. 1260-1270.
- 586 20. Barton, L.L., et al., *Sulfur Cycling and the Intestinal Microbiome*. Digestive Diseases and  
587 Sciences, 2017. 62(9): p. 2241-2257.
- 588 21. Pitcher, M.C.L., E.R. Beatty, and J.H. Cummings, *The contribution of sulphate reducing*  
589 *bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis*.  
590 Gut, 2000. 46(1): p. 64.
- 591 22. Próchnicki, T., et al., *Mitochondrial damage activates the NLRP10 inflammasome*. Nature  
592 Immunology, 2023. 24(4): p. 595-603.
- 593 23. Lewis, J.D., et al., *A Randomized Trial Comparing the Specific Carbohydrate Diet to a*  
594 *Mediterranean Diet in Adults With Crohn's Disease*. Gastroenterology, 2021. 161(3): p.  
595 837-852.e9.
- 596 24. Cohen, S.A., et al., *Clinical and mucosal improvement with specific carbohydrate diet in*  
597 *pediatric Crohn disease*. J Pediatr Gastroenterol Nutr, 2014. 59(4): p. 516-21.
- 598 25. Wastyk, H.C., et al., *Gut-microbiota-targeted diets modulate human immune status*. Cell,  
599 2021. 184(16): p. 4137-4153.e14.
- 600 26. Slomski, A., *Mediterranean Diet vs Low-fat Diet for Patients With Heart Disease*. Jama,  
601 2022. 327(24): p. 2386.
- 602 27. Longo, V.D. and R.M. Anderson, *Nutrition, longevity and disease: From molecular*  
603 *mechanisms to interventions*. Cell, 2022. 185(9): p. 1455-1470.
- 604 28. Kahleova, H., S. Levin, and N.D. Barnard, *Vegetarian Dietary Patterns and*  
605 *Cardiovascular Disease*. Prog Cardiovasc Dis, 2018. 61(1): p. 54-61.
- 606 29. Gentile, C.L. and T.L. Weir, *The gut microbiota at the intersection of diet and human health*.  
607 Science, 2018. 362(6416): p. 776-780.
- 608 30. Corbin, K.D., et al., *Host-diet-gut microbiome interactions influence human energy*  
609 *balance: a randomized clinical trial*. Nat Commun, 2023. 14(1): p. 3161.
- 610 31. Bajaj, J.S., S.C. Ng, and B. Schnabl, *Promises of microbiome-based therapies*. Journal of  
611 Hepatology, 2022. 76(6): p. 1379-1391.
- 612 32. Liwinski, T., et al., *A prospective pilot study of a gluten-free diet for primary sclerosing*  
613 *cholangitis and associated colitis*. Aliment Pharmacol Ther, 2023. 57(2): p. 224-236.
- 614 33. Wishart, D.S., et al., *HMDB 5.0: the Human Metabolome Database for 2022*. Nucleic  
615 Acids Res, 2022. 50(D1): p. D622-d631.
- 616 34. Kanehisa, M., et al., *KEGG for taxonomy-based analysis of pathways and genomes*.  
617 Nucleic Acids Res, 2023. 51(D1): p. D587-d592.
- 618 35. Kim, S., et al., *PubChem 2023 update*. Nucleic Acids Research, 2022. 51(D1): p. D1373-  
619 D1380.

