ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Urine Metabolomics during a Legume Diet Intervention Suggests Altered Metabolic Signatures and Potential New Intake Markers: First Insights

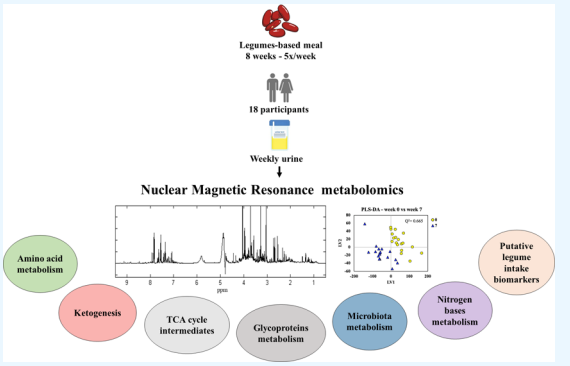
Helena Ferreira, Daniela Duarte, João A. Rodrigues, Marta W. Vasconcelos, Elisabete Pinto,*,# and Ana M. Gil*,#

Cite This: *ACS Omega* 2024, 9, 43453−43468

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Given the general increase in legume consumption worldwide, there is a need to characterize the resulting human metabolic adaptations in order to demonstrate potential legume diet/health relationships. A nuclear magnetic resonance (NMR) metabolomics urine study was carried out on a small cohort (n = 18) to characterize the excretory effects of a pilot longitudinal 8-week legume-based dietary intervention. Despite the expected high interindividual variability in the excreted metabolome, the results suggested a nonlinear metabolic response, with higher metabolic activity in the first 4 weeks and a tendency toward baseline at the end of the intervention. The excretion of isoleucine, leucine, and threonine increased, along with metabolite changes suggestive of activation of the tricarboxylic acid cycle (through anaplerosis), ketogenesis, fat catabolism, and glycoprotein biosynthesis. Gut microbiota adaptations were also suggested based on the increased excretion of 2-hydroxyisobutyrate, allantoin, and hippurate. Increased levels of trigonelline were consistent with its role as a legume intake marker, whereas malonate and pseudouridine were suggested as possible additional markers. Correlation of NMR data with nutritional parameters aided putative explanatory hypotheses to be advanced. Our results suggest a dynamic response to legume consumption, mainly through increased amino acid excretion and altered energy metabolism, while advancing potential new markers of legume intake. These results require confirmation in larger cohorts but pave the way for an informed interpretation of the effects of legume-based diets on human health.



1. INTRODUCTION

Current global dietary patterns are depleting the planet of its natural resources.[1] Transition toward more sustainable diets[2] has involved a shift toward eco-friendly foods,[3] in particular plant-based diets, with minimal consumption of animal protein, a.k.a. as "flexitarianism."[4,5] The highly nutritious protein-rich plant seeds of the Leguminosae family, such as peas, lentils, chickpeas, and beans, provide cultural, social, environmental, economic, and nutritional advantages, compared to livestock production and other feed and food crops.[6,7] Not only do they improve dietary nutritional quality, but they also reduce risk factors associated with cardiovascular diseases, *e.g.,* obesity, dyslipidemia, hyperglycemia, inflammation, and oxidative stress.[8,9] A successful transition toward a legume-rich diet requires, however, an overall scientific evaluation of its impact on individual's health so that an adequate balance between nutrition, health, and environment may be achieved.[2,10] In this respect, dietary interventions are extremely valuable; however, limitations often arise in their interpretation, for instance, regarding study design heterogeneity (*e.g.,* different trial durations, types of intervention, control diets), different target populations (*e.g.,* with distinct eating habits, lifestyle characteristics, or health status), and subjective food intake assessment (usually self-reported data).[8] Metabolomics may help reduce subjectivity and provide a clearer understanding of the link between diet and health.[11] This strategy entails the profiling of small molecules present in biological samples such as biofluids (*e.g.,* blood and urine), tissues or cells,[12] usually through mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy. In the context of legume consumption, MS or NMR metabolomic studies have included both observational-designed studies[13] and specific interventions,[14−20] the latter mostly employing MS-based strategies.[14−18] Some of these MS reports have addressed the impact of navy beans on serum/urine/stool metabolite profiles of colorectal cancer patients,[14,15,17] while others have identified

Received: May 21, 2024
Revised: September 5, 2024
Accepted: September 6, 2024
Published: October 20, 2024

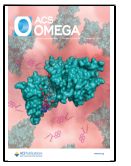


Table 1. Macronutrients, Micronutrients, and Food Groups Intake Calculated From 3-day Food Records at Baseline (Week 0) and Week 8[a]

