Precision environmental health monitoring by longitudinal exposome and multi-omics profiling

Peng Gao[1,3], Xiaotao Shen[1,3], Xinyue Zhang[1], Chao Jiang[2], Sai Zhang[1], Xin Zhou[1], Sophia Miryam Schüssler-Fiorenza Rose[1], Michael Snyder[1,*]

[1]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA

[2]Life Sciences Institute, Zhejiang University, Hangzhou, Zhejiang, China

[3]These authors contributed equally

[*]Corresponding email: mpsnyder@stanford.edu

Abstract

Conventional environmental health studies primarily focus on limited environmental stressors at the population level, which lacks the power of dissecting the complexity and heterogeneity of individualized environmental exposures. Here we integrated deep-profiled longitudinal personal exposome and internal multi-omics to systematically investigate how the exposome shapes an individual's phenome. We annotated thousands of chemical and biological components in the personal exposome cloud and found thousands of internal biomolecules were significantly correlated with the external exposome, which was further cross validated using corresponding clinical data. In particular, our results showed that agrochemicals and fungi predominated in the highly diverse and dynamic personal exposome, and the biomolecules and pathways related to the individual's immune system, kidneys, and liver were highly correlated with the personal external exposome. Overall, our findings demonstrate dynamic interactions between the personal exposome and internal multi-omics and provide insights into the impact of the environmental exposome on precision health.

1

Introduction

26

27 Human health is shaped by the personal genome, microbiome, and exposome[1]. Extensive studies have

28 been conducted on the genome and microbiome, whereas the human exposome is rarely investigated,

29 especially at the individual level. Exposomics aims to characterize all physical, chemical, and biological

30 components collectively in the human internal and external environment. Specifically, the internal

31 environment includes but is not limited to dietary components[2–4], xenobiotics and their biotransformation

32 products, foreign DNA/RNA, and bioactive molecules accumulated from exogenous sources[5]. The

33 external environment consists of all potential exposures from the near-field to the far-field sources of

34 exogenous chemical, biological, and physical exposures[6–9]. Conventional environmental health risk

35 assessments rely on environmental epidemiology within a predefined, usually large, geographical region.

36 However, recent studies revealed that personal exposome profiles are highly dynamic and

37 spatiotemporally different among individuals who live in the same geographical area. For instance,

38 studies have shown that individuals are exposed to significantly different chemical and biological

39 stressors during the same period even if they are in the same geographical region, such as the San

40 Francisco Bay Area or London[10,11]. Another limitation is that previous studies usually targeted a single

41 group of stressors, which failed to provide a holistic picture of the exposome cloud and their

42 interactions[12]. Moreover, stressor-induced physiological responses varied significantly among different

43 individuals[11]. Therefore, there is a critical need to monitor exposures at the individual level and

44 systematically integrate them with respective internal multi-omics profiles to fully characterize each

45 individual's personal responses to environmental exposures.

46

47 Multi-omics analyses enable a detailed investigation into the biological mechanisms underlying human

48 phenotypes by integrating multiple omics, such as proteomics, metabolomics, and microbiomics[13]. Multi-

49 omics profiling, together with clinical measures such as cytokines and blood tests, can comprehensively

50 assess one's health status and detect significantly correlated exposures to understand the impact of the

51 external exposome on human biology and health[10,14,15]. In addition, longitudinal profiling can avoid biases

2

52 introduced by one-time sampling and provide a molecular portrait of the effect of different exposures at

53 an individual level.

54

55 In this study, we utilized our previously published datasets to integrate thousands of longitudinally

56 measured chemical and biological components along with physical factors in the personal exposome to

57 investigate how the exposome impacted internal -omes, such as the proteome, metabolome, the gut

58 microbiome as well as cytokines and blood markers[10,14,15]. In addition, toxins and carcinogens, cytokines,

59 and blood tests were obtained or annotated from the blood samples to understand the impact of the

60 exposome on precision environmental health. Specifically, this study 1) improved the annotation of

61 biological and chemical exposures in the external exposome and human blood; 2) integrated the external

62 exposome with internal multi-omics to investigate the effect of the exposome on molecular phenotypes

63 and pathways; and 3) correlated the environmental stressors with clinical measurements to associate the

64 health effects of the external exposome.

65

66 **Results**

67 **Longitudinal profiling of the exposome and internal multi-omics to monitor personal**

68 **environmental health.** We investigated whether the external exposome is related to the internal

69 molecular and physiology profile at a comprehensive and personal level using the schematic shown in

70 **Figure 1a**. We reanalyzed deep biological and chemical exposome data collected from our previously

71 published study in which an individual had been continuously wearing a personal exposome collection

72 device "exposometer" and correlated it with the internal molecular profiles. Over the 52-day period

73 relevant for this study, the device captured organic chemicals using zeolite, followed by methanol elution

74 and liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) analysis.

75 Biological specimens were also captured using polyethersulfone filters and nucleic acids analyzed by high

76 throughput sequencing of DNA and RNA. (**Supplementary Data 1**). General environmental factors (e.g.,

77 temperature, humidity, total particulate matter) were recorded by the device, and the other environmental

3

78   factors were also obtained from the local air quality monitoring stations (**Figure 1a**). Contrary to

79   conventional exposome monitoring studies, which usually focused on the exposures at a single

80   timepoint[16–18], we captured personal exposome profiles across 18 timepoints, and annotated 1,265 genera,

81   158 known chemical stressors among 3,299 chemical features, and 10 environmental factors which

82   include physical stressors that may impact environmental health in this study (**Figure 1b**). These genera

83   and known chemical stressors were annotated from the improved microbiome and chemical annotation

84   pipelines that we developed (**Methods**).

85

86   Over the same 52 days period, we also collected stool and blood samples from the same participant to

87   profile the gut microbiome, proteome, metabolome, toxins and carcinogens, cytokines, and blood tests

88   (**Figure 1c** and **Supplementary Data 1**). Through reanalysis pipelines, we were able to annotate 60

89   toxins and carcinogens as well as 664 metabolites, 302 proteins, and 62 gut microbiome taxa. We also

90   measured 62 cytokines and 46 clinical blood parameters to longitudinally monitor personal health status[15].

91   All sample collections were performed during the first quarter of 2016 from three distinct locations in the

92   U.S. (**Figure S1**). However, not all sample types were collected at each time point, and the inter-omics

93   analyses were performed only when overlaps were available (**Figure 1b**). Despite our limited ability to

94   control all confounding variables, we searched for significant intra- and inter- exposome correlations and

95   high-degree components which have the most significant correlations in each analysis as those may play

96   important roles in the exposome-ome interactions [$|r| > 0.9$; False Discovery Rate (FDR) adjusted p-value

97   (q-value) < 0.05].

98

99   **Intra-exposome relationships in the highly dynamic and diverse personal exposome cloud.** To

100   annotate as many chemicals as possible, we searched through the 3,299 LC-HRMS raw features using a

101   combination of five public exposome related databases as well as an in house database that we assembled.

102   Using this new annotation pipeline, we were able to annotate 158 known stressors (**Figure 2a, Methods**).

103   These stressors were categorized into 13 classes, with the dominant class being agrochemicals, followed

4

104   by pharmaceuticals & personal care products (PPCPs), plasticizers, and International Agency for

105   Research on Cancer (IARC) Group 2A carcinogens, and the chemicals in each class varied dynamically

106   during the monitoring period (**Figure 2a, Figure S2** and **Supplementary Data 2**). To characterize the

107   biological exposome domain, we circumvented the limited ability of 16S rRNA/18rRNA/ITS sequencing

108   by applying metagenomic sequencing. We found 17 genera dominated during the study period, most of

109   which were fungi and bacteria, but they varied dynamically (**Figure 2b**). Ten general environmental

110   factors, measured either by personal exposometer (temperature, humidity, and total particulate matter) or

111   local air monitoring stations (atmospheric pressure, wind speed, SO_2, NO_2, O_3, CO, and air quality index),

112   were also included in the study (**Figure 2c**).

113

114   We performed intra-omics correlation analyses to investigate the potential relationships among all the

115   exposome components (**Methods, Supplementary Data 3**). We found a total of 60 statistically

116   significant correlations ($|r| > 0.9$ and q-value < 0.05) among 74 exposome components, including 41

117   chemicals, 30 genera, and 3 environmental factors (**Figure 2d** and **Figure S3**). Specifically, diisononyl

118   phthalate (a plasticizer) and butylated triphenyl (an organophosphate flame retardant) had the most

119   significant correlations, followed by various agrochemicals, PPCPs, and IARC group 2A carcinogen.

120   Among the biological components, *Tricholoma* had the highest number of significant correlations,

121   followed by *Cylindrobasidium*, *Piriformospora*, *Erysiphe*, *Schizophyllum*, *Serendipita*, and *Hirsutella*, all

122   of which are fungi (**Figure 2e**). In terms of environmental factors, only temperature and humidity

123   collected by the exposometer as well as SO_2 concentration collected by the local monitoring stations were

124   significantly correlated with other exposome components (**Figure 2d**). For example, *Paenibacillus* was

125   positively correlated with the temperature, consistent with the literature that members of *Paenibacillus* are

126   heat resistant and grow well in relative hot temperatures[19]. Azoxystrobin, ethylparaben and captan are

127   fungicides or antifungal agents[20–22] that negatively correlated with different fungi (**Figure 2f**).

128

129 **Inter-omics analyses between the exposome and multi-omics revealed physiological links to the**

130 **exposome.** To investigate how the exposome shapes one's phenome longitudinally, we investigated the

131 links between the exposome and internal multi-omics. Specifically, we found 8,986 significant

132 correlations (|r| > 0.9 and q-value < 0.05) among 1,700 factors from all -omes, and positive correlations

133 were more predominant than negative correlations (**Figure 3a, Supplementary Data 4**). Biological

134 exposome and metabolome were the most extensive -omes in the network, and they also had the greatest

135 number of significant correlations (N = 4148; **Figures 3a and 3b**). Additionally, we found that the

136 exposome and internal multi-omics networks can be divided into several subnetworks with high

137 modularity (0.819, **Figure S4a, b**).

138

139 **Personal exposome-gut microbiome interactions.** We found 1,333 significant correlations (|r| > 0.9 and

140 q-value < 0.05) between the exposome and the gut microbiome (16S rDNA data), and the number of

141 positive and negative correlations were approximately equal (**Figure 3c**). Specifically, the six highest

142 degree bacteria (each correlates with 34 exposome components) may be involved in multiple

143 physiological processes that respond to the personal exposome. For example, members from *Alistipes*

144 were shown to play essential roles in inflammation and various diseases[23], members from *Eggerthella*

145 were implicated as the causes of liver and anal abscesses, ulcerative colitis, and systemic bacteremia[24],

146 members from *Odoribacter* were found to maintain short-chain fatty acid availability and systolic blood

147 pressure[25], members from *Parasutterella* were involved in bile acid maintenance and cholesterol

148 metabolism[26], whereas members from *Roseburia* played vital roles in producing short-chain fatty acids

149 and anti-inflammatory pathways[27]. Out of the top six genera, all but *Roseburia* positively correlated with

150 chemical stressors and usually negatively correlated with biological components (**Figure 3d and S4d**).

151 As a result, members from *Alistipes*, *Eggerthella*, *Odoribacter*, and *Parasutterella* were more likely to be

152 involved in proinflammatory processes, while members from *Roseburia* were mainly involved in anti-

153 inflammatory processes. On the exposome side, *Botryosphaeria*, *Corynespora*, and *Enterobacter* were the

6

154   highest degree genera (each correlated with 18 gut bacteria) among all exposome components, indicating

155   their essential roles in interacting with the participant's gut microbiome (**Figure 3c**).

156

157   **Exposome-proteome interaction network.** We found 2,054 statistically significant correlations ($|r| > 0.9$

158   and q-value < 0.05) between the individual's exposome and internal proteome. Most of the high-degree

159   exposome components were biological components, and positive correlations were slightly more frequent

160   than negative correlations (**Figure 4a**). Specifically, we found 11 highest degree substances (nine genera

161   and two chemicals), each of which was significantly correlated with more than 22 proteins in the

162   proteome. The high-degree biological genera were fungi and primarily positively correlated with proteins;

163   in contrast, *Xeromyces* negatively correlated with proteins. Fenazaquin (a pesticide) and

164   tetrabromobisphenol A diallyl ether (a brominated flame retardant) were two high-degree chemical

165   stressors, both of which primarily negatively correlated with proteins. On the proteome side, 17 highest

166   degree proteins (each correlated with 21 exposome components) were discovered, and 14 of them were

167   directly immune-related. For instance, alpha-1-HS-glycoprotein (AHSG) promotes endocytosis,

168   complement component 3 (C3) activates the complement system, and fibrinogen alpha chain (FGA) is

169   involved in both innate and T-cell mediated pathways[28]. Additionally, we discovered significantly

170   correlated signaling pathways when queried against GO, KEGG, and Reactome databases

171   (**Supplementary Data 5 and 6**). Chemical and biological exposome shared several significantly

172   correlated pathways, such as protein activation cascade, platelet degranulation and acute inflammatory

173   response, whereas some pathways were uniquely correlated with the chemical exposome, such as platelet

174   activation, signaling, and aggregation pathway (**Figure 4b**). Moreover, immune-related pathways were

175   among the most common high-degree signaling pathways correlating with chemical and biological

176   exposome, and those pathways were about half positively and half negatively correlated with the

177   exposome (**Figure 4c**).

