Title

Unbalanced dietary patterns contribute to the pathogenesis of precocious puberty by affecting gut microbiota and host metabolites

Authors:

Ying Wang[1, 2#], Dingfeng Wu[3, 4#], Hongying Li[1 ,2], Xiangrong Liang[1, 2], Na Jiao[5, 6], Wenxing Gao[3], Lu Zhao[2], Han Yu[2], Qian Wang[2], Yongsheng Ge[2], Changying Zhao[2], Meiling Huo[2], Ruifang Cao[7], Sheng Gao[3], Liwen Tao[3], Yunchao Ling[7], Lingna Zhao[7], Xin Lv[1, 2], Yi Liu[1, 2], Lehai Zhang[1, 2], Haokui Zhou[8], Guoqing Zhang[7], Guoping Zhao[7, 9, 10], Lei Zhang[1, 10, 11*], Ruixin Zhu[3, 4, 6*], Zhongtao Gai[1, 2*]

[1]Qilu Children's Hospital, Cheeloo College of Medicine, Shandong University, Jinan 250012, P.R.China.

[2]Jinan Children's Hospital, Jinan 250022, P.R.China.

[3]Department of Gastroenterology, The Shanghai Tenth People's Hospital, Department of Bioinformatics, School of Life Sciences and Technology, Tongji University, Shanghai 200092, P.R.China.

[4]Bioinformatics Division, GloriousMed Clinical Laboratory Co., Ltd., Shanghai 201318, P. R.China.

[5]Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Department of Colorectal Surgery, the Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, P.R. China.

[6]Research Institute, GloriousMed Clinical Laboratory Co., Ltd., Shanghai 201318, P. R.China.

[7]CAS Key Laboratory of Computational Biology, Bio-Med Big Data Center, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, P.R.China.

[8]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, P.R.China.

[9]Hangzhou Institute for Advanced Study, UCAS, Hangzhou 330106, P.R.China.

[10]Microbiome-X, National Institute of Health Data Science of China & Institute for Medical Dataology, Cheeloo College of Medicine, Shandong University, Jinan 250012, P.R.China.

11. Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, P.R.China.

[#] Equal contributors

[*]Correspondence: zhanglei7@sdu.edu.cn (L.Z.), rxzhu@tongji.edu.cn (RX.Z.), gzt@etyy.com (ZT.G.)

38

39 **ABSTRACT:**

40 Precocious puberty (PP) mostly stems from endocrine disorders. However, its

41 triggering factors, especially for the early onset of partial PP, and the associated

42 pathogenic mechanisms remain ambiguous. In this study, a systematic analysis in the

43 form of a questionnaire of lifestyles, gut microbiome, and serum metabolome data was

44 carried out to examine the pathogenesis of PP in a cohort comprised of 200 girls, with

45 or without PP. The analysis revealed substantial alterations in gut microbiota, serum

46 metabolites, as well as lifestyle patterns in the PP group, which were characterized by

47 an elevated abundance of β-glucuronidase-producing and butyrate-producing bacteria,

48 and excessive lipid concentration with decreased levels of organic nitrogen compounds

49 in the serum of the participants. These differential microbes and metabolites tend to be

50 reliable non-invasive diagnostic biomarkers aiding the early diagnosis of PP and exhibit

51 a strong discriminative power (AUC = 0.93 and AUC = 0.97, respectively).

52 Furthermore, the microbial biomarkers were confirmed in an independent validation

53 cohort (n = 83, AUC = 0.85). Moreover, structural equation modeling revealed that

54 unhealthy dietary habits were the primary contributors for the alteration of gut

55 microbiota and serum metabolites, triggering the imbalance in the host hormones that

56 leads to premature physical development. Our study determines a causal relationship

57 among the gut microbiota, host metabolites, diet, and clinical characteristics of

58 preadolescent girls who experienced early onset of PP, and formulates non-invasive

59 diagnostic tools demonstrating excellent performance for the early detection of PP.

60 **Keywords:** causality, dietary pattern, gut microbiota, metabolomics, precocious

61 puberty.

62

63 INTRODUCTION

64 Precocious puberty (PP) refers to the premature occurrence of pubertal development or

65 secondary sexual characteristics, before the age of 8 in girls and 9 in boys, depicted by

66 features such as advanced breast and ovary development along with rapid bone growth

67 or maturation (Root, 2000), where the morbidity rate is also seen to rise progressively

68 (Kim et al., 2015). PP is a hormonal condition predominantly seen in females and can

69 be attributed to endocrine disorders, accompanied by an elevated sex hormone secretion

70 (Du et al., 2019). However, there still exists lack of clarity about the triggering factors

71 of this condition, especially for the premature onset of partial PP, and the pathogenic

72 mechanisms associated with it. It has been reported that dietary patterns seem to

73 considerably influence the estrogen metabolism mechanism, which is inextricably

74 linked with PP (Chen et al., 2018; Kim et al., 2011b; Merzenich et al., 1993; Rogers et

75 al., 2010). Over-nutrition or hyperalimentation, the excessive consumption of processed

76 and high-fat diet, is considered to be the principal agent responsible for the secular

77 decline in pubertal age (Muir, 2006; Soliman et al., 2014). Certain animal studies have

78 suggested that postnatal over-nutrition tends to invariably escalate the secretion of

79 luteinizing hormone (LH), follicle stimulating hormone (FSH), leptin, and insulin

80 levels in pubertal females, while the consumption of postnatal high-fat diet after

81 commencing weaning stimulates premature puberty in females (Soliman et al., 2014).

82 At the same time, harmful dietary patterns seem to significantly affect the composition

83 of human gut microbiota and metabolome (Kong et al., 2014; Sheflin et al., 2017). A

84 number of former studies conducted on adults with estrogen-mediated diseases, such as

85 experiencing menopausal symptoms, revealed that gut microbiota is capable of

86 effectively regulating metabolism and transforming estrogen-like compounds to

87 biologically active forms (Baker et al., 2017; Frankenfeld et al., 2014). Hence, PP has

88 been understood to be the outcome of early activation of hypothalamic-pituitary-

89 gonadal (HPG) axis initiated by certain pathophysiological stimuli, such as gut

90 microbiota or diet patterns (Brito et al., 1999; Cussotto et al., 2018; Qi et al., 2012).

91 Therefore, we investigated the gut microbiome and serum metabolome with respect to

92 the lifestyle information and clinical characteristics for a cohort of 200 participants,

93 where 133 girls experienced partial PP in early onset stage while 67 were healthy girls.

94 Our study registered an imperative association between the key metabolomics and

95 bacterial biomarkers, and their promising discriminatory power values by analyzing the

96   discovery and validation cohorts. Our innovative structural equation modeling (SEM)

97   analysis demonstrated the direct and/or indirect causal relationships across various

98   factors of this study for determining PP in preadolescent girls, such as gut microbiota,

99   host metabolism, lifestyles, and clinical characteristics.

100

101   **RESULTS**

102   **Baseline Characteristics of Participants**

103   200 female participants were recruited for this study (Table 1). 168 stool samples from participants

104   were collected for 16S rRNA sequencing to probe the microbiota alterations existing among girls

105   with PP (n = 105) and normal girls (n = 63). The average age of girls in the PP and the normal groups

106   at the time of stool sample collection was 6.641 (95 confidence interval, ci95 = 0.403) and 7.008

107   (ci95 = 0.515) years (P = 0.278, Table 1), respectively. 129 serum samples were collected for

108   untargeted metabolomics analysis, which included 45 PP and 84 normal girls illustrating an average

109   age of 6.662 (ci95 = 0.458) and 6.250 (ci95 = 0.571) years (P = 0.289, Table 1), respectively.

110   Additionally, lifestyle information of the 200 participants was obtained by means of a questionnaire

111   involving 117 variables of dietary patterns, living environment, maternal health, childbirth, and the

112   physical condition of participants as well as their parents. For the dietary patterns of the participants,

113   15 variables, including seafood (FDR = 2.36e-5), freshwater products (FDR = 2.91e-5), tubers (FDR

114   = 0.65e-3), and vegetables (FDR = 0.0021), showed significant differences between PP and the

115   normal group (Table S1). Likewise, 3 variables depicting the physical condition of the participants,

116   like dental care (FDR = 0.021), eczema (FDR = 0.045), and normal vaccination (FDR = 0.047), also

117   showed significant differences among the two groups under consideration (Table S1).

118

119   **Gut Microbiota Dysbiosis in Girls Suffering from PP**

120   Gut microbial composition displayed a huge variation between PP and the normal group. As

121   compared to the normal girls (P < 0.001, Fig. 1A), significantly elevated bacterial richness was

122   observed in the PP group. At the same time, the microbial diversity in the PP group appeared to be

123   substantially distinct from that of the normal group, which was further validated by PERMANOVA

124   test (pseudo-F = 3.24, P = 0.001, Fig. 1B). 45 notably differential amplicon sequence variants (ASVs)

125   were identified from the gut microbiota samples of the PP group and the normal controls (FDR <

126   0.05, Table S2). The abundance of ASVs, assigned as Bacteroidaceae, Ruminococcaceae,

127   Faecalibacterium, Enterobacteriaceae, and Escherichia-Shigella, was witnessed to have increased,

128   while those of genus *Agathobacter* and family Peptostreptococcaceae (*Romboutsia* and

129   *Intestinibacter (1)*) seemed to have decreased in the PP group (Fig. 1D, Table S2). Most of these

130   differential taxa exhibited the potential of encoding/producing β-glucuronidase (Fig. 1D, Table S2),

131   an enzyme that deconjugates estrogens into their active forms (Mcintosh et al.).

