Metabolomic profiling reveals systemic signatures of premature aging induced by Hutchinson-Gilford Progeria Syndrome

Gustavo Monnerat^{1,2} PhD, Geisa Paulino Caprini Evaristo² PhD, Joseph Albert Medeiros Evaristo² PhD, Caleb Guedes Miranda dos Santos³ PhD, Gabriel Carneiro² MsC, Leonardo Maciel¹ PhD, Vânia Oliveira Carvalho⁴ MD, PhD, Fábio César Sousa Nogueira^{2,5} PhD, Gilberto Barbosa Domont^{*5} PhD, Antonio Carlos Campos de Carvalho^{*1,6} MD, PhD

¹Institute of Biophysics Carlos Chagas Filho, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

²Laboratory of Proteomics, LADETEC, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

³Instituto de Biologia do Exército, Ministério da Defesa do Brasil

⁴Department of Pediatrics, Federal University of Paraná, Curitiba, Brazil

⁵Proteomics Unit, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

⁶National Institute of Cardiology, Rio de Janeiro, Brazil

*Correspondence and requests for materials should be addressed to acarlos@biof.ufrj.br or gilberto@iq.ufrj.br at IBCCF-UFRJ, Av. Carlos Chagas Filho 373 – CCS – Bloco G, Rio de Janeiro 21941–902, Brazil. Tel./fax: +55 21 3938-6559

Suggested running title: **Metabolic features of premature aging**

Keywords: Lamin; HGPS; Premature aging; Metabolomics; Biomarkers; Metabolic profiling.

Abstract

Hutchinson-Gilford Progeria Syndrome (HGPS) is an extremely rare genetic disorder. HGPS children present a high incidence of cardiovascular complications along with altered metabolic processes and accelerated aging process. No metabolic biomarker is known and the mechanisms underlying premature aging are not fully understood. The present study analysed plasma from six HGPS patients of both sexes (7.7 ± 1.4 years old; mean $\pm$ SD) and eight controls (8.6 ± 2.3 years old) by LC-MS/MS in high-resolution non-targeted metabolomics (Q-Exactive Plus). Several endogenous metabolites with statistical difference were found. Multivariate statistics analysis showed a clear separation between groups. Potential novel metabolic biomarkers are identified using the multivariate area under ROC curve (AUROC) based analysis, showing an AUC value higher than 0.80 using only two metabolites, and reaching 1.00 when increasing the number of metabolites in the AUROC model. Targeted metabolomics was used to validate some of the metabolites identified by the non-targeted method. Taken together, changed metabolic pathways in that panel involve sphingolipid, amino acid, and oxidation of fatty acids among others. In conclusion our data show significant alterations in cellular energy use and availability, in signal transduction, and in lipid metabolites, creating new insights on metabolic alterations associated with premature aging.

52 Introduction

53 Hutchinson-Gilford Progeria Syndrome (HGPS) is an extremely rare ge-
 54 netic disorder. Children with HGPS present a high incidence of severe cardio-
 55 vascular complications along with altered metabolic processes, associated with
 56 an accelerated aging process (1-4). Despite a great increase in the scientific
 57 knowledge about HGPS, no specific biomarker is known for HGPS and the un-
 58 derlying molecular mechanisms are not fully understood. HGPS is induced by a
 59 single mutation in the LMNA gene, creating a mutant protein isoform with dele-
 60 tion of 50 amino-acids near in the protein Lamin A. The mutated protein, known
 61 as progerin (isoform 6), is toxically accumulated in the cells. Progerin, despite
 62 being able to enter the cell nucleus, does not incorporate normally into the nu-
 63 clear membrane lamina, leading to several abnormalities in nuclear trafficking(5,
 64 6). Interestingly, unaffected aged individuals show a similar splice event, lead-
 65 ing to progerin expression that may play a role in cellular senescence(6).

66 Aging is the biological process of gradually accumulating deleterious
 67 changes in cells, decreasing the physiological capacity(7, 8). Aging is not con-
 68 sidered a disease, but it intensely rises the risk of developing chronic cardio-
 69 vascular(9) and metabolic diseases(10). It is known that metabolic systemic
 70 profiles are age-dependent, reflecting metabolism alterations, such as incom-
 71 plete fatty acid mitochondrial oxidation(11-13).

72 Metabolomics is, among other “omics” strategies, one of the most com-
 73 plete and reliable sources of information for circulatory mediator analysis, bi-
 74 omarker discovery pipeline and mechanistic disease investigation(14). In the
 75 present study, we applied metabolomics to samples from HGPS patients and
 76 identified several metabolites from different biological pathways dysregulated.
 77 Multivariate and univariate statistical analysis demonstrated metabolic pathways
 78 and potential new biomarkers that might act as central mediators in this syn-
 79 drome and in senescence.

80 Results

82

83 We used non-targeted based metabolomics to investigate metabolites
84 differentially expressed in plasma samples obtained from 6 HGPS patients and
85 8 healthy donors, as summarized in Table S1, aiming to identify new bi-
86 omarkers and novel mechanisms of the disease.