- 620 36. Pang, Z., et al., *Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics*  
621 *integration and covariate adjustment of global metabolomics data*. Nature Protocols, 2022.  
622 17(8): p. 1735-1761.
- 623 37. Xia, J. and D.S. Wishart, *MSEA: a web-based tool to identify biologically meaningful*  
624 *patterns in quantitative metabolomic data*. Nucleic Acids Research, 2010. 38(suppl\_2): p.  
625 W71-W77.
- 626 38. Weininger, D., et al., *SMILES. 2. Algorithm for generation of unique SMILES notation*.  
627 1989. 29(2): p. 97-101.
- 628 39. Wolf, P.G., et al., *Diversity and distribution of sulfur metabolic genes in the human gut*  
629 *microbiome and their association with colorectal cancer*. Microbiome, 2022. 10(1): p. 64.
- 630 40. Marcelino, V.R., et al., *Disease-specific loss of microbial cross-feeding interactions in the*  
631 *human gut*. Nat Commun, 2023. 14(1): p. 6546.
- 632 41. Heinken, A., et al., *Systematic assessment of secondary bile acid metabolism in gut*  
633 *microbes reveals distinct metabolic capabilities in inflammatory bowel disease*.  
634 Microbiome, 2019. 7(1): p. 75.
- 635 42. Lobel, L., et al., *Diet posttranslationally modifies the mouse gut microbial proteome to*  
636 *modulate renal function*. 2020. 369(6510): p. 1518-1524.
- 637 43. Blachier, F., M. Beaumont, and E. Kim, *Cysteine-derived hydrogen sulfide and gut health:*  
638 *a matter of endogenous or bacterial origin*. Curr Opin Clin Nutr Metab Care, 2019. 22(1):  
639 p. 68-75.
- 640 44. Guzior, D.V. and R.A. Quinn, *Review: microbial transformations of human bile acids*.  
641 Microbiome, 2021. 9(1): p. 140.
- 642 45. Collins, S.L., et al., *Bile acids and the gut microbiota: metabolic interactions and impacts*  
643 *on disease*. Nature Reviews Microbiology, 2023. 21(4): p. 236-247.
- 644 46. Xing, K., et al., *Roseburia intestinalis*; *generated butyrate boosts*  
645 *anti-PD-1 efficacy in colorectal cancer by activating cytotoxic*  
646 *CD8<sup>+</sup> T cells*. Gut, 2023. 72(11): p. 2112.
- 647 47. Nie, K., et al., *Roseburia intestinalis: A Beneficial Gut Organism From the Discoveries in*  
648 *Genus and Species*. 2021. 11.
- 649 48. Stummer, N., et al. *Role of Hydrogen Sulfide in Inflammatory Bowel Disease*. Antioxidants,  
650 2023. 12, DOI: 10.3390/antiox12081570.
- 651 49. Teigen, L.M., et al. *Dietary Factors in Sulfur Metabolism and Pathogenesis of Ulcerative*  
652 *Colitis*. Nutrients, 2019. 11, DOI: 10.3390/nu11040931.
- 653 50. Rowan, F.E., et al., *Sulphate-reducing bacteria and hydrogen sulphide in the aetiology of*  
654 *ulcerative colitis*. British Journal of Surgery, 2009. 96(2): p. 151-158.
- 655 51. Khalil, N.A., et al., *In vitro batch cultures of gut microbiota from healthy and ulcerative*  
656 *colitis (UC) subjects suggest that sulphate-reducing bacteria levels are raised in UC and*  
657 *by a protein-rich diet*. International Journal of Food Sciences and Nutrition, 2014. 65(1):  
658 p. 79-88.

- 659 52. Jowett, S.L., et al., *Influence of dietary factors on the clinical course of ulcerative colitis: a prospective cohort study*. *Gut*, 2004. 53(10): p. 1479.
- 660
- 661 53. Magee, E.A., et al., *Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans*<sup>123</sup>. *The American Journal of Clinical Nutrition*, 2000. 72(6): p. 1488-1494.
- 662
- 663
- 664 54. Day, A.S., et al., *Therapeutic Potential of the 4 Strategies to Sulfide-REduction (4-SURE) Diet in Adults with Mild to Moderately Active Ulcerative Colitis: An Open-Label Feasibility Study*. *The Journal of Nutrition*, 2022. 152(7): p. 1690-1701.
- 665
- 666
- 667 55. U.S. Department of Health and Human Services and U.S. Department of Agriculture. *2015 – 2020 Dietary Guidelines for Americans*. 8th Edition. December 2015.
- 668
- 669 56. Tuttle, S.G., et al., *Further Observations on the Amino Acid Requirements of Older Men: II. Methionine and Lysine*. *The American Journal of Clinical Nutrition*, 1965. 16(2): p. 229-231.
- 670
- 671
- 672 57. Gottschall, E., *Breaking the vicious cycle: intestinal health through diet*. 1994: Kirkton, Ont.: Kirkton Press.
- 673
- 674 58. Blanco-Míguez, A., et al., *Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlan 4*. *Nature Biotechnology*, 2023. 41(11): p. 1633-1644.
- 675
- 676
- 677 59. Beghini, F., et al., *Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3*. *eLife*, 2021. 10: p. e65088.
- 678
- 679 60. Friedman, J.H., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent*. *Journal of Statistical Software*, 2010. 33(1): p. 1 - 22.
- 680
- 681 61. Mallick, H., et al., *Multivariable association discovery in population-scale meta-omics studies*. *PLOS Computational Biology*, 2021. 17(11): p. e1009442.
- 682