178

179   **Exposome-metabolome interaction network.** The blood metabolome is considered the most interactive

180   -ome with the exposome since xenobiotics interact with endogenous metabolites initially after entering

181   the human body. In fact, the blood exposome overlaps with the blood metabolome from an analytical

182   perspective as current approaches cannot distinguish the sources of the molecules present in the blood.

183   Moreover, xenobiotic biotransformation is similar to that of metabolic pathways and can even involve the

184   same enzymes, such as cytochromes P450[16,29]. Therefore, it is essential to investigate the interactions

185   between the exposome and metabolome to better understand the initial health impact of the exposome.

186

187   In this study, we found 4,624 statistically significant correlations ($|r| > 0.9$ and q-value < 0.05) between

188   the exposome and internal metabolome. Specifically, positive correlations were more frequent in the

189   exposome-metabolome analysis than the exposome-proteome analysis (**Figure 5a and S5**). The high-

190   degree biological components were primarily fungi and usually positively correlated with the metabolites;

191   interesting exceptions are *Aegilops* (a grass), the bacteria *Pontibacter* and *Hymenobacter,* and

192   *Paramecium* (a ciliated protist). Salicylic acid (a PPCP), dinoseb (an herbicide), dibromoethane (an IARC

193   group 2A carcinogen) were the three highest degree chemicals, all of which primarily positively

194   correlated with endogenous metabolites. Importantly, we found 19 high-degree metabolites, each

195   significantly correlated with 21 exposome substances. Several metabolic pathways were significantly

196   correlated with both the chemical and biological exposome (**Methods**, **Figure S7**), such as protein

197   digestion and absorption and aminoacyl-tRNA biosynthesis, whereas some pathways were only correlated

198   with the biological exposome (**Figure 5b, Supplementary Data 7 and 8**). Similar to the exposome-

199   proteome analysis, we performed correlation network analysis among the exposome, metabolites, and

200   metabolic pathways. Trimethyl phosphate (a plasticizer and organophosphate flame retardant),

201   2,2',3,3',4,4',5-Heptachlorobiphenyl (a polychlorinated biphenyl), and tetrachloroethylene (an IARC

202   group 2A carcinogen) were positively correlated with all the metabolic pathways, whereas

203   tetrabromobisphenol A diallyl ether, salicylic acid, and zeranol (a mycotoxin) were negatively correlated

204   with all metabolic pathways (**Figure 5c**).

205

Monitoring precision environmental health by investigating the exposome-clinical data

correlations. Standard clinical measurements such as blood and cytokine tests directly reflect the

individual's health. Thus, clinical test results are ideal indicators to investigate the health impact of the

exposome. Based on our exposome-cytokine analysis, biological exposome had the most significant

correlations with cytokines, followed by chemical and environmental factors. 362 significant correlations

($|r| > 0.9$ and q-value < 0.05) were found between the exposome and cytokines, most of which were

positive correlations. After converting correlation coefficients to variable importance in projection scores,

we determined the contributions of all significantly correlated exposome components on cytokines

(**Supplementary Data 10**, **Methods**). Specifically, 60% of the cytokine variation was explained by the

determined factors in this study. Furthermore, the top 13 cytokines, which were almost entirely

contributed by exposome components ($> 90\%$), were all proinflammatory cytokines, such as IL-23, MCP-

1, and VCAM-1, indicating that those cytokines may play essential roles in response to the exposome.

Additionally, 14 highest degree (each correlates with > 7 exposome components) cytokines were found to

be primarily positively correlated with the exposome, whereas only MCP-1 was primarily negatively

correlated (**Figure 6a**). Most high-degree biological components were fungi, such as *Wallemia*, which are

filamentous food-borne pathogens[30]. Moreover, other than *Xeromyces,* most of the exposome

components were primarily positively correlated with cytokines, consistent with the exposome-proteome

analysis where *Xeromyces* primarily negatively correlated the proteins.

224

Similar to the exposome-cytokine analysis, biological components had the most significant correlation

with blood tests, followed by chemicals and environmental factors. However, fewer chemicals were

correlated with blood tests than those correlated with cytokines (**Figure 6a**). 513 significant correlations

were found between the exposome and blood tests, and the majority were positive correlations. Using

similar contribution determination algorithms, 77% of the blood tests variation was explained by the

determined factors in this study. Similarly, the top 13 blood tests which were almost entirely contributed

231   by the determined exposome components (contributions of the exposome > 95%), were primarily related

232   to the immune system, liver, and kidney functions. Additionally, eight highest degree (each correlated

233   with > 25 exposome components) blood tests were primarily positively correlated with the exposome

234   while only platelet was primarily negatively correlated. Interestingly, the highest degree blood test,

235   creatinine, correlated with 62 exposome components, which is a biomarker for kidney function (**Figure**

236   **6b**). Unlike cytokine profiles, where we cannot draw a clear line of the individual's health status, blood

237   tests have clinically established reference ranges facilitating the differentiation. We therefore performed

238   correlation analyses to understand the effects of the exposome on personal health using blood test results

239   with out-of-range values. Interestingly, we found the abnormal blood glucose level values were

240   significantly correlated with 3 chemical stressors and 13 microbes. For instance, salicylic acid

241   concentration was negatively correlated with glucose level, and salicylic acid has been shown to decrease

242   glucose concentration and used as a treatment for type 2 diabetic patients[31], which is consistent with our

243   findings. Similarly, abnormal values of absolute eosinophils and urea nitrogen correlated with specific

244   biological exposome components (**Figure 6b**). Previous studies were consistent with our results as

245   eosinophilic phagocytosis consumes eosinophils during allergy and inflammation[32], and various

246   pathogenic microbes can utilize urea as a nitrogen source[33].

247

248   **Discussion**

249   It has long been acknowledged that environmental factors affect one's health, but conventional

250   environmental health studies face limitations. For example, a) population or cohort studies overlook the

251   significant differences between individuals; b) single timepoint sampling fails to reflect the continuous

252   effects of stressors; c) and focusing on a single or a class of stressors does not capture the holistic health

253   impact of the exposome.

254

255   To overcome these challenges, we generated a more comprehensive precision environmental health

256   profile by longitudinally monitoring both the personal exposome and internal multi-omic profiles (**Figure**

10

257 **1**). In addition, we also measured standard clinical indices to investigate the health effects of the

258 exposome. Using Spearman correlation analysis, we discovered many significant correlated physiological

259 parameters and exposome components, indicating their interactions in the participant's responses to the

260 personal exposome. Additionally, our study provided vast testable hypotheses to further investigate the

261 underlying mechanisms using analytical and experimental approaches.

262

263 We were able to capture more than chemical 3,000 features, but only annotated 158 known chemical

264 stressors by a broad annotation method that utilizes various databases, including those containing

265 emerging contaminants[34–37]. This indicates that existing exposome databases still lack the power to

266 annotate the majority of the chemical exposome. Interestingly, we found that the concentrations of most

267 chemicals increased after January 25, 2016, when the individual transitioned from a period of residing at

268 home to a period of high travel, indicating that the chemical exposome greatly increased with travel to

269 other locations (**Figure S2f**). Agrochemicals had the highest concentrations among all annotated

270 chemicals, indicating their ubiquitous presence in the environment. An alternative view is that

271 agrochemicals are the most frequently studied chemical stressors, making them most easily identifiable. It

272 is also worth noting that high concentrations do not necessarily imply high health risks since each

273 chemical has its own safe dose, and the combined effects among them are still unclear[13].

274

275 The biological exposome revealed a number of interesting observations as well. The fungal genus

276 *Stereum* was dominant at most time points, reflecting its high abundance in the personal exposome

277 (**Figure 2a**)[38]. We found several interesting correlations of the biological exposome with chemical and

278 environmental factors, such as associations of antifungals with a decrease in fungal exposures. Many of

279 these are intended food or soil antifungal products (azoxystrobin and captan), whereas others are common

280 preservatives (ethylparaben and anthracene) that have antifungal properties[10]. Overall, nearly 100

281 significant correlations were found by intra-exposome analysis, representing the complex interactions

282 within the exposome domains. Importantly, the negative correlations of fungi with various pesticides and

11

283 herbicides indicate these agrochemicals may inhibit the fungi growth as well (e.g., *Tricholoma* versus

284 propoxur and *Erysiphe* versus bentazone). Moreover, we find several interesting tertiary relationships,

285 such as the mycotoxin fusarin C (produced by *Fusarium*) negatively correlated with *Cylindrobasidium,*

286 suggesting a possible competition among the different fungi (**Supplementary Data 3**).

287

288 A recent study identified radioprotective gut microbes and internal metabolites in mice using a multi-

289 omics analysis[39], demonstrating the potential of this approach to investigate essential components in the

290 internal -omes that respond to the external environment. To this end, we performed inter-omics analyses

291 between the exposome and gut microbiome, proteome, and metabolome, respectively. By discovering

292 high-degree components in each analysis, we identified the critical components in the exposome-internal

293 omes interactions. For instance, we found six highest degree gut bacteria that may be important in the

294 responses to the personal exposome. The high-degree biological components in both exposome-proteome

295 and exposome-metabolome analyses were mainly fungi, yet few had known human health effects.

296 However, we identified major high-degree annotated chemicals that are known human stressors; for

297 instance, the herbicide dinoseb exposure causes various developmental toxicities and loss of thyroid and

298 body weight[40], and brominated flame retardants like tetrabromobisphenol A diallyl ether are known

299 neurotoxicants[41]. In addition to the endogenous metabolites, we were able to annotate 60 toxins and

300 carcinogens in the individual's blood based on the exposome related databases (**Figure 1c**). Unlike the

301 annotated xenobiotics in the exposome samples, most of the chemical stressors annotated in the blood

302 were food and animal toxins. Furthermore, only 11 chemical stressors were annotated in both the external

303 and blood exposome. This is likely to be partly due to the limited power of the current databases, since

304 most of the databases only contain the information of parent chemicals but not their biotransformation

305 products. Additionally, persistent hydrophobic substances tend to accumulate in adipose tissues, but not

306 in the blood that we profiled, while non-persistent hydrophilic chemicals are efficiently excreted out of

307 the human body[13,16], limiting their identification. Finally, the bioavailability of chemicals in different

308 external matrices also limits the exposure, dose-response, and concentration of bioavailable fraction[36,42,43].

12

309

310 Importantly, our results indicate that the immune system, kidneys, and liver may play essential roles in

311 response to the exposome, which are all known to regulate and respond to foreign substances[17,44,45]. In the

312 exposome-proteome analysis, we found 14 out of 17 high-degree proteins were involved in immune

313 responses (e.g., complement component 3, interleukin-1 receptor accessory protein, and immunoglobulin

314 heavy chain proteins)[45], and immune-related pathways (e.g., acute inflammatory response, humoral

315 immune response, and complement and coagulation cascades) and were among the highest degree

316 signaling pathways. In the exposome-metabolome analysis, we identified 19 highest degree metabolites

317 related to protein metabolism, inflammation, kidney and liver functions (e.g., L-arginine, nutriacholic

318 acid, epsilon-(gamma-Glutamyl)-lysine, and uracil), indicating that these metabolic pathways were

319 involved in responses to the exposome. Moreover, certain high-degree metabolic pathways are both

320 protein and immune-related pathways, such as alanine aspartate and glutamate metabolism, protein

321 digestion and absorption, and beta-alanine metabolism. Specifically, particular protein metabolisms (e.g.,

322 amino acids synthesis and protein breakdown) were highly sensitive to oxidative stresses caused by the

323 exposome components[46], inflammation is often the first immunological response to foreign substances,

324 and kidneys and liver are the main detoxification organs[46], with liver and bile acids serving essential roles

325 in responding to the foreign substances[47]. To further investigate the health effects of the exposome, we

326 performed Spearman correlation analysis with cytokines and blood test results. It is important to note that

327 we do not identify many exposome factors, such as dermal and ingestion exposures, inorganic chemical

328 components[48–51], psychosocial stressors, as well as personal lifestyle, which may affect the clinical

329 measurements as well (**Figure 6a**). Proinflammatory cytokines were the most significantly correlated

330 with the external exposome components (e.g., IL-23, MCP-1, and IL-2), and they have been previously

331 shown to be elevated after exposure to external stressors[52]. Our blood test results provided further

332 evidence as fluctuations of creatinine and urea nitrogen, biomarkers of kidney and liver functions,

333 respectively, were correlated with specific exposome components. As a result, exposome-proteome

334 analysis cross-validated with exposome-cytokine analysis, indicating that the proinflammatory processes

335 play essential roles in responding to the exposome, and the exposome-metabolome analysis cross-

336 validated with exposome-blood tests analysis, showing that liver and kidneys play significant roles in

337 responding to the exposome. Furthermore, the exposome-microbiome analysis showed that the highest

338 degree gut bacteria are related to the proinflammatory processes and liver metabolism. Therefore, these

339 physiological processes and organs may be ideal candidates for testing the combined effects of multiple

340 stressors in future studies. On the exposome side, high-degree exposome components that overlapped in

341 more than one inter-omics analysis are significant health concerns. Specifically, *Isaria*, *Sporothrix*, and

342 *Tarenaya* were among the highest-degree microbes correlated with all internal -omes. Members of these

343 genera were found to be involved in complex physiological mechanisms and may exhibit adverse health

344 effects. For example, species of *Isaria* were found to induce cell death[53], species of *Sporothrix* triggered

345 skin and lung inflammatory reactions[54], and the pollen of *Tarenaya* members are allergens and generate

346 immunoglobulin-E mediated allergic reactions[55]. Also, certain types of agrochemicals, PPCPs, and flame

347 retardants that we detected are high-degree chemical stressors known to cause epigenetic alterations,

348 endocrine disruption, impaired nervous system function, oxidative stress, and inflammation[46]. Therefore,

349 those substances would be ideal candidates for investigating the underlying mechanisms of their

350 combined health effects in future studies.