132   Furthermore, alterations on microbiome-mediated functional potentials were also explored, which

133   led to the identification of 88 differential pathways (FDR < 0.05) between PP and the normal group

134   (Fig. S1A). The PP group demonstrated enhanced activity levels in most metabolic processes, such

135   as metabolizing carbohydrate, cofactor and vitamin, fatty acid and lipid, and inorganic nutrient

136   metabolism.

137

138   **Microbial Markers Could Act as Non-Invasive Tools for PP Diagnosis**

139   To explore the possibility of differential microbes functioning as prospects of novel non-invasive

140   tools for PP diagnosis, a classification model was designed by employing 45 differential ASVs via

141   random forest algorithm. The model emerged highly capable for performing the clinical diagnosis

142   of PP, with an area under the receiver operating characteristic curve (AUC) of about 0.93 (Fig. 1C).

143   Afterwards, the classification model was validated in an independent cohort (n = 83, Table S3) to

144   further measure its generalization ability. The model constituting of 13 out of 45 differential ASVs

145   that exhibited the same abundance change pattern accomplished excellent results for distinguishing

146   between PP and the normal group with an AUC of about 0.85 (Fig. S2B). These 13 ASVs primarily

147   belonged to *Bacteroides* and Enterobacteriaceae (Table S3) and illustrated critical elevation

148   regarding several metabolic processes in the PP group (Fig. S1B).

149

150   **Latent Microbial Factors Revealed Underlying Relationships of Altered Microbiome**

151   Delving into the underlying relationships among differential taxa, we performed exploratory factor

152   analysis (EFA) and identified 10 latent microbial factors, which revealed the latent patterns of the

153   change in gut microbiota. These factors were then presented in the microbial co-occurrence network

154   (Fig. 1D). The classification models designed on the basis of these 10 microbial factors achieved a

155   high accuracy (AUC = 0.88, Fig. S2A) in distinguishing the PP group from the normal controls, not

156   only indicating the efficacy of the EFA analysis but also preserving the vital information possessed

157   by differential taxa. Methodically speaking, Gammaproteobacteria, such as family

158   Enterobacteriaceae and genus *Escherichia-Shigella*, which retained a strong correlation with each

159   other, were members of Factor 1 and 6 (Fig. 1D). The chief members of Factor 2, genus

160   *Subdoligranulum* and *Faecalibacterium* of family Ruminococcaceae, are known as butyrate-

161   producing bacteria in gut microbiota (Cussotto et al., 2018). Besides, four taxa, *Romboutsia*,

162   *Intestinibacter*, *Streptococcus*, and *Haemophilus*, present in a considerably low concentration in the

163   PP group, expressed a strong loading in Factor 7 (Fig. 1D, Table S2). Notably, several species

164   (except Peptostreptocaccaceae, Factor 7), were witnessed in elevated measures in the PP group (Fig.

165   1D, Table S2). Among them, factor *Bacteroides*, butyrate-producing bacteria, Enterobacteriaceaes,

166   and Burkholderiaceae appeared to be the most relevant towards the microbial metabolic pathways

167   (Fig. S1B) and could prove to be instrumental in microbial biological functions.

168

169   **Altered Serum Metabolome in Girls Suffering from PP**

170   Untargeted metabolomic profiles were evaluated under LC-MS/MS system in 84 PP and 45 normal

171   girls (Table 1) for the purpose of investigating metabolite alterations, and 182 differential

172   metabolites were identified (FDR < 0.05, Table S4). Among them, 131 differential metabolites, such

173   as Phenylalanine, 4-Guanidinobutamide, and Lysophosphatidylcholine (LPC), were present in

174   decreased quantities in the PP group as compared to the normal controls (Table S4). Significantly,

175   these 182 differential metabolites were remarkably efficient in differentiating the PP group from the

176   normal controls and obtained an AUC of 0.97 (Fig. 2B), which appears to be higher than microbial

177   markers.

178

179   **Organic Nitrogen Compounds and Lipids Were the Characteristic Latent Metabolic Factors**

180   Analogous to the EFA analysis for microbiota, 12 latent metabolic factors were identified based on

181   their differential metabolites (Fig. 2A, Table S4). Preserving the foremost information, the metabolic

182   factors exhibited the detection capability comparable to that of the differential metabolites for

183   detecting PP (AUC = 0.97, Fig. S2C). Specifically, Factor 1 constituted of organic nitrogen

184   compounds (Fig. 2A), including oligopeptide and nitrogen-containing alkaloids (such as

185   Ethylmorphine, Levallorphan, and Alangicine), and their abundances were seen consistently

186   declining in the PP group (Table S4). The serum levels of lipids in the PP group (Factor 2), such as

187   Oleoylethanolamide (Fold-change, FC = 158.85), 17beta-Nitro-5alpha-androstane (FC = 59.90),

188   Thiofanox (FC = 54.32), and 17-Methylstearate (FC = 17.33) were observed to be considerably

189   elevated (Fig. 2A, Table S4). In addition, an appreciable reduction in several other differential

190   metabolites, such as organosulfur compound (Factor 3), phenylalanine (Factor 5), and terpenoid

191   (Factor 6) were detected in the PP group as compared to the normal group (Table S4).

192

193   **Associations Between Microbial Factors and Metabolic Factors**

194   Subsequently, the associations between microbial and metabolic factors were evaluated to discover

195   the potential key drivers of such modifications. Relatively, organic polycyclic compound

196   expressed vital positive correlations with *Lachnospira*, Enterobacteriaceae, *Dialister*,

197   Peptostreptococcaceae, and Burkholderiaceae (Fig. 2C). Among them, Enterobacteriaceae was

198   positively associated with aferine, phenylalanine, and with LPC as well. The overlapping metabolic

199   pathways, such as amino acid, carbohydrate, aromatic compound metabolic pathway, and others

200   (Glycolysis II, Glyoxylate cycle, and Incomplete reductive TCA cycle), could prove to be the

201   foundation of the strong relationships between microbiota and metabolites (Fig. S1B). Interestingly,

202   even though *Bacteroides* and butyrate-producing bacteria exhibited striking correlations to most

203   metabolic pathways (Fig. S1B), they were not significantly associated with the differential

204   metabolites of PP (Fig. 2C). Furthermore, organic nitrogen compounds and lipids revealed no

205   substantial connection with most of the gut microbiota factors, except for factor *Alistipes* and factor

206   *Dialister* (Fig. 2C), suggesting that they may be directly affected by other factors, such as lifestyles.

207

208   **Three PP Subtypes Were Revealed by Differential Metabolites**

209   In comparison to the gut microbiota, more evident modifications were seen in serum metabolism

210   (Fig. S3A, B). The girls in the PP group were classified into three subgroups based on the expression

211   of 182 differential metabolites, all of which demonstrated varied metabolite patterns, implying that

212   PP could be categorized into 3 different subtypes. Specifically, latent metabolic factors, organic

213   nitrogen compounds, and organic polycyclic compounds emerged critically different among the

214   three defined subtypes (Fig. S3C and Table S5). Moreover, some latent microbial factors, such as

215   Peptostrptococcaceae and Lachnospira, also expressed varying abundance patterns among these

216   subtypes. Additionally, clinical laboratory tests, such as for LH, testosterone (TES), Zn, and Ca,

217   presented a similar trend (Fig. S3C). Although the widely utilized standard suggests no difference

218   between the various phenotypes (Table S5), subtype 2 showed a tendency to be distinct from the

219   other two subtypes and appeared to be more analogous to the normal group, which highlights the

220   individual differences among the PP group, especially in the metabolic changes.

221

222   **Dietary Pattern Varied Significantly between Groups**

223   Detailed lifestyle information that may potentially affect the PP group was obtained, including

224   dietary patterns, living environment, maternal health, childbirth, and physical condition of the

225   participants and their parents (Table S1). Dietary patterns presented considerable differences among

226   the patients and displayed good discriminative ability for distinguishing the PP group from the

227   normal controls, obtaining an AUC of about 0.87 (Fig. S2D). However, other lifestyle variables

228   failed to express any significant variations between the normal and the PP groups (Table S1) along

229   with a poor discriminative ability (Fig. S2E-H).

230   Furthermore, EFA facilitated the identification of 9 latent dietary factors derived from the dietary

231   patterns recorded (Fig. 3A, Table S6), which preserved the foremost information of dietary patterns

232   (AUC = 0.86, Fig. 3C). As expected, these latent dietary factors were an amalgamation of the various

233   aspects of the children's diet, such as healthy foods, junk foods, items containing monosodium

234   glutamate (MSG), and the balance between meat and vegetables. The most critical latent dietary

235   factor, healthy foods, appeared to have a significantly lower intake rate in the PP group (FDR =

236   0.30e-4, Fig. 3B), which entails preferences for seafood, freshwater products, tubers, vegetables,

237   bean products, fruits, nuts, etc. (Table S6). Several noteworthy correlations were noticed between

238   dietary, microbial, and metabolic factors (Fig. 3D). The intake of healthy foods expressed a highly

239   negative correlation with the abundance of serum lipid ($P < 0.01$). On the other hand, the intake of

240   snacks and drinks were witnessed to be positively linked with the organic nitrogen compounds (P

241   < 0.01), whereas negatively correlated with butyrate-producing bacteria ($P < 0.05$). These results

242   were indicative of the unbalanced dietary patterns influencing the PP progression through

243   microbiota and metabolites.