	BASELINE (week 0)			WEEK 8			
Nutrients and Food Groups (n = 13)	Median	P_{25}	P_{75}	Median	P_{25}	P_{75}	p
Nutrients							
Energy (kcal/kJ)	1691.7/7097.4	1579.2/6635.9	1994.1/8380.9	1884.9/7920.7	1614.6/6777.6	2168.2/9118.4	0.196
Total fat (%E)	31.6	27.5	32.8	31.2	29.2	32.5	0.972
Saturated fatty acids (%E)	9.9	8.1	11.6	10.3	9.2	11.2	0.650
Polyunsaturated fatty acids (%E)	5.0	3.9	5.9	4.6	4.4	5.5	0.972
Trans fatty acids (%E)	0.6	0.4	0.9	0.7	0.4	1.0	0.807
Total carbohydrates (%E)	48.4	42.9	54.0	46.4	44.3	51.2	0.116
Total protein (%E)	17.6	13.0	19.8	18.9	15.7	21.1	0.152
Total fat (g)	59.0	51.5	71.1	66.9	54.2	76.5	0.507
Saturated fatty acids (g)	21.6	15.4	24.7	23.0	17.4	25.7	0.382
Polyunsaturated fatty acids (g)	9.5	7.1	11.7	10.4	7.5	12.9	0.600
Trans fatty acids (g)	1.4	0.8	1.6	1.4	0.8	2.4	0.272
Cholesterol (mg)	258.5	164.5	381.0	295.9	248.2	328.3	0.600
Total carbohydrates (g)	220.3	158.7	251.6	213.6	170.8	270.0	0.463
Sugars (g)	97.6	68.0	123.6	92.5	66.1	121.0	0.753
Total dietary fiber (g)	17.2	14.1	24.0	20.2	16.1	25.5	0.046*
Total protein (g)	74.5	56.8	92.6	84.0	73.3	102.7	0.023*
Vitamin A (μg; RE)	577.6	444.8	992.7	765.8	607.2	1011.3	0.552
Alfa-tocopherol (mg)	6.4	5.0	7.8	6.4	4.9	8.3	0.650
Thiamin (mg)	1.1	0.9	1.3	1.2	0.9	1.4	0.861
Riboflavin (mg)	1.3	0.9	1.8	1.3	1.1	1.6	0.480
Niacin (mg; NE)	32.0	22.5	42.9	32.2	29.5	38.4	0.807
Vitamin B_6 (mg)	1.6	1.3	2.1	1.7	1.4	1.9	0.861
Vitamin B_{12} (μg)	3.0	1.7	8.0	2.8	1.7	3.5	0.249
Vitamin C (mg)	98.8	73.1	154.1	90.2	61.5	146.0	0.196
Folate (μg)	182.2	167.0	209.2	207.7	151.3	243.6	0.753
Sodium (mg)	2366.7	1986.9	2523.2	2534.2	2234.3	3352.0	0.064
Potassium (mg)	2953.6	2664.3	3475.0	2974.9	2079.6	3547.6	0.650
Calcium (mg)	607.5	458.5	765.9	625.1	489.0	821.0	0.173
Phosphorus (mg)	1159.4	823.3	1350.1	1142.3	906.5	1343.9	0.221
Magnesium (mg)	256.8	219.9	330.6	243.1	219.7	332.0	0.807
Iron (mg)	10.1	7.8	12.0	8.8	7.0	10.8	0.279
Zinc (mg)	8.8	7.2	9.9	8.4	7.2	10.8	0.625
Food groups							
Fruits and vegetables (g)	473.6	316.5	629.2	487.5	365.2	677.2	0.345
Meat (g)	121.7	57.3	134.6	93.8	66.9	138.2	0.972
Fish and seafood (g)	39.2	23.3	70.5	19.9	12.2	50.3	0.173
Sugar-rich foods and beverages (g)	112.5	55.0	215.8	130.5	77.4	179.9	0.753
Legumes (g)	10.0	0.0	28.0	54.8	47.5	73.5	0.007*

[a]These data were available for 13 of 18 participants. †P value for the comparison of median nutrient intake between the two periods (Wilcoxon test); E: total daily energy intake; RE: retinol equivalents; NE: niacin equivalents; P25: 25th percentile; P75: 75th percentile; *p < 0.05.

possible intake biomarkers of specific legumes, *e.g.*, peas, lentils, chickpeas, and white beans.[16,18] Despite the lower sensitivity of NMR compared to MS, its holistic nature, along with minimal sample preparation, higher reproducibility, and the possibility of sample preservation,[21] has made it a popular alternative approach in nutrition research,[22−24] potentially supporting larger-scale epidemiological diet/health studies.[25] The human metabolic response to plant-based diets in general,[26−32] or to legume intake in particular,[13,19,20] has been investigated by NMR, mainly using urine,[13,19,20,26−30,32] (easier and non-invasive collection), followed by blood[31] and fecal extracts.[32] In particular, the response of healthy individuals (n = 12) to "high meat," "low meat," or "vegetarian" diets was evaluated through urine samples collected in a longitudinal setup. A metabolite signature of "high meat consumers" (Western-type phenotype) was proposed, comprising increased excretion of taurine, carnitine, acetylcarnitine, 1-methylhistidine, 3-methylhistidine, and trimethylamine-*N*-oxide (TMAO).[30] Other NMR-based dietary interventions and observational studies[26−29,32] revealed excretory metabolite fingerprints associated with vegetarian diets (compared to omnivorous diets) comprising elevated citrate, succinate, dimethylamine (DMA), glycine, mannitol, *p*-hydroxyphenylacetate, hippurate, and *N*-acetyl-*S*-methyl-cysteine-sulfoxide.

Legume intake was, to our knowledge, first studied by urine NMR metabolomics within a large *Primary Prevention of Cardiovascular Disease with a Mediterranean Diet* (PREDIMED) study,[13] revealing that individuals (n = 50) consuming legumes (chickpeas, lentils, and beans) exhibited differences in choline metabolism, protein-related compounds, and energy metabolism. Glutamine, DMA, and 3-methylhistidine were then suggested as biomarkers of legume intake,[13]

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

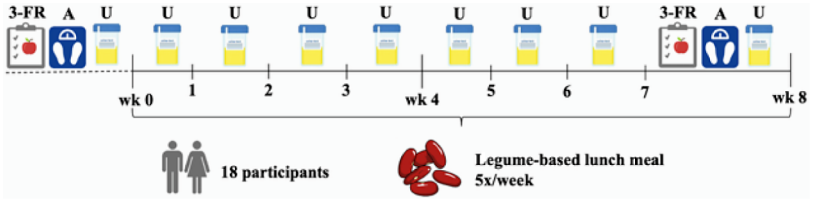


Figure 1. Schematic representation of the study design. 3-FR : 3-day food records; A: anthropometry assessment; U: urine collection; wk: week. The diet intervention involved a set of legumes and not just one type, as schematically depicted in this figure (please check the Experimental Section for more specific information).

while dimethylglycine, lysine, trimethylamine (TMA), and trigonelline were later proposed as additional markers.[19] It was suggested that gluconeogenesis activity may decline in legume consumers, consistently with reported glucose-lowering effects and a possible beneficial role in hyperglycemia-related conditions, such as type 2 diabetes.[8] To our knowledge, most urine NMR metabolomic studies of legume diets have been based on observational designs[13,26,28,29,32] and/or short-term randomized controlled dietary interventions (<15 days),[19,20,27,30] with a low number of assessment points per individual (*e.g.*, single-spot sample collection). This reveals a need for experimental designs to accommodate longer intervention periods, preferably of a longitudinal nature (i.e., involving multisampling for each individual), to better account for interindividual variability, a particularly important feature in urine metabolomics.

In this work, untargeted (hypothesis-generating) NMR metabolomics was used, for the first time to our knowledge, with the aim of investigating the urine metabolite profile adaptations of traditionally omnivorous free-living healthy young adults (n = 18, predominantly female) subjected to a longitudinal 8-week legume-based dietary intervention. The intervention included legumes typically present in the Portuguese diet and less extensively studied before (*e.g.*, as has been the case for soybean), namely, beans, chickpeas, lentils, and peas, and we hypothesize that such a diet will impact the individual's excretory profile as a reflection of their organism's metabolic adaptation. The preliminary results presented here for a small cohort suggest a possible dynamic picture of metabolic adaptations triggered by a long-term legume-based diet, thus justifying subsequent testing in larger cohorts and comparison with an additional free-diet control group.

