

1 **The plasma metabolome of long COVID-19 patients two years after infection**

2 Yamilé López-Hernández<sup>1</sup>, Joel Monárrez Aquino<sup>2</sup>, David Alejandro García López<sup>3</sup>,  
3 Jiamin Zheng<sup>4</sup>, Juan Carlos Borrego<sup>5</sup>, Claudia Torres-Calzada<sup>4</sup>, José Pedro Elizalde-Díaz<sup>6</sup>,  
4 Rupasri Mandal<sup>4</sup>, Mark Berjanskii<sup>4</sup>, Eduardo Martínez-Martínez<sup>6</sup>, Jesús Adrián López<sup>7</sup>,  
5 David S. Wishart<sup>4\*</sup>

6

7 <sup>1</sup>CONACyT- Metabolomics and Proteomics Laboratory, Academic Unit of Biological  
8 Sciences, Autonomous University of Zacatecas, Zacatecas, 98000, Mexico.

9 <sup>2</sup>Department of Health Research. Christus Muguerza del Parque Hospital – University of  
10 Monterrey. Chihuahua, 31125, México.

11 <sup>3</sup>Academic Unit of Biological Sciences, Autonomous University of Zacatecas, Zacatecas,  
12 98000, Mexico.

13 <sup>4</sup>The Metabolomics Innovation Centre, University of Alberta, Edmonton, AB T6G 1C9,  
14 Canada.

15 <sup>5</sup>Departamento de Epidemiología, Hospital General de Zona #1 “Emilio Varela Luján”,  
16 Instituto Mexicano del Seguro Social, Zacatecas, 98000, México.

17 <sup>6</sup>Laboratory of Cell Communication & Extracellular Vesicles, Division of Basic Science,  
18 Instituto Nacional de Medicina Genómica, Ciudad de México, 14610, Mexico.

19 <sup>7</sup>MicroRNAs and Cancer Laboratory, Academic Unit of Biological Sciences, Autonomous  
20 University of Zacatecas, Zacatecas, 98000, Mexico.

21

22 **\*Corresponding author: [dwishart@ualberta.ca](mailto:dwishart@ualberta.ca)**

23

24 **ABSTRACT**

25 **Background:** One of the major challenges currently faced by global health systems is the  
26 prolonged COVID-19 syndrome (also known as “long COVID”) which has emerged as a  
27 consequence of the SARS-CoV-2 epidemic. The World Health Organization (WHO)  
28 recognized long COVID as a distinct clinical entity in 2021. It is estimated that at least 30%  
29 of patients who have had COVID-19 will develop long COVID. This has put a tremendous  
30 strain on still-overstretched healthcare systems around the world. **Methods:** In this study,  
31 our goal was to assess the plasma metabolome in a total of 108 samples collected from  
32 healthy controls, COVID-19 patients, and long COVID patients recruited in Mexico  
33 between 2020 and 2022. A targeted metabolomics approach using a combination of LC-  
34 MS/MS and FIA MS/MS was performed to quantify 108 metabolites. IL-17 and leptin  
35 concentrations were measured in long COVID patients by immunoenzymatic assay.  
36 **Results:** The comparison of paired COVID-19/post-COVID-19 samples revealed 53  
37 metabolites that were statistically different (FDR < 0.05). Compared to controls, 29  
38 metabolites remained dysregulated even after two years. Notably, glucose, kynurenine, and  
39 certain acylcarnitines continued to exhibit altered concentrations similar to the COVID-19  
40 phase, while sphingomyelins and long saturated and monounsaturated LysoPCs,  
41 phenylalanine, butyric acid, and propionic acid levels normalized. Post-COVID-19 patients  
42 displayed a heterogeneous metabolic profile, with some showing no symptoms while others  
43 exhibiting a variable number of symptoms. Lactic acid, lactate/pyruvate ratio,  
44 ornithine/citrulline ratio, sarcosine, and arginine were identified as the most relevant  
45 metabolites for distinguishing patients with more complicated long COVID evolution.  
46 Additionally, IL-17 levels were significantly increased in these patients. **Conclusions:**  
47 Mitochondrial dysfunction, redox state imbalance, impaired energy metabolism, and

48 chronic immune dysregulation are likely to be the main hallmarks of long COVID even two  
49 years after acute COVID-19 infection.

50 **Keywords:** long COVID, PACS, targeted metabolomics, mass spectrometry, COVID-19

## 51 **Background**

52 Historically, highly pathogenic beta-coronaviruses have been associated with severe  
53 respiratory diseases. According to the WHO, the severe acute respiratory syndrome  
54 coronavirus (SARS-CoV), and the Middle East respiratory syndrome coronavirus (MERS-  
55 CoV) were responsible for epidemics in 2002-2003 and 2015, respectively. During the  
56 SARS-CoV epidemic, the virus was reported in 29 countries with 8,437 cases and 813  
57 fatalities [1]. On the other hand, MERS-CoV was reported in 27 countries with 2,519  
58 laboratory-confirmed cases between 2012 and 2020, resulting in 866 deaths [2]. In 2019,  
59 exactly 100 years after the last pandemic caused by an H1N1 influenza A virus (the  
60 Spanish flu), a new pandemic affected almost every country around the world. As of  
61 February 26, 2023, over 758 million confirmed cases of SARS-CoV-2 and over 6.8 million  
62 deaths have been reported globally. To date, around 653 million patients have recovered  
63 [3]. However, as early as spring 2020, people began describing their experiences of not  
64 fully recovering from SARS-CoV-2 infection [4]. This extended version of the disease has  
65 been called “long COVID”. Interestingly, the term long COVID is a patient-created term  
66 promoted in Twitter by Elsa Perego, an archeologist at University College London.

67       It has been widely described that some viruses lead to persistent physiological  
68 alterations even a decade after infection. The term “post-viral syndrome” has been in use  
69 for over a century [5]. Chronic symptoms such as fatigue, joint pain, and cardiovascular  
70 problems have been reported after recovery from other infections such as the West Nile,  
71 Polio, Dengue, Zika, seasonal flu, Epstein-Barr, Ebola, MERS, and SARS [6, 7]. However,  
72 none of these viruses have affected so many people in the same time window as SARS-  
73 CoV-2, which offers the scientific community a unique opportunity to understand the  
74 etiology of post-viral syndromes such as long COVID.

75 Long COVID (also known as post-COVID-19 syndrome or post-acute sequelae of  
76 COVID-19 (PACS)) is a condition characterized by long-term or persistent health problems  
77 appearing after the initial recovery from COVID-19 infection. The WHO has described  
78 long COVID as a condition “that occurs in individuals with a previous history of probable  
79 or confirmed SARS-CoV-2 infection, usually three months after the onset, with symptoms  
80 lasting at least two months that cannot be explained by an alternative diagnosis” [8]. It is  
81 estimated that 30-60% of recovered patients, even after a mild disease, will experience long  
82 COVID or symptoms persistence with varying durations [9]. Based on a conservative  
83 estimated incidence, at least 65 million individuals worldwide could be experiencing long  
84 COVID [3].

85 Similar to COVID-19, long COVID affects multiple organ systems, including the  
86 respiratory, cardiovascular, nervous, and gastrointestinal systems. More than 50 symptoms  
87 have been reported associated with long COVID [10]. Observational studies have reported  
88 that some symptomatic conditions are resolved within three months of hospitalization in  
89 50% of patients [11], but the rate of full recovery drops to 35% between three and six  
90 months after hospitalization, and to 15% between six to nine months. Importantly, a high  
91 proportion of that population has residual lung tissue injury, with detectable radiological  
92 abnormalities on chest computed tomography (CT) scans [12]. Fatigue, loss of  
93 concentration, headaches, shortness of breath, anosmia, muscle weakness, joint pain are  
94 some of the symptoms most reported. Therefore, more than a homogeneous entity, long  
95 COVID could be considered as a spectrum of disorders, that affects individuals with  
96 complications directly linked to the virus (long-term residual damage in the lungs, brain, or  
97 heart), and individuals manifesting systemic unspecific signs/symptoms (fatigue, headache,  
98 and arthromyalgias) [13]. The increasing number of patients with long COVID poses a

99 challenge for public health systems around the world, but, currently, there are no guidelines  
100 for accurately diagnosing patients with long COVID and classification is still  
101 underestimated and subjective.

102 In the present work we used quantitative targeted metabolomics to evaluate the  
103 metabolic reversion of patients with persistent sequelae due to confirmed SARS-CoV-2  
104 infection. Comparison with negative controls allowed us to identify those metabolites  
105 persistently dysregulated after two years of the initial infection. Number, type of symptoms  
106 as well as metabolic signatures were different in patients experiencing long COVID.  
107 Besides, IL-17 level was increased in patients with the worst disease evolution. To the best  
108 of our knowledge, this is the first targeted metabolomics study of long COVID patients  
109 conducted beyond twenty months post-infection.