351

352 In conclusion, we found both time and location impacted the personal exposome, especially the biological

353 components and environmental factors (**Figure S2f**), and the biological and chemical exposome was

354 highly dynamic. These results emphasize the significance of individual precision environmental health

355 over traditional environmental epidemiology studies. Undoubtedly, due to the limitation of the single

356 participant and annotation power in this study, future research should monitor the precision environmental

357 health of more individuals and increase the annotation confidence of the chemical and biological

358 components in the exposome. Nonetheless, this study demonstrates the power of using a holistic approach

359 of monitoring the exposome on personal environmental health using inter-omics analyses, and serves as a

360 useful model to scale to the other individuals and locations. Our study also identified high-degree

14

361   components as essential components among the exposome-internal omes interactions and provided

362   abundant testable hypotheses to further investigate their underlying mechanisms of impacting individual

363   health.

364

365 **Materials and Methods**

366 **Personal exposome and internal multi-omics samples collection**

367 **Exposome samples collection.** The participant in the study is enrolled under Stanford University's IRB

368 protocols IRB-23602 and IRB-34907. Modified RTI MicroPEM V3.2 personal exposure monitor (RTI

369 International, Research Triangle Park, North Carolina, USA), termed exposometer, was used to collect

370 chemical and biological components exposed by individuals from January 2016 to March 2016. Also,

371 temperature, humidity, and total particulate matter were simultaneously collected by the exposometer in a

372 real-time manner. The original sequential oiled frit impactor was removed to maximize the collection of

373 biological components. A 0.8 mm pore-size polyethersulfone with a diameter of 25 mm filter (Sterlitech,

374 Kent, Washington, USA) was placed in filter cassettes to collect particulates for DNA and RNA extraction.

375 An in-house designed, 3D printed cartridge was placed at the end of the airflow, which contained 200 mg

376 of zeolite adsorbent beads (Sigma 2-0304) to collect chemicals. Before deployment to the participant, the

377 MicroPEM was calibrated to a flow rate of 0.5 L/min ($\pm$ 5%) using a mass flow meter (TSI 4140, Shoreview,

378 Minnesota, USA). During the study, the participant was instructed to place the exposometer on his arm or

379 within a radius of 2 meters. Samples were collected after 1 to 3 days of use and stored at -80 °C until

380 analysis. To minimize the potential contamination, filters and related components were handled in sterile

381 biological safety cabinets and cleaned by ethanol before use. Clean polyethersulfone filter and zeolite

382 adsorbent beads were included before extraction as background controls. MicroPEM log files were

383 downloaded using Docking Station software (RTI International, Research Triangle Park, North Carolina,

384 USA). The participant used the MOVES App to track geographic locations through GPS coordinates and

385 daily activities[10]. General environmental data were collected from the exposometer or National Oceanic

386 and Atmospheric Administration's National Climatic Data Center or National Centers for Environmental

387 Information.

388

389 **Analysis of chemical exposome by LC-HRMS.** LC-HRMS was performed on a platform composed of a

390 Waters UPLC coupled with Exactive Orbitrap mass spectrometer (Thermo, Waltham, Massachusetts, USA)

16

391   using a mixed-mode OPD2 HP-4B column (4.6 mm x 50 mm) with a guard column (4.6 mm x 10 mm;

392   Shodex, Showa Denko, Tokyo, Japan). The column temperature was maintained at 45 °C and the sample

393   chamber at 4 °C. The binary mobile phase solvents were: A, 10 mM ammonium acetate in 50:50 (vol/vol)

394   acetonitrile: water; B, 10 mM ammonium acetate in 90:10 (vol/vol) acetonitrile: water. Both solvents were

395   modified with 10 mM acetic acid (pH 4.75) for positive mode acquisition or 10 mM ammonium acetate

396   (pH 9.25) for negative mode. The flow rate was set as follows: flow rate, 0.1 ml/min; gradient, 0–15 min,

397   99% A, 15–18 min, 99% to 1% A; 18–24 min, 1% A; 24–25 min, 1% to 99% A; 25–30 min, 99% A. The

398   MS acquisition was set as full scan mode with an ESI probe or APCI probe. The capillary temperature was

399   275 °C, the sheath gas was 40 units, the positive mode spray voltage was 3.5 kV, and 3.1 kV for the negative

400   mode. The capillary voltage was 30 V, the tube lens voltage was 120 V, and the skimmer voltage was 20

401   V. The mass spectrum scan used 100,000 mass resolution, high dynamic range for AGC target, maximum

402   injection time of 500 ms, and a scan range of 70-1,000 m/z. The details of quality assurance and quality

403   control were described in the previous study[53].

404

405   **Post-acquisition analysis of the chemical exposome.** Analysis of chemical exposome was performed as

406   previously described[10,53]. In brief, feature detection was performed with XCMS. For a conservative

407   assessment of the number of unique chemical features, a customized Python script was used to remove

408   potential isoforms, isotopes, and adducts from the 3,299 features enriched at least 10-fold as compared with

409   the blank control. The annotation was based on various exposome related databases that are publicly

410   available as well as in-house databases[54–56]. The annotation confidence levels of all the chemicals in the

411   exposome were at least level 5, with at least one chemical at level 1 or 2 in each chemical class[53,57].

412

413   **Sequencing and analysis of biological exposome.** DNA and RNA sequencing and analysis were

414   performed as previously described[10,53]. In brief, DNA and RNA were extracted from filters and linearly

415   amplified for sequencing. Libraries were sequenced by Illumina HighSeq 4000 platform (2 x 151bp) that

416   yields an average of ~50 M unique reads per sample. Sequenced reads were deduplicated, and adapters

417 were trimmed using Trim Galore! (version 0.4.4). Human related reads were identified using BWA mapped

418 to the hg19 human genome and removed. Following dehumanization, non-human reads were used for de-

419 novo assembly using Megahit (1.1.1), and contigs were queried against our in-house database with

420 BLASTN (2.3.9+) wrapper. The extensive in-house reference genome database included more than 40,000

421 species covering all domains of life[10]. Taxonomy classification and abundance were determined using a

422 customized Lowest Common Ancestor (LCA) algorithm.

423

424 **Blood sample collection.** At the designated time point, blood was drawn from the overnight fasted

425 participant in the Clinical and Translational Research Unit at Stanford University. Aliquots of blood were

426 condensed at room temperature to coagulate, and clots were subsequently pelleted. The serum supernatant

427 was then immediately frozen at -80 °C. The blood in the EDTA tubes was immediately layered onto the

428 Ficoll medium and spun with gradient centrifugation. Then the top layers were removed, and plasma was

429 aliquoted and immediately frozen at –80 °C. Subsequently, blood mononuclear cells (PBMCs) were

430 collected and counted using a cell counter. Aliquots of PBMCs were further pelleted and frozen with

431 DMSO/FBS. For the later multi-omics analyses, PBMCs were thawed on ice and then lysed to protein

432 fraction using Allprep Spin Columns (Qiagen) according to the manufacturer's instructions with the

433 QIAshredder lysis option. Upon receipt of samples, blood samples were then stored at –80 °C for clinical

434 tests. The details of the blood and cytokines tests can be found in **Supplementary Data 1**[17].

435

436 **Collection and analysis of the gut microbiome.** Stool samples were collected according to the Human

437 Microbiome Project-Core Microbiome Sampling Protocol A (https://www.hmpdacc.org/). Following the

438 Human Microbiome Project-Core Microbiome Sampling Protocol A (HMP Protocol #07-001, v12.0), DNA

439 extraction was performed. We used the MOBIO PowerSoil DNA extraction kit and proteinase K to isolate

440 DNA in a clean fume hood. Samples were then treated with lysozyme and staphylococcal hemolysin. For

441 16S (bacterial) rRNA gene amplification, primers 27F and 534R (27F: 5′-

442 AGAGTTTGATCCTGGCTCAG-3′ and 534R: 5′-ATTACCGCGGCTGCTGG-3′) were used to amplify

18

443   the 16S hyper-variable regions V1–V3. Unique barcode amplicons were used, and samples were sequenced

444   on the Illumina MiSeq platform (V3; 2×300 bp). Illumina software handled the initial processing of all

445   raw sequencing data. Reads were further processed by removing low-quality (average quality <35) and

446   ambiguous base (Ns) sequences. UChime was used to remove chimeric amplicons, cluster the amplicon

447   sequences, and select the operational taxonomic unit by Usearch based on the GreenGenes database

448   (version in May 2013). The final biological classification assignment was performed using the RDP-

449   classifier in QIIME with custom scripts[58].

450

451   **Untargeted proteomics by LC-HRMS.** Preparation and analysis of plasma samples were performed as

452   previously described[58]. In short, tryptic peptides from plasma samples were separated on the NanoLC 425

453   system (SCIEX). 0.5×10 mm ChromXP (SCIEX) was used for trap-elution settings, and the flow rate was

454   set to 5 µl/min. The LC gradient was a 43-minute gradient with mobile phase A: 0.1% formic acid in 100%

455   water and mobile phase B: 0.1% formic acid in 100% acetonitrile. During the gradient, mobile phase B was

456   4–32%. Then 8 µg of undepleted plasma were loaded on LC. SWATH acquisition was performed on a

457   TripleTOF 6600 system equipped with a DuoSpray source and a 25 µm ID electrode (SCIEX). The variable

458   Q1 window SWATH acquisition mode (100 windows) was constructed in the high-sensitivity MS^2 mode.

459   PyProphet[59] was used to score the peak groups in each run statistically, and TRIC[60] was used to align all

460   runs. Finally, a matrix with a peptide level of 1% FDR and a protein level of 10% FDR was generated for

461   subsequent analysis. The protein abundances were the sum of the first three most abundant peptides. Perseus

462   (v 1.4.2.40) was applied to subtract the main components showing the main batch deviation to reduce batch

463   effect[58].

464

465   **Cytokine profiling.** The levels of circulating cytokines in the blood were measured by the 62-plex Luminex

466   antibody-conjugated magnetic bead capture assay (Affymetrix), which has been extensively characterized

467   and benchmarked by the Stanford Human Immunological Monitoring Center. The human 62-plex

468   (eBiosciences/Affymetrix) was utilized with the modifications described below. Briefly, beads were added

469   to a 96-well plate and washed using Biotek ELx405. Samples were added to the plate containing mixed

470   antibody-linked beads, incubated at room temperature for 1 hour, and then overnight at 4 °C with shaking

471   (500-600 r.p.m, orbital shaker). After overnight incubation, the plate was washed, and then the biotinylated

472   antibody was added. The plate was incubated at room temperature with shaking for 75 minutes. The plate

473   was washed, and streptavidin-PE was added for detection. After incubating for 30 minutes at room

474   temperature, the plate was washed once, and then the reading buffer was added to the wells. The plate was

475   read by a Luminex 200 instrument, and the lower limit of each cytokine per sample was set to 50 beads.

476   Radix Biosolutions custom assay control beads were added to all wells. The batch effect was corrected

477   using replicates and controls shared between batches[58].