2. EXPERIMENTAL SECTION

2.1. Subjects and Study Design. This pilot study consisted of a one-group comparison, quasi-experimental dietary intervention. All procedures followed the Declaration of Helsinki on ethical principles for medical research involving humans and were approved by the Institute of Bioethics of the Portuguese Catholic University (Ethics Screening Report 11/2017). Initially, 20 volunteers were recruited at the university campus and screened for the following eligibility criteria: (1) men or women between 18 and 45 years of age; (2) daily intake of animal-protein food sources, as part of both lunch and dinner; and (3) usual intake of <25 g (dry weight) legumes per day. The subjects were fully informed about the study protocol and signed an informed consent form. One subject dropped out at week 1 and another was later excluded for failure to provide urine sample under fasting. Hence, a total of 18 nonvegetarian participants were finally considered (1 man and 17 women; 19−43 years of age, median: 28.0 years, P_{25} = 23.8, P_{75} = 35.5, average: 29.0 ± 7.2; body mass index (BMI) 17.5−38.5 kg/m^2, median: BMI 22.3 kg/m^2, P_{25} = 21.2, P_{75} = 25.8, average: 23.8 ± 5.0) (Table S1). Considering the final 1/17 M/F ratio, the authors opted to not exclude the male subject as this would impact negatively on an already small cohort (thus subsequently analyzed as predominantly female). Minor baseline health conditions were admitted for inclusion as they were considered of generally common occurrence in the population: depressive disorders (n = 1), allergic skin disorders (n = 1), allergic respiratory tract disorders (n = 4), and thyroid disorders (n = 2). Regarding medication, the most prevalent drug was oral birth control (n = 12), in addition to antidepressants (n = 1), thyroid drugs (n = 2), antiandrogen therapy drugs (n = 1), and asthma control or antiallergic drugs (n = 1). Four participants reported a baseline use of vitamin or mineral supplements. Almost all participants were nonsmokers (17/18), and 11 reported regular physical activity. Exclusion criteria included the following: (1) being vegetarian or vegan; (2) suffering from severe food allergies or food intolerances; (3) suffering from severe chronic inflammatory, infectious, endocrine, or metabolic diseases, including gastrointestinal disorders; (4) intake of antibiotic drugs or probiotic foods/supplements within the 2 weeks before intervention; and (5) being pregnant or breastfeeding. The participants' socio-demographic data and lifestyle information (diet, food restrictions, nutritional supplements, smoking, and physical activity habits) were recorded in an individual interview. At weeks 0 and 8 (last intervention week), 3-day food records[33] were retrieved for 13 of the 18 subjects (Table 1), and anthropometric assessments were carried out (Table S1), following standard guidelines for anthropometric data collection.[34,35]

2.2. Food Intervention. The dietary intervention ran for 8 consecutive weeks (Figure 1). Volunteers replaced a typical omnivorous lunch meal with a vegetarian-legume-based meal, 5 consecutive times per week (Monday to Friday), maintaining regular food habits throughout the rest of the day.

Study meals were supplied by a catering company (https://www.eurest.pt/), and meal composition was rigorously ascertained. The intervention diet included four basic elements: (i) vegetable soup (non-legume based), (ii) legume-based main course, (iii) dessert (fruit), and (iv) water as the only available beverage. The meals' food and nutritional composition were defined based on reference values for the general Portuguese adult population[36] and on dietary reference values (DRVs) for adults, indicated by the European Food Safety Authority (EFSA).[37] The legume-based main courses were standardized in terms of food composition, whereas all other meal items were made available *ad libitum* to mimic free-living conditions. Legumes or legume-based foods

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

were included in all vegetarian main courses, substituting animal-protein food sources. Meals could contain eggs or dairy products, as long as legumes remained the main protein food source. All main dish components were weighed before and after the meal to calculate legume intake. During holiday periods, the participants were provided with frozen legume-based main dishes and asked to provide a photograph of the meals before and after consumption, together with a description of the amount eaten using weight and/or household measures.

A total of 658 ovo-lacto-vegetarian legume-based lunch meals were delivered, containing one of the following types of legumes: chickpea (n = 197), bean (n = 163), lentil (n = 132), pea (n = 131) or soybean (n = 35). Over the 8-week trial, the participants consumed 38.0 (P_{25} = 34.75; P_{75} = 39.00) meals, resulting in an intake of 79.1 g (P_{25} = 74.85; P_{75} = 92.48) of cooked legumes per meal. Food records (obtained for 13 of the 18 subjects) revealed that daily baseline intake (before intervention) of cooked legumes was approximately 10.0 g (P_{25} = 0.00; P_{75} = 28.0), significantly increasing to 54.8 g (P_{25} = 47.5; P_{75}= 73.5; p = 0.007) at the end of the 8 weeks (Table 1, bottom section). Increases in total fiber intake [17.2 g (P_{25} = 14.1, P_{75} = 24.0, week 0), compared to 20.2 g (P_{25} = 16.1, P_{75} = 25.5, week 8); p = 0.046], and protein intake [74.5 g (P_{25} = 56.8, P_{75} = 92.6, week 0), compared to 84.0 g (P_{25} = 73.3, P_{75} = 102.7, week 8); p = 0.023], were also registered. No differences were observed in other macronutrients or micronutrients.

2.3. Urine Sample Collection. Participants were requested to provide urine samples under fasting conditions and weekly, from week 0 (baseline) to week 8 (Figure 1), amounting to 162 samples. However, 5 participants failed to deliver their samples at 1−2 time points (but were maintained in the study), resulting in a total of 156 urine samples. The sample collection protocol followed standard operating procedures established for urine NMR metabolomics,[38,39] with samples self-collected in the morning after a minimum of 8 h (and a maximum of 10 h) of overnight fasting. Midstream collection was recommended to avoid contamination from epithelial cells and bacteria in the urinary tract. Samples were collected into sterile, leak-proof, tightly sealed polyethylene containers and kept at 4 °C for up to 2 h, after which samples were stored at −80 °C until NMR analysis.