110

## 111 **Methods**

### 112 **Patient recruitment**

113 For the aims of this study, COVID-19 patient survivors (with confirmed diagnostic based  
114 on a positive PCR for SARS-CoV-2) who developed a mild, severe, or critical disease, and  
115 were admitted (or hospitalized) in the Instituto Mexicano de Seguridad Social (Zacatecas  
116 city, Mexico) and Christus Muguerza del Parque Hospital (Chihuahua city, Mexico)  
117 between March and November 2020, were recruited. Participants were contacted for a  
118 face-to face interview. They were invited to respond to a questionnaire and to donate a  
119 blood sample. Plasma was isolated from the donated blood. COVID-19 patients from the  
120 Instituto Mexicano de Seguridad Social were recruited from an initial set of 124 COVID-19  
121 patients enrolled in a previous research study [14]. Of these, 44 (35.6%) passed away  
122 during hospitalization and in the following months after hospital discharge. From 80

123 survivors, it was possible to contact 36 (by their Social Security Number or  
124 personal/relative phone number kept in hospital records), and 15 agreed to participate. For  
125 these 15 patients, paired plasma samples from the first diagnosis of the acute disease  
126 (COVID-19 group) and post-COVID phase were available.

127         Additionally, from a cohort of patients that were hospitalized in 2020 in Christus  
128 Muguerza del Parque Hospital, 33 were randomly selected by age stratification. For those  
129 patients, a basal blood sample was not available; however, all clinical information and chest  
130 computed tomography (CT) scans were recorded in the hospital archive.

131         For the neuropsychological assessment, the validated Hamilton Anxiety Rating  
132 Scale (HAM-A) [15] was used. For depression assessment, the Hamilton Depression scale  
133 (HAM-D) was used [16]. The Montreal Cognitive Assessment (MoCA) was employed for  
134 cognitive impairment [17]. For dyspnea assessment, the modified Medical Research  
135 Council (mMRC) dyspnea scale was implemented [18]. Basic blood biochemical markers  
136 were performed (i.e., hemoglobin, platelets, leukocytes, lymphocytes, and creatinine) for all  
137 enrolled patients.

138         To assess for differences in the severity of long COVID patients, our own  
139 classification was made (arbitrarily) considering the frequency of concomitant symptoms.  
140 Recovered patients were classified as those who did not report persistent symptoms. Long  
141 COVID was considered if patients reported at least one persistent neurologic, psychiatric,  
142 gastrointestinal, cardiologic, respiratory, or systemic symptom. The Class A long COVID  
143 patients were those reporting less than five persistent symptoms (17 patients), while class B  
144 long COVID patients were those reporting five or more persistent symptoms (13 patients).  
145 As negative controls and an indicator of normal population, stored plasma samples from 37  
146 individuals who tested negative for SARS-CoV-2 in 2020 were used.

147 This study was conducted in accordance with the Declaration of Helsinki (1976). It  
148 was also revised and approved by the Research and Ethics Committees of the Instituto  
149 Mexicano de Seguridad Social, with the registration number R-2022-3301-038, and  
150 Christus Muguerza del Parque Hospital (HCMP-CEI-15042020-3, and HCMP-CEI-  
151 28022022-A01). Informed consent was obtained from all participants. All patients included  
152 in this study were informed in writing regarding the collection of their samples for research  
153 aims and were given the right to refuse participation.

#### 154 **Metabolomics analysis**

155 A combination of direct injection mass spectrometry with a reverse-phase LC-MS/MS  
156 custom assay was used, as previously described [14]. Briefly, metabolites were measured  
157 using a locally developed LC-MS/MS metabolomics assay called The Metabolomics  
158 Innovation Centre (TMIC) Prime (TMIC PRIME®) Assay. This assay provides  
159 quantitative results for up to 143 endogenous metabolites, including biogenic amines,  
160 amino acids, organic acids, lipids, and lipid-like compounds.

161 The method combines the derivatization and extraction of analytes, and the selective  
162 mass-spectrometric detection using multiple reaction monitoring (MRM) pairs. Isotope-  
163 labeled internal standards and other internal standards were used for metabolite  
164 quantification. The custom assay uses a 96 deep-well plate with a filter plate attached via  
165 sealing tape, and reagents and solvents used to prepare the plate assay. The first 14 wells of  
166 the 96-well plate were used for calibration and quality control with one double blank, three  
167 zero samples, seven calibration standards and three quality control samples. To measure all  
168 metabolites except organic acids, samples were first thawed on ice and were vortexed. 10  
169  $\mu$ L of each sample was loaded onto the center of the filter on the upper 96-well plate and  
170 dried under a stream of nitrogen. Subsequently, phenyl-isothiocyanate (PITC) was added



171 for derivatization. After incubation, the filter spots were dried again using an evaporator.  
172 Extraction of the metabolites was then achieved by adding 300  $\mu\text{L}$  of extraction solvent.  
173 The extracts were obtained by centrifugation into the lower 96-deep well plate, followed by  
174 a dilution step with the mass spectrometry running solvent.

175 For organic acid analysis, 150  $\mu\text{L}$  of ice-cold methanol and 10  $\mu\text{L}$  of isotope-labeled  
176 internal standard mixture was added to 50  $\mu\text{L}$  of each plasma sample for overnight protein  
177 precipitation. Each sample was then centrifuged at 13000 x g for 15 min. 50  $\mu\text{L}$  of  
178 supernatant was loaded into the center of wells of a 96-deep well plate, followed by the  
179 addition of  $^{13}\text{C}$  labeled 3-nitrophenylhydrazine (3-NPH) as an isotopic labeling reagent (for  
180 quantification). After incubation for 2h, butylated hydroxytoluene (as a stabilizer) and  
181 water were added before LC-MS injection.

182 Mass spectrometric analysis for the PITC-derivatized and 3-NPH-derivatized  
183 samples was performed on an ABSciex 4000 Qtrap® tandem mass spectrometry instrument  
184 (Applied Biosystems/MDS Analytical Technologies, Foster City, CA) equipped with an  
185 Agilent 1260 series UHPLC system. Organic acids, biogenic amines, amino acids, and  
186 amino acid derivatives were detected and quantified via LC-MS, while lipids,  
187 acylcarnitines, and glucose were detected and quantified via a direct injection (DI) method.

188 Analyst 1.6.2 and MultiQuant 3.0.3 was used for quantitative analysis. An  
189 individual seven-point calibration curve was generated to quantify organic acids, amino  
190 acids, biogenic amines, and derivatives. Ratios for each analyte's signal intensity to its  
191 corresponding isotope-labelled internal standard were plotted against the specific known  
192 concentrations using quadratic regression with a  $1/x^2$  weighting. For lipids, acylcarnitines,  
193 and glucose, a single point calibration of a representative analyte was built using the same

194 group of compounds that share the same core structure assuming a linear regression  
195 through zero.

### 196 **Plasma IL-17 and leptin determinations**

197 ELISA kits were used for the quantification of IL-17 (Catalog Number RAB0262, Sigma-  
198 Aldrich, St. Louis, MO, USA) and leptin (catalog number ab108879, Abcam, Cambridge,  
199 UK), following manufacturer's instructions. Briefly, standard solutions (or plasma  
200 samples), were added to each type of pre-coated 96-well plate and incubated overnight at  
201 4°C. The plates were then incubated with the corresponding detection antibodies (100  
202 µL/well) for 1 h at room temperature. Streptavidin solution (100 µL) was then added to  
203 each well and the plates were incubated for 45 min. After the antibody-HRP incubation,  
204 TMB one-step substrate reagent (100 µL) was added to the wells and the plates were  
205 incubated for another 30 min before the addition of a stop solution (50 µL/well).  
206 Absorbance values (at 450 nm) were used for the calculation of the protein concentrations  
207 (pg/mL) by comparing the absorbance to an appropriate standard curve.

### 208 **Statistical analysis**

209 To describe baseline characteristics of negative controls (non-COVID-19), COVID-19 or  
210 post-COVID-19 patients, medians with interquartile ranges (IQRs) or means [with standard  
211 deviations (s.d.)] and frequencies (%) were used for continuous and categorical data,  
212 respectively. Normality was assessed using the D'Agostino-Pearson normality test.  
213 Student's t-test or Mann-Whitney tests were used for continuous data. For categorical  
214 variables (e.g., sex, smoking, symptoms, and comorbidities) Pearson Chi<sup>2</sup> tests or Fisher's  
215 exact tests were used. All p-values less than 0.05 (p<0.05) were considered statistically  
216 significant. Analyses were conducted using SPSS (IBM, version 24).

217 Metabolite analysis was performed with MetaboAnalyst 5.0 [19]. Those metabolites  
218 with more than  $\square$ 20% of missing values were removed from further analysis. For the  
219 remaining metabolites, values below the limit of detection (LOD) were imputed using 1/5  
220 of the minimum positive value of each variable. The data were then subject to median  
221 normalization, log-transformed and Pareto-scaled to generate appropriate Gaussian  
222 metabolite concentration distributions. Differences in mean metabolic values between  
223 controls, COVID-19, post-COVID-19, and long COVID patients were assessed using a  
224 parametric t-test or one-way ANOVA [adjusted p-value (FDR) cut-off = 0.05]. For the  
225 paired study, t-test, and volcano plots of log-transformed p-values were generated to  
226 address significant metabolites. Principal component analysis (PCA) and two-dimensional  
227 partial least squares discriminant analysis (2-D PLS-DA) scores plots were used to compare  
228 plasma metabolite data across and between study groups; 2000-fold permutation tests were  
229 used to assess statistical significance and minimize the possibility that the observed  
230 separation of the PLS-DA clusters was due to chance. Differentiated metabolites were  
231 identified by a variable importance in projection (VIP) using a score cutoff of  $>1.5$ . Heat  
232 maps of the top 50 significant metabolites (via t-test or ANOVA) were created via  
233 MetaboAnalyst.