478

479   **Untargeted metabolomics by LC–HRMS.** All blood samples were prepared and analyzed for

480   metabolomics as previously described[61]. In short, plasma samples were extracted with acetone: acetonitrile:

481   methanol (1:1:1 vol/vol/vol), evaporated to dryness under nitrogen and reconstituted in methanol: water

482   (1:1 vol/vol) for LC-HRMS analysis. HILIC and RPLC separations were used to analyze the extractants

483   four times in positive and negative modes, respectively. HILIC metabolomics data was obtained on a

484   Thermo Q Exactive plus, and RPLC metabolomics data was obtained on a Thermo Q Exactive. Both

485   instruments were equipped with HESI-II probes and operated in the full MS scan mode. We only used the

486   combined quality control samples from the study to obtain MS^2 data. We used a ZIC-HILIC column (2.1 ×

487   100 mm, 3.5 µm, 200Å; Merck Millipore) and mobile phases composed of 10 mM ammonium acetate in

488   acetonitrile: water (50:50 vol/vol) (A) and 10 mM ammonium acetate in acetonitrile/water (95:5 vol/vol)

489   (B), and a Zorbax SB-Aq column (2.1 × 50 mm, 1.7 µm, 100Å; Agilent Technologies) and mobile phases

490   composed of 0.06% acetic acid in water (A) and 0.06% acetic acid in methanol (B) to perform HILIC and

491   RPLC analyses, respectively. All raw metabolomics data were processed using Progenesis QI (Nonlinear

492   Dynamics, Waters). We also removed features that did not show sufficient linearity when diluted. Only

493   features presented in more than ⅓ samples were retained for further analysis, and the KNN method was

494   used to estimate missing values. To normalize the data, locally assessed scatter plot smoothness analysis

495 was applied[62]. Metabolic signatures were identified by matching retention time and fragmentation spectra

496 to corresponding standards or comparing fragmentation patterns to public repositories, as previously

497 reported[58]. Toxin and carcinogens were annotated out of the metabolome features if the feature cannot be

498 annotated as a human metabolite. The annotations of toxins and carcinogens were based on various blood

499 exposome related databases that are publicly available as well as in-house databases[54–56]. The confidence

500 levels of blood chemical annotation were at least level 5, with the majority at least level 3[57].

501

502 **Statistical analyses**

503 **General statistical analysis and data visualization.** All statistical analysis and data visualization were

504 performed using R (v3.6.0, https://www.r-project.org/) and RStudio (v 1.2.5019). Most of the R packages

505 and their dependencies used in this study were deployed in CRAN (https://cran.r-project.org/) or

506 Bioconductor (https://bioconductor.org/), and some of them are deployed on Github. Session information

507 for this study is provided in **Supplementary Note 1**. All scripts to reproduce analysis and data visualization

508 for   this   study   is   available   on   Github

509 (https://github.com/jaspershen/precision_exposome/tree/main/R/20200511). All data from the exposome

510 and internal -omes data were log2 transformed before analysis. According to the participant's food log,

511 fiber intake was adjusted for all internal -omes data to reduce fiber intake biases (**Supplementary Note 2**).

512

513 **Exposome and internal multi-omics correlation networks.** Spearman correlation was used to build the

514 correlations in the intra/inter-omics analyses. In general, for each two -omes pair, the correlation matrix

515 was calculated as below: for each variable in one -omes, Spearman correlations and FDR adjusted p-values

516 were generated with all features in the other -omes. Only correlations between each pair variable with

517 absolute correlation > 0.9 and FDR adjusted p-value < 0.05 were kept to construct the final correlation

518 networks.

519

520   **Community analysis.** Community analysis was performed based on edge betweenness embedded in R

521   package **igraph** (https://igraph.org/). Briefly, this is an iterative process, the edges with the highest edge

522   betweenness score were removed in each iteration, and the process was repeated until only individual nodes

523   remain. At each iteration, modularity was calculated, and communities were analyzed at the iteration that

524   maximized this quantity. A visualization of iteration community versus modularity is shown in **Figure S6**

525   **a, b**. To ensure the robustness and reliability of our findings, only communities (or clusters) with at least 3

526   nodes were kept for subsequent analysis. All the networks were visualized using R package **igraph**, **ggraph**

527   and **tidygraph**.

528

529   **GO, KEGG and Reactome pathway enrichment for proteome.** The R package **ClusterProfiler** (v

530   3.18.0, https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html) was used for GO, KEGG

531   and Reactome pathway enrichment for proteomics. In general, UNIPROT and ENTREZID were obtained

532   for proteins that connect with the exposome in correlation networks. Then the GO, KEGG and Reactome

533   pathway databases were used for pathway enrichment (hypergeometric distribution test, p-values are

534   adjusted by the FDR method, and the cutoff was set as 0.05). Only pathways with hitting protein number >

535   3 were retained for subsequent analysis.

536

537   **Metabolic feature based dysregulated module detection.** Applying the same concept from mummichog[63]

538   and PIUMet[64], metabolic networks from KEGG and community analysis were utilized to detect

539   dysregulated modules based on metabolic features connecting the exposome, respectively[63]. In general, the

540   metabolic network (MN) was downloaded from KEGG, which contains 1,377 nodes (metabolites) and 1561

541   edges (reactions). The brief workflow is described below:

542   1.   All the metabolic features connecting the exposome (L_{sig}) were matched with the KEGG metabolite

543   database based on different adducts (**Supplementary Table 1**). Then all matched metabolites

544   (significant metabolites) are mapped in the metabolic network to get the subnetwork (SN). Non-

545   significant metabolites (hidden metabolites) that can connect significant metabolites within 3

22

546   reactions were also included in the subnetwork. Then the modules (M) were detected in the

547   subnetwork via random walks[65]. Only modules with at least 3 nodes were kept. These modules

548   were named significant modules (M_{sig}) from real biological-related metabolic features.

549   2.   For each module, the activity score (S) was calculated to measure both the modularity and

550   enrichment of input metabolites (I). Activity score (S) of the module (M) was calculated as follows:

551   For a module M:

552      $S = Q * \frac{N_{I,M}}{N_M}$

553   where S is the activity score, N_M is the metabolite number in module M, and $N_{I,M}$ is the input

554   metabolite number in module M. Q is the adjusted Newman-Girvan modularity calculated as

555   below:

556      $Q = \sqrt{\frac{N_I}{N_M}} * \left(\frac{E_M}{m} - \sum_{i,j} \frac{k_i}{2m} * \frac{k_j}{2m} \right), i, j \in M$

557   where k_i is the degree of metabolite i in module M, m is the total number of edges in the metabolic

558   network MN, E_M is the total number of edges in module M, and N_I is the number of input

559   metabolites. The original Newman-Girvan modularity has a bias towards larger modules, and $\sqrt{\frac{N_I}{N_M}}$

560   was used to reduce this bias.

561   3.   Then the NULL metabolic features (L_{null}, the same number with L_{sig}) were selected from all

562   metabolic features (exclude L_{sig}) and then steps 1 and 2 were repeated 100 times to generate a list

563   of NULL modules (M_{null}) and their activity score (S_{null}).

564   4.   Using maximum likelihood estimation, S_{null} was modeled as a Gamma distribution, and a

565   cumulative distribution function (CDF) was calculated. The p-value for each significant module

566   was then calculated, and only modules with the p-value < 0.05 remained.

567 The annotation results from this method were also compared with the annotation results from the

568 "**Untargeted metabolomics by LC-HRMS**" section, provided in **Supplementary Data 9**. These results

569 showed that annotations from this method have high specificity.

23

570

571 **KEGG pathway enrichment analysis for metabolomics data.** The KEGG pathway database was

572 downloaded from KEGG (https://www.genome.jp/kegg/) using the R package **KEGGREST**. Pathway

573 enrichment analysis was used in the hypergeometric distribution test, and *p*-values were adjusted by the

574 FDR method, and only pathways with FDR adjusted p-value < 0.05 were kept.

575

576 **Exposome contributions to cytokine and blood test.** To calculate the contributions of the exposome on

577 each cytokine and blood test, principal components (PCs) were first extracted for each exposome

578 component, and only PCs with cumulative explained variation $> 80\%$ were kept. Then the linear regression

579 model was constructed using each cytokine/blood test as y and corresponding exposome component's PCs

580 as x. R^2 was extracted and used to represent the contributions of the exposome to each cytokine/blood test.

581 To calculate the contribution of the exposome components, partial least squares (PLS) and variable

582 important projection (VIP) were calculated. Finally, $R^2 * \frac{VIP_i}{sum(VIP)}$ (i $\in$ chemical, biological, and

583 environment) was used to represent the contributions of the exposome components on cytokine/blood tests.

584

585 **Data availability**

586 The biological exposome data generated in this and previous studies were deposited to the National Center

587 of Biotechnology Information under bioproject ID PRJNA421162. Some raw data utilized in this study are

588 presented on the NIH Human Microbiome 2 project site (https://portal.hmpdacc.org). Some other raw and

589 processed data are shown on the Stanford iPOP site (http://med.stanford.edu/ipop.html). The processed data

590 used for reproductive analysis can also be found on Github

591 (https://github.com/jaspershen/precision_exposome) and were provided in **Supplementary Data 1**.

592

593 **Code availability**

24

594 All codes used in this study can be found on Github (https://github.com/jaspershen/precision_exposome).

595 Certain in-house tools for this study can also be found on Github (https://github.com/jaspershen/metID).

596

603

Author contributions

P.G. and M.S. conceived the study. P.G., X.S., and C.J. analyzed the chemical exposome and

environmental factors data. P.G., X.S., X.Z., and C.J. analyzed the biological exposome data. P.G., X.S.,

and X.Z. analyzed multi-omics data. M.S. supervised the study and provided funding support. P.G.

drafted and revised the manuscript. X.S., X.Z., C.J., S.Z., X.Z., S.M.S.-F.R., and M.S. revised the

manuscript.

610

References

612 1. Topol, E. J. Individualized medicine from prewomb to tomb. *Cell* **157**, 241–253 (2014).

613 2. Zhang, J. *et al.* Ionic liquid-based totally organic solvent-free emulsification microextraction coupled

614   with high performance liquid chromatography for the determination of three acaricides in fruit juice.

615   *Talanta* **115**, 556–562 (2013).

616 3. de Oliveira, L. M. *et al.* Metal concentrations in traditional and herbal teas and their potential risks to

617   human health. *Sci. Total Environ.* **633**, 649–657 (2018).

618 4. Jin, Z. *et al.* Catabolism of (2E)-4-hydroxy-2-nonenal via ω- and ω-1-oxidation stimulated by

619   ketogenic diet. *J. Biol. Chem.* **289**, 32327–32338 (2014).

620   5.   Vermeulen, R., Schymanski, E. L., Barabási, A.-L. & Miller, G. W. The exposome and health:

621   Where chemistry meets biology. *Science* vol. 367 392–396 (2020).

622   6.   Liu, Y. *et al.* PAHs in urban soils of two Florida cities: Background concentrations, distribution, and

623   sources. *Chemosphere* **214**, 220–227 (2019).

624   7.   Li, S. *et al.* Sequential dispersive liquid-liquid microextraction for the determination of

625   aryloxyphenoxy-propionate herbicides in water. *J. Sep. Sci.* **35**, 3389–3395 (2012).

626   8.   Li, S. *et al.* Slow-Injection Ultrasound-Assisted Emulsification–Microextraction for Determination

627   of Phthalate Esters in Water†. *Journal of Chromatographic Science* vol. 52 1127–1134 (2014).

628   9.   Guo, Y. *et al.* Polycyclic aromatic hydrocarbons in processed yard trash. *Waste Manag. Res.* **38**,

629   825–830 (2020).

630   10.   Jiang, C. *et al.* Dynamic Human Environmental Exposome Revealed by Longitudinal Personal

631   Monitoring. *Cell* **175**, 277–291.e31 (2018).

632   11.   Sinharay, R. *et al.* Respiratory and cardiovascular responses to walking down a traffic-polluted road

633   compared with walking in a traffic-free area in participants aged 60 years and older with chronic

634   lung or heart disease and age-matched healthy controls: a randomized, crossover study. *The Lancet*

635   vol. 391 339–349 (2018).

636   12.   Gauthier, P. T., Norwood, W. P., Prepas, E. E. & Pyle, G. G. Metal-PAH mixtures in the aquatic

637   environment: a review of co-toxic mechanisms leading to more-than-additive outcomes. *Aquat.*

638   *Toxicol.* **154**, 253–269 (2014).

639   13.   Gao, P. The Exposome in the Era of One Health. *Environ. Sci. Technol.* (2021)

640   doi:10.1021/acs.est.0c07033.

641   14.   Zhou, W. *et al.* Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* **569**,

642   663–671 (2019).

643   15.   Schüssler-Fiorenza Rose, S. M. *et al.* A longitudinal big data approach for precision health. *Nat.*

644   *Med.* **25**, 792–804 (2019).

645   16. Gao, P. *et al.* Human exposure to polycyclic aromatic hydrocarbons: Metabolomics perspective.

646   *Environ. Int.* **119**, 466–477 (2018).

647   17. Xiang, P. *et al.* Cellular responses of normal (HL-7702) and cancerous (HepG2) hepatic cells to dust

648   extract exposure. *Chemosphere* **193**, 1189–1197 (2018).

649   18. Xiang, P. *et al.* Water extract of indoor dust induces tight junction disruption in normal human

650   corneal epithelial cells. *Environ. Pollut.* **243**, 301–307 (2018).

651   19. Kaur, N. *et al.* Paenibacillus xerothermodurans sp. nov., an extremely dry heat resistant spore

652   forming bacterium isolated from the soil of Cape Canaveral, Florida. *Int. J. Syst. Evol. Microbiol.* **68**,

653   3190–3196 (2018).

654   20. Rodrigues, E. T., Lopes, I. & Pardal, M. Â. Occurrence, fate and effects of azoxystrobin in aquatic

655   ecosystems: a review. *Environ. Int.* **53**, 18–28 (2013).