2.4. Sample Preparation and NMR Spectroscopy Analysis. Urine samples were prepared according to protocols described elsewhere.[40] Briefly, frozen samples were left to thaw at room temperature (approximately 30 min *per* sample) and homogenized before use. 700 μL of each sample was centrifuged (14,000 *g*, 5 min, room temperature) and 630 μL of the supernatant was added to 70 μL of 1.5 M phosphate buffer (KH_2PO_4/D_2O) at pH 7.4 (in 99.9% D_2O and containing 2 mM sodium azide and 0.1% 3-(trimethylsilyl)-propionic acid-d_4 (TSP) for chemical shift referencing) and homogenized. Sample pH was adjusted to 7.40 ± 0.02 using KOD (potassium hydroxide (Sigma-Aldrich) in D_2O) (4 M) or DCl (hydrochloric acid (Sigma-Aldrich) in D_2O) (4 M), and 600 μL of each mixture was transferred into 5 mm NMR tubes. All NMR spectra were acquired on a Bruker AVANCE III 500 spectrometer (Bruker, Rheinstetten, Germany), operating at a 500.13 MHz frequency for ^{1}H observation, using a 5 mm inverse probe, at 300 K. For each sample, a standard one-dimensional (1D) ^{1}H NMR spectrum was acquired with water peak suppression using a NOESY-1D pulse sequence ("noesypr1d"; Bruker library). Acquisition parameters were as follows: 128 scans, 64 k data points 12 019 Hz spectral width, 2.7 s acquisition time, 4 s relaxation delay, and 0.01 s mixing time.[40] Free induction decays were multiplied by a 0.3 Hz line-broadening factor prior to Fourier transformation. The spectra were manually phased, baseline corrected, and referenced to TSP at δ 0.00 ppm. Peak assignments were performed based on 1D and 2D NMR experiments (namely, total correlation spectroscopy (TOCSY), heteronuclear single-quantum coherence spectroscopy (HSQC), and J-resolved experiments), consultation of databases BBIORFCODE-2−0−0 database (Bruker Biospin, Rheinstetten, Germany) and Human Metabolome Database,[41] and literature.[42,43] Statistical total correlation spectroscopy (STOCSY) was also used to aid the assignment.[44] Unfortunately, 2D NMR and STOCSY could not provide further information on the possible nature of the unassigned features, which were observed to vary during the intervention.

2.5. Statistical Analysis. Descriptive statistics (IBM SPSS software, version 27) were used to describe the sociodemographic data, health-related features, and legume intake by the participants. For NMR data analysis, the spectral regions corresponding to water (4.45−5.05 ppm) and urea (5.49−6.16 ppm) resonances were excluded from the data matrix (Amix 3.9.14, Bruker BioSpin, Rheinstetten, Germany). Spectra alignment was performed using recursive segment-wise peak alignment (Matlab 7.12.0, The MathWorks, Inc.).[45] Normalization to total spectral area was carried out (since profile differences did not lead to largely distinct total area values, as revealed by visual inspection and supported by the use of an identical received gain for all spectra), and spectra were scaled to unit variance (UV) (SIMCA-P 11.5, Umetrics, Umea, Sweden). Multivariate analysis (MVA) comprised principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA) (SIMCA-P 11.5, Umetrics, Umea, Sweden). For all PLS-DA models, classification power (expressed by predictive power Q^2, specificity, sensitivity and classification rate) was further assessed by Monte Carlo Cross Validation (MCCV) (7 blocks, 500 runs).[46,47] For variable selection, spectral variables were selected as previously described,[48,49] namely through the intersection of three conditions: higher variable importance to projection (VIP) and lower standard errors in relation to both VIP (VIPcvSE) and b-coefficients (bcvSE), given by VIP > 1, VIP/VIPcvSE > 1 and |b/bcvSE| > 1, respectively. Then, PLS-DA was reapplied, and models underwent MCCV again. Models were considered valid in case the median predictive power (Q^2_{median}) was higher than 0.5. The peaks contributing to class discrimination were identified by integration not only of the resonances suggested by PLS-DA loadings, but also of all clear nonoverlapped resonances throughout the spectra. Peak integration was carried out in the original spectra, followed by normalization to the total spectral area. Effect size (ES) values were calculated,[50] and integral areas presenting ES > 0.5 and ES > ES_{error} were selected for statistical comparison (IBM SPSS software, version 27). The Shapiro−Wilk test was used to assess variable normality and the paired Student's t test or Wilcoxon test was applied accordingly. Statistical significance of metabolite variations was considered for p values <0.05. Dependence of urine metabolic signature on gender, age, body mass index (BMI), fat mass, and legume intake was assessed by PLS regression.[51] The (nonparametric) Spearman rank correlation analysis and its significance were calculated and

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

visualized using the g "*corrplot*" package, R software,[52] with a threshold of |r| > 0.6 and strong correlations considered for |r| > 0.75.[53,54]

3. RESULTS

3.1. ^{1}H NMR Spectra of Urine. Figure 2a illustrates the high resolution and complexity of urine spectra, with the

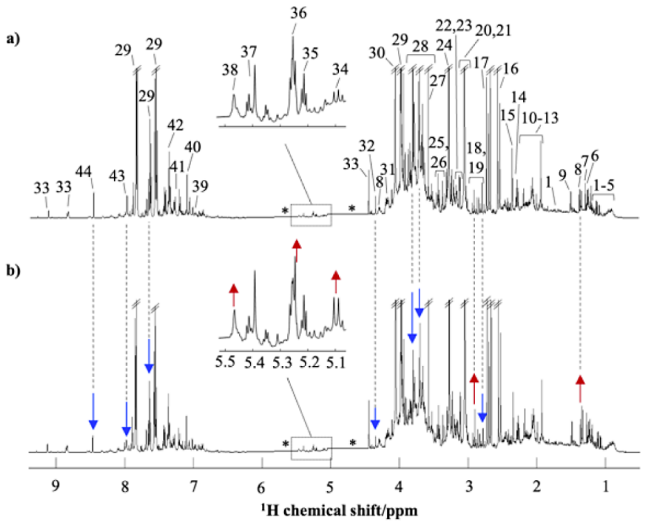


Figure 2. Average ^{1}H NMR spectra of urine samples from (a) baseline (week 0) and (b) week 7. Arrows and dashed vertical lines indicate visible spectral alterations: blue arrows indicate a visible decrease and red arrows indicate a visible increase. 1: leucine; 2: isoleucine; 3: valine; 4: 3-hydroxyisobutyrate (3-HIBA); 5: 4-deoxyerythronic acid (4- DEA); 6: 4-deoxythreonic acid (4-DTA); 7: 3-hydroxyisovalerate (3-HIVA); 8: threonine; 9: alanine; 10acetate; 11: *N*-acetyl glycoproteins (NAG); 12: isovalerylglycine; 13: acetone; 14: *N*-acetylglutamine; 15: *p*-cresol sulfate (*p*-cs); 16: citrate; 17: dimethylamine (DMA); 18: unassigned resonance at δ 2.78 (singlet); 19: unassigned resonance at δ 2.88 (singlet); 20: α-ketoglutarate (α-KG); 21:creatine; 22: malonate; 23: *cis*-aconitate; 24: trimethylamine-*N*-oxide (TMAO); 25: *scyllo*-inositol; 26: taurine; 27: glycine; 28: several overlapped resonances (including those arising from sugars); 29: hippurate; 30: creatinine; 31: 3-hydroxybutyrate (3-HBA); 32: tartrate; 33: trigonelline; 34: unassigned resonance at δ 5.11 (doublet); 35: xylose; 36: glucose; 37: allantoin; 38: anhydroglucose; 39: 4-hydroxyhippurate; 40: histidine; 41: 3-indoxylsulfate; 42: PAG, phenylacetylglutamine; 43: unassigned resonance at δ 7.95 (singlet); 44: formate. * Excluded regions: water (δ 4.45−5.05) and urea (δ 5.49−6.16).