683

684



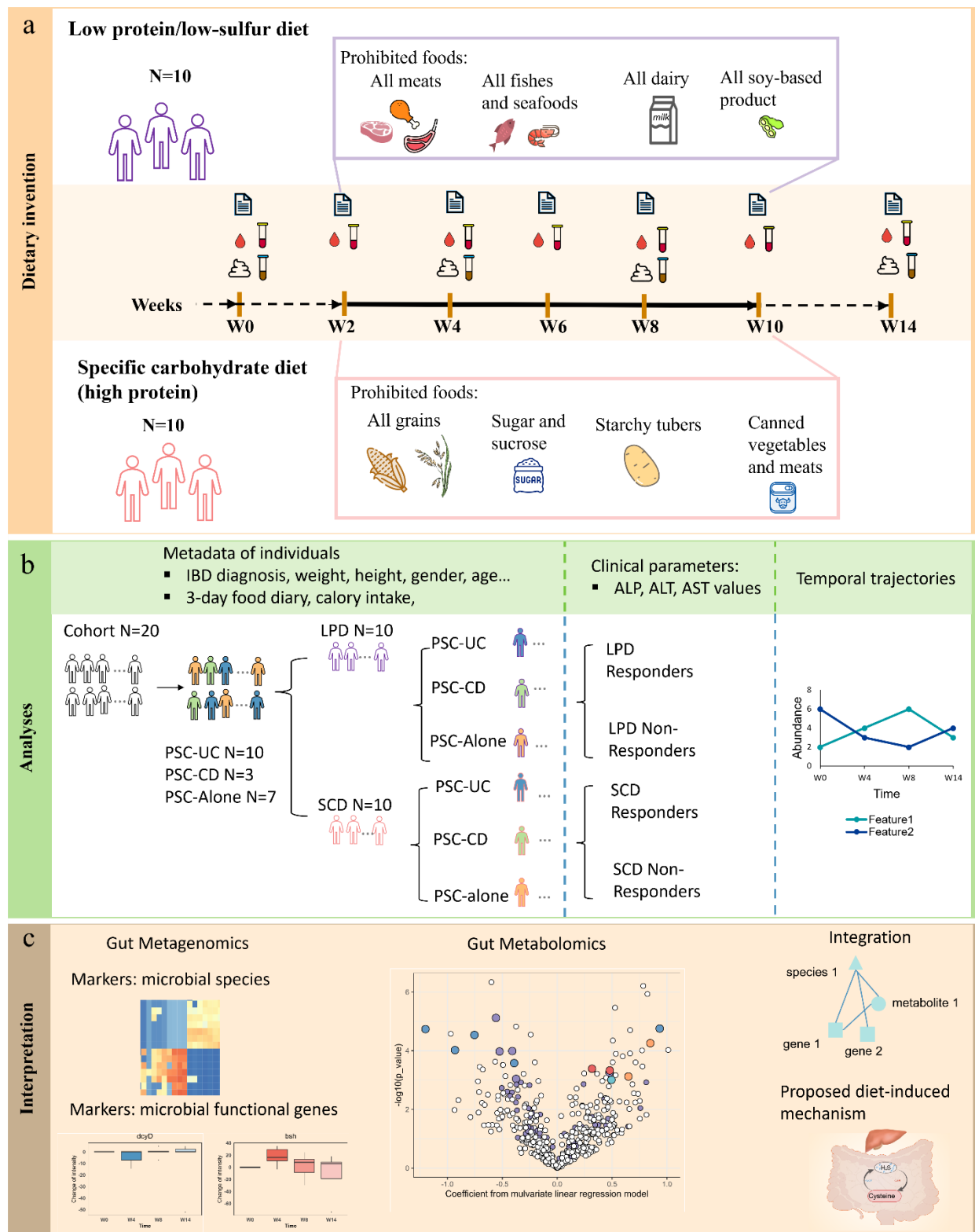
<b>Table 1 Patients' characteristics at inclusion.</b>		
	LPD (n=10)	SCD (n=10)
Age at inclusion (years) <sup>a</sup>	43.79 ± 10.05	43.95 ± 8.99
Gender (male/female)	10 (7/3)	10 (4/6)
Diagnosis <sup>b</sup>		
PSC-UC	6 (60%)	4 (40%)
PSC-CD	1 (10%)	2 (20%)
PSC-alone	3 (30%)	4 (40%)
BMI <sup>a</sup>	26.89 ± 6.69	23.53 ± 3.77
Smoker=yes <sup>b</sup>	3 (30%)	2 (20%)

685 <sup>a</sup> Mean ± SD

686 <sup>b</sup> n (%)