656   21. Final amended report on the safety assessment of Methylparaben, Ethylparaben, Propylparaben,

657   Isopropylparaben, Butylparaben, Isobutylparaben, and Benzylparaben as used in cosmetic products.

658   *Int. J. Toxicol.* **27 Suppl 4**, 1–82 (2008).

659   22. European Food Safety Authority (EFSA) *et al.* Peer review of the pesticide risk assessment of the

660   active substance captan. *EFSA J* **18**, e06230 (2020).

661   23. Parker, B. J., Wearsch, P. A., Veloo, A. C. M. & Rodriguez-Palacios, A. The Genus Alistipes: Gut

662   Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. *Frontiers in*

663   *Immunology* vol. 11 (2020).

664   24. Lau, S. K. P. *et al.* Eggerthella hongkongensis sp. nov. and eggerthella sinensis sp. nov., two novel

665   Eggerthella species, account for half of the cases of Eggerthella bacteremia. *Diagn. Microbiol.*

666   *Infect. Dis.* **49**, 255–263 (2004).

667   25. Gomez-Arango, L. F. *et al.* Increased Systolic and Diastolic Blood Pressure Is Associated With

668   Altered Gut Microbiota Composition and Butyrate Production in Early Pregnancy. *Hypertension* **68**,

669   974–981 (2016).

670   26. Ju, T., Kong, J. Y., Stothard, P. & Willing, B. P. Defining the role of Parasutterella, a previously

671   uncharacterized member of the core gut microbiota. *ISME J.* **13**, 1520–1534 (2019).

672 27. Tamanai-Shacoori, Z. *et al.* Roseburia spp.: a marker of health? *Future Microbiol.* **12**, 157–170

673   (2017).

674 28. Safran, M. *et al.* GeneCards Version 3: the human gene integrator. *Database* vol. 2010 baq020–

675   baq020 (2010).

676 29. Walker, D. I. *et al.* The metabolome: A key measure for exposome research in epidemiology. *Curr*

677   *Epidemiol Rep* **6**, 93–103 (2019).

678 30. Zajc, J. & Gunde-Cimerman, N. The Genus Wallemia—From Contamination of Food to Health

679   Threat. *Microorganisms* vol. 6 46 (2018).

680 31. Rumore, M. M. & Kim, K. S. Potential role of salicylates in type 2 diabetes. *Ann. Pharmacother.* **44**,

681   1207–1221 (2010).

682 32.Shamri, R., Xenakis, J. J. & Spencer, L. A. Eosinophils in innate immunity: an evolving story.

683   *Cell Tissue Res.* **343**, 57–83 (2011).

684 33. Rutherford, J. C. The emerging role of urease as a general microbial virulence factor. *PLoS Pathog.*

685   **10**, e1004062 (2014).

686 34. Li, K. *et al.* Molecular mechanisms of PFOA-induced toxicity in animals and humans: Implications

687   for health risks. *Environ. Int.* **99**, 43–54 (2017).

688 35. Gao, P. *et al.* Emerging and legacy PAHs in urban soils of four small cities: Concentrations,

689   distribution, and sources. *Sci. Total Environ.* **685**, 463–470 (2019).

690 36. Gao, P., da Silva, E. B., Townsend, T., Liu, X. & Ma, L. Q. Emerging PAHs in urban soils:

691   Concentrations, bioaccessibility, and spatial distribution. *Sci. Total Environ.* **670**, 800–805 (2019).

692 37. Xiang, P. *et al.* Effects of organophosphorus flame retardant TDCPP on normal human corneal

693   epithelial cells: Implications for human health. *Environ. Pollut.* **230**, 22–30 (2017).

694 38. Kirk, P. M., Minter, D. W., Stalpers, J. A. & Cannon, P. F. *Dictionary of the Fungi*. (Cab

695   International, 2011).

696 39. Guo, H. *et al.* Multi-omics analyses of radiation survivors identify radioprotective microbes and

697   metabolites. *Science* **370**, (2020).

698   40. Matsumoto, M., Poncipe, C. & Ema, M. Review of developmental toxicity of nitrophenolic

699   herbicide dinoseb, 2-sec-butyl-4,6-dinitrophenol. Reprod. Toxicol. 25, 327–334 (2008).

700   41. Eriksson, P., Jakobsson, E. & Fredriksson, A. Brominated flame retardants: a novel class of

701   developmental neurotoxicants in our environment? Environ. Health Perspect. 109, 903–908 (2001).

702   42. Zhang, S. *et al.* Bioaccessibility of PAHs in contaminated soils: Comparison of five in vitro methods

703   with Tenax as a sorption sink. *Sci. Total Environ.* **601-602**, 968–974 (2017).

704   43. Li, C. *et al.* Relative bioavailability and bioaccessibility of PCBs in soils based on a mouse model

705   and Tenax-improved physiologically-based extraction test. *Chemosphere* **186**, 709–715 (2017).

706   44. Bajaj, P., Chowdhury, S. K., Yucha, R., Kelly, E. J. & Xiao, G. Emerging Kidney Models to

707   Investigate Metabolism, Transport, and Toxicity of Drugs and Xenobiotics. *Drug Metab. Dispos.* **46**,

708   1692–1702 (2018).

709   45. Shibutani, S. T., Saitoh, T., Nowag, H., Münz, C. & Yoshimori, T. Autophagy and autophagy-

710   related proteins in the immune system. *Nature Immunology* vol. 16 1014–1024 (2015).

711   46. Peters, A., Nawrot, T. S. & Baccarelli, A. A. Hallmarks of environmental insults. *Cell* (2021)

712   doi:10.1016/j.cell.2021.01.043.

713   47. Orešič, M., McGlinchey, A., Wheelock, C. E. & Hyötyläinen, T. Metabolic Signatures of the

714   Exposome-Quantifying the Impact of Exposure to Environmental Chemicals on Human Health.

715   *Metabolites* **10**, (2020).

716   48. Xu, M. *et al.* Biochar impacts on phosphorus cycling in rice ecosystem. *Chemosphere* **225**, 311–319

717   (2019).

718   49. Gao, P. *et al.* Spatial and temporal changes of P and Ca distribution and fractionation in soil and

719   sediment in a karst farmland-wetland system. *Chemosphere* **220**, 644–650 (2019).

720   50. Xu, M. *et al.* Biochar impact on chromium accumulation by rice through Fe microbial-induced redox

721   transformation. *J. Hazard. Mater.* **388**, 121807 (2020).

722   51. da Silva, E. B. *et al.* Background concentrations of trace metals As, Ba, Cd, Co, Cu, Ni, Pb, Se, and

723   Zn in 214 Florida urban soils: Different cities and land uses. *Environ. Pollut.* **264**, 114737 (2020).

724 52. Schaue, D., Kachikwu, E. L. & McBride, W. H. Cytokines in radiobiological responses: a review.

725   *Radiat. Res.* **178**, 505–523 (2012).

726 53. Chhetri, D. R. *et al.* Isaria tenuipes Peck, an entomopathogenic fungus from Darjeeling Himalaya:

727   Evaluation of in-vitro antiproliferative and antioxidant potential of its mycelium extract. *BMC*

728   *Complement Med Ther* **20**, 185 (2020).

729 54. Barros, M. B. de L., de Almeida Paes, R. & Schubach, A. O. Sporothrix schenckii and

730   Sporotrichosis. *Clin. Microbiol. Rev.* **24**, 633–654 (2011).

731 55. Danella Figo, D. *et al.* Cashew Tree Pollen: An Unknown Source of IgE-Reactive Molecules. *Int. J.*

732   *Mol. Sci.* **20**, (2019).

733 56. Jiang, C., Zhang, X., Gao, P., Chen, Q. & Snyder, M. Decoding personal biotic and abiotic airborne

734   exposome. *Nat. Protoc.* (2021) doi:10.1038/s41596-020-00451-8.

735 57. Barupal, D. K. & Fiehn, O. Generating the Blood Exposome Database Using a Comprehensive Text

736   Mining and Database Fusion Approach. *Environ. Health Perspect.* **127**, 97008 (2019).

737 58. Wishart, D. *et al.* T3DB: the toxic exposome database. *Nucleic Acids Res.* **43**, D928–34 (2015).

738 59. Neveu, V., Nicolas, G., Salek, R. M., Wishart, D. S. & Scalbert, A. Exposome-Explorer 2.0: an

739   update incorporating candidate dietary biomarkers and dietary associations with cancer risk. *Nucleic*

740   *Acids Res.* **48**, D908–D912 (2020).

741 60. Shen, X. *et al.* Metabolic reaction network-based recursive metabolite annotation for untargeted

742   metabolomics. *Nat. Commun.* **10**, 1516 (2019).

743 61. Röst, H. L., Schmitt, U., Aebersold, R. & Malmström, L. pyOpenMS: a Python-based interface to

744   the OpenMS mass-spectrometry algorithm library. *Proteomics* **14**, 74–77 (2014).

745 62. Röst, H. L. *et al.* TRIC: an automated alignment strategy for reproducible protein quantification in

746   targeted proteomics. *Nat. Methods* **13**, 777–783 (2016).

747 63. Contrepois, K., Jiang, L. & Snyder, M. Optimized Analytical Procedures for the Untargeted

748   Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC)

749   and Reverse-Phase Liquid Chromatography (RPLC)-Mass Spectrometry. *Mol. Cell. Proteomics* **14**,

750   1684–1695 (2015).

751 64. Dunn, W. B. *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas

752   chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–

753   1083 (2011).

754 65. Li, S. *et al.* Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* **9**,

755   e1003123 (2013).