example of the average ^{1}H NMR spectrum of urine at baseline (week 0), in which a total of 53 metabolites were identified (Table S2), consistently with previous assignments in other urine NMR spectra[42,43,55] (notably, assignments of 3-hydroxyisobutyrate (3-HIBA), pseudouridine, and trigonelline remain tentative, potentially benefiting from further spiking experiments for added certainty).

As week 7 was one of the time points where more changes were noted in urine composition (as discussed below), the corresponding average ^{1}H NMR spectrum is also shown (Figure 2b). Visual comparison suggests apparent increases in threonine (peaks 8, δ 1.33), glucose (overlapped peaks 36, δ 5.25), anhydroglucose (peaks 38, δ 5.46), and some unassigned resonances (peak 19, δ 2.88 and peak 34, δ 5.11) (Figure 2b, red arrows). In addition, possible decreases relate to hippurate (peaks 29, δ 7.65 and 7.83), tartrate (peaks 32, δ 4.35), and formate (peak 44, δ 8.47), several resonances in the sugar region (region 28), and unassigned resonances (peaks 43, δ 7.95 and peaks 18, δ 2.78) (Figure 2b, blue arrows). However, such apparent changes necessarily require statistical validation, particularly due to the high interindividual variability noted (as discussed below).

3.2. Multivariate and Univariate Statistical Analyses of NMR Data. An initial approach of pairwise comparison for consecutive weeks, both by PCA and PLS-DA, provided models with no visible group separation in PCA score plots and with low Q^2 values for PLS-DA models, thus indicating no statistical relevance (namely, with Q^2 values between −0.2 and 0.2). This meant that, considering the expected high interindividual variability in urine profile (given the small cohort size and that no variables other than one meal/day were controlled/matched), any consecutive diet-related changes were either absent or minimal. Hence, each week was compared to week 0 (baseline) to circumvent possible small gradual changes not picked up in consecutive pairwise comparisons. Table S3 indicates that the predictive power (expressed by Q^2) of the PLS-DA models obtained with the original NMR spectra remained under 0.5 for all comparisons, again reflecting weak or no group separation. However, upon variable selection, a method used to identify spectral regions more consistently related to group classes,[48,49] the apparent predictive power increased for all pairwise comparisons (Table S3), with Q^2 achieving values >0.5 for weeks 2 and 4−8, compared to baseline. The effect of variable selection is graphically illustrated for weeks 0 and 7, using the full original spectra (no variable selection; Figure S1a) and variable-selected spectra (Figure S1b). Variable selection improved group separation in unsupervised multivariate analysis through PCA (Figure S1, left) and increased the predictive power of the corresponding PLS-DA models (Q^2 0.665 compared to 0.481) (Figure S1, right, and Table S3). Hence, using reduced data matrixes may unveil meaningful group separations, although none of the resulting PLS-DA models showed classification ability, as viewed by MCCV analysis (all Q^2_{median} values <0.5). Despite this, group separation is hinted at in some PCA score plots and, in such cases, further clarified by PLS-DA (Figure S2). For instance, Figure 3 illustrates the clearer separation of week 0 vs week 7 (Q^2 0.665) compared to week 0 vs week 1 (Q^2 0.483) (although none of these models exhibited statistical robustness through MCCV) despite the random effects of several possible confounders (which are reflected in the high interindividual variability).

In order to identify statistically meaningful metabolite changes underlying group separation, the relevant resonances (chosen through PLS-DA loadings but also including all additional clear nonoverlapping resonances) were integrated, normalized to total area, and tested for significance. Table 2 lists all the identified metabolites (and still unassigned resonances) found to vary with statistical relevance in each week of intervention compared to week 0.

Overall, these included 20 identified compounds and 12 still unassigned resonances or spectral regions, confirming the increase in threonine within only the previously noted apparent changes detected visually (Figure 1). Varying metabolites comprised amino acids and derivatives (creatine, glycine, leucine, lysine, isoleucine, threonine, valine), organic acids (4-deoxyerythronic acid (4-DEA), 3-hydroxybutyrate (3-HBA), 2-hydroxyisobutyrate (2-HIBA), 3-HIBA, hippurate, malonate, pyruvate, succinate), *N*-acetylated glycoproteins

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

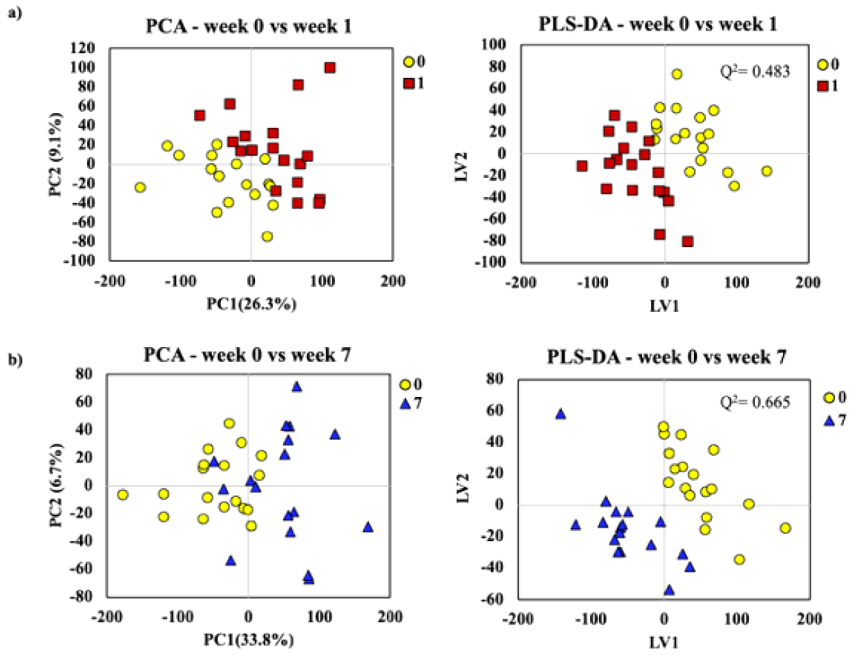


Figure 3. PCA and PLS-DA scores scatterplots obtained for the ^{1}H NMR spectra of urine for the baseline samples (week 0, yellow circles) versus (a) week 1 (red squares) and (b) week 7 (blue triangles) samples. All models were obtained with 2 principal components (PCA) and 2 latent variables (PLS-DA), using variable selection (see Experimental Section). Q^2: predictive power. For both PLS-DA models, MCCV resulted in $Q^2_{median} < 0.5$ and low specificity and sensitivity (approximately 50−70%), therefore indicating no/low classification capability.