87 In order to avoid the inclusion of exogenous compounds in our analysis,
88 contaminants, medications, and their metabolites, food and flavouring com-
89 pounds were excluded from the metabolite list, resulting in a final feature list of
90 40 known molecules of endogenous origin presenting the statistical difference
91 between the two groups. Each of the identified metabolites was found to have
92 false discovery rates (FDRs) of less than 10%. Information regarding the me-
93 tabolites identified in the present study is available in Table 1. Figure S1 shows
94 typical total extracted ion chromatograms of all analyzed samples, demonstrat-
95 ing the efficient separation of the plasma compounds and reproducibility. The
96 deuterated internal standard spiked in during sample preparation was used to
97 calculate the coefficient of variation (CV) of our method. Figure S2 demon-
98 strates that our CVs were <15% among samples.

99 A data matrix including the average area values of the uniquely identified
100 analyzed compounds in each sample was generated. Multivariate statistics us-
101 ing both unsupervised and supervised strategies were then applied to the data.
102 Unsupervised PCA of the metabolomics data demonstrated a clear separation
103 between groups (Figure 1a). Percent Variance Captured by PCA Model for the
104 Principal Component 1 (PC1) was 55.4%, and for Principal component 2 (PC2)
105 was 6.9%. The green (patients) and red (controls) areas in Figure 1a represent
106 the 95% confidence intervals for each group. The application of supervised
107 PLS-DA also permits a detailed group separation between HGPS and control
108 cohorts as shown in Figure 1b. The PLS-DA model captured 55.4% of the va-
109 riance in component 1 and 5.1% in component 2. The components of the PLS-
110 DA models were used to predict the accuracy (Accuracy) based on the cross-
111 validation, the sum of squares captured by the model (R²), and the cross-
112 validated R² (Q²). The PLS-DA cross-validation data are summarized together
113 with a set of permutation tests demonstrating statistical significance in the PLS
114 model (Figure S3).

115 In order to discover potential biomarkers for HGPS, ROC curves were
116 constructed. The area under the ROC curve (AUC) is a well-described strategy

117 for biomarker potential performance analysis, where the higher the AUC the
118 more accurate is the model. As demonstrated in figure 2a, 6 AUC models were
119 created, including different numbers of metabolites, varying from 2 to 40. The
120 results demonstrate that the classification model using only two variables for the
121 AUC resulted in a 0.803 value and 95% confidence interval (CI) ranging from
122 0.5~1. Increasing the number of variables to 5 in the classification model, the
123 AUC value increases to 0.912 and the 95% confidence interval ranges from
124 0.625~1.

125 Metabolites were ranked according to their capacity to distinguish be-
126 tween HGPS and Control subjects and the result is summarized in figure 2b.
127 Metabolites are shown as either downregulated (green) or upregulated (red) in
128 patients with HGPS. Furthermore, we performed classical univariate ROC curve
129 analysis for individual biomarkers. Figure 3 shows the ranked metabolites
130 based on area under ROC curve (AUROC), suggesting that both arginine and
131 5-hydroxytryptophol are robust upregulated candidates for HGPS biomarkers.
132 Choline and phosphatidylcholine (16:0/16:0) on the other hand are robust
133 downregulated candidates for HGPS biomarkers, showing potent diagnostic
134 power. Aiming to validate the non-targeted metabolomic analysis, we per-
135 formed targeted metabolomics based on LC-MS on a triple quadrupole using
136 metabolite standards for prior calibration. Figure 4a shows ion chromatograms
137 for arginine and ISTD. Each line represents one sample analyzed showing in-
138 tensity and retention time in minutes. Figure 4b shows the calibration curve for
139 arginine quantification. Our targeted method demonstrates an increase in argi-
140 nine levels in samples from HGPS patients in the same manner as in the non-
141 targeted approach as summarized in figure 4c. Data from arginine and other
142 metabolites analyzed by triple quadrupole are summarized in table S3.

143 Aiming to evaluate the most relevant metabolic pathways altered in pa-
144 tients with HGPS, pathway analysis was applied. Figure 5 shows an overview of
145 Pathway Analysis, using only annotated metabolites identified to be significantly
146 altered by HGPS. Figure 5a highlights pathways related to upregulated metabo-
147 lites, suggesting alterations in fatty acids metabolism, glucose metabolism, and
148 mitochondrial function. Figure 5b shows the metabolic pathways related to the
149 downregulated metabolites in HGPS patients, demonstrating alterations related

150 to phospholipids, phenylacetate and phosphatidylcholine metabolism, among
151 other alterations.

152

153 **Discussion**

154 Aging is a complex biological process poorly understood at the molecular
155 level. HGPS is a rare fatal disease where an extremely accelerated aging pro-
156 cess is observed leading to premature death mainly related to heart complica-
157 tions. Despite great effort to increase knowledge about HGPS, biological bi-
158 omarkers for this disease are not yet available, detailed disease mechanisms
159 are still being investigated and there is currently no cure(19).