756
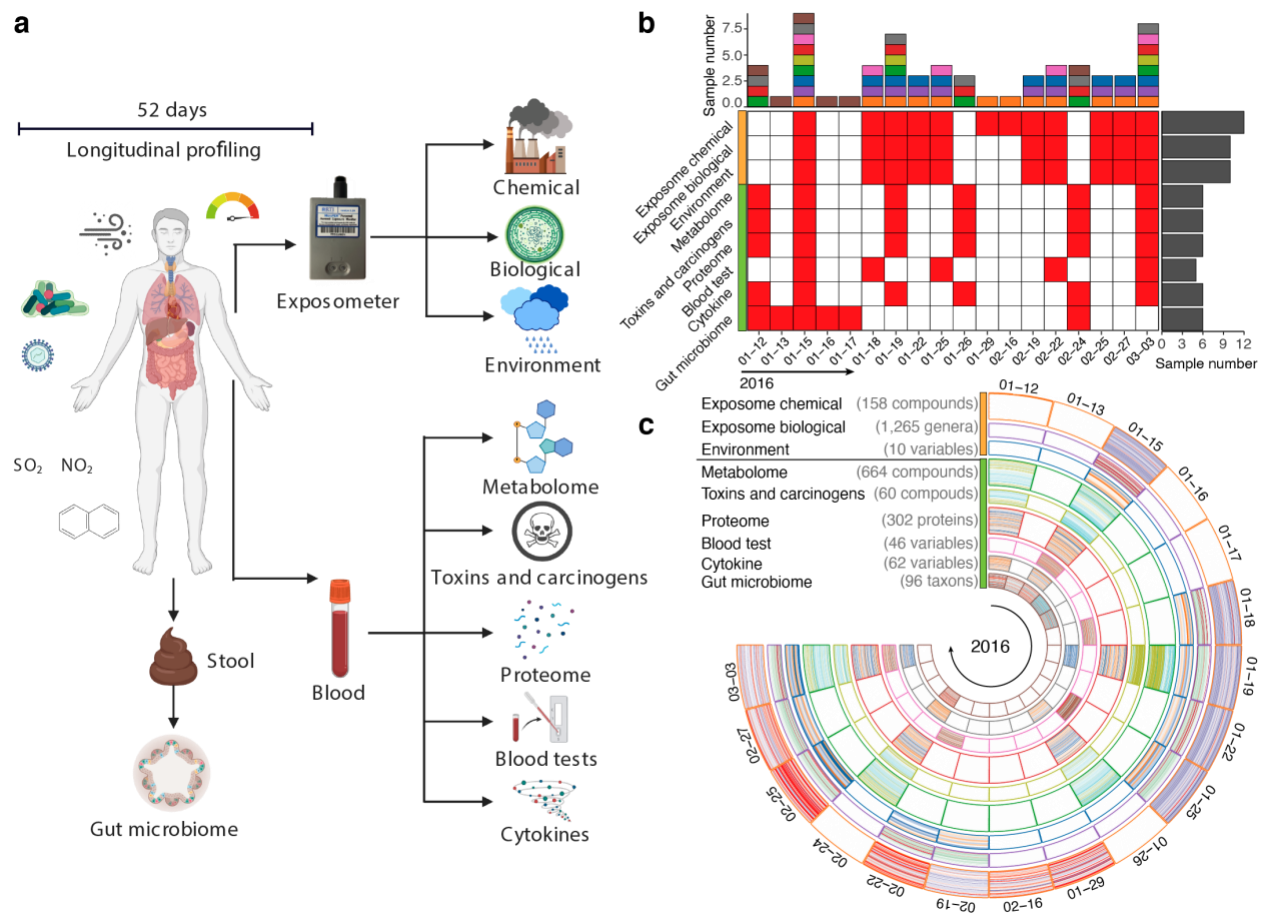
Figure 1. Overview of longitudinal sample collections for personal exposome and multi-omics profiling. (a) Personal exposome characterized by the exposometer includes environmental factors, biological components, and chemical stressors. Internal multi-omes include gut microbiome, metabolome, proteome, toxins and carcinogens, cytokines, and blood tests. (b) The amount and collection time of each type of multi-omics and exposome samples. (c) Sample distribution and constitution of the exposome and internal multi-omics for monitoring precision environmental health.
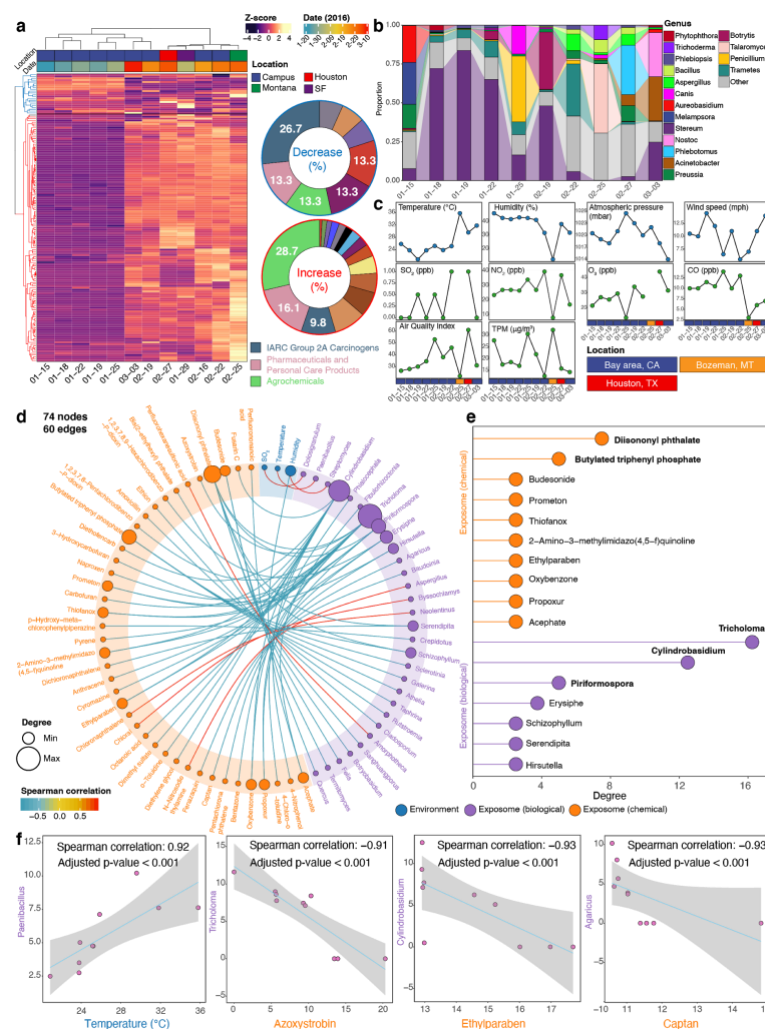
Figure 2. The dynamic and diverse personal exposome cloud. (a) Heat map of the annotated chemical stressors in the exposome ordered by concentrations, with the sector diagrams indicating the increased and decreased chemical groups. The abrupt concentration increase after the January 25 sample indicates the approach can monitor dramatic changes of the chemical exposome. (b) Heat map of the top abundant genera annotated in the exposome during each collection period. (c) Environmental factors collected by either the personal exposometer (temperature, humidity, total particulate matter) or local monitoring stations. TPM: total particulate matter. (d) Spearman correlation analyses within the personal exposome [|r| > 0.9; False Discovery Rate (FDR) adjusted p-value (q-value) < 0.05]. (e) Chemical and biological components that have the most significant correlations with the other substances in the exposome. (f) Representative Spearman correlation analyses between fungi and temperature/antifungal chemicals.
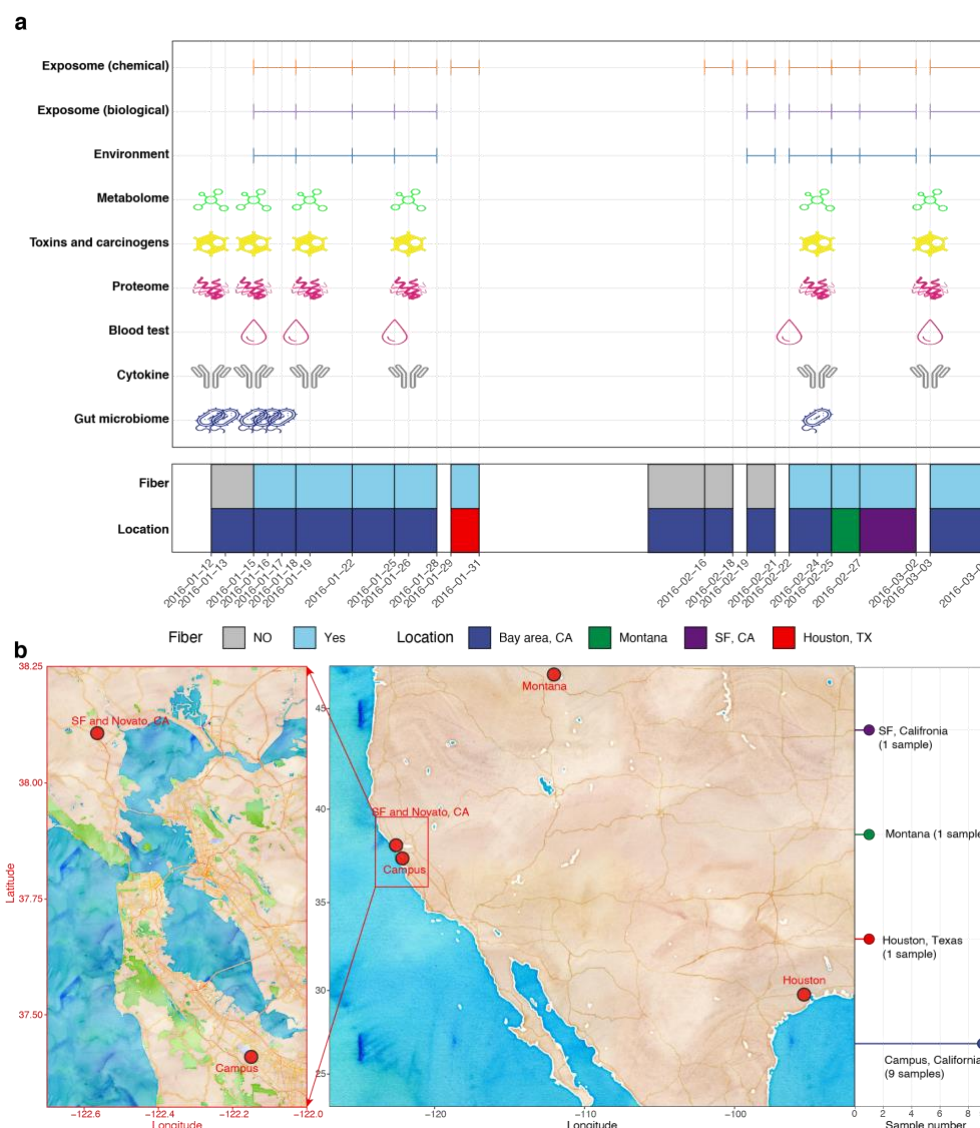
33

Figure 3. Precision environmental health network revealed by inter-omics analyses between the exposome and internal multi-omics. (a) Spearman correlation network of all longitudinally profiled exposome and internal -omes. (b) Significant Spearman correlations between the exposome and the internal multi-omics ($|r| > 0.9$; q-value < 0.05). (c) Spearman correlation analysis between the individual's exposome and the gut microbiome. Only gut bacteria with degrees > 20 are shown, and the highest degree bacteria are named. The complete network is provided in **Figure S4c**. (d) Heat map of the highest degree personal gut bacteria and significant correlations with the exposome components ($|r| > 0.9$; q-value < 0.05). Additional high-degree gut bacteria are provided in **Figure S4d**.

Figure 4. Exposome-proteome interactions: proteins and signaling pathways that were significantly correlated with the exposome. (a) Spearman correlation analysis between the exposome and proteome ($|r| >$ 0.9; q-value < 0.05). Only the proteins with degree > 5 are shown. The complete network is provided in **Figure S5**. (b) Signaling pathways that significantly correlated with the exposome revealed by pathway analysis using KEGG, GO, and Reactome databases. Immune-related pathways are shown in bold. (c) Spearman correlation networks between chemicals, top twenty biological exposome components, immune-related proteins and signaling pathways ($|r| > 0.9$; q-value < 0.05), with positive correlations shown in blue and negative correlations shown in red. A detailed network for each pathway was provided in **Figure S6**.

Figure 5. Exposome-metabolome interactions: metabolites and metabolic pathways that were significantly correlated with the exposome. (a) Spearman correlation analysis between the exposome and metabolome. (b) Significantly correlated metabolic pathways revealed by pathway analysis using KEGG database (|r| > 0.9; q-value < 0.05). (c) Significant correlations between chemicals and top twenty biological exposome components and metabolites (represented by KEGG compound entry) and metabolic pathways revealed by Spearman correlation networks (|r| > 0.9; q-value < 0.05), with positive correlations shown in blue and negative correlations shown in red. The complete network was provided in **Figure S8**.

801

Figure 6. Effects of the exposome on precision environmental health. (a) Relative contributions of various exposome components on the alterations of personal cytokines and blood tests ($|r| > 0.9$; q-value < 0.05). (b) Representative blood test results with corresponding reference ranges (green areas) and their significantly correlated exposome components.

807 **Supplementary Materials**

809 **Figure S1.** A detailed overview of sample collection. (a) Sample collection time points/periods for all

810 exposome and internal multi-omics datasets. Corresponding fiber intake and geographical locations were

811 also provided. (b) Collection locations for the exposome samples.

812

Figure S2. Overview of the exposome data. (a) PCA plot shows the data quality of chemical exposome data. For each sample, 3 repeats were acquired. (b) The number and percentage of biological exposome data at different taxonomic ranks. (c) Correlation plots between all environmental factors. (d) Sample matching for 3 exposome domains. (e) Intra correlation network for the chemical exposome. (f) PCA plots of the 3 exposome domains based on the sampling time and locations.

818

Figure S3. Representative Spearman correlation plots between environment factors and microbes.

820
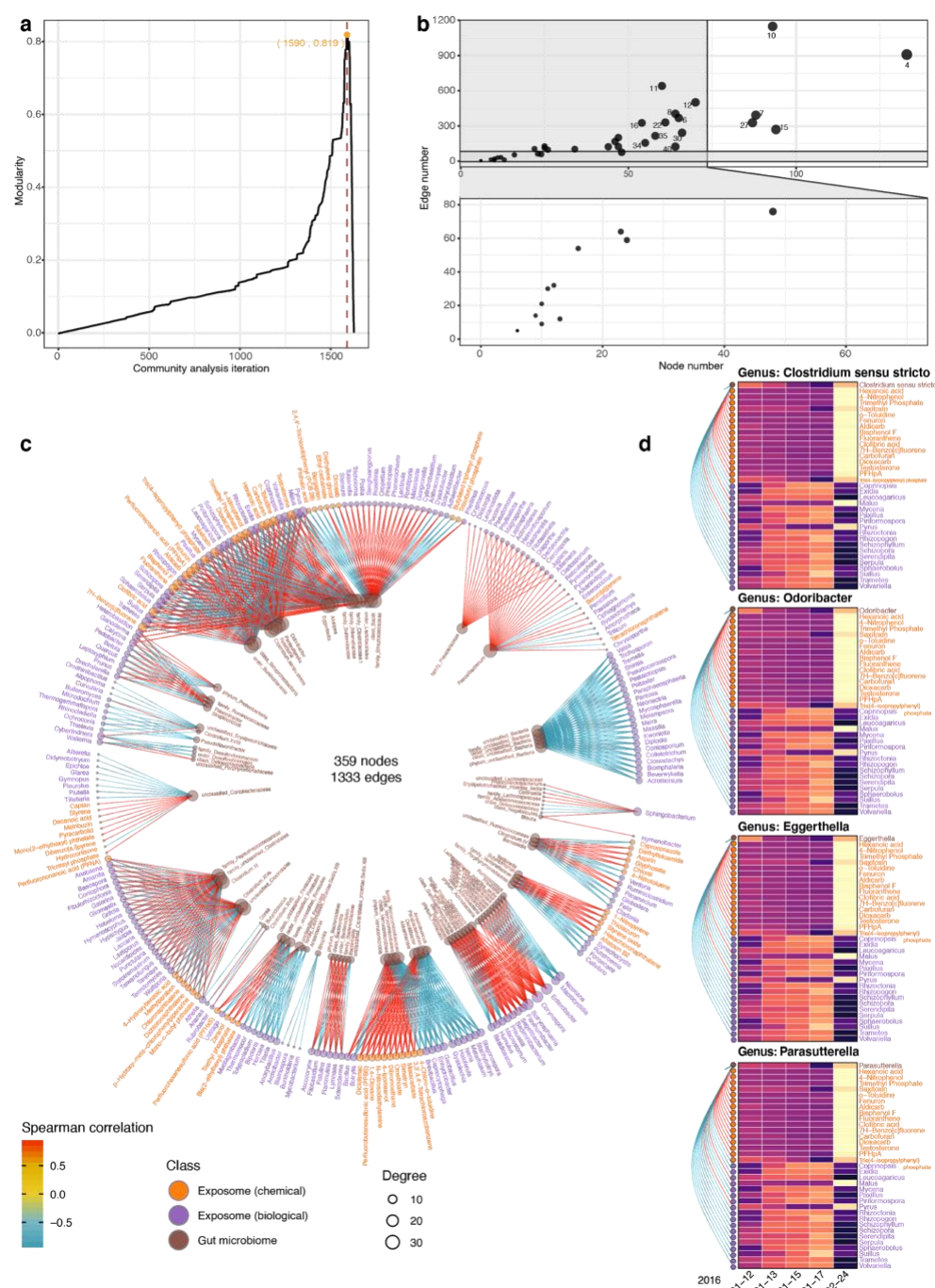
Figure S4. The exposome and internal multi-omics analyses. (a) The maximum modularity in the

exposome and internal-omes correlation network community analysis was 0.819. (b) Node and edge

numbers for all the subnetworks in the community analysis. (c) The complete correlation network

between the exposome and the gut microbiome ($|r| > 0.9$; q-value < 0.05). (d) Representative personal gut

bacteria that has 34 significant correlations with the exposome components ($|r| > 0.9$; q-value < 0.05).