234 Pathway analysis was done using Metabolite Set Enrichment Analysis (MSEA) and  
235 Metabolomic Pathway Analysis (MetPA) modules as found in MetaboAnalyst 5.0 [19]. The  
236 Homo sapiens pathway library was used for pathway analysis. The global test was used for  
237 the selected pathway enrichment analysis method, whereas the node importance measure  
238 for topological analysis was used to assess the relative betweenness centrality.

239 The metabolites with the highest VIP scores and LASSO frequencies were used to  
240 create metabolite panels for predicting long COVID using multivariate logistic regression.

241 Additionally, models were adjusted for relevant potential confounders such as sex, age,  
242 relevant comorbidities (i.e., DM-II, HTN, and obesity), so that only statistically significant  
243 variables ( $p < 0.05$ ) remained in the final models. Logistic regression analysis was  
244 performed with the Pareto-scaled data. K-fold cross-validation (CV) was used to ensure  
245 that the logistic regression models were robust. To determine the performance of each  
246 generated model, the area under the receiver operating characteristics curve (AUROC or  
247 AUC) was calculated, as was sensitivity and specificity.

248

## 249 **Results**

### 250 **Demographic, clinical data and symptoms description**

251 **Table 1** shows baseline characteristics of patients enrolled in the study. Age was  
252 statistically different between negative (healthy) controls and COVID-19 patients.  
253 However, differences were not found in the self-reported comorbidities. Six patients  
254 (12.5%) developed mild disease; 37 patients developed (77%) moderate/severe disease  
255 while five patients (10.4%) developed critical disease. Six patients (12.5%) were reinfected  
256 during 2021 and 2022. All patients were fully vaccinated during the period of 2021-2022.  
257 The questionnaire answered by the patients revealed the most persistent symptoms which  
258 were grouped into five broad categories: systemic, neurologic, psychiatric, cardiologic, and  
259 respiratory. The most predominant symptoms were loss of memory (73.3%), sleep  
260 disorders, arthralgia, fatigue, exercise intolerance, myalgia (66.7%), and anxiety (60.0%).

261 **Table 1:** Baseline characteristics of participants in the study

Variable	Controls (N=38)	COVID-19 (N=15)	post-COVID-19 (N=48)	long COVID (N=29)	p-value
Age, median (Q1-	40.5 (37-	51 (44.3-59.3)	51.5 (43.5-60.8)	50.5 (44.3-62)	0.0079 <sup>*a, b, c</sup>

Q3)	53.3)				
Male gender, n (%)	17 (44.7)	12 (54.5)	28 (58.3)	9 (34.6)	0.2226
Smoking, n (%)	4 (10.5)	3 (13.6)	7 (14.6)	3 (11.5)	0.7231
<b>Comorbidities (self-reported), n (%)</b>					
Diabetes	3 (7.9)	3 (13.6)	7 (14.6)	3 (11.5)	0.7833
Hypertension	9 (23.7)	6 (27.3)	20 (41.7)	9 (34.6)	0.1035
Obesity	3 (7.9)	0 (0)	5 (10.4)	2 (7.7)	0.3823
<b>Laboratory data, median (Q1-Q3)</b>					
Hemoglobin (g/dL)	15.3 (14.4-16.2)	14.4 (12.6-16.2)	15.4 (14.5-16.1)	15.5 (14.68-16.23)	0.7833
Platelets (thousands/ mL)	276 (236.5-325.5)	177 (68.2-281.8)	242.5 (219.8-290.5)	242.3 +/- 69.7	0.1035
Leukocytes ( $\times 10^3$ )	7.1 (6.0-8.3)	6.7 (4.4-11.1)	7.0 (6.2-7.8)	7.3 (6.5-7.7)	0.3823
Lymphocyte counts (%)	31.8 (25.5-36.4)	12.1 (4.9-22.1)	33.8 (29-38.9)	34 +/-7.1	<0.0001 *a, d, e
Creatinine (mg/dL)	0.9 (0.7-1.1)	0.7 (0.6-0.9)	0.8 (0.7-1.0)	0.8 (0.7-0.9)	0.2452

262 a: negative controls vs. COVID-19; b: COVID-19 vs. post-COVID-19; c: negative controls  
 263 vs. post-COVID-19; d: COVID-19 vs. Post-COVID-19; e: COVID-19 vs. long COVID; f:  
 264 Post-COVID-19 vs. long COVID

265

266 **Paired analysis: COVID-19/post-COVID-19 phases**

267 When paired samples (COVID-19/post-COVID-19) from 15 patients were compared  
 268 metabolically, 53 plasma metabolites were found to be significantly different (FDR < 0.05).

269 **Supplementary Table 1** shows the results of the t-test (autoscaling normalization). The  
 270 volcano plot (**Figure 1A**) shows that 13 metabolites were significantly upregulated in the  
 271 post-COVID-19 phase, while 32 were downregulated with a fold change (FC) threshold >

272 1.3 (FDR <0.05). Heatmap analysis (**Figure 1B**) shows a clustering of patients  
273 corresponding to their COVID-19 and post-COVID phases, revealing that lysoPCs (except  
274 LysoPC 18:2) and SMs were downregulated in the post-COVID phase. Multivariate  
275 analysis (via PLS-DA) demonstrated a clear separation between both COVID phases  
276 (accuracy: 0.97,  $R^2$ : 0.94,  $Q^2$ : 0.77) (**Figure 1C**). The VIP plot (**Figure 1D**) shows that  
277 phenylalanine, taurine, and spermidine had lower plasma concentrations in the post-  
278 COVID-19 phase, while the glutamine/glutamate ratio was increased in the post-COVID-  
279 19 phase.

#### 280 **Post-COVID-19 patients compared with controls**

281 In order to know if the altered metabolites (and those that remained insignificant) were  
282 dysregulated with respect to normal values, a group of negative SARS-CoV-2 controls (i.e.,  
283 healthy controls) collected from 2020 was added to the analysis. The heatmap (**Figure 2A**)  
284 shows the metabolites with significant differences across the three study groups (one way  
285 ANOVA). Volcano plots (**Figure 2B**) were used to identify the dysregulated metabolites in  
286 the post-COVID-19 patients with respect to the healthy controls. Two-sample t-tests &  
287 Wilcoxon rank-sum tests shows that in comparison with the healthy controls, 29  
288 metabolites were still dysregulated in post-COVID patients (**Supplementary Table 2**).  
289 C18:1(adjusted  $p = 4 \times 10^{-3}$ ), C18:2 (adjusted  $p = 6.1 \times 10^{-6}$ ), C10:2 (adjusted  $p = 7 \times 10^{-3}$ ),  
290 C10:1(adjusted  $p = 3 \times 10^{-2}$ ), carnitine (adjusted  $p = 4 \times 10^{-2}$ ) as well as glutamine (adjusted  
291  $p = 4.2 \times 10^{-5}$ ), choline (adjusted  $p = 5.7 \times 10^{-5}$ ), glucose (adjusted  $p = 4.2 \times 10^{-5}$ ),  
292 kynurenine (adjusted  $p = 7 \times 10^{-4}$ ), pyruvic acid (adjusted  $p = 1.0 \times 10^{-3}$ ),  
293 kynurenine/tryptophan ratio (adjusted  $p = 6.0 \times 10^{-03}$ ), threonine (adjusted  $p = 7 \times 10^{-3}$ ),  
294 putrescine (adjusted  $p = 0.01$ ), ornithine (adjusted  $p = 0.04$ ), PC ae 36:0 (adjusted  $p = 0.03$ ),

295 and PC aa 40:2 (adjusted  $p = 0.02$ ) were found in higher concentrations in post-COVID-19  
296 patients relative to the healthy controls.

297 In addition, lysoPC 14:0 (adjusted  $p = 0.01$ ), lysoPC 16:0 (adjusted  $p = 1.4 \times 10^{-11}$ ),  
298 lysoPC 16:1 (adjusted  $p = 1.0 \times 10^{-3}$ ), lysoPC 18:0 (adjusted  $p = 3.0 \times 10^{-6}$ ), LysoPC 17:0  
299 (adjusted  $p = 1.0 \times 10^{-4}$ ), LysoPC 20:4 (adjusted  $p = 2.0 \times 10^{-3}$ ), SM(OH)22:1 (adjusted  $p =$   
300 0.04), PC aa 32:2 (adjusted  $p = 0.01$ ), sarcosine (adjusted  $p = 1.04 \times 10^{-5}$ ), taurine (adjusted  
301  $p = 4.2 \times 10^{-5}$ ), and glutamic acid (adjusted  $p = 1.0 \times 10^{-3}$ ) were found downregulated in  
302 post-COVID-19 patients relative to the healthy controls. Likewise, the glutamine/glutamate  
303 ratio (adjusted  $p = 2.7 \times 10^{-6}$ ) was increased in post-COVID-19 patients, while the  
304 lactate/pyruvate ratio (adjusted  $p = 0.04$ ) was found to be decreased.

305 Glucose, kynurenine, C10:2, C18:1, C10:1, lysoPC14:0, lysoPC16:1, PC ae 36:0,  
306 and PC aa 32:2, PC aa 40:2, were found to be in similar concentration levels as for those in  
307 the COVID-19 phase.