(NAG) and other compounds (acetone, allantoin, pseudouridine, and trigonelline). It is interesting to note that weeks 2, 4, and 7 seem to be characterized by a higher number of statistically relevant variations, compared to the remaining weeks, whereas weeks 6 and 8 are the less eventful, only with increases in pseudouridine and unassigned region at δ 8.25. No significant changes were observed in body composition parameters upon the intervention (Table S1) and no significant correlations were found between BMI, or % fat mass and the NMR-measured excretory profile, as viewed by PLS regression (R^2 values <0.5). Therefore, we hypothesize that the set of 20 identified metabolites and additional unassigned compounds that varied significantly, at some point during the 8-week long intervention (despite the high intersubject variability), may arise from metabolic adaptations taking place as a response to the legumes diet ingestion.

The noted variations may be organized in terms of metabolic features or pathways (Table 3), suggesting that ingestion of a daily legume diet may impact important amino acid metabolism, with threonine, isoleucine, and, to a lesser extent, leucine showing more frequent significant increases along the intervention (Table 3 and Figure 4). Other amino acids are seen to increase in only one of the 8 weeks of the experiment. In particular, creatine and glycine were both increased in week 4 alone, whereas 4-DEA, lysine, and valine increased in week 7 (Table 3 and Figure 4). Notably, no amino acids varied significantly either in week 6 or in week 8, which (after verification of any coincidental interferences such as weekends or holidays) suggests a nonlinear metabolic response of the organism, perhaps tending to basal levels at the end of the intervention. This recovery of metabolite basal levels at week 8 is observed for all other compounds, with the exception of pseudouridine and an unassigned region at δ 8.25 (possibly arising from a nitrogen-containing compound) (Table 3).

The ketone body acetone increases early (weeks 1 and 2), subsequently recovering basal levels (Figure 5). The ketone bodies 3-HBA and 3-HIBA (Figure 6), are elevated in later weeks (Table 3 and Figure 5). No changes were noted in the third ketone body, acetoacetate. Together with ketone bodies, increases in tricarboxylic acid (TCA) cycle intermediates pyruvate (weeks 1, 2 and 7) and succinate (week 4) are also consistent with energy metabolism adaptations. Gut microbiota metabolites 2-HIBA (increased in weeks 2 and 7), allantoin (increased in weeks 2 and 5), and hippurate (decreased in week 2) suggest an early higher overall impact (week 2), in tandem with an increase in *N*-acetylated glycoproteins (increased in weeks 2 and 7) (Table 3 and Figure 5).

To the best of our knowledge, an apparently consistent increase in pseudouridine was newly observed, compared to week 0 (from week 3 and despite intersubject variability) (Table 3 and Figure 5). Slight changes in hippurate, malonate, and trigonelline were noted early on in the intervention (weeks 1−3), all quickly returning to average values, largely affected by high interindividual variability (Figure 5). The bottom section of Table 3 (and Figure S3) lists the still unassigned resonances arising from urine metabolites, which change significantly during the intervention, either consistently throughout the intervention (U3 and U4, at δ 1.45 and 1.79, respectively) (Figure S3), or as descriptive features of the initial intervention impact (namely, U6, U7 and U8, at δ 2.20, 2.21 (probably an acetate species), and 6.49, respectively). This illustrates the importance and need for further peak identification in urine NMR spectra, a recurrent challenge in urine metabolomics. Overall, it seems possible that a vast set of metabolite variations take place up to week 4, seemingly winding down to week 6, in which only pseudouridine and an unassigned resonance (U11) are varying, exactly as seen for week 8. Week 7 stands out, as described by a high number of apparent

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

Table 2. Univariate Statistical Analysis of Metabolite Variations in Each Week of the Intervention, Compared to Week 0[a]