160 . In the present work, we created ROC curve models to identify metabolic
161 HPGS biomarkers. As shown in results, biomarkers found to be statistically dif-
162 ferent between the HGPS and control groups are highly likely to be associated
163 with premature aging based on performance in terms of both specificity and
164 sensitivity. Furthermore, targeted analysis using high purity standards of the
165 metabolites of interest previously identified in the non-targeted experiments are
166 in accordance with the non-targeted strategies, showing similar results in terms
167 of statistical significance.

168 A number of investigations used this methodology to study the mecha-
169 nisms underlying the aging progression, and whether strategies such as exer-
170 cise training and hormonal treatment can revert the metabolic changes induced
171 by aging. In a recent study by Houtkooper et al, metabolomic hallmarks of aging
172 were demonstrated, including affected pathways in both liver and muscle tis-
173 sues, indicating a significant modification in fatty acid metabolism(20). In the
174 present study, we found several compounds up or downregulated in the plasma
175 of HGPS patients, highlighting a profound metabolic alteration compared to our
176 control cohort. Aging metabolomic studies showed an increase in lactate and
177 glucose suggesting changes in glucose/pyruvate and glycogen metabolism(20),
178 in accordance with our data using HGPS plasma, where we observed an in-
179 crease in glucose and lactide, a dimer of lactic acid (Table 1). Metabolomic
180 studies in diabetic patients also demonstrate glucose and lactate increase(21).
181 In addition to the glucose/pyruvate pathway alteration, we observed an increase
182 in a long chain carnitine family molecule, Acetyl-L-Carnitine, associated with

183 fatty oxidation. Interestingly, children in early-stage type 1 diabetes present ele-
 184 vated Acyl-Carnitine. Adult patients with type 2 diabetes may also present
 185 dysregulation of fatty acid oxidation, characterized by glucolipotoxicity(22). In
 186 this context, it is interesting to note that a recent study demonstrated that met-
 187 formin, a popular anti-diabetic biguanide, alleviates the nuclear defects and
 188 premature aging phenotypes in HGPS fibroblasts, perhaps constituting a prom-
 189 ising therapeutic approach for life extension in HGPS(23). Furthermore, insulin
 190 resistance has been described in children with HGPS(24).

191 Mitochondria play a key role in several metabolic inborn errors as well as
 192 in the aging process, highlighting a decline in mitochondrial respiration(20, 25,
 193 26). In this regard, carnitine metabolites are important during fatty acid oxidation
 194 in the mitochondria. We found an 11 fold increase in Acetyl-L-Carnitine and
 195 also in L-carnitine in HGPS patients reflecting a broad dysfunction in β -
 196 oxidation, indicating a diminished lipid transport capacity in the mitochondria(27,
 197 28). On the other hand, we found some carnitine metabolites decreased in the
 198 plasma of HGPS, such as Decanoylcarnitine. Interestingly, fetal congenital dis-
 199 orders are associated with decreases in some carnitines, such as
 200 Decanoylcarnitine among others(29, 30). Collectively, these findings highlight
 201 the multiplicity of perturbations in lipid metabolism related to mitochondrial dys-
 202 functions in HGPS. These metabolic alterations may be related to the growth
 203 abnormalities observed in HGPS children(31).

204 During aging as well as in systemic metabolic dysfunction, amino-acid
 205 metabolism is significantly modified(32, 33). Previous publications demonstrate
 206 that branched-chain amino acids (BCAA) as well as methionine content in the
 207 diet changes mice lifespan. In the present work, we identified altered amino ac-
 208 id availability in HGPS patients' plasma that was further investigated by targeted
 209 metabolomics. Our experiments showed a decrease in the levels of methionine
 210 and histidine, but in contrast levels of arginine and cystine were increased. In-
 211 terestingly, Cheng and coworkers demonstrated that amino acids concentra-
 212 tions, such as Histidine, might be related to human longevity (34). Regarding
 213 BCAA no changes were observed, as well as in other important amino-acids
 214 such as proline and alanine. In agreement with our findings, Houtkooper et al
 215 showed that methionine is decreased in the plasma of aged mice and no
 216 changes were observed in BCAA(20). Interestingly, choline supplementation

217 seems to improve cognitive function and is an important strategy to ameliorates
218 Alzheimer's disease pathology, a pathological process typically associated with
219 advanced age(35, 36). Furthermore, choline converts homocysteine, a neuro-
220 toxic amino acid in methionine(37, 38), also found to be decreased in the HGPS
221 patients in the present work.

222

223 Altered metabolic processes can lead to the formation of toxic metabo-
224 lites as well as alterations in acid-base equilibrium. Interestingly, the 4,6-
225 dioxoheptanoic acid, also known as Succinylacetone, a medium-chain keto ac-
226 id, and derivative metabolite, was found in higher levels in HGPS plasma.
227 Succinylacetone can rise due to abnormal activity of the enzyme
228 fumarylacetoacetase, being suggested as an acidogenic, oncometabolite and a
229 metabotoxin. Of note, aging and progeria course with a hypertrophic cardiac
230 process, that dramatically increases the risk of severe cardiac complications. In
231 this context, patients with hypertrophic cardiomyopathy are reported to have an
232 increased level of this metabolite(39, 40).