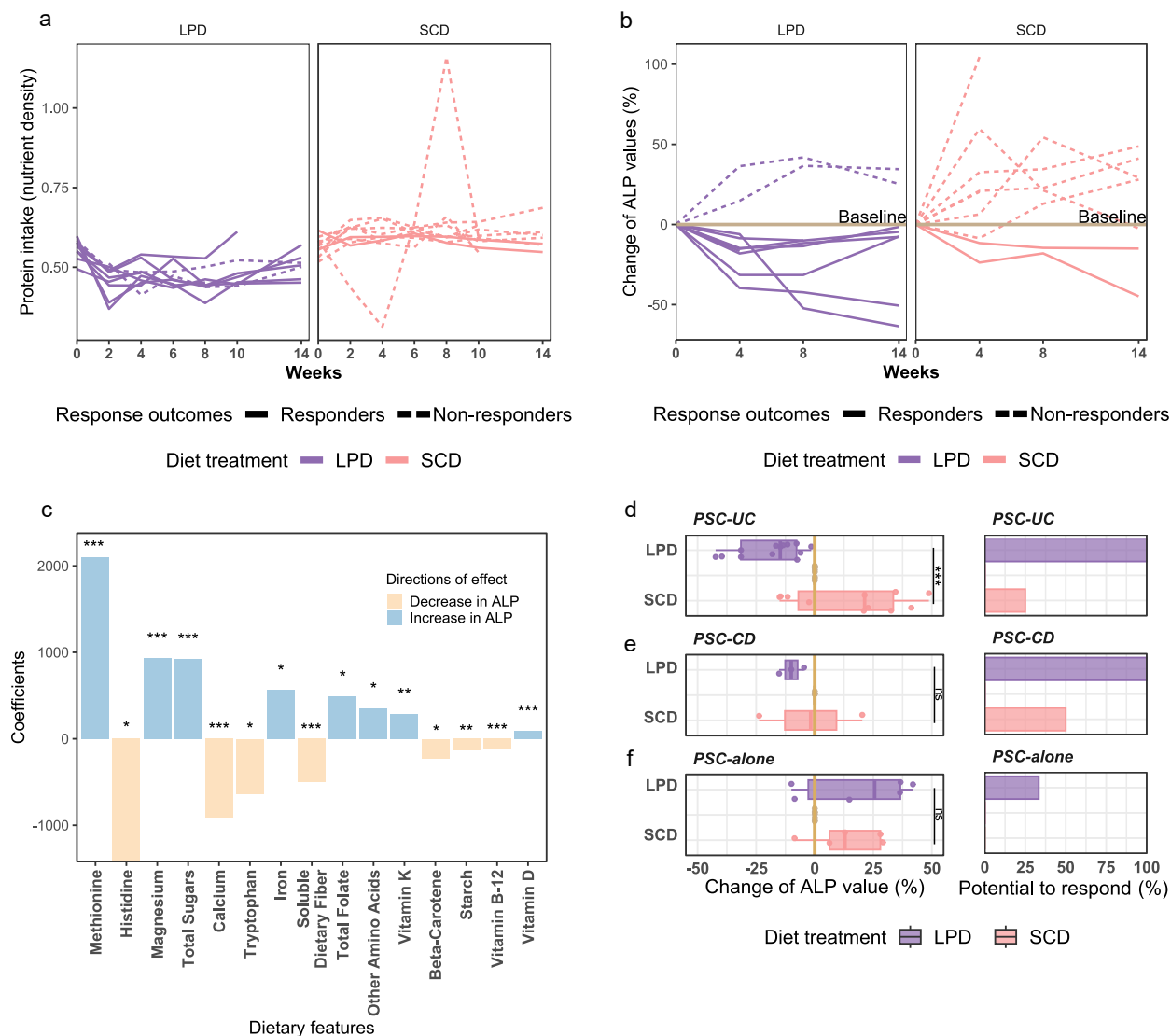
687 UC, ulcerative colitis; CD, Crohn's disease; IBD, inflammatory bowel disease; PSC, primary  
688 sclerosing cholangitis; BMI, Body mass index.