308 Several other metabolites previously related with severity in COVID-19 tend  
309 towards normal or healthy levels (kynurenine/tryptophan ratio, C18:2, glutamic acid,  
310 glutamine, spermidine) in the post-COVID-19 group. Of note, a group of sphingomyelins  
311 (SM(OH)14:0, SM16:0, SM(18:0), SM(OH)16:1, SM(OH)24:1, SM(18:1), SM(16:1)) were  
312 found to be normalized, as well as lysoPC 26:0, lysoPC 26:1, lysoPC 28:1, and lysoPC  
313 28:0, phenylalanine, butyric acid, and propionic acid.

314 The multivariate analysis (PLS-DA) showed a clear separation between both classes  
315 (accuracy: 1;  $R^2$ : 0.98;  $Q^2$ : 0.89) (**Figure 2C**). The VIP score plot (**Figure 2D**) shows that  
316 the most important variables that can be used to differentiate negative controls from post-  
317 COVID-19 patients are glutamine/glutamate ratio, sarcosine, C18:2, taurine and  
318 LysoPC16:0.

319 A logistic regression model was built using both symptoms and plasma metabolites  
320 to distinguish the development of long COVID among COVID-19 patients. The AUC,  
321 sensitivity, and specificity values with 95% CI for this model are shown in Supplementary  
322 Table 3. A logistic regression model had the following equation:  $\text{logit}(P) = \log(P / (1 - P))$   
323  $= 1.13 - 1.724 \text{ myalgia} - 2.763 \text{ ornithine/citrulline} + 3.076 \text{ lactate/pyruvate}$ , where the  
324 numeric value of each named metabolite in the equation is the concentration after auto-  
325 scaling.

### 326 **Investigating post-COVID-19 patients.**

327 Differences were found in patients from the post-COVID-19 group, both in the frequency  
328 of symptoms reported and in the plasma levels of some metabolites such as lactic acid, with  
329 a bimodal distribution across the group (**Supplementary Figure 1**). Therefore, these  
330 patients were subclassified according to our own scale as a surrogate for disease severity.  
331 18 patients did not report any symptoms (recovered). 17 patients reported less than five  
332 persistent symptoms (class A long COVID), while 13 reported more than five symptoms  
333 (class B long COVID).

334 **Figure 4** shows the box and whisker plots based on one way ANOVA for class A,  
335 class B, and fully recovered patients. The lactate/pyruvate ratio (adjusted p value =  $5.8 \times$   
336  $10^{-7}$ ), lactate (adjusted p value =  $4.8 \times 10^{-6}$ ), arginine ( $1.8 \times 10^{-3}$ ), ornithine/citrulline ratio  
337 (adjusted p value  $5. \times 10^{-3}$ ), and sarcosine (adjusted p value = 0.02) were the variables best  
338 able to differentiate long COVID patients with more than five symptoms from patients with  
339 less than five symptoms. Arginine and sarcosine negatively correlated with the number of  
340 symptoms.

341 For differentiating class B long COVID patients from all other post-COVID-19  
342 patients, the lactate/pyruvate ratio had the best performance (AUC: 0.95 (0.92 -0.97),



343 sensitivity: 0.92 (0.87-0.97), specificity: 0.94 (0.91-0.98)), followed by the combination of  
344 the ornithine/citrulline ratio and uric acid (AUC: 0.92 (0.89-0.95), sensitivity: 0.83 (0.77-  
345 0.90), specificity: 0.84 (0.79-0.90)).

346 Class B patients reported symptoms falling in all the categories, while class A  
347 patients reported mainly neuropsychiatric symptoms. We also wanted to know whether the  
348 classification based on the group of symptoms would be most appropriate rather than the  
349 classification based on the number of symptoms. 18 patients had systemic symptoms, while  
350 13 patients reported mostly neuropsychiatric associated disorders. Lactic acid, arginine, the  
351 lactate/pyruvate ratio, the ornithine/citrulline ratio and sarcosine were common in both  
352 types of long COVID classifications. Long saturated or monounsaturated LysoPCs were  
353 found to be increased in patients with neuropsychiatric disorders (**Supplementary Figure**  
354 **2**).

### 355 **Pathway analysis**

356 Our pathway enrichment analysis (**Figure 5**) shows that the top five metabolic pathways  
357 significantly dysregulated (FDR<0.05) in post-COVID patients (relative to controls) were:  
358 phospholipids biosynthesis, gluconeogenesis, the glucose-alanine cycle, the Warburg effect,  
359 and taurine and hypotaurine metabolism. When comparing class B patients with those  
360 recovered, the top five metabolic pathways (FDR<0.05) were: pyruvate metabolism,  
361 gluconeogenesis, glycine and serine metabolism, urea cycle metabolism, and the Warburg  
362 effect.

### 363 **Plasma IL-17 and leptin**

364 **Figure 6** shows plasma concentrations of IL-17 and leptin as measured by ELISA. IL-17  
365 was significantly increased in class B patients relative to class A patients (Mann-Whitney

366 test,  $p = 0.0073$ ) and recovered patients (Mann-Whitney test,  $p = 0.002$ ). Leptin did not  
367 show any statistically significant differences in the three-group comparison.

368

## 369 **Discussion**

370 Cumulative evidence from the last three years supports the dysregulation of metabolic and  
371 immune markers due to SARS-CoV-2 infection [20]. A retrospective cohort study has  
372 demonstrated that COVID-19 patients have a significantly higher risk to develop  
373 subsequent autoimmune diseases such as rheumatoid arthritis, ankylosing spondylitis,  
374 systemic sclerosis, type I diabetes mellitus, among others [21]. In the present work, our aim  
375 was to evaluate the persistence of long-term metabolic alterations in post-COVID-19  
376 patients, as well as to measure immune markers that, when chronically produced, can  
377 trigger autoimmune diseases.

378         Since well-defined classification or diagnostic criteria are not available for the long  
379 COVID assessment, there is an urgent need for molecular methods able to stratify patients  
380 according to the severity of the symptoms they are experiencing. Quantitative and validated  
381 scales, such as HAM-A, HAM-D, MoCA and mMRC are considered gold standards for  
382 neurocognitive impairment and for dyspnea assessment. However, their practical utility  
383 could be limited for complex conditions such as long COVID where a broader range of  
384 self-reported symptoms with different severity and duration are present. It has been  
385 reported that some post-COVID-19 patients complain about extreme cognitive disorders  
386 (self-reported symptoms) but without any objective alterations, while others do not report  
387 symptoms but exhibit severe cognitive disorders after six to nine months following SARS-  
388 CoV-2 infection [22]. In our study, several symptoms were corroborated through objective

389 measures, but with lower rates when using validated scales. Therefore, molecular markers  
390 are urgently needed for the correct classification of patients.

391 Our results revealed that 50% of analyzed plasma metabolites showed statistical  
392 differences between COVID-19 and post-COVID-19 phases. One of the most dysregulated  
393 metabolites was glucose. Montefusco *et al.* [23] reported glycemc abnormalities in  
394 recovered patients two months after the onset of disease. The hyperglycemic state has been  
395 reported to be even worse in hospitalized patients, pointing to a possible causal role of  
396 administered drug regimens, including remdesivir and corticosteroids. These drugs  
397 stimulate hepatic gluconeogenesis from amino acids released from muscles, which then  
398 inhibits glucose uptake [24].

399 A number of other metabolites were also found to be dysregulated. Increased  
400 plasma pyruvate levels could be both a consequence of glycolytic dysregulation and protein  
401 degradation. The increase in putrescine levels in the post-COVID phase may be an  
402 indicator of increased protein degradation to help fuel pyruvate metabolism.

403 Taurine and spermidine were found significantly decreased in the post-COVID  
404 phase, although a trend towards normalization was observed when compared with controls.  
405 Decreased levels of serum taurine have been observed in patients with Myalgic  
406 Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) [25]. The depleted levels  
407 observed in post-COVID-19 phase could explain at least in part the fatigue, since taurine  
408 has multiple roles in skeletal muscle, the central nervous system, and energy metabolism.  
409 Nevertheless, based on our results, we did not find any correlation between fatigue or  
410 myalgia and taurine concentrations in post-COVID-19 patients. Holmes *et al.* [26] found  
411 that taurine levels were increased in post-COVID-19 patients, suggesting hepatic injury,  
412 hepatotoxicity, or muscle damage. However, the cohort evaluated in the Holmes study had

413 a three-month follow-up after the initial infection, which is much shorter than the follow-up  
414 used in this study.

415 Furthermore, we observed increased levels of kynurenine (similar to levels in the  
416 COVID-19 acute phase), and a trend towards normalization in tryptophan and the  
417 kynurenine/tryptophan ratio in post-COVID-19 patients. This indicates that, although lower  
418 in magnitude, the inflammatory conditions attributable to the hyperactivation of this  
419 metabolic pathway are still present and may account for some persistent physiological  
420 symptoms in these patients.

421 The increase in glutamine (and decrease in glutamate levels) indicates a partial  
422 reestablishment of critical processes that took place during the COVID-19 infection phase,  
423 such as severe immunometabolic dysregulation. It is well known that SARS-CoV-2 induces  
424 metabolic reprogramming in host cells, similar to the Warburg effect in cancer, and a  
425 depletion of glutamine has been associated with its consumption for feeding TCA, and as a  
426 nitrogen source for nucleotide (ATP) synthesis [27]. On the other hand, glutamate is the  
427 most abundant neurotransmitter in the brain. A disruption within the glutamatergic  
428 pathway can lead to important neurological consequences, such as cognitive deficits [28].  
429 The glutaminergic dysfunction could be associated with some psychiatric and neurologic  
430 symptoms like those reported in the present work.