233 Our study has limitations imposed by the cohort size used. As indicated
234 in methods/results, we analyzed only 6 HGPS patients' samples, a small num-
235 ber for a biomarker investigation and disease mechanism comprehension.
236 However, HGPS is an extremely rare disease, as emphasized by the fact that in
237 a 200 million people country like Brazil, only one donor was recruited. The 5
238 other samples from our cohort were donated by The Progeria Research Foun-
239 dation which collects patients' samples worldwide. These samples come from
240 children with different genetic backgrounds, most probably contain different
241 contaminants, were subject to distinct sample handling procedures and time of
242 storage. Human genome databases show that the interindividual differences
243 are very extensive between distinct populations. From the 40.000.000 variant
244 polymorphic DNA sites predicted, some are rare and present only in a person or
245 his family, ethnicity or country, which may reflect in their plasma
246 metabolome(41). Remarkably, in view of the expected variability and the great
247 possibility that the diverse genetic backgrounds might influence the metabolic
248 plasma levels, our approach based in the multivariate analysis of multiple me-
249 tabolites was capable to clearly separate patients from controls, generating an

important biomarker profile related to the disease, even using a very small sample size.

In summary, the present work applied a powerful metabolomics pipeline based in liquid chromatography coupled to high-resolution mass spectrometry along with multivariate statistics and pathway analysis. We were able to identify putative circulating biomarkers for the disease that may be interesting targets for pharmacological treatment, nutritional supplementation and for diagnosing and follow up of HGPS patients.

Conclusions

The present study reports for the first time a metabolic profiling with LC-MS based metabolomics of premature aging in patients with HGPS. We identified a total of 40 known metabolites differentially expressed between HGPS and age and sex-matched controls. Creating a panel with the most distinct metabolites, we identified circulating putative biomarkers candidates with high accuracy for group classification based on ROC curve models. Changed metabolic pathways involved fatty acids, amino-acids, and sphingolipids, among other metabolic pathways. Taken together these alterations impact the cellular energy use, enzyme activities, and cell signalling, creating new insights into the molecular mechanisms underlying premature aging associated with HGPS.

Methods

Sample preparation

Plasma samples were obtained from 5 HGPS patients kindly donated by The Progeria Research Foundation (www.progeriaresearch.org), including 3 females (2.3, 4.7, 12.2 years old) and 2 males (8.5, 10.2 years old). An additional sample was obtained from a Brazilian HGPS patient (female 8.4 years old) at the Federal University of Paraná as summarised in Table S1. The average age for HGPS patients was 7.7 ± 1.4 years (mean $\pm$ SD). As controls, we used 8 healthy donors of both genders (4 males and 4 females) with a mean age of 8.6 ± 2.3 years ($p = 0.4154$).

283 Ethics statement

284 Parents or the legal guardians of all controls and of the Brazilian patient
285 have given full written informed consent for participation in the study. The study
286 has been approved by the Ethics Committee of the Instituto Nacional de
287 Cardiologia, number - 27044614.3.0000.5272 and the Department of Pediatrics
288 from the University Hospital of the Federal University of Paraná. All procedures
289 were in accordance with the ethical standards of the responsible local Ethics
290 Committees and with the Helsinki Declaration of 1975, as revised in 2000.

291

292 Metabolomics

293 Blood samples in Brazil were collected using EDTA tubes and plasma
294 was obtained by centrifugation for 10 min at 10.000g (Megafuge 8R, Thermo
295 Scientific, USA). The Progeria Research Foundation disposed frozen plasma
296 samples. For metabolomic experiments, plasma proteins were precipitated with
297 methanol (3:1 (v/v)) at -20°C for 1h. After protein precipitation, samples were
298 centrifuged for 10 min at 14.000g, at 4°C, supernatants were collected and
299 dried in a SpeedVac Concentrator (SPD111v, Thermo Scientific, USA). The
300 metabolites were then reconstituted with a dilution factor of 3 in methanol/water
301 (1:9 (v/v)). 5nM of deuterated testosterone (D3-Testosterone, purchased from LGC
302 Standards; London, England) was spiked and used as an internal standard
303 (ISTD) for coefficient of variance (CV) calculation. For quality control (QC), a
304 pool of all the analyzed samples was prepared.

305

306 Non-targeted metabolomics

307 For liquid chromatography-tandem mass spectrometry (LC-MS) analysis,
308 5µl volumes of each sample were analyzed in triplicate. As a blank control,
309 methanol/water followed the same steps and was used as background for data
310 analysis as previously described(15). Between samples, a washing protocol
311 was performed. Samples were analyzed in a random sequence and the QC
312 sample was analyzed 5 different times along the experiment.