244

Unbalanced Dietary Patterns Affecting PP Progression Through Microbiota and Metabolites

Supported by the above results, potential causal relationships among gut microbiota, metabolites, dietary patterns, and the characteristics of the disease were investigated. For this purpose, we introduced the innovative SEM path analysis to construct a credible model (Fisher's $C = 815.85$ with $P = 0.341$) in accordance with the correlation results (Fig. 2, 3 and Fig S2I) and our prior knowledge about the subject, to reveal the internal connections among them (Fig. 4). Hormones (estradiol (E2), prolactin (PRL), LH and FSH, and trace elements (Zn, Ca, My, Cu, Fe) were significantly regulated by gut microbiota and serum metabolism. Among the serum metabolism agents, lipids elevated in the most dramatic fashion in the PP group (Table S4), positively affecting the ovarian volume ($P < 0.05$), while producing a negative effect on the breast volume ($P < 0.05$). The reduced intake of healthy foods ($P < 0.001$) and elevated intake of all-meat diet ($P < 0.001$) were known to be the primary factors causing the surge of serum lipids. Similarly, organic nitrogen compounds, the principal metabolic factor, produced a significantly positive effect on the level of serum E2 and further affected the development of ovarian volume. As an important sex hormone, LH was appreciably decreased in the PP group ($P < 0.01$, Table S7) and generated a considerably positive effect on the bone age ($P < 0.05$) and uterine volume ($P < 0.001$) in the PP group (Fig. 4). In serum metabolism, LPC may be the chief influencing factor of LH ($P < 0.05$, Fig. S2I). Furthermore, butyrate-producing bacteria, including genus *Subdoligranulum* and *Faecalibacterium* of family Ruminococcaceae, produced stronger positive effects on the follicle size ($P < 0.05$). The up-regulation of butyrate-producing bacteria may possibly be able to explain the premature development of follicles in the PP group.

The SEM analysis revealed that dietary patterns were the most vital of all the catalysts of change in microbiota and metabolism. Simultaneously, the dysbiosis of the gut bacteria taxa and metabolites produced a remarkable effect on the host hormone levels and PP progression (Fig. 4).

DISCUSSION

Mounting evidence suggests that gut microbiota and metabolism are the major decisive forces of the growth in children (Tamburini et al., 2016a; Yatsunenko et al., 2012). Nevertheless, no definite explanation is available about the effects of gut microbiota and serum metabolism in the

275 pathogenesis of precocious puberty. In this study, we performed a systematic analysis investigating

276 the lifestyle patterns, altered microbiota, metabolome, and their relationship with PP.

277 It was found that the PP group showed a significant dysbiosis of gut microbiota and serum

278 metabolome, which could be mainly attributed to the unhealthy dietary habits, directly affecting the

279 progression of PP. At the same time, gut microbiota and serum metabolome could prove to be non-

280 invasive as well as reliable diagnostic biomarkers for the early detection of PP, expressing an AUC

281 = 0.93 and 0.97, respectively, and can help evade the time-consuming and painful gonadotropin-

282 releasing hormone (GnRH) stimulation test (Kim et al., 2011a).

283 Diet shapes our gut microbiota in the early stages of our life (Tamburini et al., 2016b; Zmora et

284 al., 2019). A large number of studies have suggested that unhealthy dietary habits shift the gut

285 microbiota and may very well contribute to the pathogenesis of various metabolic diseases,

286 including overweight, obesity, type 2 diabetes, non-alcoholic liver disease, cardio-metabolic

287 diseases, and malnutrition (Fan and Pedersen, 2020; Tamburini et al., 2016b). An extensively shifted

288 microbial composition was detected in the PP group, implying a dysbiosis of gut bacterial

289 community. β-glucuronidase-producing bacteria, including *Alistipes*, *Bacteroides*, *Escherichia,* and

290 *Faecalibacterium* were noticed to be significantly increased in the PP group, possibly triggering the

291 elevated levels of circulating estrogen and increased estrogenic burden, which actuate the onset of

292 PP (Baker et al., 2017; Mcintosh et al.; Sultan et al., 2012). The shift in gut microbiota, especially

293 *Bacteroides* and butyrate-producing bacteria is capable of affecting various metabolic processes

294 (Fan and Pedersen, 2020), such as carbohydrate, fatty acid, and lipid metabolism activity (Fig. S1).

295 Interestingly, butyrate-producing bacteria, which was considered to be a beneficial bacteria for

296 maintaining the gut health in several preceding studies (Cheng et al., 2016; Valles-Colomer et al.,

297 2019), may emerge as a pathogenic agent for PP. Considering the promoting effect of butyrate to

298 the levels of LH (Ruddon et al., 1979) and FSH (Ghosh and Cox; Liang et al., 2020), we propose

299 that overproducing butyrate induced by gut bacteria may produce a detrimental effect on the health

300 condition during puberty, especially for follicular development (Fig. 4). Consistent with the previous

301 study on the gut microbiota of girls suffering from PP conducted by Dong et al., the gut bacteria

302 related to the production of short-chain fatty acids (SCFAs) are known to be present in increased

303 concentrations in girls suffering from PP, which promote the expression of the leptin gene, activate

304 the HPG axis through a high concentration of SCFAs, and trigger the early onset of puberty (Dong

305   et al., 2019).

306   Moreover, PP gives rise to a more serious metabolic disturbance. Differential metabolites exhibit

307   a strong ability for PP diagnosis (AUC = 0.97) and present three different PP subtypes. A large

308   amount of organic nitrogen compound down-regulation and lipid up-regulation with high fold-

309   change were seen to be the chief characteristics of PP serum metabolome. Decreased intake of

310   healthy foods ($P < 0.001$), the unique differential dietary pattern of PP, illustrated a direct effect on

311   the serum levels of lipids (2-158 fold increase), which was previously verified in high-fat diet mice

312   (Walker et al., 2017). Overnutrition and excessive intake of processed and high-fat food leads to

313   obesity at the beginning of the PP pathogenesis (Latronico et al., 2016; Mugo et al., 2007). In

314   addition, animal studies have also indicated that postnatal overnutrition consistently increases the

315   LH, FSH, leptin, and insulin levels in pubertal females, and postnatal high fat diet after commencing

316   weaning tends to trigger advanced puberty in females [31,32]. The dysbiosis of the serum

317   metabolites indicates an influence on the level of serum metallic elements (Ca, Zn, Cu, Mg, and Pb),

318   which could be potential endocrine disrupters that are capable of modulating estrogenic activity of

319   endogenous hormones (Arjmandi et al., 1993).

320   Unlike the traditional correlation research, this study adopted the causal inference method based

321   on the SEM analysis, which has been progressively implemented for varied purposes in the

322   microbiome field (Chen et al., 2019; Mamet et al., 2019). This method can effectively derive and

323   comprehend the causal relationships between dietary patterns, gut microbiota, serum metabolome,

324   and PP. Being constrained by prior knowledge and data integrity, the explanation for some results

325   remains unclear, and requires further in-depth research.

326   In summary, it was found that unhealthy dietary habits could disrupt the homeostasis of gut

327   microbiota and serum metabolism, and consequently trigger the imbalance of hormones, leading to

328   the excessive change of physical development progress and PP genesis. Therefore, the intestinal

329   microbiota may be regarded as a prospective therapeutic target for the prevention and treatment of

330   PP.

331   **Limitations of Study**

332   The etiology of PP remains complicated; hence it requires further validation by employing larger

333   samples and effectively designing disease prediction models. In addition, although this study

334   included the validation cohorts, the multi-center studies and big sizes of the validation cohorts will

335   be needed to further validate the biomarkers found in this study.

336

337   **Materials and methods**

338   **Study participants**

339   133 girl participants with partial PP and 67 normal girls were recruited in the discovery cohort at

340   Qilu Children's Hospital, Shandong University (Table 1) for this study. 66 normal and 17 girls

341   suffering from PP (age < 8) were included in the independent validation cohort to externally verify

342   our findings. The PP group was diagnosed in accordance with the criteria defined by Lawson

343   Wilkins Pediatric Endocrine Society. Exclusion criteria includes (1) other organic etiologies with

344   presence of isointense tumor on magnetic resonance imaging (MRI); (2) usage of antibiotics,

345   probiotics, or prebiotics within 3 months before enrolling; (3) associated endocrine, gastrointestinal,

346   metabolic disease (including obesity and diabetes, among others), mental disease, or hepatobiliary

347   disease. Recruited age-matched normal controls were girls below 8 years of age without PP. The

348   study protocol was maintained in accordance with the Declaration of Helsinki and was approved by

349   the Ethics Committee of Qilu Children's Hospital (ETYY-2016-202). Written informed consents

350   and questionnaires were obtained from the children's parents.

351

352   **Sample collection**

353   Stool and blood samples were collected from each participant and stored at -80 °C before analysis.

354   200 mg stool was preserved in sterile 2 ml tubes containing pure ethanol, aliquoted (Tinygene

355   Biological Company, China) and stored at -80 °C for 16S rRNA sequencing. Blood samples were

356   thawed at 4 °C, 3000 rpm, and centrifuged at 4 °C for 10 min. Serum aliquots were immediately

357   frozen at -80 °C for further untargeted metabolomics analysis. The study was approved by local

358   ethics committees (Qilu Children's Hospital of Shandong University, IRB Number EYY-2016-202)

359   and informed consent was obtained from all the participating subjects.

360

361   **DNA extraction and illumina sequencing**

362   Total DNA extraction from fecal samples (250 mg, wet weight) was performed using a Fast DNA

363   SPIN Kit for feces (MP Biomedicals, Santa Ana, CA, USA), as per the manufacturer's instructions.

364   The V1-V2 hypervariable region was amplified with the universal primer pair F27 (5' -

365   AGAGTTTGATCMTGGCTCAG-3') and R355 (5'- GCTGCCTCCCGTAGGAGT -3').

366   Sequencing was conducted on Illumina HiSeq 2500 System (Illumina Inc., San Diego, CA, USA)

367   using the 2×250 paired-end mode following the standard Illumina platform protocols. All

368   sequencing data is available at NODE (http://www.biosino.org/node) with the accession number

369   OEP000731.

370

371   **16S rRNA gene sequencing data analysis**

372   16S rRNA sequencing data was analyzed using Quantitative Insights Into Microbial Ecology

373   (QIIME2 V.2019.07). In brief, raw sequence data was demultiplexed and DADA2 (SP) was

374   employed to denoise sequencing reads for quality control and the identification of amplicon

375   sequence variants (ASVs) via q2-dada2 plugin. Taxonomy classification was carried out by utilizing

376   classify-sklearn based on a Naiva Bayes classifier against the Silva-132-99 reference sequences.