431 Alterations in lipid metabolism are evident in most post-COVID-19 patients. These  
432 patients exhibited significantly higher levels of carnitine and some short, medium, and long  
433 acylcarnitines. These alterations have been largely associated with altered fatty acid  
434 metabolism, dysfunctional mitochondria-dependent lipid catabolism, and immune  
435 processes or the lysis of white blood cells. Similar results have been reported by Guntur *et*  
436 *al.* [29], pointing to mitochondrial dysfunction, as was also recognized during COVID-19

437 acute phase. Besides, decreased levels of LysoPC 16:0, LysoPC 17:0, LysoPC 18:0, and  
438 LysoPC 20:4, were found with respect to negative controls. These reductions have been  
439 reported in other inflammatory conditions [30] and other septic processes [31]. Depleted  
440 levels of lysophosphatidylcholines and phospholipid ethers, as well as depleted levels of  
441 PCs, can impede mitochondrial respiration, as has been also demonstrated in ME/CFS [32].  
442 In line with the lipid dysregulation demonstrated by the targeted metabolomic analysis,  
443 routine clinical laboratory tests exhibited elevated levels of total cholesterol, triglycerides,  
444 and VLDL, as well as normal levels of HDL and LDL. Xu *et al.* [33] found increased LDL,  
445 triglycerides, total cholesterol, and decreased HDL in survivors of COVID-19, based on a  
446 large observational study with participants from the US Department of Veterans Affairs  
447 database compared to controls who had never tested positive for COVID-19.

448 As positive findings for the metabolic state of post COVID patients, we found that  
449 30 metabolites fell within normal levels. Phenylalanine, which has been widely associated  
450 with sepsis and COVID disease severity [20] decreased to normal levels. Beta-  
451 hydroxybutyric acid and citric acid were also normalized, indicating partial recovery of the  
452 tricarboxylic acid cycle [20, 34]. Butyric acid and propionic acid, two short-chain fatty  
453 acids that were found to be altered during COVID-19 phase, also fell within normal levels  
454 in post-COVID-19 patients, probably indicating that the leaky gut phenomenon and gut  
455 dysbiosis detected during COVID infection could be partially reestablished [35].  
456 Spermidine was also normalized. The decrease in spermidine levels could reflect a trend for  
457 normalization in overall redox balance. Although excessive levels of spermidine (as those  
458 reported in COVID-19 patients) trigger the production of superoxide radicals, optimal  
459 concentrations mitigate oxidative stress and diminish overall ROS production [36].

460 In addition, sphingomyelins and long-chain monounsaturated and saturated  
461 LysoPCs were found to be within normal levels. We previously noted altered sphingolipids  
462 levels during COVID-19 infection [20]. Sphingolipids play a crucial role in the regulation  
463 of signal transduction pathways and in certain pathological conditions, such as  
464 inflammation-associated illnesses and innate immune response.

465 In a recent report, Holmes *et al.* [26] found a high degree of interindividual  
466 variability in follow-up patients, reflecting the heterogeneity of post-COVID-19 patients  
467 and the fact that long COVID is a spectrum of disorders. Indeed, computationally modeling  
468 of the long COVID phenotype data based on electronic healthcare records found six distinct  
469 clusters, each with distinct profiles of phenotypic abnormalities [37]. Since symptom  
470 classification is still highly subjective, we decided to arbitrarily classify long COVID  
471 patients as: class A (less than five symptoms, mainly neuropsychiatric disorders), and class  
472 B (more than five symptoms, with a broad spectrum of systemic disorders).

473 We believe that metabolic information may complement, and partially explain the  
474 phenotypic differences among post-COVID-19 patients, and especially in long COVID  
475 patients. Xu *et al.* [38] classified recovered patients based on abnormal pulmonary  
476 functions, finding increased levels of triacylglycerols, phosphatidylcholines, prostaglandin  
477 E2, arginine, and decreased levels of betaine and adenosine in patients with abnormal  
478 pulmonary function.

479 In our work, lactic acid levels were increased in patients with more than five  
480 symptoms and systemic disorders (class B patients). Ghali *et al.* [39] found that patients  
481 with ME/CFS exhibited elevated blood lactate at rest. Mitochondrial dysfunction, with  
482 increased blood lactate, low levels of ATP, and increased levels of oxidative stress markers  
483 have been associated to these alterations [40], as well as relative deficiency of mitochondria

484 type I fibers on muscle biopsies, and low intracellular pH during recovery phase [41–43].  
485 De Boer *et al.* [44] also reported altered lactate levels in long COVID patients, suggesting  
486 that long COVID patients have significant impairment in fat beta-oxidation and  
487 increased blood lactate accumulation even during low-intensity exercise. In contrast,  
488 Guntur *et al.* [29] reported low levels of lactic acid and pyruvate in long COVID patients.  
489 However, this study was conducted in non-hospitalized patients who had recovered from  
490 COVID-19 in March 2020.

491 Increased level of the lactate/pyruvate ratio in class B patients is another important  
492 indicator of mitochondrial dysfunction. The lactate/pyruvate ratio has been proposed as a  
493 marker for mitochondrial disorders since it indirectly reflects the NADH/NAD<sup>+</sup> redox  
494 state[45], lipid metabolism (fat oxidation), and ATP generation. In our study, both markers  
495 (lactate and the lactate/pyruvate ratio) were found positively correlated with fatigue,  
496 myalgia and arthralgias (Spearman correlation,  $R > 0.6$ ,  $p < 0.05$ ) (**Supplementary Figure**  
497 **3**).

498 The increased ornithine/citrulline ratio level in class B patients reflects abnormal  
499 metabolic activity in the urea cycle. It is notable that Yamano *et al.* [46] reported a  
500 similarly increased ornithine/citrulline ratio in CFS patients. An adequate balance of  
501 citrulline and ornithine is vital for the clearance of ammonia via urea cycle [47]. If  
502 ammonia accumulates intracellularly, the aerobic utilization of pyruvate to feed the TCA  
503 cycle is inhibited, resulting in lactate production, which further contributes to fatigue.

504 In addition, class B patients had decreased levels of arginine. The reduced  
505 bioavailability of arginine to produce adequate levels of NO in endothelial cells and  
506 vascular tissues leads to the impairment of multiple physiological functions of skeletal  
507 muscles, including contractile functions, and muscle repair. This decreased level is not a

508 residual effect of COVID-19, as the paired study (COVID-19/post-COVID-19) showed  
509 normal arginine levels. Arginine is also a substrate for ornithine production by arginase. It  
510 is well known that under certain inflammatory conditions, arginase activity is increased  
511 [48], producing an excess of ornithine and an imbalance in the urea cycle.

512 Sarcosine was found decreased in class B patients. Previously, Fraser *et al.* [49]  
513 found that sarcosine was depressed in COVID-19 patients. Sarcosine plays a vital role in  
514 immune functions, as it activates autophagy and the removal of damaged cells. A reduced  
515 amount of plasma sarcosine could in part lead to a sustained inflammatory process.

516 Previous studies have pointed to the persistent immune dysregulation following  
517 COVID-19 infection [50]. We found increased levels of monocytes in class B patients.  
518 Nuber-Champier *et al.* [51] found that monocyte percentage in the acute phase of the  
519 disease allowed them to distinguish between patients with anosognosia for memory deficits  
520 in the chronic phase (6–9 months after SARS-CoV-2 infection) and nosognosic patients.

521 We also measured IL-17 levels in post-COVID-19 patients since it is well known  
522 that this cytokine is persistently altered in several chronic inflammatory and autoimmune  
523 diseases [52], and previous reports have indicated an increased risk of such diseases in  
524 COVID-19 patients [21]. IL-17 is a proinflammatory cytokine mainly produced by T helper  
525 type 17 cells, playing a vital role in the regulation of host immune response against SARS-  
526 CoV-2. IL-17-induced dysregulated immune responses have been shown to potentially  
527 cause hyperinflammatory COVID-19 disease [53]. It has been reported that IL-17  
528 downregulates protein phosphatase 6, resulting in increased arginase-1 expression in  
529 psoriatic keratinocytes [54]. IL-17A has been found to be associated with neurological  
530 sequelae and pulmonary fibrosis in post-COVID-19 patients [55, 56]. Fluctuations in IL-17  
531 have been associated with fatigue and fatigue severity in ME/CFS patients [57].



532 Metabolomics is not only useful in providing a snapshot of transient physiological  
533 or pathophysiological processes taking place in a living organism, but it has also proven to  
534 be a powerful tool for proposing and monitoring therapeutic interventions. In the case of  
535 long COVID, a common situation worldwide is that patients have reported an absence of  
536 adequate support and a poor recognition of their condition, initially attributed to psychiatric  
537 issues. People with long COVID have tried a vast range of self-prescribed medicines,  
538 supplements, remedies, and dietary changes to manage the disease and to overcome the  
539 effects it has on their quality of life and work capacity. Based on our findings, some  
540 interventions could be tested for treating long COVID patients: 1) supplementation of  
541 taurine (reducing musculoskeletal disorders); 2) supplementation of arginine and/or  
542 citrulline (enhancing ammonia clearance and reducing blood lactate, as well as increasing  
543 arginine bioavailability for adequate NO production); 3) supplementation of glutamine  
544 (primary source for neurotransmitters and immune function balancing); 4) supplementation  
545 of antioxidants such as N-acetylcysteine or NAD<sup>+</sup> (redox balance). Similarities found in  
546 our results with the ME/CFS pathophysiology may pave the way to common therapeutic  
547 interventions for both diseases.