313 Samples were analyzed by Dionex Ultimate 3000 UHPLC coupled to a
314 Q-Exactive Plus high-resolution mass spectrometer (Thermo Scientific, USA).
315 LC separations were obtained using a 2.1 x 50mm ZORBAX 1.8µm C18 column
316 (Agilent, USA). Mobile phases used were: phase A) water with 0.1% formic acid

and 5mM ammonium formate, and phase B) methanol with 0.1% formic acid. Total run time was 30 min. The first 4 min of the run consisted of a linear gradient from 10% to 60% of B phase, followed by a 20 min linear gradient from 60% to 98% of B. After reaching 98%, a stable run with 98% of B was sustained for 3 minutes followed by 3 minutes of 10% of B solution to regenerate the column pumped at 450 μ L/min with a column temperature of 55°C and the sample chamber held at 7°C acquired in positive mode. The data were obtained with the MS detector in full-scan mode (Full-MS) with the data-dependent acquisition (dd-MS2) for the top-10 most abundant ions per scan(15), with settings: In-source CID 0.0 eV, micro scans = 1, resolution = 70,000, AGC targeted 1e6, maximum IT = 50 ms, scan range 67 to 1000 m/z, spectrum data = Profile. Detector setting for dd-MS2 were: micro scans = 1, resolution = 17,500, AGC targeted 1e5, maximum IT = 100 ms, loop count = 10, isolation window 2.0 m/z, NCE 15, 35, 50, spectrum data = profile, underfill ratio = 1.5%, charge exclusion = unassigned, dynamic exclusion = 6s.

332

333 Data analysis and statistics

334 Data were analyzed by Compound Discoverer 2.1 (Thermo Fischer, USA). For compound detection a mass tolerance of 5 ppm was accepted to extract ions with a minimum of 1.000.000 peak intensity; for compound consolidation, a 0.2 min of retention time tolerance was employed. The ChemSpider search including BioCyc and Human metabolome database (HMDB)(16) was used with 5ppm mass tolerance as well as the mzCloud search. In the non-targeted method, the identification of non-novel metabolites was based on accurate mass and tandem mass spectra, without chemical standards references, providing a level 2 identification (putatively annotated). The samples were analyzed in triplicate. A principal component analysis (PCA) was performed to evaluate the experimental reproducibility and the QC samples were used to identify the reproducibility throughout experiments. Data of the triplicate injection experiments were unified and the average was used as a unique value. Data were scaled by auto-scaling. For statistical analysis, group area data from control vs patient data fold change was calculated and the p-value per group was calculated by t-test. Compounds that presented $p < 0.05$ after adjustment using p-value (FDR) cutoff of 0.1 were considered statistically different. Chromatogram

351 visualization and base peak chromatogram figure generation were performed
352 using MZmine 2.26 software(17) and metabolomics statistics data was per-
353 formed using MetaboAnalyst (18). The curated data matrix was used to gener-
354 ate a model for sample class discrimination via PCA and Partial Least Squares -
355 Discriminant Analysis (PLS-DA) using online MetaboAnalyst (18). The model
356 quality was analyzed by the goodness-of-fit parameter (R²) and the goodness-
357 of-prediction parameter (Q²). For biomarker analysis, multivariate ROC curve
358 based exploratory analysis was performed using a classification method (SVM)
359 and feature ranking method (SVM built-in) applied to the statistically different
360 metabolites between the two groups.

361

362 Targeted metabolomics

363 Amino-acid mixture standard was purchased from Sigma-Aldrich (São Paulo,
364 Brazil) and D3-testosterone (ISTD) from LGC Standards (London, England).
365 Amino acid quantification was carried out using a TSQ Quantiva from Thermo
366 Scientific (San Jose, USA) with a Dionex Ultimate 3000 HPLC system
367 (Germering, Germany). Chromatographic separation was achieved using a re-
368 versed phase column (C18 Zorbax, 50 × 3 mm, 1,7 µm, Agilent, Santa Clara,
369 USA). The analyte was eluted from the column using a gradient with the eluent
370 changing from 5% to 100% methanol in water within 3 min. The column was
371 washed for 1.2 min in 100% methanol and equilibrated for 3 min at the initial
372 eluent composition. All solvents contained 0.1% formic acid. The flow rate, col-
373 umn temperature, and injection volume were 300 µL/min, 40°C and 5 µL, re-
374 spectively.

375 Amino-acids were monitored by selected reaction monitoring (SRM) in the posi-
376 tive ion mode. The transitions selected for amino-acid quantification and ISTD
377 are listed in Table S2. The curve was constructed using a mix of amino-acids in
378 triplicate at 1; 2,5; 5; 10 and 20 nmol/mL. All samples were spiked with D3-
379 Testosterone (ISTD) at 5 ng/mL. The area ratios of the total extracted ion of the
380 product ions and the product ion of the IS were plotted versus the concentra-
381 tion.

382

383 Statistical analysis

384

385 Data are presented as mean $\pm$ SEM. Two-tailed Student's *t*-test was used. We
386 did not use statistical methods to predetermine sample size; samples sizes
387 were determined on the basis of sample availability. The non-targeted metabo-
388 lomics statistics is described in detail with the metabolomics data analysis
389 methods above. Values of $P < 0.05$ were considered statistically significant us-
390 ing GraphPad Prism 6.0 (GraphPad Software, USA).

391

392 **Acknowledgments**

393 We are grateful to The Progeria Research Foundation for the availability of
394 plasma samples, to Edna Aleixo from the Federal University of Rio de Janeiro
395 for assistance with the importation process and to the Laboratório de Apoio ao
396 Desenvolvimento Tecnológico (LADETEC) of the Institute of Chemistry of the
397 Federal University of Rio de Janeiro for providing high quality infrastructure for
398 the LC-MS analysis.