689



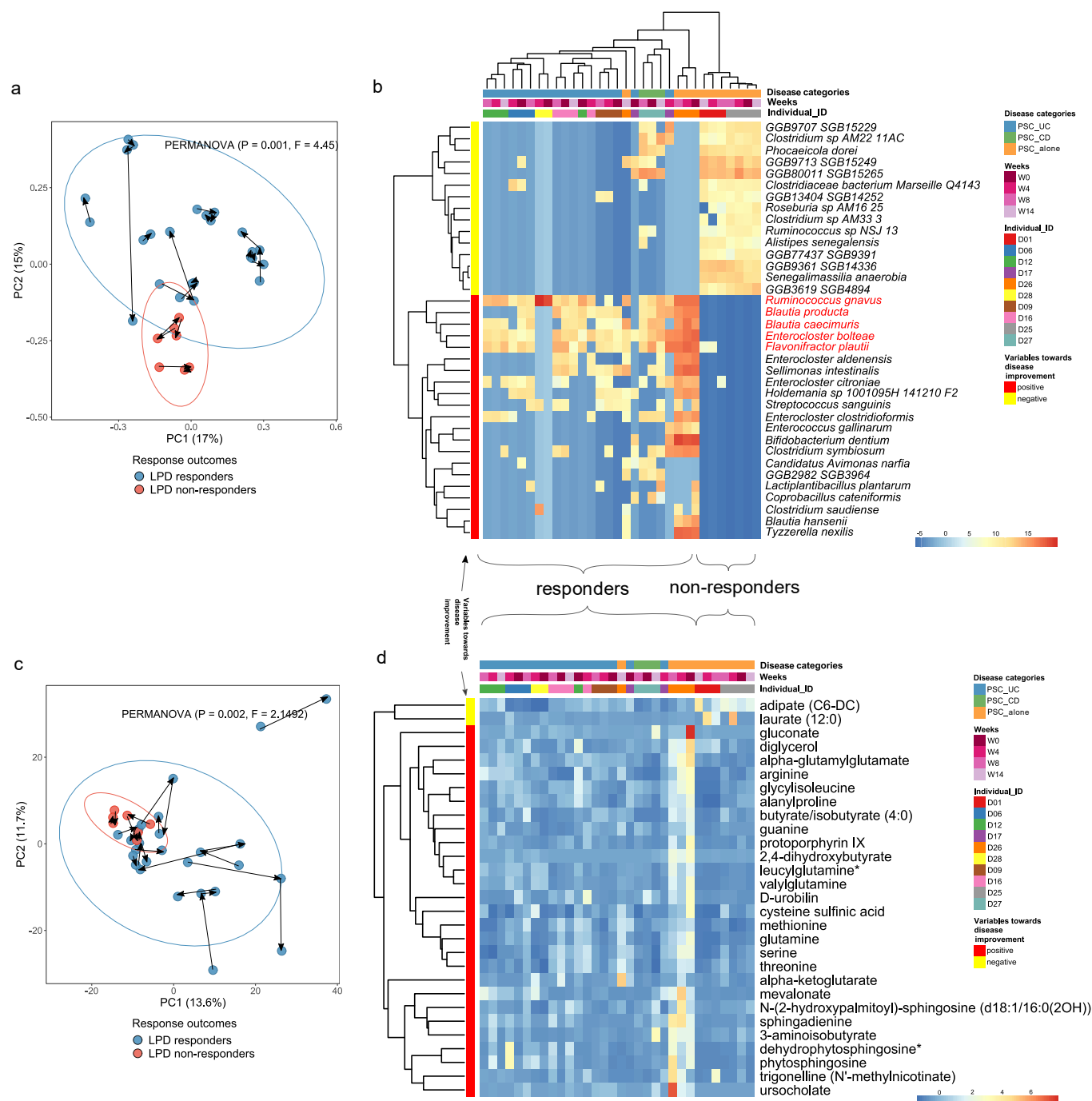
692 **Figure 1. Overview of the study design.** Two dietary interventions: Low protein/low sulfur diet  
693 (LPD) and Specific carbohydrate diet (SCD) were compared. We profiled fecal metagenomic and  
694 metabolomic data at four time points. 3-day dietary records, co-diagnosis of IBD (UC, CD, or non-  
695 IBD), and clinical parameters were collected. After dietary interventions, the response outcomes  
696 and temporal trajectory were analyzed and interpreted through multiple perspectives.

697



698

699 **Figure 2. LPD dietary intervention benefits patients with PSC, whereas SCD exacerbates the**  
 700 **condition.** **a)** Protein intake in LPD and SCD. Each line represents an individual. Protein intake  
 701 amount was adjusted by the total energy intake to estimate the nutrient density. **b)** Percent change  
 702 in alkaline phosphatase (ALP) from baseline before, during and after dietary intervention. Each  
 703 line represents an individual. **c)** Coefficients of selected nutrient features after Lasso regression,  
 704 indicating that reduced methionine intake significantly lowers ALP values. Nutrients with  
 705 significance in the regression model were plotted, with \*\*\* denoting strong significance (p-value  
 706 < 0.001), \*\* very significant (0.001 < p-value < 0.01), \* significant (0.01 < p-value < 0.05). **d-f)**  
 707 Variations in ALP values and the potential rates to respond indicated differing responses among  
 708 three disease categories, **d)** PSC-UC, **e)** PSC-CD, and **f)** PSC-alone. The Wilcoxon rank-sum test  
 709 was employed to assess significance between LPD and SCD, with \*\*\* denoting strong significance  
 710 (p-value < 0.001), and 'ns' indicating no significant difference (p-value > 0.05).