548 We need to acknowledge several limitations with this study. The small sample size  
549 was due to the limited number of patients who agreed to participate. While several  
550 objective measures of mood and cognition (HAM-A, HAM-D, MoCA mMRC) were used,  
551 the sample size did not allow for stratification of patients according to the different test  
552 scores obtained, and only self-reported symptoms were used for sub-group classification.  
553 Furthermore, we were unable to have a detailed tracking of treatments, medications or  
554 alternative therapies during the period evaluated. This limited our interpretation with regard  
555 to the impact of pharmacological interventions on the metabolome. Moreover, patients

556 from two different hospitals participated: one private hospital located in Chihuahua city,  
557 and one public hospital located in Zacatecas city. In general, private hospital patients have  
558 relatively high incomes while public hospital patients have lower incomes. Logistic  
559 regression models showed no effects of sex, age, comorbidities, vaccination status or  
560 severity during the acute phase in the metabolomic profile associated with long COVID.  
561 However, patients from the public hospital reported more systemic symptoms in general,  
562 while patients from the private hospital reported principally neuropsychiatric symptoms. A  
563 recent study found that patients diagnosed with a post-COVID-19 condition were more  
564 likely to be unemployed or on public health insurance, illustrating racial and social  
565 disparities in access to and experience with healthcare, at least in the USA [58]. Whether  
566 the socioeconomic conditions and lifestyles, along with causes of biological origin  
567 influence the metabolic phenoreversion of patients recruited in our study, needs to be  
568 further investigated. This is particularly important in countries with significant health  
569 system disparities and significant differences in population life conditions.

570

## 571 **Conclusions**

572 To our knowledge, this study is the first describing quantitative metabolic perturbations two  
573 years after the initial acute COVID-19 infection using targeted metabolomics. The  
574 evolution of post-COVID-19 patients is different, and symptoms are associated to  
575 distinctive metabolic patterns resembling, to some extent, the ME/CSF condition.  
576 Moreover, the differences observed between the phenotypes of post-COVID-19 patients  
577 reveals potential biomarkers that could enable a more accurate and precise molecular  
578 classification of long COVID patients beyond classification via self-reported symptoms.

## 579 **List of abbreviations**

- 580 COVID-19: Coronoravirus Disease 2019
- 581 LC-MS: Liquid Chromatography-mass spectrometry
- 582 FIA: Flow Injection Analysis
- 583 LysoPCs: Lysophosphatidylcholines
- 584 PACS: Post-Acute COVID-19 Syndrome
- 585 ME/CFS: Myalgic Encephalomyelitis/Chronic Fatigue Syndrome
- 586 IL-17: Interleukin 17
- 587 PC: Phosphatidylcholine
- 588 SM: Sphingomyelin
- 589 TCA: Tricarboxylic acid
- 590 HAM-A: Hamilton Anxiety Rating Scale
- 591 HAM-D: Hamilton Depression scale
- 592 MoCA: The Montreal Cognitive Assessment
- 593 mMRC: Medical Research Council

594

595 **Availability of data and materials:** The datasets generated and analyzed during the  
596 current study are available in the Mendeley repository (<https://data.mendeley.com>) (doi:  
597 10.17632/8zfdjsypd8.1).

598 **Ethics approval and consent to participate:** The study was revised and approved by the  
599 Research and Ethics Committees of the Instituto Mexicano de Seguridad Social, with the  
600 registration number R-2022-3301-038, and Christus Muguerza del Parque Hospital  
601 (HCMP-CEI-15042020-3, and HCMP-CEI-28022022-A01).

602 **Consent for publication:** Not required.

603 **Competing interests:** The authors declare that they have no competing interests.

604 **Funding:** CONACyT grant number 311880 “Identificación y validación de marcadores  
605 inmuno-metabólicos que incrementan la susceptibilidad a desarrollar formas graves de  
606 infección por SARS-CoV-2”.

607 CONACyT grant number 319503 “Tamizaje masivo de pequeñas moléculas para la  
608 identificación de biomarcadores y candidatos terapéuticos con alto potencial de efectividad  
609 contra nuevas variantes de SARS-CoV-2 circulantes a nivel mundial”.

610 Christus Muguerza del Parque Hospital, Chihuahua, Mexico.

611 Instituto Mexicano de Seguridad Social.

612 Genome Alberta (a division of Genome Canada) grant number TMIC MC4, the Canadian  
613 Institutes of Health Research (CIHR) grant number FS 148461, the Canada Foundation for  
614 Innovation (CFI) grant number MSIF 35456.

#### 615 **Authors' contributions**

616 Study concept and design: YLH and JMA. JCB, DAGL recruited study subjects and  
617 collected samples. Data acquisition: JZ, JPS, RM and EMM. Statistical analysis: MB, RM.  
618 Data interpretation: YLH, MB, JMA, JAL, CTC, DSW. Drafting the manuscript: YLH.  
619 Manuscript editing: CTC. DSW supervised the whole research and revised the original  
620 manuscript. All the authors contributed to the discussion and gave approval to the final  
621 version of the manuscript.

622

#### 623 **Acknowledgements**

624 JPE-D is a postdoctoral researcher at the Instituto Nacional de Medicina Genómica and  
625 received a fellowship from CONACYT in the program “Estancias Posdoctorales por  
626 México en Apoyo por SARS CoV2 (COVID-19, 2166969)”.

627

628

629

## 630 References

- 631 1. Lam W, Zhong N, Tan W. Overview on SARS in Asia and the World. *Respirology*.  
632 2003;8:S2–5.
- 633 2. Al Mutair A, Ambani Z. Narrative review of Middle East respiratory syndrome  
634 coronavirus (MERS-CoV) infection: updates and implications for practice. *Journal of*  
635 *International Medical Research*. 2020;48:030006051985803.
- 636 3. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings,  
637 mechanisms and recommendations. *Nat Rev Microbiol*. 2023.  
638 <https://doi.org/10.1038/s41579-022-00846-2>.
- 639 4. Chopra V, Flanders SA, O'Malley M, Malani AN, Prescott HC. Sixty-Day Outcomes  
640 Among Patients Hospitalized With COVID-19. *Ann Intern Med*. 2021;174:576–8.
- 641 5. Wallace PG. Epidemiology: A critical review. *Br Med Bull*. 1991;47:942–51.
- 642 6. Hickie I, Davenport T, Wakefield D, Vollmer-Conna U, Cameron B, Vernon SD, et al.  
643 Post-infective and chronic fatigue syndromes precipitated by viral and non-viral pathogens:  
644 prospective cohort study. *BMJ*. 2006;333:575.
- 645 7. Kukla M, Skonieczna-Żydecka K, Kotfis K, Maciejewska D, Łoniewski I, Lara LuisF, et  
646 al. COVID-19, MERS and SARS with Concomitant Liver Injury—Systematic Review of  
647 the Existing Literature. *J Clin Med*. 2020;9:1420.
- 648 8. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz J V. A clinical case definition of post-  
649 COVID-19 condition by a Delphi consensus. *Lancet Infect Dis*. 2022;22:e102–7.
- 650 9. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The  
651 prevalence and long-term health effects of Long Covid among hospitalised and non-  
652 hospitalised populations: a systematic review and meta-analysis. *EClinicalMedicine*.  
653 2023;55:101762.
- 654 10. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio  
655 A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-  
656 analysis. *Sci Rep*. 2021;11:16144.
- 657 11. Wu X, Liu X, Zhou Y, Yu H, Li R, Zhan Q, et al. 3-month, 6-month, 9-month, and 12-  
658 month respiratory outcomes in patients following COVID-19-related hospitalisation: a  
659 prospective study. *Lancet Respir Med*. 2021;9:747–54.
- 660 12. Schlemmer F, Valentin S, Boyer L, Guillaumot A, Chabot F, Dupin C, et al.  
661 Respiratory recovery trajectories after severe-to-critical COVID-19: a 1-year prospective  
662 multicentre study. *European Respiratory Journal*. 2023;61:2201532.
- 663 13. Pintos-Pascual I, Moreno-Torres V, Ibáñez-Estélez F, Corrales-Rodríguez P, Treviño  
664 A, Corpas M, et al. Is SARS-CoV-2 the only cause of long-COVID? *AIDS Rev*. 2022;24.
- 665 14. López-Hernández Y, Monárrez-Espino J, Oostdam A-SH, Delgado JEC, Zhang L,  
666 Zheng J, et al. Targeted metabolomics identifies high performing diagnostic and prognostic  
667 biomarkers for COVID-19. *Sci Rep*. 2021;11:14732.
- 668 15. Shear MK, Vander Bilt J, Rucci P, Endicott J, Lydiard B, Otto MW, et al. Reliability  
669 and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-  
670 A). *Depress Anxiety*. 2001;13:166–78.
- 671 16. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*.  
672 1960;23:56–62.
- 673 17. Nasreddine ZS, Phillips NA, BÃ©dirian V, Charbonneau S, Whitehead V, Collin I, et  
674 al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild  
675 Cognitive Impairment. *J Am Geriatr Soc*. 2005;53:695–9.