399 **Author contributions**

400 GM and ACCC conceptualized the study and wrote the manuscript; GM,
401 CGMS, JAME, GPCE, FCSN, GC, GBD, LM, ACCC acquired and analyzed the
402 data; VOC, GBD, FCSN, ACCC critically revised the study and the manuscript.

403 **Funding**

404 This work was funded by the Brazilian National Research Council (CNPq), the
405 Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ) and
406 National Institutes of Science and Technology for Regenerative Medicine.

407 **Competing Interests**

408

409 The authors declare no competing interests.

410

411

412

413

414

415

416

417 **Figure legends**

418 **Figure 1. Multivariate analysis of the metabolomics data**

419 a) Principal component analysis (PCA) 2D score plot and b) Partial Least
420 Squares - Discriminant Analysis (PLS-DA) 2D score plot from the dataset with
421 all of the features expected to be endogenous metabolites with statistical differ-
422 ence. The green (patients) and red (controls) areas represent the 95% confi-
423 dence interval regions for each group.

424 **Figure 2. Potential biomarkers in diagnosing HGPS with metabolomics**

425 a) Areas under the ROC curve (AUC) for different numbers of variables used to
426 construct the ROC curves. The inset shows the number of variables used for
427 the AUCs models. b) The most significant features of the ROC model
428 downregulated (green) or upregulated (red) in patients with HGPS compared to
429 controls.

430 **Figure 3. ROC curve analysis of individual biomarkers for HGPS based** 431 **metabolomics**

432 Top ranked metabolites based on area under ROC curve (AUROC) identified by
433 the non-targeted metabolomics analysis of HGPS plasma samples. The rectan-
434 gle to the right of the ROC curves show the individual values determined for
435 each metabolite in the control and HGPS groups. The red line represents the
436 mean value for the control group + two standard deviations.

437

438 **Figure 4. Targeted metabolomics performed by LC-MS/MS for biomarker** 439 **validation**

440 a) Extracted ion chromatograms for arginine and internal standard (ISTD). Each
441 colored line represents one sample analyzed showing intensity and retention
442 time (RT) in minutes. b) Calibration curve for arginine quantification. c) Graph
443 shows arginine quantification in each group and bars represent SEM. * indi-
444 cates $P < 0.05$.

445

446 **Figure 5. Metabolic pathways altered by HGPS**
447 Overview of Pathway Analysis highlighting potential functional relationships be-
448 tween the set of annotated metabolites identified to be significantly altered by
449 HGPS. a) Pathways related to upregulated and b) downregulated metabolites in
450 the HGPS patients.

451

452 Table 1. Significant dysregulated metabolites of HGPS

Metabolite	Class	HMDB ID	Fold Change	p-value	FDR
LysoPC(17:0)	Glycerophospholipid	HMDB0012108	0.57	0.0179	0.12
LysoPC(16:0)		HMDB0010382	0.50	0.0007	0.02
PC(16:0/18:2)		HMDB0011211	0.45	0.0004	0.01
Glucose	Carbohydrates and carbohydrate conjugates	HMDB0000122	3.66	0.0001	0.01
N(6)-Methyllysine	Carboxylic acids and derivatives	HMDB0002038	7.74	0.0238	0.13
Arginine		HMDB0000517	2.48	0.0056	0.06
gamma-Aminobutyric acid		HMDB0000112	1.80	0.0413	0.19
Aminolevulinic acid		HMDB0001149	0.77	0.0331	0.17
Phenylalanine		HMDB0000159	0.54	0.0003	0.01
N-Phenylacetylglutamine		HMDB06344	0.45	0.0172	0.12
Pyroglutamylglycine		HMDB0061890	0.32	0.0016	0.02
Aceglutamide		HMDB0006029	0.29	0.0011	0.02
Threoninyl-Aspartate		HMDB0029057	0.27	0.0005	0.01
Oleoylcarnitine	Fatty Acyl	HMDB0005065	0.44	0.0071	0.07
L-Carnitine		HMDB0000062	1.72	0.0087	0.08
Acetyl-L-carnitine		HMDB0000201	11.23	0.0312	0.16
3-Dehydroxycarnitine		HMDB0006831	0.75	0.0275	0.15
Palmitoylcarnitine		HMDB0000222	0.64	0.0263	0.15
Decanoylcarnitine		HMDB0000651	0.47	0.0411	0.19