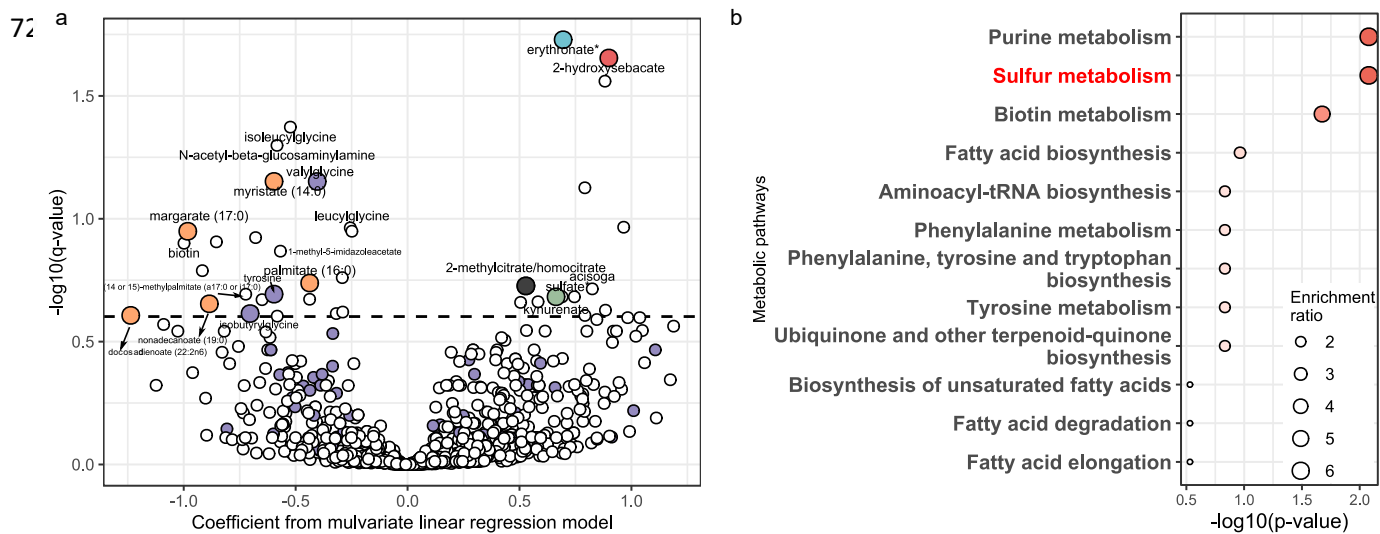


711

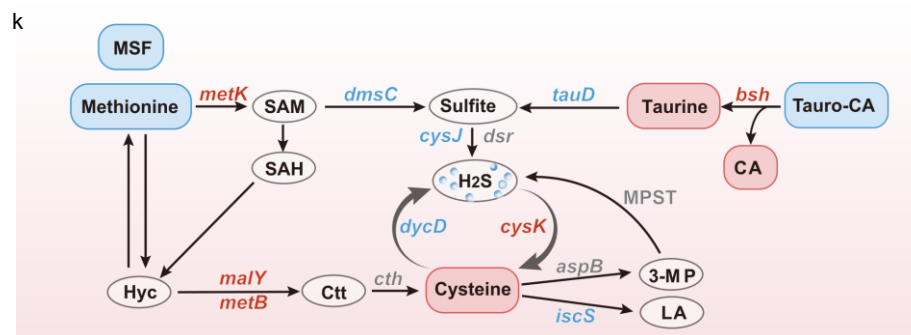
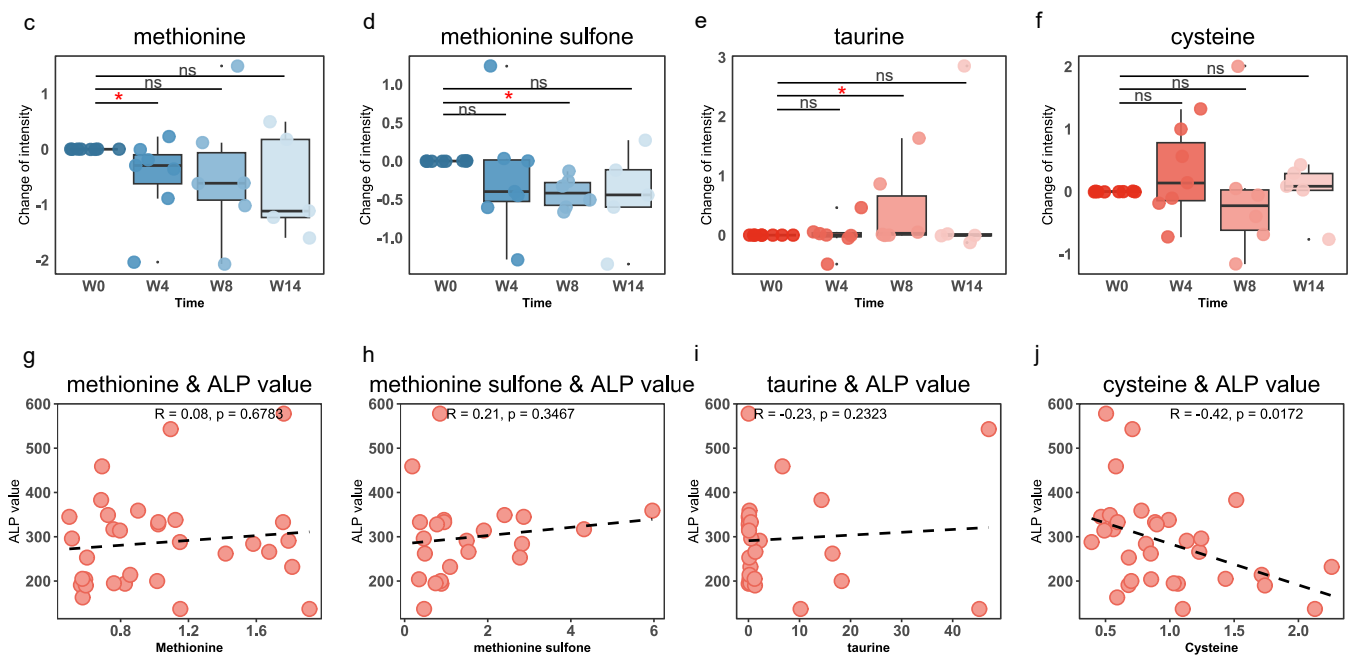
712 **Figure 3. Fecal metagenomics predicts response outcomes.** a) PCoA plot of the gut microbial  
 713 compositions of LPD samples, differentiated by response outcomes (blue for responders:  
 714 individuals showing decreased ALP, and pink for non-responders: individuals showing increased  
 715 ALP, shows that LPD responders have distinct microbial structure from LPD non-responders.  
 716 Arrows were drawn to connect samples from the same individuals in chronological order. b)  
 717 Heatmap of bacterial species markers that can distinguish LPD responders from non-responders  
 718 and the abundance distribution of these markers correlate with disease categories, as determined