Week no.	Metabolites	δ_{ppm} (multiplicity)[b]	ES ± ES_{error}	p value
Week 1	Acetone	2.24 (s)	0.73 ± 0.675	0.011*
	Malonate	3.11 (s)	0.75 ± 0.676	0.015*
	Pyruvate	2.38 (s)	0.70 ± 0.673	0.002**
	Threonine	4.26 (dd)	0.91 ± 0.686	<0.001***
	U4	1.79 (m)	0.77 ± 0.677	0.002**
	U5	2.18 (s)	0.67 ± 0.672	0.022*
	U6	2.20 (s)	1.04 ± 0.696	<0.001***
	U7	2.21 (s)	1.02 ± 0.694	0.001**
	U8	6.49 (d)	0.71 ± 0.674	0.035*
Week 2	2-HIBA	1.36 (s)	0.85 ± 0.692	0.015*
	Acetone	2.24 (s)	0.78 ± 0.687	0.036*
	Allantoin	5.39 (s)	0.98 ± 0.701	0.004**
	Hippurate	8.52 (br)	-0.82 ± 0.690	0.001**
	Isoleucine	0.98 (d)	0.88 ± 0.694	0.007**
	3-HIBA	1.08 (d)	0.74 ± 0.685	0.002**
	NAG	2.06 (br)	0.73 ± 0.685	0.042*
	Pyruvate	2.38 (s)	0.85 ± 0.692	0.013*
	U3	1.45 (d)	1.13 ± 0.714	0.004**
	U4	1.79 (m)	1.22 ± 0.721	<0.001***
	U6	2.20 (s)	1.16 ± 0.716	0.001**
	U7	2.21 (s)	1.12 ± 0.713	0.003**
	U9	2.39 (reg.)	0.98 ± 0.702	0.012*
	U8	6.49 (d)	0.84 ± 0.691	0.026**
	U11	8.25 (reg.)	-0.90 ± 0.696	0.031*
	U12	8.32 (reg.)	-0.69 ± 0.682	0.017*
Week 3	3-HBA	4.15 (m)	0.82 ± 0.713	0.024*
	Isoleucine	0.98 (d)	0.71 ± 0.706	0.041*
	Pseudouridine	7.68 (s)	0.80 ± 0.712	0.020*
	Trigonelline	9.12 (s)	0.81 ± 0.712	0.005**
	U1	1.13 (s)	0.75 ± 0.709	0.006**
	U4	1.79 (m)	0.80 ± 0.712	0.004**
	U6	2.20 (s)	0.71 ± 0.706	0.020*
	U7	2.21 (s)	0.72 ± 0.707	0.017*
Week 4	Creatine	3.04 (s)	0.68 ± 0.672	0.016*
	Glycine	3.57 (s)	0.73 ± 0.675	0.001**
	Isoleucine	0.98 (d)	0.76 ± 0.676	0.007**
	Leucine	0.96 (t)	0.73 ± 0.675	0.002**
	3-HIBA	1.08 (d)	0.70 ± 0.673	0.006**
	Pseudouridine	7.68 (s)	1.30 ± 0.719	0.001**
	Succinate	2.41 (s)	0.94 ± 0.688	0.001**
	Threonine	4.26 (dd)	0.78 ± 0.678	0.036*
	U3	1.45 (d)	0.91 ± 0.686	0.010*
	U4	1.79 (m)	0.81 ± 0.680	0.022*
	U6	2.20 (s)	0.99 ± 0.692	0.003**
	U8	6.49 (d)	0.71 ± 0.674	0.016*
	U11	8.25 (reg.)	-1.15 ± 0.705	0.002**
Week 5	3-HBA	4.15 (m)	0.93 ± 0.688	0.018*
	Allantoin	5.39 (s)	0.70 ± 0.673	0.022*
	Isoleucine	0.98 (d)	0.73 ± 0.675	0.025*
	Leucine	0.96 (t)	0.68 ± 0.672	0.025*
	Pseudouridine	7.68 (s)	0.91 ± 0.686	0.012*
	Threonine	4.26 (dd)	0.84 ± 0.681	0.013*
	U4	1.79 (m)	0.99 ± 0.692	0.001**
	U11	8.25 (reg.)	-1.50 ± 0.740	<0.001***
Week 6	Pseudouridine	7.68 (s)	0.94 ± 0.688	0.008**
	U11	8.25 (reg.)	-1.47 ± 0.736	<0.001***
Week 7	2-HIBA	1.36 (s)	0.94 ± 0.698	0.007**
	3-HBA	4.15 (m)	0.96 ± 0.700	0.007**
	3-HIBA	1.08 (d)	0.85 ± 0.692	0.012*
	4-DEA	1.11 (d)	0.90 ± 0.696	0.004**
	Isoleucine	0.98 (d)	1.12 ± 0.713	<0.001***

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

Table 2. continued

Week no.	Metabolites	δ_{ppm} (multiplicity)[b]	ES ± ES_{error}	p value
	Leucine	0.96 (t)	0.74 ± 0.685	0.001**
	Lysine	1.73 (m)	0.86 ± 0.693	<0.001***
	NAG	2.06 (br)	0.80 ± 0.689	0.013*
	Pseudouridine	7.68 (s)	0.83 ± 0.691	0.007**
	Pyruvate	2.38 (s)	1.06 ± 0.708	0.001**
	Threonine	4.26 (dd)	1.12 ± 0.713	<0.001***
	Valine	1.05 (d)	0.79 ± 0.688	0.018*
	U1	1.13 (s)	0.79 ± 0.688	0.012*
	U2	1.25 (d)	0.79 ± 0.688	0.007**
	U3	1.45 (d)	0.99 ± 0.702	0.006**
	U4	1.79 (m)	1.31 ± 0.730	<0.001***
	U6	2.20 (s)	0.97 ± 0.701	0.013*
	U7	2.21 (s)	0.91 ± 0.696	0.012*
	U9	2.39 (reg.)	0.90 ± 0.696	0.011*
Week 8	Pseudouridine	7.68 (s)	1.07 ± 0.709	0.010*
	U11	8.25 (reg.)	-1.00 ± 0.703	0.017*

[a]Abbreviations: 2-HIBA: 2-hydroxyisobutyrate; 3-HBA: 3-hydroxybutyrate; 3-HIBA: 3-hydroxyisobutyrate; 3-HIVA: 3-hydroxyisovalerate; 4-DEA: 4-deoxyerythronic acid; NAG: *N*-acetyl glycoproteins; Ui: unassigned resonance i, numbered by increasing chemical shift (as shown in Table 3); s: singlet; d: doublet; t: triplet; dd: doublet of doublets; m: multiplet; reg.: complex spectral region; br: broad NMR resonance; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. [b]Peak is used for integration, within the whole spin system of the metabolite.

metabolite changes, although differing in profile in relation to week 4. Given the high interindividual variability characterizing this small cohort of 18 subjects, the presentation of metabolite variations (Tables 2 and 3) was carried out in a qualitative manner rather than computing % variation or effect size, which should indeed be attempted for larger cohorts.

Attempting to further understand the human biochemistry underlying the apparent above changes, correlations between nutrient intake parameters, and the varying NMR resonances were calculated for (i) weeks 4 and 0, (ii) weeks 7 and 0, and (iii) weeks 8 and 0 (Figure S4), assuming that intragroup variability within week 0 samples (no legume intake) may be descriptive of the cohort at baseline. Despite the high data variability, which demands careful interpretation of these results, some correlations ($|r| > 0.6$) were identified during the intervention, while absent at baseline (Table S3). We hypothesize that these correlations may arise as a response to the intervention, although, notably, no correlations were observed with legume intake (last line in plots in Figure S4). First, it is noted that in each of weeks 4, 7, and 8 (Figure S4a–c, respectively), correlations are found between some of the main nutrients and TCA cycle intermediates: saturated fat and α-ketoglutarate (α-KG) (−) in week 4; polyunsaturated fat and citrate (+) in week 7 (in tandem with energy intake correlated to energy metabolites *cis*-aconitate and creatine); and total carbohydrates and *cis*-aconitate (−) in week 8 (and energy intake to *cis*-aconitate). The remaining correlations (Figure S4) involve two vitamins (riboflavin and vitamin B_6) and several minerals (calcium, iron, magnesium, phosphorus, and potassium) (Table S3). Although potentially informative on the detailed metabolism of these nutrients, we chose to only highlight the most consistent correlations, namely those between (i) vitamin B_6 and the energy-related metabolites creatine, creatinine (weeks 4 and 7), and threonine (both glucogenic and ketogenic amino acid) (week 8); (ii) potassium levels and creatine and glycoproteins (NAG) (weeks 4 and 8); and (iii) iron levels and lysine (4, 7, and 8 weeks).