9-Decenoylcarnitine		HMDB0013205	0.44	0.0005	0.01
Phosphatidylcholine(16:0/16:0)	Glycerophospholipids	HMDB0011206	0.67	0.0015	0.02
1-octadecylglycero-3-phosphocholine		HMDB0062195	0.57	0.0066	0.07
Glycerolphosphorylcholine		HMDB0000086	0.51	0.0049	0.06
Dipalmitoylphosphatidylcholine		HMDB0000564	0.47	0.0029	0.04
Lactide	Hydroxy acids and derivatives	not found	3.62	0.0001	0.01
5-Hydroxytryptophol	Indoles and derivatives	HMDB0001855	6.65	0.0121	0.10
4,6-Dioxoheptanoic acid	Keto acids and derivatives	HMDB00635	2.04	0.0032	0.04
LysoPC(P-16:0)	Lysophospholipid	HMDB0010407	0.35	0.0003	0.01
Choline	Organonitrogen compounds	HMDB0000097	0.57	0.0002	0.01
PC(o-18:1(11Z)/16:0)	Phosphatidylcholine	HMDB0013424	0.77	0.0342	0.17
PC(O-16:0/18:2(9Z,12Z))		HMDB0011151	0.51	0.0074	0.07
N-Methylethanolaminium phosphate	Phosphoethanolamine	HMDB0060173	0.46	0.0367	0.18
(3beta,19alpha)-3,19,23,24-Tetrahydroxy-12-oleanen-28-oic acid	Prenol lipids	HMDB0040784	0.23	0.0168	0.12
Pyridoxamine	Pyridines and derivatives	HMDB0001431	0.56	0.0142	0.11
Citicoline	Pyrimidine nucleotides	HMDB0001413	0.55	0.0055	0.06
13-cis retinol	Retinoids	HMDB0006221	2.16	0.0156	0.12
Palmitoyl sphingomyelin	Sphingolipids	HMDB0061712	0.70	0.0136	0.11
Pregnenolone	Steroids and steroid derivatives	HMDB0000253	2.10	0.0188	0.12
(3beta,24R,24'R)-fucosterol epoxide	Sterol Lipids	not found	0.43	0.0389	0.19
Bilirubin	Tetrapyrroles and derivatives	HMDB0000054	6.34	0.0128	0.10

References

1. R. C. Hennekam, Hutchinson-Gilford progeria syndrome: review of the phenotype. *Am J Med Genet A* **140**, 2603-2624 (2006).
2. M. A. Merideth *et al.*, Phenotype and course of Hutchinson-Gilford progeria syndrome. *N Engl J Med* **358**, 592-604 (2008).
3. A. Prakash *et al.*, Cardiac Abnormalities in Patients With Hutchinson-Gilford Progeria Syndrome. *JAMA Cardiol* **3**, 326-334 (2018).
4. J. Xu *et al.*, Omi/HtrA2 Participates in Age-Related Autophagic Deficiency in Rat Liver. *Aging Dis* **9**, 1031-1042 (2018).
5. M. Eriksson *et al.*, Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423**, 293-298 (2003).
6. D. McClintock *et al.*, The mutant form of lamin A that causes Hutchinson-Gilford progeria is a biomarker of cellular aging in human skin. *PLoS One* **2**, e1269 (2007).
7. R. Dahse, W. Fiedler, G. Ernst, Telomeres and telomerase: biological and clinical importance. *Clin Chem* **43**, 708-714 (1997).
8. C. J. Kenyon, The genetics of ageing. *Nature* **464**, 504-512 (2010).
9. B. J. North, D. A. Sinclair, The intersection between aging and cardiovascular disease. *Circ Res* **110**, 1097-1108 (2012).
10. F. Bonomini, L. F. Rodella, R. Rezzani, Metabolic syndrome, aging and involvement of oxidative stress. *Aging Dis* **6**, 109-120 (2015).
11. R. C. Noland *et al.*, Carnitine insufficiency caused by aging and overnutrition compromises mitochondrial performance and metabolic control. *J Biol Chem* **284**, 22840-22852 (2009).
12. Z. Yu *et al.*, Human serum metabolic profiles are age dependent. *Aging Cell* **11**, 960-967 (2012).
13. D. Nguyen, S. L. Samson, V. T. Reddy, E. V. Gonzalez, R. V. Sekhar, Impaired mitochondrial fatty acid oxidation and insulin resistance in aging: novel protective role of glutathione. *Aging Cell* **12**, 415-425 (2013).
14. G. Carneiro, A. L. Radcenco, J. Evaristo, G. Monnerat, Novel strategies for clinical investigation and biomarker discovery: a guide to applied metabolomics. *Horm Mol Biol Clin Investig*, (2019).
15. G. Monnerat *et al.*, Aging-related compensated hypogonadism: Role of metabolomic analysis in physiopathological and therapeutic evaluation. *J Steroid Biochem Mol Biol*, (2018).
16. D. S. Wishart *et al.*, HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res* **41**, D801-807 (2013).
17. T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).
18. J. Xia, D. S. Wishart, Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr Protoc Bioinformatics* **55**, 14.10.11-14.10.91 (2016).
19. L. B. Gordon *et al.*, Association of Lonafarnib Treatment vs No Treatment With Mortality Rate in Patients With Hutchinson-Gilford Progeria Syndrome. *JAMA* **319**, 1687-1695 (2018).
20. R. H. Houtkooper *et al.*, The metabolic footprint of aging in mice. *Sci Rep* **1**, 134 (2011).