- 676 18. Mahler DA, Wells CK. Evaluation of Clinical Methods for Rating Dyspnea. *Chest*.  
677 1988;93:580–6.
- 678 19. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for  
679 metabolomic data analysis and interpretation. *Nucleic Acids Res*. 2009;37 Web  
680 Server:W652–60.
- 681 20. Herrera-Van Oostdam AS, Castañeda-Delgado JE, Oropeza-Valdez JJ, Borrego JC,  
682 Monárrez-Espino J, Zheng J, et al. Immunometabolic signatures predict risk of progression  
683 to sepsis in COVID-19. *PLoS One*. 2021;16:e0256784.
- 684 21. Chang R, Yen-Ting Chen T, Wang S-I, Hung Y-M, Chen H-Y, Wei C-CJ. Risk of  
685 autoimmune diseases in patients with COVID-19: a retrospective cohort study.  
686 *EClinicalMedicine*. 2023;56:101783.
- 687 22. Almeria M, Cejudo JC, Sotoca J, Deus J, Krupinski J. Cognitive profile following  
688 COVID-19 infection: Clinical predictors leading to neuropsychological impairment. *Brain*  
689 *Behav Immun Health*. 2020;9:100163.
- 690 23. Montefusco L, Ben Nasr M, D’Addio F, Loretelli C, Rossi A, Pastore I, et al. Acute and  
691 long-term disruption of glycometabolic control after SARS-CoV-2 infection. *Nat Metab*.  
692 2021;3:774–85.
- 693 24. Negahdaripour M. Post-COVID-19 Hyperglycemia: A Concern in Selection of  
694 Therapeutic Regimens. *Iran J Med Sci*. 2021;46:235–6.
- 695 25. Germain A, Ruppert D, Levine SM, Hanson MR. Metabolic profiling of a myalgic  
696 encephalomyelitis/chronic fatigue syndrome discovery cohort reveals disturbances in fatty  
697 acid and lipid metabolism. *Mol Biosyst*. 2017;13:371–9.
- 698 26. Holmes E, Wist J, Masuda R, Lodge S, Nitschke P, Kimhofer T, et al. Incomplete  
699 Systemic Recovery and Metabolic Phenoreversion in Post-Acute-Phase Nonhospitalized  
700 COVID-19 Patients: Implications for Assessment of Post-Acute COVID-19 Syndrome. *J*  
701 *Proteome Res*. 2021;20:3315–29.
- 702 27. Aydın H, Tekin YK, Korkmaz İ, Tekin G, Yurtbay S, Keleş S, et al. Glutamine-Driven  
703 Metabolic Adaptation to COVID-19 Infection. *Indian Journal of Clinical Biochemistry*.  
704 2023;38:83–93.
- 705 28. Yesilkaya UH, Sen M, Balcioglu YH. COVID-19-related cognitive dysfunction may be  
706 associated with transient disruption in the DLPFC glutamatergic pathway. *Journal of*  
707 *Clinical Neuroscience*. 2021;87:153–5.
- 708 29. Guntur VP, Nemkov T, de Boer E, Mohning MP, Baraghoshi D, Cendali FI, et al.  
709 Signatures of Mitochondrial Dysfunction and Impaired Fatty Acid Metabolism in Plasma of  
710 Patients with Post-Acute Sequelae of COVID-19 (PASC). *Metabolites*. 2022;12:1026.
- 711 30. Engel KM, Schiller J, Galuska CE, Fuchs B. Phospholipases and Reactive Oxygen  
712 Species Derived Lipid Biomarkers in Healthy and Diseased Humans and Animals – A  
713 Focus on Lysophosphatidylcholine. *Front Physiol*. 2021;12.
- 714 31. Drobnik W, Liebisch G, Audebert F-X, Fröhlich D, Glück T, Vogel P, et al. Plasma  
715 ceramide and lysophosphatidylcholine inversely correlate with mortality in sepsis patients.  
716 *J Lipid Res*. 2003;44:754–61.
- 717 32. Che X, Brydges CR, Yu Y, Price A, Joshi S, Roy A, et al. Metabolomic Evidence for  
718 Peroxisomal Dysfunction in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Int J*  
719 *Mol Sci*. 2022;23:7906.
- 720 33. Xu E, Xie Y, Al-Aly Z. Risks and burdens of incident dyslipidaemia in long COVID: a  
721 cohort study. *Lancet Diabetes Endocrinol*. 2023;11:120–8.



- 722 34. Albóniga OE, Jiménez D, Sánchez-Conde M, Vizcarra P, Ron R, Herrera S, et al.  
723 Metabolic Snapshot of Plasma Samples Reveals New Pathways Implicated in SARS-CoV-  
724 2 Pathogenesis. *J Proteome Res.* 2022;21:623–34.
- 725 35. Kim HS. Do an Altered Gut Microbiota and an Associated Leaky Gut Affect COVID-  
726 19 Severity? *mBio.* 2021;12.
- 727 36. Firpo MR, Mastrodomenico V, Hawkins GM, Prot M, Levillayer L, Gallagher T, et al.  
728 Targeting Polyamines Inhibits Coronavirus Infection by Reducing Cellular Attachment and  
729 Entry. *ACS Infect Dis.* 2021;7:1423–32.
- 730 37. Reese JT, Blau H, Casiraghi E, Bergquist T, Loomba JJ, Callahan TJ, et al.  
731 Generalisable long COVID subtypes: findings from the NIH N3C and RECOVER  
732 programmes. *EBioMedicine.* 2023;87:104413.
- 733 38. Xu J, Zhou M, Luo P, Yin Z, Wang S, Liao T, et al. Plasma Metabolomic Profiling of  
734 Patients Recovered From Coronavirus Disease 2019 (COVID-19) With Pulmonary  
735 Sequelae 3 Months After Discharge. *Clinical Infectious Diseases.* 2021;73:2228–39.
- 736 39. Ghali A, Lacout C, Ghali M, Gury A, Beucher A-B, Lozac’h P, et al. Elevated blood  
737 lactate in resting conditions correlate with post-exertional malaise severity in patients with  
738 Myalgic encephalomyelitis/Chronic fatigue syndrome. *Sci Rep.* 2019;9:18817.
- 739 40. Kennedy G, Spence VA, McLaren M, Hill A, Underwood C, Belch JJJ. Oxidative  
740 stress levels are raised in chronic fatigue syndrome and are associated with clinical  
741 symptoms. *Free Radic Biol Med.* 2005;39:584–9.
- 742 41. Lane RJM, Barrett MC, Woodrow D, Moss J, Fletcher R, Archard LC. Muscle fibre  
743 characteristics and lactate responses to exercise in chronic fatigue syndrome. *J Neurol*  
744 *Neurosurg Psychiatry.* 1998;64:362–7.
- 745 42. Lane RJM, Barrett MC, Taylor DJ, Kemp GJ, Lodi R. Heterogeneity in chronic fatigue  
746 syndrome: evidence from magnetic resonance spectroscopy of muscle. *Neuromuscular*  
747 *Disorders.* 1998;8:204–9.
- 748 43. Petrović V, Buzadžić B, Korać A, Vasiljević A, Janković A, Mićunović K, et al.  
749 Antioxidative defence alterations in skeletal muscle during prolonged acclimation to cold:  
750 role of  $\text{L-arginine/NO}$ -producing pathway. *Journal of Experimental Biology.*  
751 2008;211:114–20.
- 752 44. de Boer E, Petrache I, Goldstein NM, Olin JT, Keith RC, Modena B, et al. Decreased  
753 Fatty Acid Oxidation and Altered Lactate Production during Exercise in Patients with Post-  
754 acute COVID-19 Syndrome. *Am J Respir Crit Care Med.* 2022;205:126–9.
- 755 45. Debray F-G, Mitchell GA, Allard P, Robinson BH, Hanley JA, Lambert M. Diagnostic  
756 Accuracy of Blood Lactate-to-Pyruvate Molar Ratio in the Differential Diagnosis of  
757 Congenital Lactic Acidosis. *Clin Chem.* 2007;53:916–21.
- 758 46. Yamano E, Sugimoto M, Hirayama A, Kume S, Yamato M, Jin G, et al. Index markers  
759 of chronic fatigue syndrome with dysfunction of TCA and urea cycles. *Sci Rep.*  
760 2016;6:34990.
- 761 47. Walker V. Severe hyperammonaemia in adults not explained by liver disease. *Annals of*  
762 *Clinical Biochemistry: International Journal of Laboratory Medicine.* 2012;49:214–28.
- 763 48. Saligan LN, Lukkahatai N, Jin Zhang Z, Cheung CW, Min Wang X. Altered Cd8+ T  
764 lymphocyte Response Triggered by Arginase 1: Implication for Fatigue Intensification  
765 during Localized Radiation Therapy in Prostate Cancer Patients. *Neuropsychiatry.* 2018;08.
- 766 49. Fraser DD, Slessarev M, Martin CM, Daley M, Patel MA, Miller MR, et al.  
767 Metabolomics Profiling of Critically Ill Coronavirus Disease 2019 Patients: Identification  
768 of Diagnostic and Prognostic Biomarkers. *Crit Care Explor.* 2020;2:e0272.