41

a

826
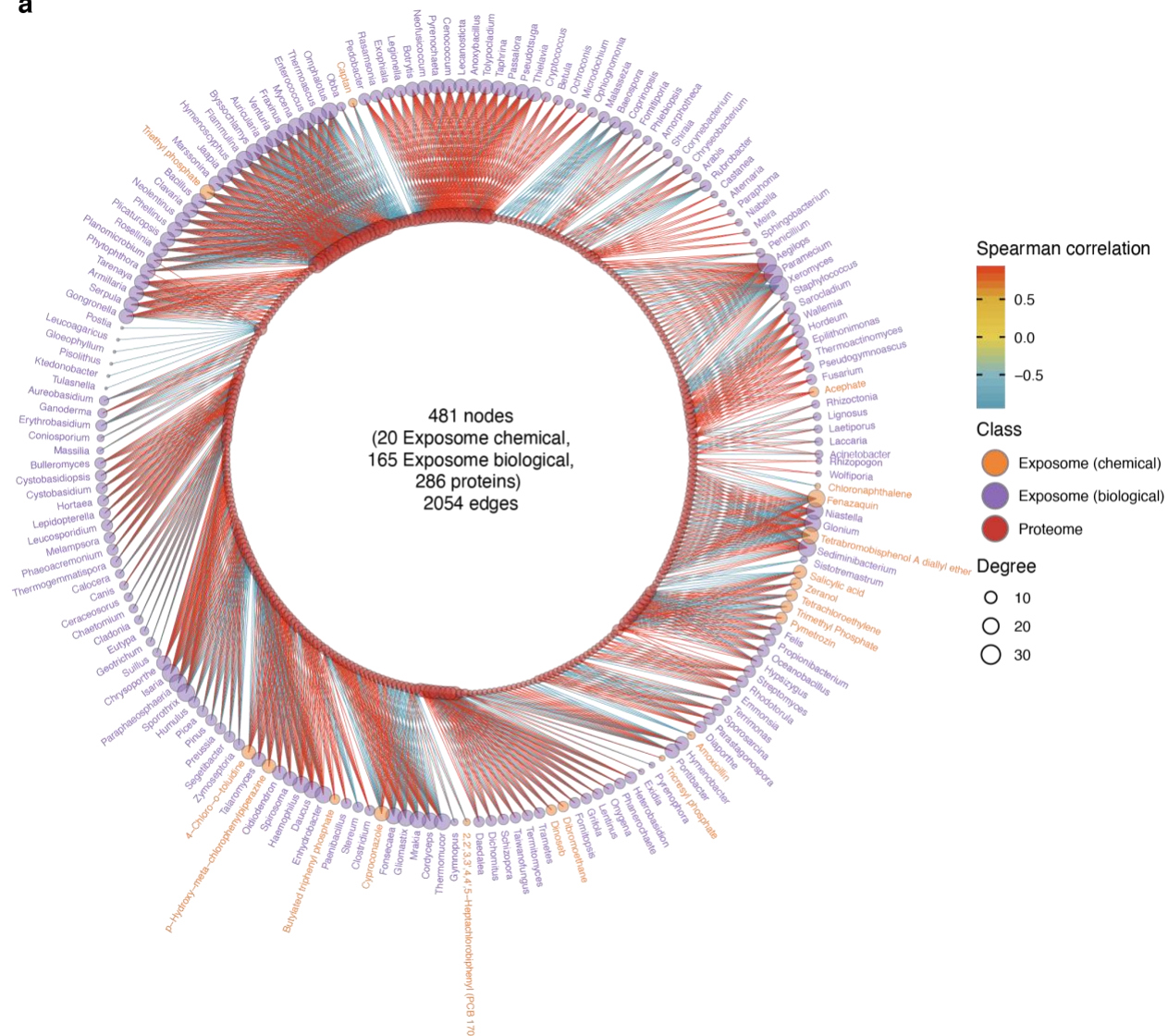
827 **Figure S5.** The complete correlation network between the exposome and proteome ($|r| > 0.9$; q-value <

828 0.05).
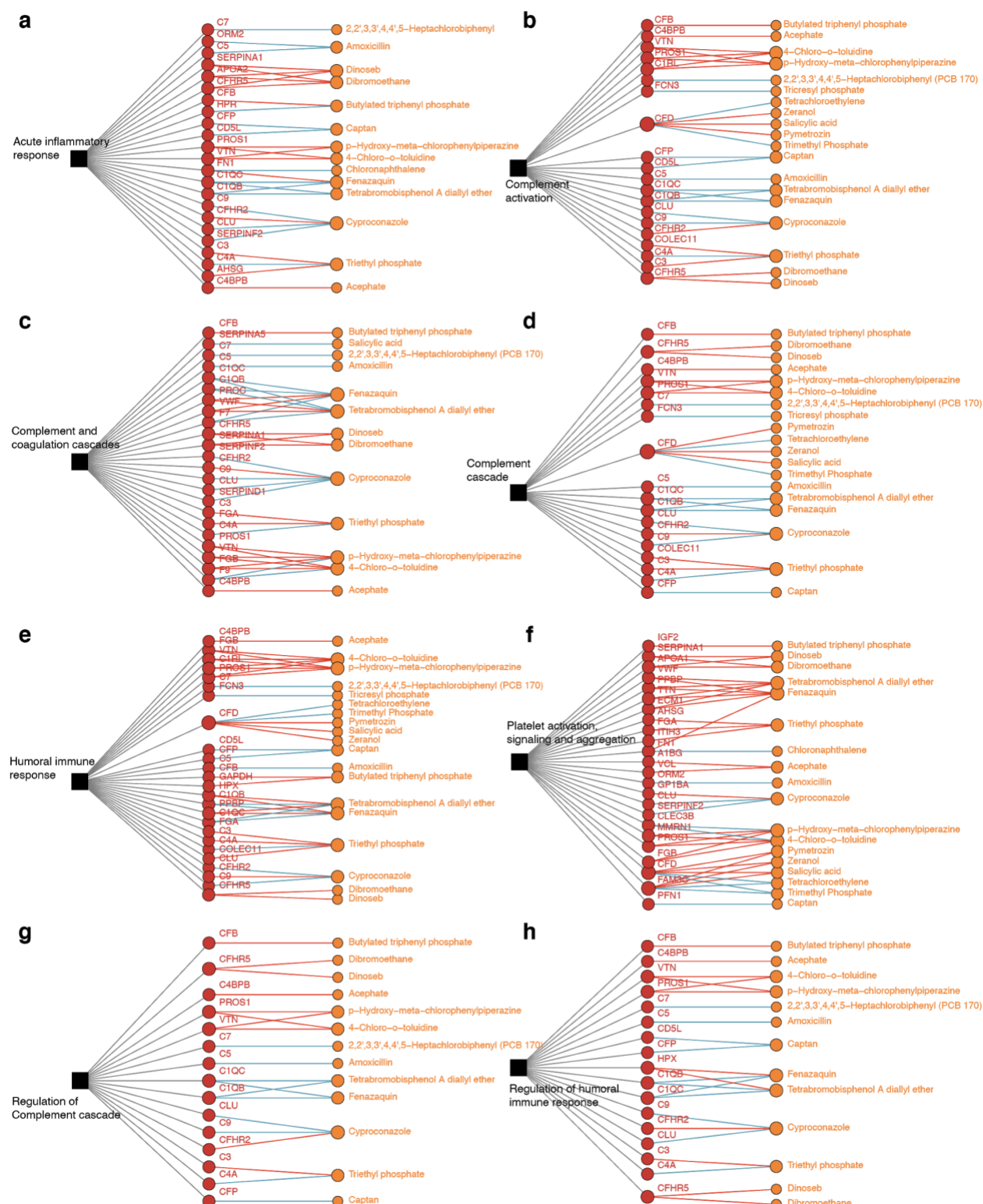
Figure S6. The detailed networks between the exposome and proteins for each signaling pathway (|r| > 0.9; q-value < 0.05).

832
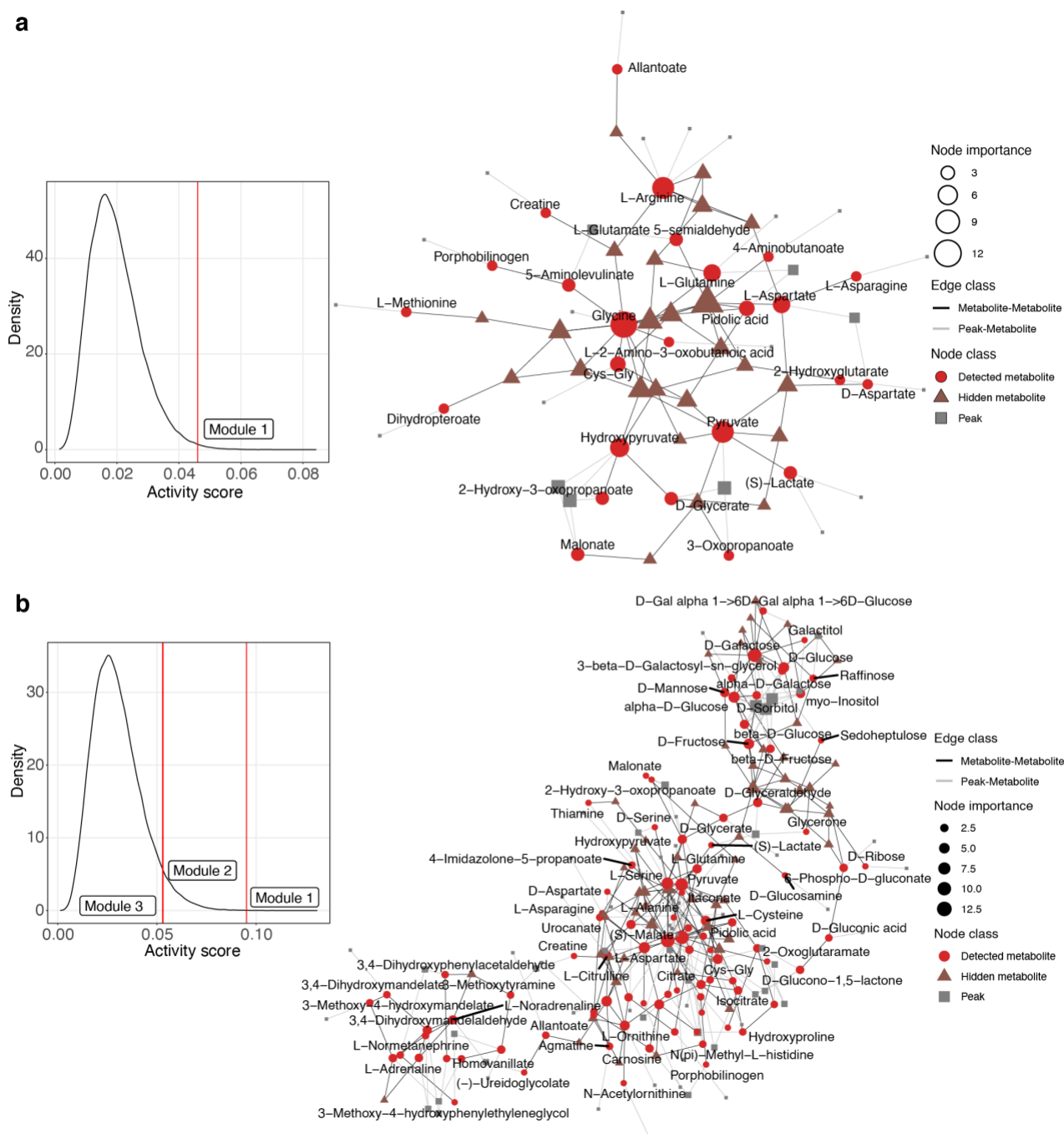
Figure S7. Feature-based network analysis for metabolic features significantly correlated with the

exposome (**Methods**). (a) Chemical exposome. Left panel, the null distribution of activity scores and only

module 1 was significant. Right panel, network constructed by significant modules. (b) Biological

exposome. Left panel, the null distribution of activity scores and modules 1, 2 and 3 were significant.