377   Respective sequences of each ASV were aligned with Multiple Alignment using Fast Fourier

378   Transform (MAFFT) (Katoh et al., 2002) (via q2-alignment) and the phylogenetic tree was

379   constructed with Fast-Tree (Price et al., 2010) (via q2-phylogeny). Chao1, an index of richness

380   estimator, was calculated to assess the community alpha diversity. Principal coordinate analysis

381   (PCoA) was performed based on the Bray-Curtis distance; and PERMANOVA test was conducted to

382   evaluate the significant differences present among the microbial communities, with 9999

383   permutations.

384   Additionally, the functions of gut microbiota were inferred based on 16S rRNA sequencing data

385   using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States

386   (PICRUSt2), as previously described (Douglas et al., 2020).

387

388   **Untargeted metabolomics and analysis**

389   100 μL of serum sample was transferred to an EP tube. After the addition of 300 μL of methanol

390   (containing internal standard 1 μg/mL), the samples were vortexed for 30 s, followed by sonication

391   for 10 min in ice-water bath, and incubation for 1 h at -20 °C to precipitate the proteins. The sample

392   was then centrifuged at 12000 rpm for 15 min at 4 °C. The resulting supernatants were then

393   transferred to LC-MS vials and stored at -80 °C until the UHPLC-QE Orbitrap/MS analysis. The

394   quality control sample was prepared by mixing an equal aliquot of the supernatants from all the

395   samples collected.

396   LC-MS/MS analyses were performed using an UHPLC system (1290, Agilent Technologies) with

397   a UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 μm) coupled to Q Exactive (Orbitrap MS, Thermo).

398   The mobile phase A was 0.1% formic acid in water for positive, and 5 mmol/L ammonium acetate

399   in water for negative, and the mobile phase B was acetonitrile. The elution gradient was set as

400   follows: 0 min, 1% B; 1 min, 1% B; 8 min, 99% B; 10 min, 99% B; 10.1 min, 1% B; 12 min, 1%

401   B. The flow rate was 0.5 mL/min. The injection volume was 2 μL. The QE mass spectrometer was

402   utilized due to its ability to acquire MS/MS spectra on an information-dependent basis (IDA) during

403   an LC/MS experiment. In this mode, the acquisition software (Xcalibur 4.0.27, Thermo)

404   continuously examines the full scan survey MS data as it collects and triggers the acquisition of

405   MS/MS spectra depending on preselected criteria. ESI source conditions were set as follows: Sheath

406   gas flow rate as 45 Arb, Aux gas flow rate as 15Arb, Capillary temperature at 400 °C, Full ms

407   resolution as 70000, MS/MS resolution as 17500, Collision energy as 20/40/60 eV in NCE model,

408   Spray Voltage as 4.0 kV (positive) or -3.6 kV (negative), respectively.

409   Raw data was converted to mzXML format using ProteoWizard and processed using MAPS

410   software (version 1.0). Preprocessed results were employed to generate a data matrix which

411   consisted of the retention time (RT), mass-to-charge ratio (m/z) values, and peak intensity. In-house

412   MS2 database was applied for metabolites identification.

413

414   **Questionnaire survey**

415   Questionnaires were distributed to the participants and their parents in presence of trained doctors,

416   providing professional guidance throughout the whole process. From the questionnaire survey, 117

417   variables were collected, involving dietary pattern, living environment, maternal health, childbirth,

418   and personal physical condition of the recruited girls (PP and normal group), as well as the physical

419   condition of their parents.

420   The dietary pattern section in our questionnaire was based on the most frequently consumed foods

421   that were clinically considered to be closely related to PP, including vegetables, fruits, seafood, meat,

422   cereal, tubers, eggs, milk, bean products, nuts, fungi, greasy food, beverages, fried food, sweets,

423   barbecued food, puffed food, pickles, gourmet powder, and tonic, on a 4-level intake frequency

424   scale. The dietary section also entailed personal food preferences, including preferences for fruit,

425   vegetables, meat, pasta, bean products, fried food, snacks, beverage, meat, etc.

426   The variables regarding living environment included presence of pets, use of mineral water,

427   existence of bowel dysfunction of close contacts, presence of plastic foam products, pesticides,

428   fertilizers, insecticide, etc. The questionnaire also separately covered the physical condition of

429   participants' parents, including stomachaches, abdomen distension, diarrhea, gastric acid

430   regurgitation, constipation, hypertension, hyperglycemia, anemia, rheumatism, urticaria,

431   immunodeficiency, and irregular menstruation and menarche age (only for mothers). The data about

432   maternal conditions during pregnancy and delivery was derived, including reproductive age,

433   medication during pregnancy, secretory disorders during pregnancy, folic acid supplement, dietary

434   patterns, alcohol consumption, abnormal fetal movement, oxytocin, dystocia, or fetal hypoxia,

435   cesarean delivery, or spontaneous delivery, etc. The physical condition of the recruited girls was

436   investigated thoroughly, including sleep disorder, poor weight gain, jaundice, eczema, diarrhea, oral

437   malodor, sediment in urine, smelly urine, perianal red, dental caries, etc.

438

439   **Clinical laboratory tests**

440   Clinical parameters were determined at the clinical lab of Qilu Children's Hospital. The trace

441   element (Cu, Zn, Ca, Mg, Fe, and Pb) levels from serum samples were measured using the flame

442   atomic absorption method (BH5100, Bohui, China). The thyroid function test was conducted by

443   analyzing the serum levels of follicle stimulating hormone (FSH), luteinizing hormone (LH), and

444   thyroid stimulating hormone (TSH), employing the chemi-luminescence immunoassay methods

445   (Abbott, Architect I2000, US). GnRH stimulation test was conducted for the luteinizing hormone

446   (LH) and follicle-stimulating hormone (FSH) utilizing the chemi-luminescence immunoassay

447   methods (Abbott, Architect I2000, US). Cortisol (COR), adrenocortieotropic hormone (ACTH),

448   alpha fetoprotein (AFP), and carcinoembryonic antigens (CEA) were measured by chemi-

449   luminescence immunoassay methods (LIAISON, type 2210, Germany). The plasma levels of the

450   insulin and insulin-like growth factors were quantified by adopting immunoluminescence method

451   (Siemens immulite 2000, USA). The sizes of uterus, breasts, and ovaries, as well as the number of

452   ovarian follicles were determined by employing B-ultrasonic examination (EPIQ5, L12-5, Philips,

453   Holland). The hand-wrist radiographs were used for bone-age assessment through nuclear magnetic

454   resonance (MRI) examination (Digital Dianost3, Philips, Holland).

455

Co-occurrence analysis

457   Co-occurrence analysis was applied for microbial, metabolic, or dietary network by using

458   correlations (Spearman, Spearman or Kendall). Correlations with $P \leq 0.05$ (permutation test with

459   1000 permutations) were included in the co-occurrence networks. Network visualization was

460   conducted using Gephi software (M et al., 2009).

461

Disease diagnosis model

463   Classification model of different samples was constructed using Random Forest classifier in Scikit-

464   learn package of Python (3.6.0)(F et al., 2011). The AUC of 5-fold cross-validation was utilized to

465   measure the discriminative ability of the differential biomarkers.

466

Exploratory factor analysis

468   For differential microbiota, metabolites, and dietary patterns, exploratory factor analysis (EFA) was

469   employed to identify the latent factors with FactorAnalyzer in Python (C, 2016). The number of

470   factors, solutions (minimum residual, maximum likelihood or principal factor), and rotations

471   (varimax or promax) of EFA were determined by minimizing information loss after dimension

472   reduction, which was evaluated by the discriminative ability of the classification of normal and PP

473   samples. Latent factors with high loading were explained and labeled based on the observed

474   variables and former knowledge.

475

Structural equation model

477   Based on the latent factors from EFA, path analysis of the SEM (D et al., 2019) was employed to

478   discover the causal relationships between lifestyle, gut microbiota, metabolism, and clinical

479   characteristics of disease. Considering that the samples cannot match perfectly, piecewise SEM

480   (Shipley, 2013) was used for confirmatory path analysis in our study, in which each set of

481   relationships was determined independently (or locally). The p-value of Fisher's C statistic was

482   adopted to prove the overall rationality of the model and to facilitate the model comparison and

483   selection.

484

Statistical analysis

485

486 All statistical analyses were conducted using Python (3.6.0). Statistical significance was determined

487 by two-sided Wilcoxon rank-sum test, Permutation test or one-sided Fisher's exact test, and

488 Benjamini-Hochberg test was applied to control the false positive rate (FDR) under multiple

489 comparisons. Differences were considered statistically significant when FDR < 0.05.

490

REFERENCES:

491

492 Arjmandi, B.H., Salih, M.A., Herbert, D.C., Sims, S.H., and Kalu, D.N. (1993). Evidence for
493 estrogen receptor-linked calcium transport in the intestine. Bone and mineral 21, 63-74.

494 Baker, J.M., Al-Nakkash, L., and Herbst-Kralovetz, M.M. (2017). Estrogen–gut microbiome axis:
495 physiological and clinical implications. Maturitas 103, 45-53.