719 by multivariate linear regression models (target p-value < 0.05 and FDR < 0.2). Each column in  
720 the heatmap is a sample, and all the samples including both LPD responders and LPD non-  
721 responders are clustered based on the similarity of the bacterial composition. Each row is a species,  
722 and species are clustered based on the similarity of the distribution pattern across samples of the  
723 specified species. **c)** PCA plot of the gut metabolite compositions of LPD samples, differentiated  
724 by response outcomes (blue and pink), with arrows to connect samples from the same individuals  
725 in chronological order. **d)** Heatmap of metabolite markers that can distinguish LPD responders  
726 from non-responders. The columns in this heatmap are arranged to correspond directly with the  
727 order of the columns in the microbial profile heatmap in Fig. 3b.



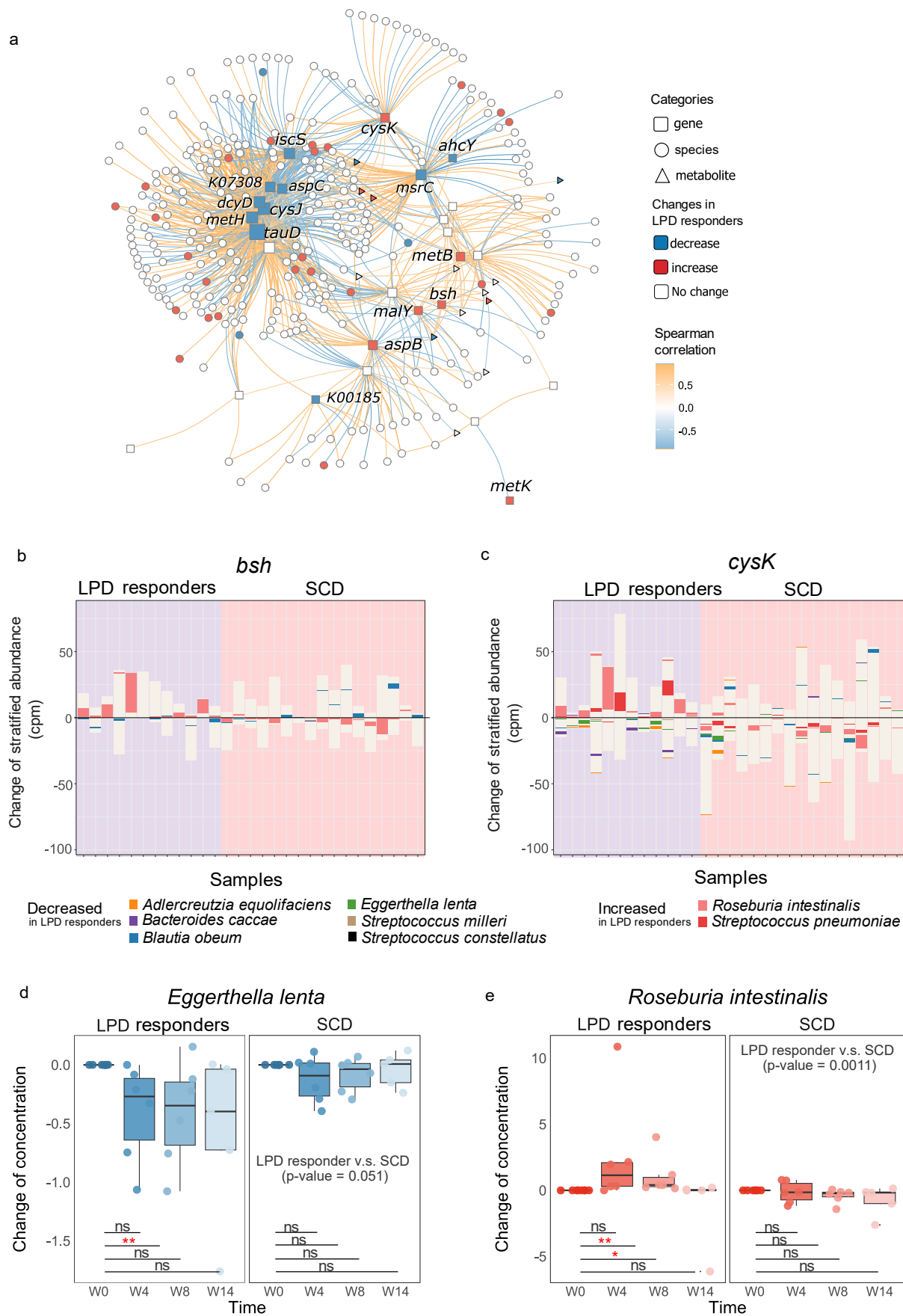


- Amino acids, peptides, and analogues
- Medium-chain hydroxy acids and derivatives
- Others
- Carbohydrates and carbohydrate conjugates
- Quinoline carboxylic acids
- Fatty acids and conjugates
- Tricarboxylic acids and derivatives



729 **Figure 4. The fecal metabolomics predicts the disease improvement trajectory under LPD in**  
730 **those LPD responders. a)** Volcano plot derived from multivariate linear regression models, using  
731 time series metabolite intensity profiles from W0, W4, and W8 and adjusting for covariates  
732 including disease categories. The metabolites were colored according to the class subclasses. **b)**  
733 Metabolite set enrichment analysis (MSEA) of selected metabolites with significant diet-induced  
734 changes after the regression model in Fig.3a (MetaboAnalyst v5.0, p-value < 0.05). **c-f)** Temporal  
735 changes in four key sulfur metabolism metabolites, including methionine, methionine sulfone,  
736 taurine, and cysteine. Bars are colored blue to represent metabolites that decreased after LPD, and  
737 orange for those that increased. The Wilcoxon signed-rank test was employed to assess  
738 significance between sampling time points, with \* denoting significance ( $0.01 < p\text{-value} < 0.05$ ),  
739 and 'ns' indicating no significant difference ( $p\text{-value} > 0.05$ ). **g-j)** The relationship between each  
740 of the four metabolites and ALP values tested using Spearman correlation. Cysteine shows a  
741 negative association with ALP values, whereas the other three metabolites have no significant  
742 association ( $p\text{-value} < 0.05$ ). **k)** Characterized and proposed sulfur-related metabolic pathways in  