- 504 21. S. O. Crawford *et al.*, Association of blood lactate with type 2 diabetes:
505 the Atherosclerosis Risk in Communities Carotid MRI Study. *International*
506 *journal of epidemiology* **39**, 1647-1655 (2010).
- 507 22. S. J. Mihalik *et al.*, Increased levels of plasma acylcarnitines in obesity
508 and type 2 diabetes and identification of a marker of glucolipotoxicity.
509 *Obesity (Silver Spring)* **18**, 1695-1700 (2010).
- 510 23. S. K. Park, O. S. Shin, Metformin alleviates ageing cellular phenotypes in
511 Hutchinson-Gilford progeria syndrome dermal fibroblasts. *Exp Dermatol*
512 **26**, 889-895 (2017).
- 513 24. A. L. Rosenbloom *et al.*, Progeria: insulin resistance and hyperglycemia.
514 *J Pediatr* **102**, 400-402 (1983).
- 515 25. J. Thompson Legault *et al.*, A Metabolic Signature of Mitochondrial
516 Dysfunction Revealed through a Monogenic Form of Leigh Syndrome.
517 *Cell Rep* **13**, 981-989 (2015).
- 518 26. D. J. Dietzen, M. J. Bennett, S. F. Lo, V. L. Grey, P. M. Jones, Dried
519 Blood Spot Reference Intervals for Steroids and Amino Acids in a
520 Neonatal Cohort of the National Children's Study. *Clin Chem* **62**, 1658-
521 1667 (2016).
- 522 27. S. M. Houten, R. J. Wanders, A general introduction to the biochemistry
523 of mitochondrial fatty acid β -oxidation. *J Inherit Metab Dis* **33**, 469-477
524 (2010).
- 525 28. S. H. Adams, Emerging perspectives on essential amino acid
526 metabolism in obesity and the insulin-resistant state. *Adv Nutr* **2**, 445-456
527 (2011).
- 528 29. R. O. Bahado-Singh *et al.*, Metabolomic prediction of fetal congenital
529 heart defect in the first trimester. *Am J Obstet Gynecol* **211**, 240.e241-
530 240.e214 (2014).
- 531 30. M. Yu *et al.*, Discovery and Validation of Potential Serum Biomarkers for
532 Pediatric Patients with Congenital Heart Diseases by Metabolomics. *J*
533 *Proteome Res*, (2018).
- 534 31. L. B. Gordon *et al.*, Clinical trial of a farnesyltransferase inhibitor in
535 children with Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci*
536 *U S A* **109**, 16666-16671 (2012).
- 537 32. T. J. Wang *et al.*, Metabolite profiles and the risk of developing diabetes.
538 *Nat Med* **17**, 448-453 (2011).
- 539 33. D. H. Katz, R. E. Gerszten, A Role for Branched-Chain Amino Acids in
540 the Pathophysiology of Diabetes: Using Data to Guide Discovery. *Clin*
541 *Chem* **64**, 1250-1251 (2018).
- 542 34. S. Cheng *et al.*, Distinct metabolomic signatures are associated with
543 longevity in humans. *Nat Commun* **6**, 6791 (2015).
- 544 35. R. Velazquez *et al.*, Maternal choline supplementation improves spatial
545 learning and adult hippocampal neurogenesis in the Ts65Dn mouse
546 model of Down syndrome. *Neurobiol Dis* **58**, 92-101 (2013).
- 547 36. R. Velazquez *et al.*, Maternal choline supplementation ameliorates
548 Alzheimer's disease pathology by reducing brain homocysteine levels
549 across multiple generations. *Mol Psychiatry*, (2019).
- 550 37. M. R. Olthof, E. J. Brink, M. B. Katan, P. Verhoef, Choline supplemented
551 as phosphatidylcholine decreases fasting and postmethionine-loading
552 plasma homocysteine concentrations in healthy men. *Am J Clin Nutr* **82**,
553 111-117 (2005).

- 554 38. V. Shukla *et al.*, TFP5, a Peptide Inhibitor of Aberrant and Hyperactive
555 Cdk5/p25, Attenuates Pathological Phenotypes and Restores Synaptic
556 Function in CK-p25Tg Mice. *J Alzheimers Dis* **56**, 335-349 (2017).
- 557 39. D. Cyr, R. Giguère, G. Villain, B. Lemieux, R. Drouin, A GC/MS validated
558 method for the nanomolar range determination of succinylacetone in
559 amniotic fluid and plasma: an analytical tool for tyrosinemia type I. *J*
560 *Chromatogr B Analyt Technol Biomed Life Sci* **832**, 24-29 (2006).
- 561 40. B. W. Adam, T. H. Lim, E. M. Hall, W. H. Hannon, Preliminary proficiency
562 testing results for succinylacetone in dried blood spots for newborn
563 screening for tyrosinemia type I. *Clin Chem* **55**, 2207-2213 (2009).
- 564 41. C. Bouchard, Exercise genomics--a paradigm shift is needed: a
565 commentary. *Br J Sports Med* **49**, 1492-1496 (2015).

566

a)

b)

Arginine

5-Hydroxytryptophol

Choline

Phosphatidylcholine160/160

a)

b)

c)

a)

Pathways up-regulated

b)

Pathways down-regulated