496 Brito, V., Batista, M., Borges, M., Latronico, A., Kohek, M., Thirone, A., Jorge, B., Arnhold, I., and
497 Mendonca, B. (1999). Diagnostic value of fluorometric assays in the evaluation of precocious
498 puberty. The Journal of Clinical Endocrinology & Metabolism 84, 3539-3544.

499 C, H.M. (2016). A review of exploratory factor analysis decisions and overview of current practices:
500 What we are doing and how can we improve? International Journal of Human-Computer Interaction
501 32, 51-62.

502 Chen, C., Chen, Y., Zhang, Y., Sun, W., Jiang, Y., Song, Y., Zhu, Q., Mei, H., Wang, X., and Liu, S.
503 (2018). Association between dietary patterns and precocious puberty in children: a population-based
504 study. International journal of endocrinology 2018.

505 Chen, L., Jiang, Y., Liang, C., Luo, Y., Xu, Q., Han, C., Zhao, Q., and Sun, B.J.M. (2019).
506 Competitive interaction with keystone taxa induced negative priming under biochar amendments.
507 7, 1-18.

508 Cheng, J., Ringel-Kulka, T., Heikamp-de Jong, I., Ringel, Y., Carroll, I., de Vos, W.M., Salojarvi, J.,
509 and Satokari, R. (2016). Discordant temporal development of bacterial phyla and the emergence of
510 core in the fecal microbiota of young children. ISME J 10, 1002-1014.

511 Cussotto, S., Sandhu, K.V., Dinan, T.G., and Cryan, J.F. (2018). The neuroendocrinology of the
512 microbiota-gut-brain axis: a behavioural perspective. Frontiers in neuroendocrinology 51, 80-101.

513 D, M.S., E, R., M, B., G, L.E., L, H.B., K, S., and D, S.S. (2019). Structural equation modeling of
514 a winnowed soil microbiome identifies how invasive plants re-structure microbial networks. The
515 ISME journal 13, 1988-1996.

516 Dong, G., Zhang, J., Yang, Z., Feng, X., Li, J., Li, D., Huang, M., Li, Y., Qiu, M., and Lu, X. (2019).
517 The association of gut microbiota with idiopathic central precocious puberty in girls. Frontiers in
518 Endocrinology 10.

519 Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower,
520 C., and Langille, M.G.J.N.B. (2020). PICRUSt2 for prediction of metagenome functions. 1-5.

521 Du, G., Hu, J., Huang, Z., Yu, M., Lu, C., Wang, X., and Wu, D. (2019). Neonatal and juvenile
522 exposure to perfluorooctanoate (PFOA) and perfluorooctane sulfonate (PFOS): Advance puberty
523 onset and kisspeptin system disturbance in female rats. Ecotoxicology and environmental safety 167,

524   412-421.

525   F, P., G, V., A, G., V, M., B, T., O, G., M, B., P, P., R, W., V, D., et al. (2011). Scikit-learn: Machine
526   learning in Python. the Journal of machine Learning research 12, 2825-2830.

527   Fan, Y., and Pedersen, O. (2020). Gut microbiota in human metabolic health and disease. Nature
528   Reviews Microbiology, 1-17.

529   Frankenfeld, C.L., Atkinson, C., Wähälä, K., and Lampe, J.W. (2014). Obesity prevalence in relation
530   to gut microbial environments capable of producing equol or O-desmethylangolensin from the
531   isoflavone daidzein.   68, 526-530.

532   Ghosh, N.K., and Cox, R.P. Induction of human follicle-stimulating hormone in HeLa cells by
533   sodium butyrate. Nature 267, 435-437.

534   Katoh, K., Misawa, K., Kuma, K.i., and Miyata, T. (2002). MAFFT: a novel method for rapid
535   multiple sequence alignment based on fast Fourier transform. Nucleic acids research 30, 3059-3066.

536   Kim, H.K., Kee, S.J., Seo, J.Y., Yang, E.M., Chae, H.J., and Kim, C.J. (2011a). Gonadotropin-
537   releasing hormone stimulation test for precocious puberty. The Korean journal of laboratory
538   medicine 31, 244-249.

539   Kim, J., Kim, S., Huh, K., Kim, Y., Joung, H., and Park, M. (2011b). High serum isoflavone
540   concentrations are associated with the risk of precocious puberty in Korean girls. Clinical
541   endocrinology 75, 831-835.

542   Kim, S.H., Huh, K., Won, S., Lee, K.-W., and Park, M.-J. (2015). A significant increase in the
543   incidence of central precocious puberty among Korean girls from 2004 to 2010. PLoS One 10.

544   Kong, L.C., Holmes, B.A., Cotillard, A., Habi-Rachedi, F., Brazeilles, R., Gougis, S., Gausserès, N.,
545   Cani, P.D., Fellahi, S., and Bastard, J.-P. (2014). Dietary patterns differently associate with
546   inflammation and gut microbiota in overweight and obese subjects. PloS one 9.

547   Latronico, A.C., Brito, V.N., and Carel, J.-C. (2016). Causes, diagnosis, and treatment of central
548   precocious puberty. The Lancet Diabetes & Endocrinology 4, 265-274.

549   Liang, Y., Ming, Q., Liang, J., Zhang, Y., Zhang, H., and Shen, T. (2020). Gut microbiota dysbiosis
550   in Polycystic ovary syndrome (PCOS): association with obesity - A preliminary report. Canadian
551   Journal of Physiology and Pharmacology.

552   M, B., S, H., and M, J. (2009). Gephi: an open source software for exploring and manipulating
553   networks. Third international AAAI conference on weblogs and social media.

554   Mamet, S.D., Redlick, E., Brabant, M., Lamb, E.G., Helgason, B.L., Stanley, K., and Siciliano,
555   S.D.J.T.I.j. (2019). Structural equation modeling of a winnowed soil microbiome identifies how
556   invasive plants re-structure microbial networks.   13, 1988-1996.

557   Mcintosh, F.M., Maison, N., Holtrop, G., Young, P., Stevens, V.J., Ince, J., Johnstone, A.M., Lobley,
558   G.E., Flint, H.J., and Louis, P. Phylogenetic distribution of genes encoding β-glucuronidase activity
559   in human colonic bacteria and the impact of diet on faecal glycosidase activities. Environmental
560   Microbiology 14, 0-0.

561   Merzenich, H., Boeing, H., and Wahrendorf, J. (1993). Dietary fat and sports activity as
562   determinants for age at menarche. American journal of epidemiology 138, 217-224.

563   Mugo, M.N., Bachrach, B.E., and Jacobson, J.D. (2007). Partial and Gonadotropin-Dependent
564   Precocious Puberty.

565   Muir, A. (2006). Precocious puberty. Pediatrics in Review 27, 373.

566   Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2–approximately maximum-likelihood
567   trees for large alignments. PloS one 5.

568   Qi, Y., Li, P., Zhang, Y., Cui, L., Guo, Z., Xie, G., Su, M., Li, X., Zheng, X., and Qiu, Y. (2012).

569   Urinary metabolite markers of precocious puberty. Molecular & Cellular Proteomics 11.

570   Rogers, I.S., Northstone, K., Dunger, D.B., Cooper, A.R., Ness, A.R., and Emmett, P.M. (2010).

571   Diet throughout childhood and age at menarche in a contemporary cohort of British girls. Public

572   health nutrition 13, 2052-2063.

573   Root, A.W. (2000). Precocious puberty. Pediatrics in review 21, 10-19.

574   Ruddon, R.W., Anderson, C., Meade, K.S., Aldenderfer, P.H., and Neuwald, P.D. (1979). Content

575   of gonadotropins in cultured human malignant cells and effects of sodium butyrate treatment on

576   gonadotropin secretion by HeLa cells. Cancer Research 39, 3885-3892.

577   Sheflin, A.M., Melby, C.L., Carbonero, F., and Weir, T.L. (2017). Linking dietary patterns with gut

578   microbial composition and function. Gut Microbes 8, 113-129.

579   Shipley, B. (2013). The AIC model selection method applied to path analytic models compared

580   using a d-separation test. Ecology 94, 560-564.

581   Soliman, A., De Sanctis, V., and Elalaily, R. (2014). Nutrition and pubertal development. Indian

582   journal of endocrinology and metabolism 18, S39.

583   SP, D. High-resolution sample inference from Illumina amplicon data.   (Nat).

584   Sultan, C., Gaspari, L., Kalfa, N., and Paris, F. (2012). Clinical Expression of Precocious Puberty

585   in Girls. Endocr Dev 22, 84-100.

586   Tamburini, S., Shen, N., Wu, H.C., and Clemente, J.C. (2016a). The microbiome in early life:

587   implications for health outcomes. Nat Med 22, 713-722.

588   Tamburini, S., Shen, N., Wu, H.C., and Clemente, J.C.J.N.m. (2016b). The microbiome in early life:

589   implications for health outcomes.   22, 713-722.

590   Valles-Colomer, M., Falony, G., Darzi, Y., Tigchelaar, E.F., Wang, J., Tito, R.Y., Schiweck, C.,

591   Kurilshikov, A., Joossens, M., Wijmenga, C., et al. (2019). The neuroactive potential of the human

592   gut microbiota in quality of life and depression. Nat Microbiol 4, 623-632.

593   Walker, A., Pfitzner, B., Harir, M., Schaubeck, M., Calasan, J., Heinzmann, S.S., Turaev, D., Rattei,

594   T., Endesfelder, D., and zu Castell, W. (2017). Sulfonolipids as novel metabolite markers of Alistipes

595   and Odoribacter affected by high-fat diets. Scientific reports 7, 1-10.

596   Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris,

597   M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., et al. (2012). Human gut microbiome viewed

598   across age and geography. Nature 486, 222-227.

599   Zmora, N., Suez, J., Elinav, E.J.N.r.G., and hepatology (2019). You are what you eat: diet, health

600   and the gut microbiota.   16, 35-56.