743 response to LPD intervention benefiting patients with PSC. Metabolites and genes that increased  
744 are colored red, those that decreased are blue, while those without significant changes are gray  
745 (Wilcoxon signed-rank test,  $p\text{-value} < 0.05$ ). Abbreviations used are as follows: MSF: methionine-  
746 sulfone; Hyc: homocysteine; Ctt: cystathionine; LA: L-alanine; 3-MP: 3-Mercaptopyruvate;  
747 Tauro-CA: taurocholic acid; CA: cholic acid.

748



750 **Figure 5. Microbial sulfur metabolism and associated bacterial species.** a) Interactions among  
751 microbial sulfur metabolism genes, related metabolites, and relevant microbial species. Nodes  
752 were colored blue to represent features that decreased after LPD and red for those features that  
753 increased. Edge gradients represent the coefficients of Spearman correlation between the  
754 abundance (or intensity) of nodes. Only associations deemed significant, with  $FDR < 0.05$ , are  
755 shown here. **b-c)** Shifts in the stratified abundance of bacterial species for two represent genes (*bsh*  
756 and *cysK*) after dietary treatments. Each column is a sample, with the left panel on a purple  
757 background representing LPD responders under intervention, and the right panel on a pink  
758 background corresponding to samples from the SCD treatments. We subtracted the stratified  
759 abundance in each sample from its corresponding baseline to calculate the gene's change amount.  
760 Each column contains stacked bars, with each bar representing the change in stratified abundance  
761 of a specific bacterial species identified as hosts of the specified genes. Values below the  $y=0$  line  
762 indicate a decrease, while those above indicate an increase. Species are colored only if their  
763 stratified abundances showed significant changes (either increases or decreases) at W4 (or W8)  
764 compared to the baseline in LPD responders. The unit 'cpm' represents counts per million. **d-e)**  
765 Temporal changes of two key bacteria involved in sulfur metabolism: d) *Eggerthella lenta* and e)  
766 *Roseburia intestinalis*. Bars are colored blue to represent metabolites that decreased after LPD,  
767 and orange for those that increased. The Wilcoxon rank-sum test was employed to assess  
768 significance between sampling time points, with \*\* denoting very significant changes ( $0.001 < p$ -  
769 value  $< 0.01$ ), \* significant ( $0.01 < p$ -value  $< 0.05$ ), and 'ns' indicating no significant difference ( $p$ -  
770 value  $> 0.05$ ).

771