- 769 50. Ryan FJ, Hope CM, Masavuli MG, Lynn MA, Mekonnen ZA, Yeow AEL, et al. Long-  
770 term perturbation of the peripheral immune system months after SARS-CoV-2 infection.  
771 *BMC Med.* 2022;20:26.
- 772 51. Nuber-Champier A, Voruz P, Jacot de Alcântara I, Breville G, Allali G, Lalive PH, et  
773 al. Monocytosis in the acute phase of SARS-CoV-2 infection predicts the presence of  
774 anosognosia for cognitive deficits in the chronic phase. *Brain Behav Immun Health.*  
775 2022;26:100511.
- 776 52. McGeachy MJ, Cua DJ, Gaffen SL. The IL-17 Family of Cytokines in Health and  
777 Disease. *Immunity.* 2019;50:892–906.
- 778 53. Garmendia JV, García AH, De Sanctis CV, Hajdúch M, De Sanctis JB. Autoimmunity  
779 and Immunodeficiency in Severe SARS-CoV-2 Infection and Prolonged COVID-19. *Curr*  
780 *Issues Mol Biol.* 2022;45:33–50.
- 781 54. Lou F, Sun Y, Xu Z, Niu L, Wang Z, Deng S, et al. Excessive Polyamine Generation in  
782 Keratinocytes Promotes Self-RNA Sensing by Dendritic Cells in Psoriasis. *Immunity.*  
783 2020;53:204-216.e10.
- 784 55. Bazdyrev E, Rusina P, Panova M, Novikov F, Grishagin I, Nebolsin V. Lung Fibrosis  
785 after COVID-19: Treatment Prospects. *Pharmaceuticals.* 2021;14:807.
- 786 56. Saini L, Krishna D, Tiwari S, Goyal JP, Kumar P, Khera D, et al. Post-COVID-19  
787 Immune-Mediated Neurological Complications in Children: An Ambispective Study.  
788 *Pediatr Neurol.* 2022;136:20–7.
- 789 57. Montoya JG, Holmes TH, Anderson JN, Maecker HT, Rosenberg-Hasson Y, Valencia  
790 IJ, et al. Cytokine signature associated with disease severity in chronic fatigue syndrome  
791 patients. *Proceedings of the National Academy of Sciences.* 2017;114.
- 792 58. Pfaff ER, Madlock-Brown C, Baratta JM, Bhatia A, Davis H, Girvin A, et al. Coding  
793 long COVID: characterizing a new disease through an ICD-10 lens. *BMC Med.*  
794 2023;21:58.  
795
- 796

797 **Figure Legends**

798 **Figure 1:** Most common symptoms remaining after 24 months in 48 post-COVID-19  
799 patients. A) Distribution of symptoms according to organ systems. B) Results obtained via  
800 in-person questionnaire.

801

802 **Figure 2:** Multivariate analysis. A) The volcano plot of the plasma metabolomics between  
803 the COVID-19 phase and the post-COVID-19 phase (red represents the up-regulated  
804 metabolites compared with COVID-19 phase, green represents the downregulated  
805 metabolites compared with COVID-19 phase, and gray represents the metabolites with no  
806 difference between both groups. Fold change threshold  $\geq 1.3$  and p-value = 0.05 (FDR  
807 adjusted). B) Representative heatmap of top 50 significant metabolites (t-test) in the  
808 comparison of COVID-19 and post-COVID-19 phases (red: post-COVID-19 phase; green:  
809 COVID-19 phase). C) Score scatter plot based on the two-dimensional PLS-DA (red: post-  
810 COVID-19 phase; green: COVID-19 phase) D) Rank of the different metabolites (the top  
811 15) identified by the PLS-DA according to the VIP score on the x-axis. The most  
812 discriminating metabolites are shown in descending order of their coefficient scores. The  
813 color boxes indicate whether metabolite concentration is increased (red) or decreased  
814 (blue).

815

816 **Figure 3:** Multivariate analysis. A) Representative heatmap of top 50 significant  
817 metabolites (ANOVA) in the comparison of controls, COVID-19 and post-COVID-19  
818 patients. B) The volcano plot of the plasma metabolomics between controls and post-  
819 COVID-19 patients (red represents the up-regulated metabolites compared with controls,  
820 green represents the down-regulated metabolites compared with controls, and gray

821 represents the metabolites with no difference between both groups. Fold change  
822 threshold  $\geq 1.3$  and p-value  $\leq 0.05$  (FDR adjusted). C) Score scatter plot based on the  
823 two-dimensional PLS-DA (red: controls; green: post-COVID-19 patients) D) Rank of the  
824 different metabolites (the top 15) identified by the PLS-DA according to the VIP score on  
825 the x-axis. The most discriminating metabolites are shown in descending order of their  
826 coefficient scores. The color boxes indicate whether metabolite concentration is increased  
827 (red) or decreased (blue).

828

829 **Figure 4:** Box plots for some significantly altered metabolites ( $p < 0.05$ ) in plasma of class  
830 A patients (less than five symptoms), class B patients (more than five symptoms), and  
831 recovered patients. The bar plots show the original and normalized values (mean  $\pm$  one  
832 standard deviation). Medians are indicated by horizontal lines within each box.

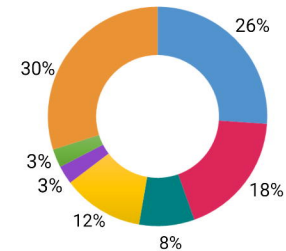
833

834 **Figure 5:** Metabolic pathway analysis. Predicted metabolic pathways with p-  
835 value  $\leq 0.05$  are listed. A) controls vs. post-COVID-19 patients. B) class B patients vs.  
836 recovered patients.

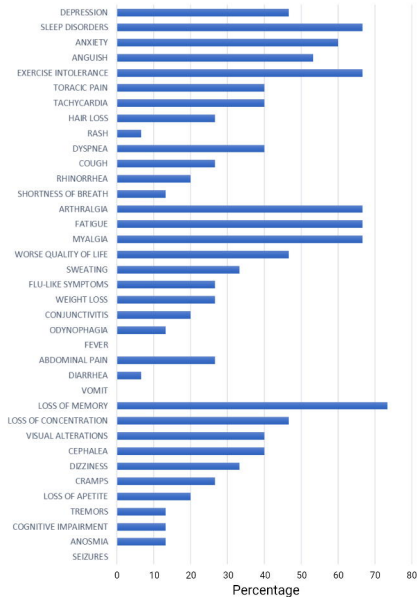
837

838 **Figure 6:** Concentrations of IL-17 and leptin measured by ELISA in post-COVID-19  
839 patients. Graphs were constructed in GraphPad Prism v8.0. The \*\* p value  $< 0.01$  was  
840 calculated using Kruskal-Wallis tests with Dunn's post-tests.

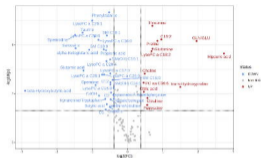
A



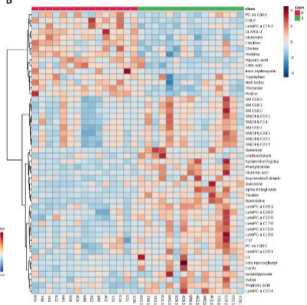
B



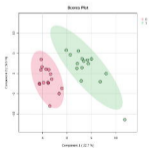
A



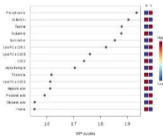
B



C

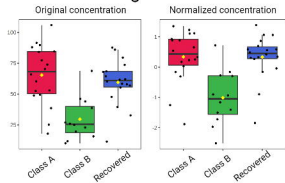


D

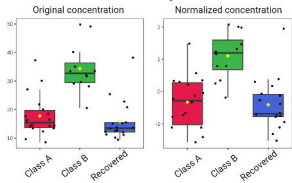




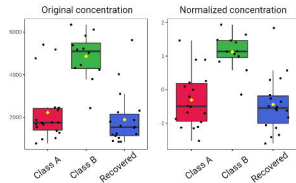
## Arginine



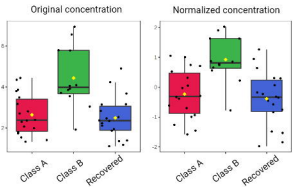
## Lactate/Pyruvate



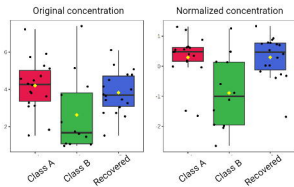
## Lactic acid



## Ornithine/Citrulline

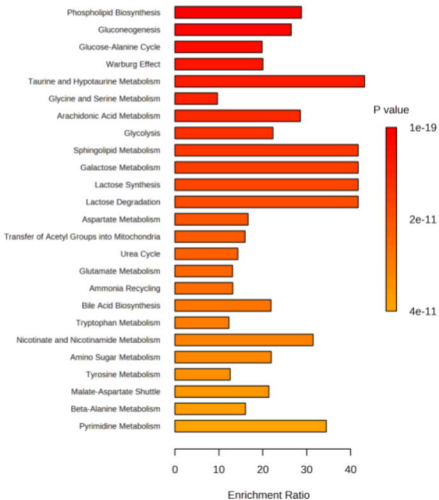


## Sarcosine



## A) Controls vs. post-COVID patients

### Enrichment Overview (top 25)



## B) Class B patients vs. recovered patients

### Enrichment Overview (top 25)