601

609   The sponsors had no role in the study design, data collection, and analysis, decision to publish, or

610   preparation of the manuscript.

611

612   **Author contributions:** Zhongtao Gai, Ruixin Zhu, Lei Zhang, Guoping Zhao and Guoqing Zhang

613   conceived and designed the project. Each author has contributed significantly to the work submitted.

614   Hongying Li, Xiangrong Liang, Qian Wang, Meiling Huo, Ying Wang, Lu Zhao, Yongsheng Ge,

615   and Changying Zhao performed the clinical diagnosis, designed and collected the clinical setting,

616   underwent the ethical evaluation process, discussed informed consent and questionnaire data sheets

617   with patients. Ying Wang, Lu Zhao, Yongsheng Ge, and Changying Zhao collected clinical samples

618   and patient information, and consolidated and organized the whole data. Xin Lv completed the

619   clinical examination work and provided the relevant data. Dingfeng Wu and Na Jiao analyzed the

620   data. Ying Wang and Dingfeng Wu drafted the manuscript. Na Jiao, Wenxing Gao, Sheng Gao,

621   Liwen Tao, Ruifang Cao，Yunchao Ling，Lingna Zhao, Haokui Zhou, Guoqing Zhang, Guoping

622   Zhao, Lei Zhang, Ruixin Zhu and Zhongtao Gai revised the manuscript. All the authors read and

623   approved the final manuscript.

624

625   **Declaration of interests**

626   The authors declare no competing interests.

627

628

629   **Table 1. The number of sample and age distribution in this study.**

	16S (n=168)	Metabolism (n=129)	Overlap (n=97)
Normal (n=67)	7.008 ± 0.515 (n=63)	6.250 ± 0.571 (n=45)	6.211 ± 0.629 (n=41)
PP (n=133)	6.641 ± 0.403 (n=105)	6.662 ± 0.458 (n=84)	6.890 ± 0.530 (n=56)
P (t-test)	0.278	0.289	0.108

630   Age Statistics (mean ± 95 confidence interval) and difference analysis (t-test)
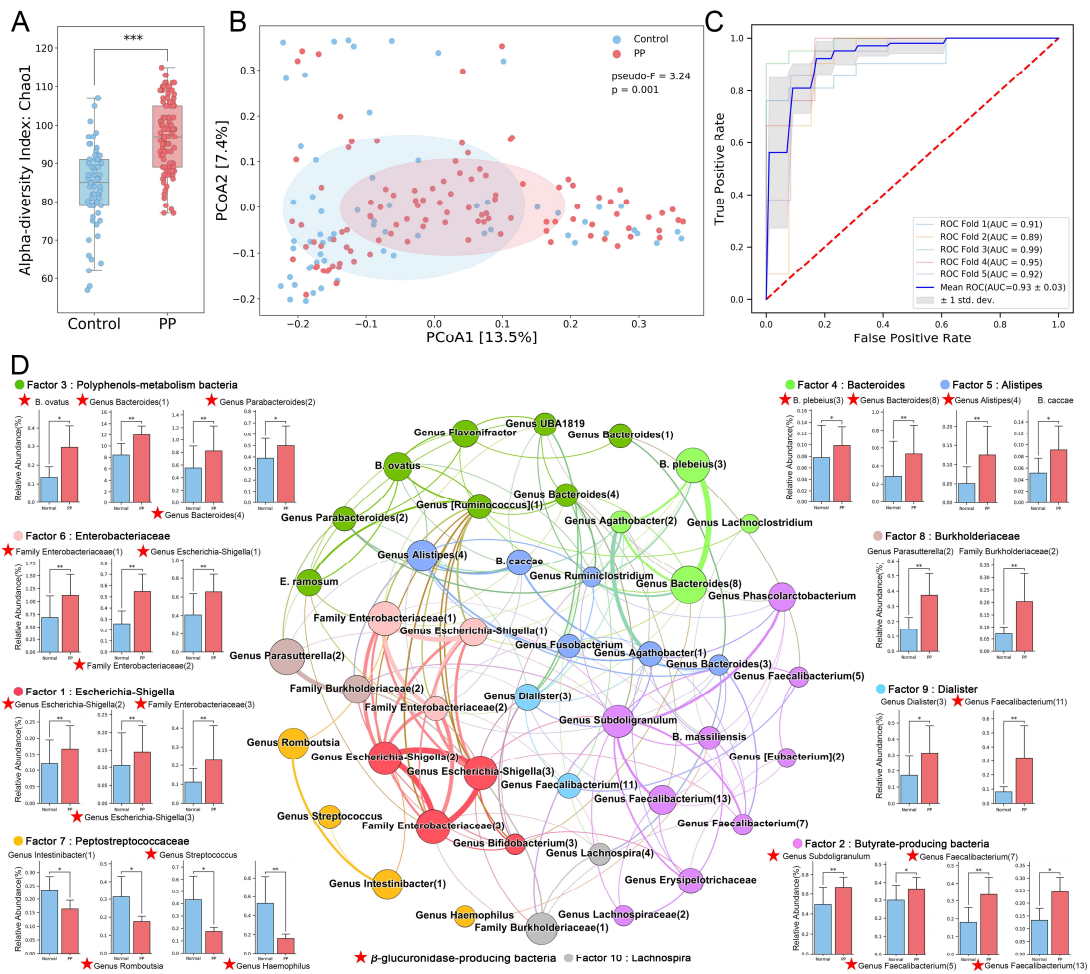
631

632

633

634

635

636

637

638 **Figures**

639

640 **Fig. 1. Gut microbiota dysbiosis in girls with partial PP.** A. The α diversity of gut microbiota based on Chao1 for

641 PP and normal group (***: *P*< 0.001). B. PCoA of bacterial beta diversity based on the Bray-Curtis distance between

642 PP and normal groups. C. The ROC curve of the disease discriminating ability with 45 differential ASVs. D. Co-

643 occurrence network of differential microbiota through Spearman's rank correlation analysis with *P*<0.05. Microbiota

644 are colored by their main latent microbial factor and the strength of correlation is represented by line thickness. The

645 abundance changes of the representative microbiota is displayed as bar plot (*: FDR < 0.05, **: FDR < 0.01). β-