Right panel, network constructed by significant modules.
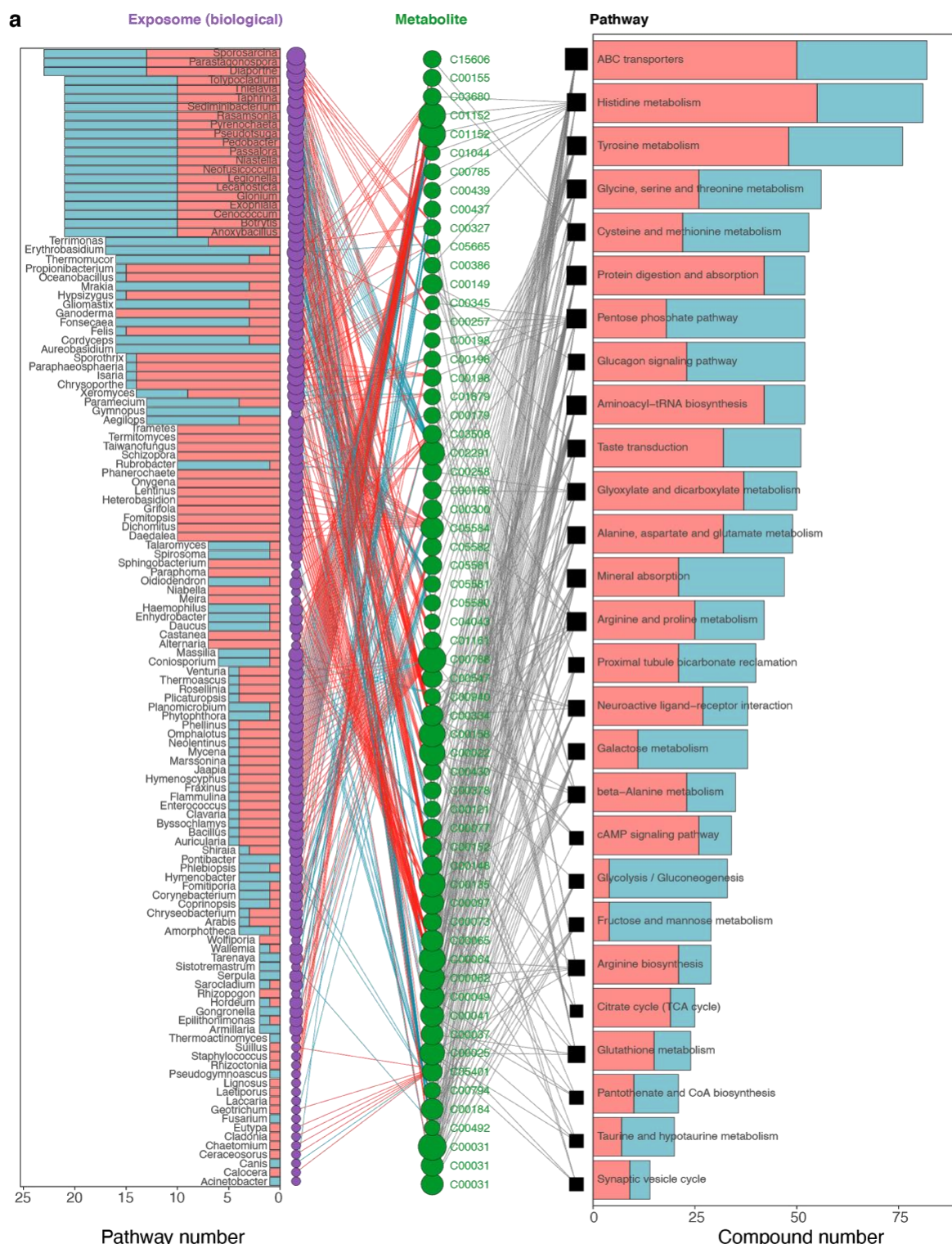
Figure S8. The complete network between the exposome and metabolite for each metabolic pathway

($|r| > 0.9$; q-value < 0.05).

842 **Supplementary Note**

843 **Note 1. The session information in the study**

844 R version 4.0.3 (2020-10-10)

845 Platform: x86_64-apple-darwin17.0 (64-bit)

846 Running under: macOS Big Sur 10.16

847

848 Matrix products: default

849 LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib

850

851 locale:

852 [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

853

854 attached base packages:

855 [1] parallel stats4   grid   stats   graphics grDevices utils   datasets methods

856 [10] base

857

858 other attached packages:

859 [1] ReactomePA_1.34.0   org.Hs.eg.db_3.10.0   AnnotationDbi_1.52.0   IRanges_2.24.1

860 [5] S4Vectors_0.28.1   Biobase_2.50.0   BiocGenerics_0.36.0   clusterProfiler_3.18.0

861 [9] circlize_0.4.11   ComplexHeatmap_2.6.2 patchwork_1.1.1   plsVarSel_0.9.6

862 [13] pls_2.7-3   openxlsx_4.2.3   plyr_1.8.6   ggalluvial_0.12.3

863 [17] tidygraph_1.2.0   ggraph_2.0.4   igraph_1.2.6   forcats_0.5.0

864 [21] stringr_1.4.0   dplyr_1.0.2   purrr_0.3.4   readr_1.4.0

865 [25] tidyr_1.1.2   tibble_3.0.4   ggplot2_3.3.3   tidyverse_1.3.0

866

867 loaded via a namespace (and not attached):

868   [1] readxl_1.3.1   shadowtext_0.0.7   backports_1.2.1

869   [4] fastmatch_1.1-0   sxtTools_0.99.01   splines_4.0.3

870   [7] BiocParallel_1.24.1   crosstalk_1.1.0.1   digest_0.6.27

871   [10] htmltools_0.5.0   GOSemSim_2.16.1   rsconnect_0.8.16

872   [13] magick_2.5.2   viridis_0.5.1   GO.db_3.10.0

873   [16] wesanderson_0.3.6   fansi_0.4.1   magrittr_2.0.1

874   [19] checkmate_2.0.0   memoise_1.1.0   cluster_2.1.0

875   [22] graphlayouts_0.7.1   modelr_0.1.8   matrixStats_0.57.0

876   [25] bdsmatrix_1.3-4   prettyunits_1.1.1   enrichplot_1.10.1

877   [28] colorspace_2.0-0   blob_1.2.1   rvest_0.3.6

878   [31] rappdirs_0.3.1   ggrepel_0.9.0   xfun_0.19

879   [34] haven_2.3.1   crayon_1.3.4   jsonlite_1.7.2

880   [37] graph_1.68.0   scatterpie_0.1.5   glue_1.4.2

881   [40] polyclip_1.10-0   gtable_0.3.0   webshot_0.5.2

882   [43] genalg_0.2.0   GetoptLong_1.0.5   graphite_1.36.0

883   [46] shape_1.4.5   scales_1.1.1   DOSE_3.16.0

884   [49] mvtnorm_1.1-1   DBI_1.1.0   miniUI_0.1.1.1

885   [52] Rcpp_1.0.5.4   progress_1.2.2   xtable_1.8-4

886   [55] viridisLite_0.3.0   clue_0.3-58   gridGraphics_0.5-1

887   [58] bit_4.0.4   reactome.db_1.70.0   htmlwidgets_1.5.3

888   [61] httr_1.4.2   fgsea_1.16.0   RColorBrewer_1.1-2

889   [64] ellipsis_0.3.1   pkgconfig_2.0.3   farver_2.0.3

890   [67] dbplyr_2.0.0   utf8_1.1.4   manipulateWidget_0.10.1

891   [70] later_1.1.0.1   ggplotify_0.0.5   tidyselect_1.1.0

892   [73] labeling_0.4.2   rlang_0.4.10   reshape2_1.4.4

893   [76] munsell_0.5.0   cellranger_1.1.0   tools_4.0.3

894   [79] downloader_0.4   cli_2.2.0   generics_0.1.0

895   [82] RSQLite_2.2.1   broom_0.7.3   evaluate_0.14

896   [85] fastmap_1.0.1   yaml_2.2.1   knitr_1.30

897   [88] bit64_4.0.5   fs_1.5.0   zip_2.1.1

898   [91] rgl_0.103.5   mime_0.9   praznik_8.0.0

899   [94] DO.db_2.9   xml2_1.3.2   compiler_4.0.3

900   [97] rstudioapi_0.13   MSQC_1.0.2   png_0.1-7

901   [100] reprex_0.3.0   tweenr_1.0.1   stringi_1.5.3

902   [103] lattice_0.20-41   Matrix_1.2-18   ggsci_2.9

903   [106] vctrs_0.3.6   pillar_1.4.7   lifecycle_0.2.0

904   [109] BiocManager_1.30.10   GlobalOptions_0.1.2   data. table_1.13.6

905   [112] cowplot_1.1.1   httpuv_1.5.4   qvalue_2.22.0

906   [115] R6_2.5.0   promises_1.1.1   gridExtra_2.3

907   [118] sessioninfo_1.1.1   MASS_7.3-53   assertthat_0.2.1

908   [121] rjson_0.2.20   withr_2.3.0   hms_0.5.3

909   [124] rmarkdown_2.6   rvcheck_0.1.8   Cairo_1.5-12.2

910   [127] ggnewscale_0.4.4   ggforce_0.3.2   shiny_1.5.0

911   [130] lubridate_1.7.9.2

912

913   **Note 2. The food log (fiber intake) of the participant in this study**

914   From 1/15/2016 to 1/31/2016, the participant took 20 grams of arabinoxylan daily. From 2/22/2016 to

915   3/17/2016, the participant took 10 grams of guar gum daily.

916

917   **Supplementary Table 1. Adduct list for metabolite annotation in this study.**

Mode	Adduct

RPLC positive mode	$(M+H)^+$, $(M+H-H_2O)^+$, $(M+H-2H_2O)^+$, $(M+NH_4)^+$, $(M+Na)^+$, $(M-H+2Na)^+$, $(M-2H+3Na)^+$, $(M+K)^+$, $(M-H+2K)^+$, $(M-2H+3K)^+$, $(M+CH_3CN+H)^+$, $(M+CH_3CN+Na)^+$, $(2M+H)^+$, $(2M+NH_4)^+$, $(2M+Na)^+$, $(2M+K)^+$, $(M+HCOO+2H)^+$
RPLC negative mode	$(M-H)^-$, $(M-H_2O-H)^-$, $(M+Na-2H)^-$, $(M+K-2H)^-$, $(M+NH_4-2H)^-$, $(2M-H)^-$, $(M+F)^-$
HILIC positive mode	$(M+H)^+$, $(M+H-H_2O)^+$, $(M+H-2H_2O)^+$, $(M+NH_4)^+$, $(M+Na)^+$, $(M-H+2Na)^+$, $(M-2H+3Na)^+$, $(M+K)^+$, $(M-H+2K)^+$, $(M-2H+3K)^+$, $(M+CH_3CN+H)^+$, $(M+CH_3CN+Na)^+$, $(2M+H)^+$, $(2M+NH_4)^+$, $(2M+Na)^+$, $(2M+K)^+$, $(M+HCOO+2H)^+$
HILIC negative mode	$(M-H)^-$, $(M-H_2O-H)^-$, $(M+Na-2H)^-$, $(M+K-2H)^-$, $(M+NH_4-2H)^-$, $(2M-H)^-$, $(M+CH_3COO)^-$

918

919 **Supplementary Data**

920 **Supplementary Data 1.** All exposome and internal multi-omics data used in this study.

921 **Supplementary Data 2.** The detailed information of the intra-correlation network of the chemical

922 exposome.

923 **Supplementary Data 3.** The detailed information of the personal exposome cloud.

924 **Supplementary Data 4.** The detailed information of the inter-correlation network of the exposome and

925 internal-multi-omics data.

926 **Supplementary Data 5.** Pathway enrichment results for proteins connected with the chemical exposome.

927 **Supplementary Data 6.** Pathway enrichment results for proteins connected with the biological

928 exposome.

929 **Supplementary Data 7.** Pathway enrichment results for metabolic peaks connected with the chemical

930 exposome.

931 **Supplementary Data 8.** Pathway enrichment results for metabolic peaks connected with the biological

932 exposome.

933 **Supplementary Data 9.** The comparison between metabolite annotations from traditional methods and

934 metabolic feature-based network analysis.

935 **Supplementary Data 10.** The contributions of the exposome on blood tests and cytokines.