646 glucuronidase-producing bacteria are labeled with a red star.
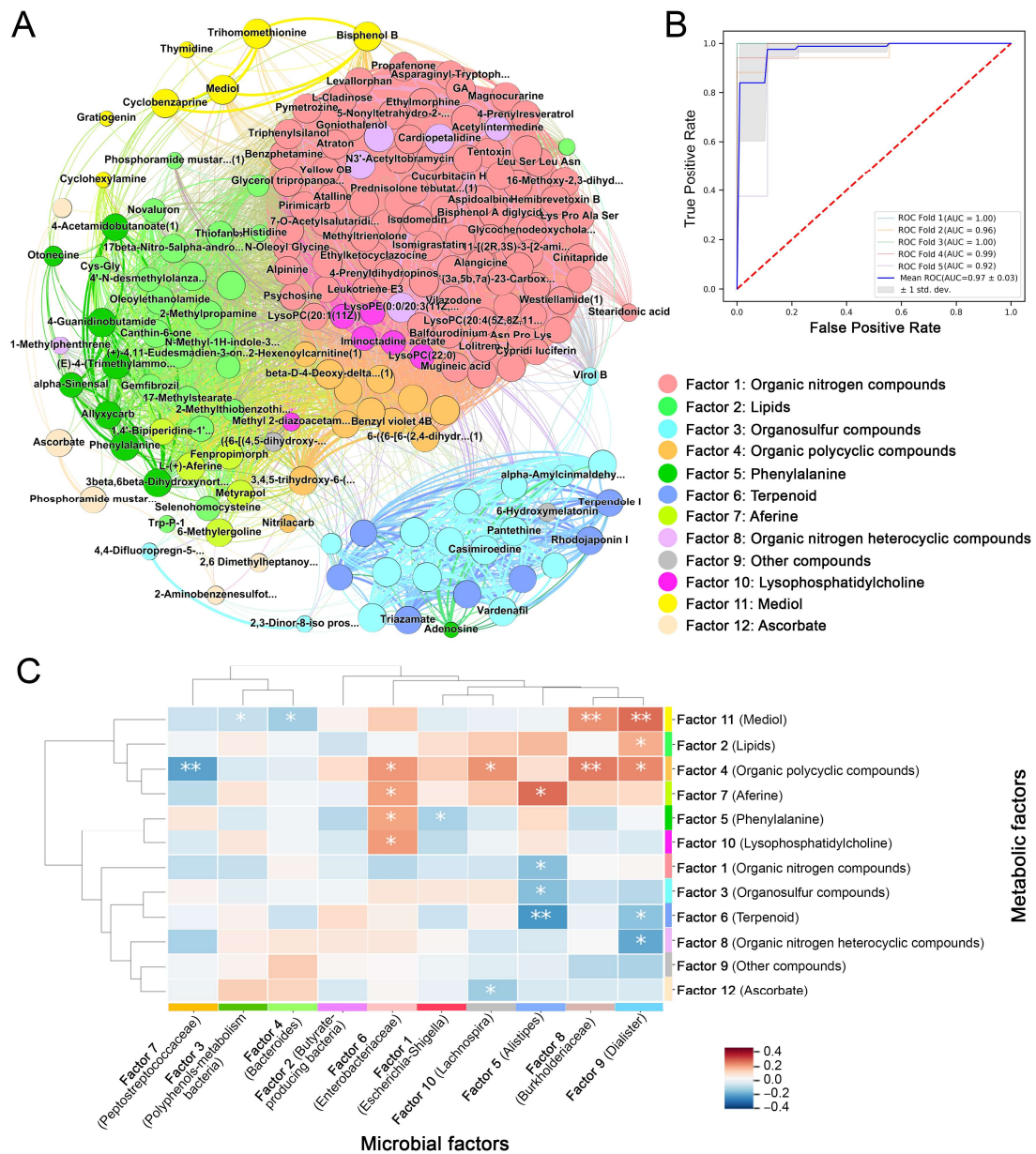
Fig. 2. The change of serum metabolome in girls with partial PP. A. Co-occurrence network of differential metabolites through Spearman's rank correlation analysis with $P<0.05$. Metabolites are colored by their main latent metabolic factor and the strength of correlation is represented by line thickness. B. The ROC curve of the disease discriminating ability with 182 differential metabolites. C. Heatmap plot of Pearson correlation between microbial and metabolic latent factors. (*: $P < 0.05$, **: $P < 0.01$.)
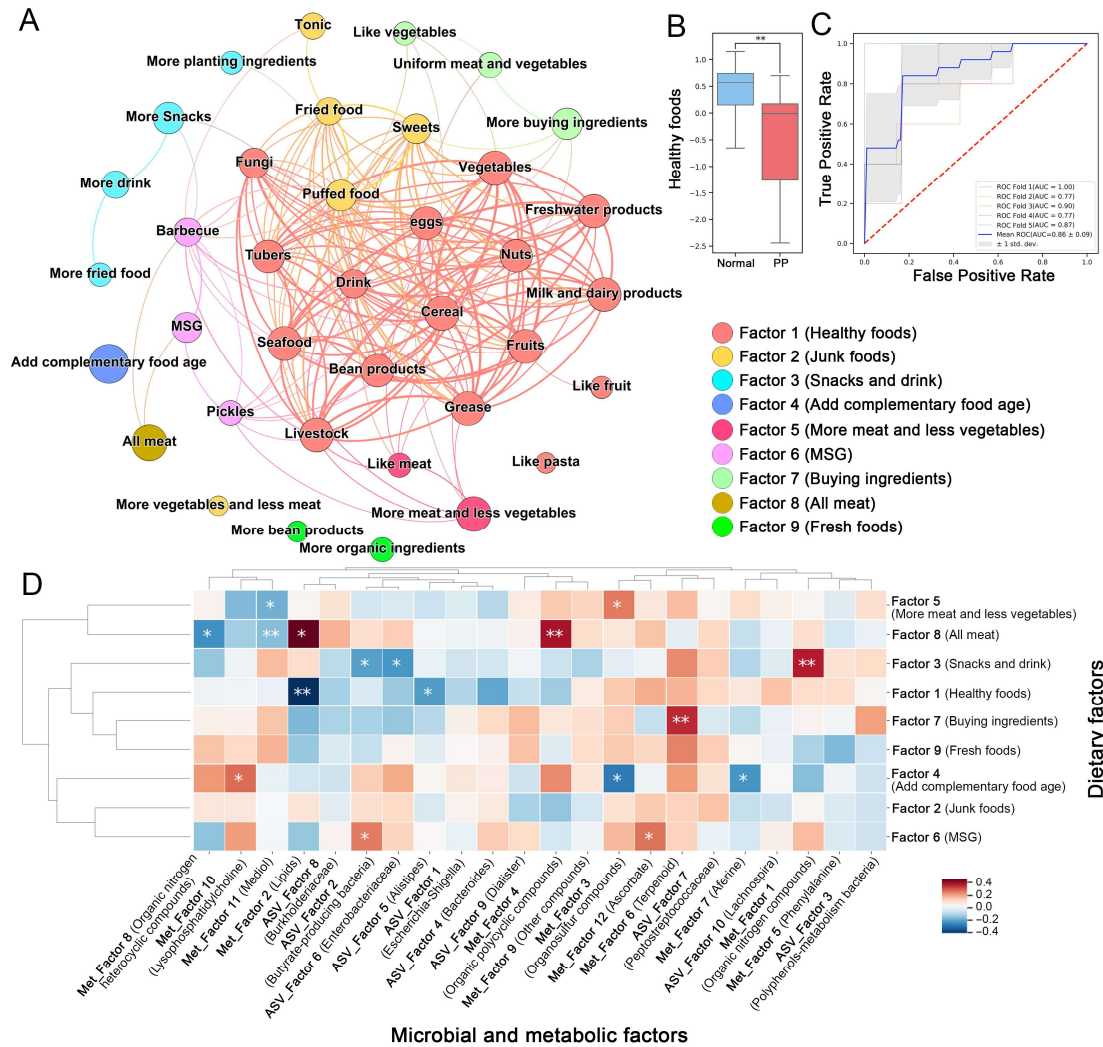
Fig. 3. The change of dietary patterns in girls with partial PP. A. Co-occurrence network of dietary patterns through Spearman's or Kendall's rank correlation analysis with *P*<0.05. Dietary patterns are colored by their main latent metabolic factor and the strength of correlation is represented by line thickness. B. The difference of healthy food (Factor1) intake between normal and PP groups. C. The ROC curve of the disease discriminating ability with 9 latent dietary factors. D. Heatmap plot of Pearson correlation between latent dietary factors and microbial or metabolic latent factors. (*: $P < 0.05$, **: $P < 0.01$.)
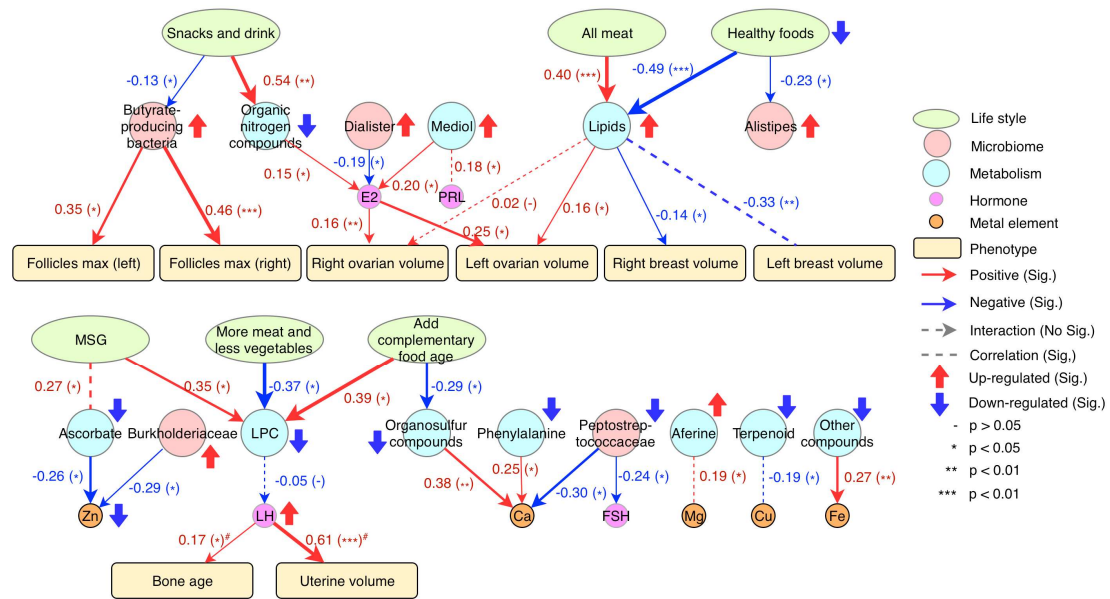
Fig 4. Causal relationships between the latent factors of the dietary, gut microbiota, metabolism, and clinical characteristics. Structural equation model is implemented using piecewise SEM in R. Different types of causal effect are colored by different colors (red: positive, blue: negative) and the strength of effect is represented by line thickness. (-:P>0.05, *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, #: the regression uses glm in R with Poisson distribution).
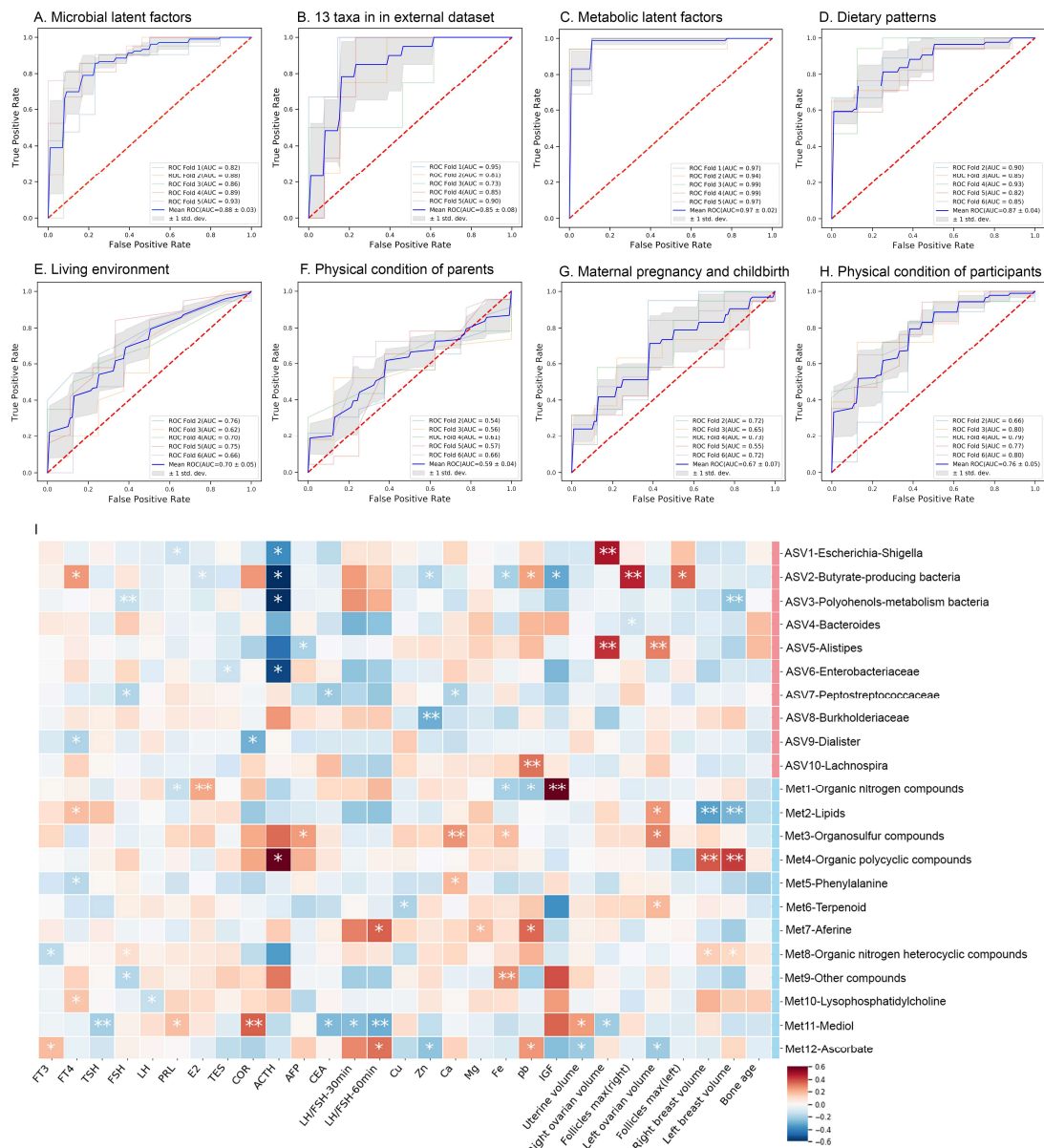
692   **Supplementary materials:**

693
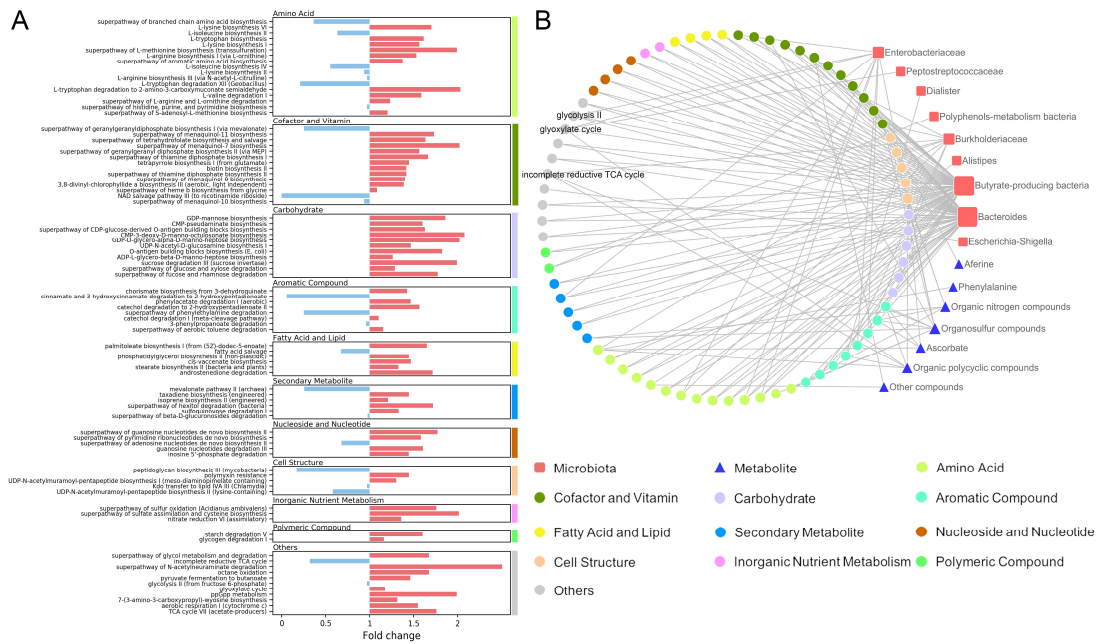
694

695   **Fig. S1. The dysbiosis of metabolic pathways in gut microbiota.** A. The differential abundance changes of gut

696   microbial metabolic pathways. B. The significant correlation links between gut microbiota, serum metabolome, and

697   pathways.

698

Fig. S2. The ROC curve of the disease discriminating ability and Correlation between latent microbial factors, latent metabolic factors, hormones, and phenotypes. The ROC curve of the disease discriminating ability with 10 latent microbial factors (A), 13 taxa in external dataset (B), 12 metabolic latent factors (C), dietary patterns (D), living environment (E), physical condition of parents (F), maternal health and childbirth (G) and physical condition of participants (H). I. Correlation between laten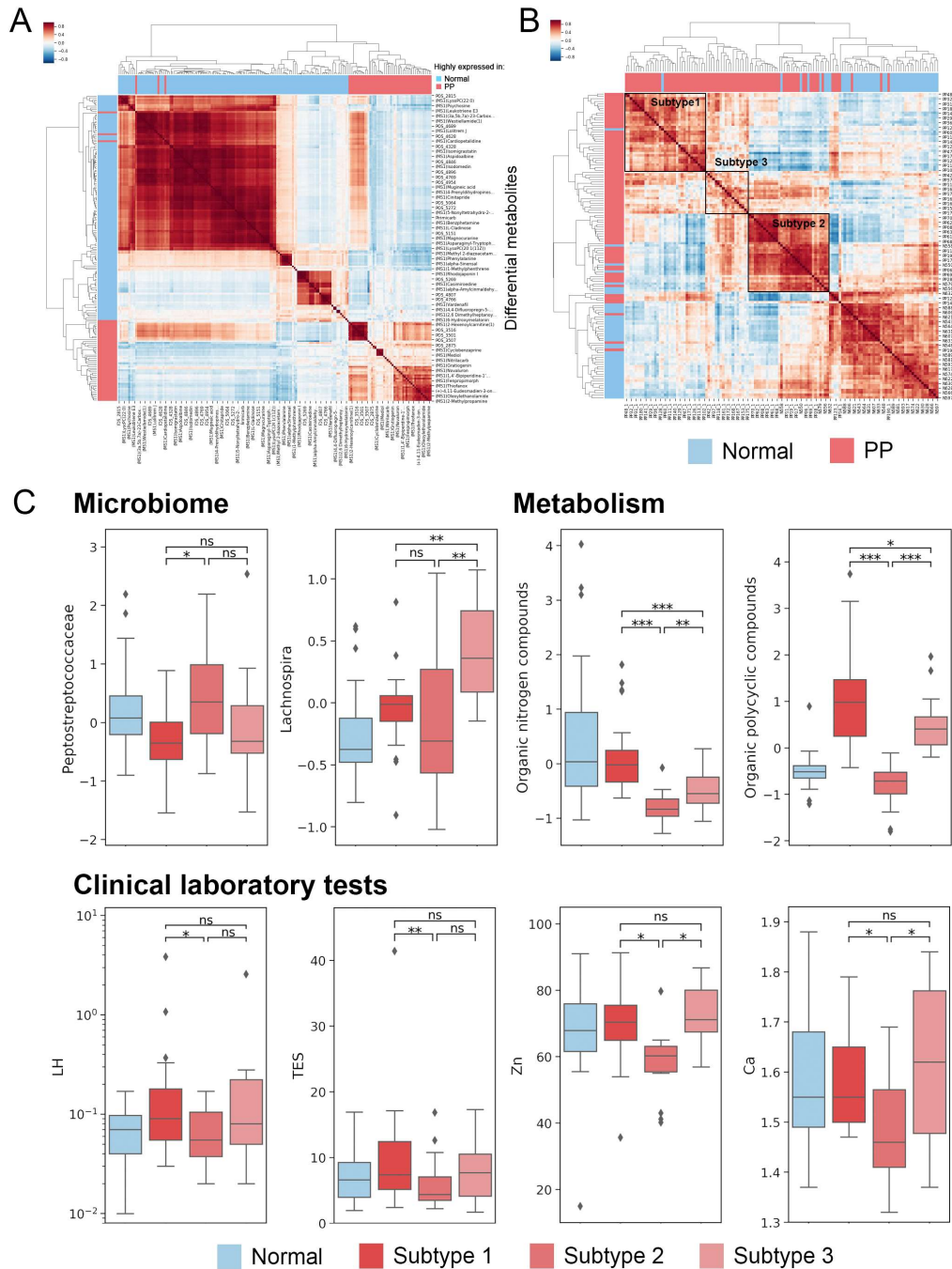t microbial factors, latent metabolic factors, hormones, and phenotypes. All relationships were calculated by Pearson correlation in Python with 1000 permutation test (*:$P < 0.05$, **: $P < 0.01$.)

Fig. S3. Three different subtypes in PP group. A. Pearson correlation between differential metabolites. B. Pearson correlation of samples with differential serum metabolites. Three subtypes were defined through hierarchical clustering algorithm with Euclidean distance. C. The difference between subtypes in gut microbiota, serum metabolome, and clinical laboratory tests. (ns: $P>0.05$, *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$.).

718   **Table S1. Differential life styles between normal and PP group.**

719   **Table S2. Differential gut microbiota between normal and PP group and latent factors.**

720   **Table S3. 13 taxa presented same differential abundance change in external dataset.**

721   **Table S4. Differential serum metabolites between normal and PP group and latent factors.**

722   **Table S5. Gut microbiota, serum metabolome, clinical laboratory tests, and phenotype difference between PP**

723   **subtypes.**

724   **Table S6. Latent factors of life styles.**

725   **Table S7. Differential clinical laboratory variables between normal and PP group.**