4. DISCUSSION

To our knowledge, this work reports the first NMR metabolomics study of the impact of a legume-based long-term dietary intervention (8 weeks, compared to 2 days in previous reports[19] on the metabolome of urine collected longitudinally, although of a preliminary nature due to the small cohort size ($n = 18$). The weekly urinary profiles of the subjects illustrated well the high interindividual variability (undoubtedly related to intersubject different characteristics, as described in the Experimental Section). However, despite such variability, some statistically relevant changes could be detected. These changes unveiled some potential new markers of legume intake, both by accounting for associated metabolic effects triggered by the legume diet and by directly reflecting legume composition.

Prior to discussing such metabolic markers in detail, it is important to briefly address the issue of confounders, in relation to the urinary profile. Namely, as the cohort was mainly composed of female subjects, the possibility of the menstrual cycle contributing to the observed changes was considered, although only lysine was found to change in common with a reported urinary signature of menstrual cycle evolution.[56] Also, as a random confounder (as evaluated through personal questionnaires), its effects are not expected to lead to consistent metabolite changes. We suggest that similar reasoning may apply to other random individual features across the cohort under study. As to the nutritional outcomes of the intervention, the 3-day food records at baseline (week 0) and week 8 showed that the dietary transition increased total dietary fiber and protein consumption, as well as legume content, as expected (Table 1). The higher daily intake of fiber should reflect the higher amounts of fiber in the intervention meals, compared to baseline lunch meals, namely, + 2.7 g ($P_{25} = 1.3$, $P_{75} = 6.4$) with $p = 0.002$. Indeed, an 8-fold increase in legume consumption was noted during the trial, with fiber expected to make up to 30% of their dry weight (considering the most commonly consumed legumes worldwide.[57,58] The baseline intake of legumes was similar to what has been reported for the Portuguese

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

Table 3. Statistically Relevant Urine Metabolite Variations as a Function of Time Compared to Week 0 (Arrows Indicate Direction of Variation), Grouped by Putative Specific Metabolic Features/Pathways[a]

Metabolites	δ_{ppm} (multiplicity)	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
AMINO ACID METABOLISM									
4-DEA	1.11 (d)							↑**	
Creatine	3.04 (s)				↑*				
Glycine	3.57 (s)				↑**				
Isoleucine	0.98 (d)		↑**	↑*	↑**	↑*		↑***	
Leucine	0.96 (t)				↑**	↑*		↑**	
Lysine	1.73 (m)							↑***	
Threonine	4.26 (dd)	↑***			↑*	↑*		↑***	
Valine	1.05 (d)							↑*	
KETOGENESIS									
3-HBA	4.15 (m)			↑*		↑*		↑**	
3-HIBA[b]	1.08 (d)		↑**		↑**			↑*	
Acetone	2.24 (s)	↑*	↑*						
TCA CYCLE INTERMEDIATES									
Pyruvate	2.38 (s)	↑**	↑*					↑**	
Succinate	2.41 (s)				↑**				
GLYCOPROTEINS METABOLISM									
N-acetyl glycoproteins	2.06 (br)		↑*					↑*	
MICROBIOTA METABOLISM									
2-HIBA	1.36 (s)		↑*					↑**	
Allantoin	5.39 (s)		↑**			↑*			
Hippurate	8.52 (br)		↓**						
NITROGEN BASES METABOLISM									
Pseudouridine[b]	7.68 (s)			↑*	↑**	↑*	↑**	↑**	↑*
LEGUME INTAKE BIOMARKERS (or components)									
Hippurate[c]	8.52 (br)		↓**						
Malonate[e]	3.11 (s)	↑*							
Trigonelline[b,f]	9.12 (s)			↑**					
Unassigned resonances									
U1	1.13 (s)			↑**				↑*	
U2	1.25 (d)							↑**	
U3	1.45 (d)		↑**		↑*			↑**	
U4	1.79 (m)	↑**	↑***	↑**	↑*	↑**		↑***	
U5	2.18 (s)	↑*							
U6	2.20 (s)	↑***	↑**	↑*				↑*	
U7	2.21 (s)	↑**	↑**	↑*				↑*	
U8	6.49 (d)	↑*	↑**		↑*				
U9	2.39 (reg)		↑*					↑*	
U10	2.90 (reg)				↑**				
U11	8.25 (reg)		↓*		↓**	↓***	↓***		↓*
U12	8.32 (reg)		↓*						

[a]Abbreviations: 2-HIBA: 2-hydroxyisobutyrate; 3-HBA: 3-hydroxybutyrate; 3-HIBA: 3-hydroxyisobutirate; 4-DEA: 4-deoxyerythronic acid; NAG: *N*-acetyl glycoproteins; TCA: tricarboxylic acid cycle; Ui: unassigned resonance i; s: singlet; d: doublet; t: triplet; dd: doublet of doublets; m: multiplet; reg.: complex spectral region; br: broad resonance. [b]Tentative assignment. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. [c]Metabolite reported in relation to plant-rich diets.[27,29] [e]Metabolite identified in chickpea and soybean plant tissues and root nodules.[78−81] [f]Biomarker of legume intake.[13,14,17,19]

population and several other European countries (<20g/day). However, in the *ad libitum* setting of the intervention, many participants were able to achieve the higher Portuguese dietary guidelines for daily legume intake. During the intervention, 8 subjects (44.4%) showed a mean daily intake that met the Portuguese guidelines of ≥80 g *per* day,[36] whereas 2 subjects (11.1%) reached international recommendations of ≥100 g *per* day.[58] Interestingly, recent studies have shown an increased willingness of Portuguese adults to consume legumes as a substitute for meat and fish [(15.0% in 2014 versus 31.7% in 2020 ($p < 0.001$)],[59,60] although also identifying a general lack of awareness in relation to the recommended daily intake.[59] This makes the results of the present intervention even more relevant, given the ability of participants to easily reach the dietary guidelines for daily legume intake. As to the increase in daily protein intake, although it may reflect the higher protein content in the trial meals, some contribution from reported intake of morning snacks may be accountable (median of 3.7 g at week 8 vs 2.0 g at week 0, p = 0.012). A more detailed nutritional discussion of these results may be found in a recent report.[61] Furthermore, the food intervention caused no variation in body composition parameters (Table S1), which was expected and desirable. Indeed, the intervention diet was planned to be isocaloric, to minimize variations in anthro-

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

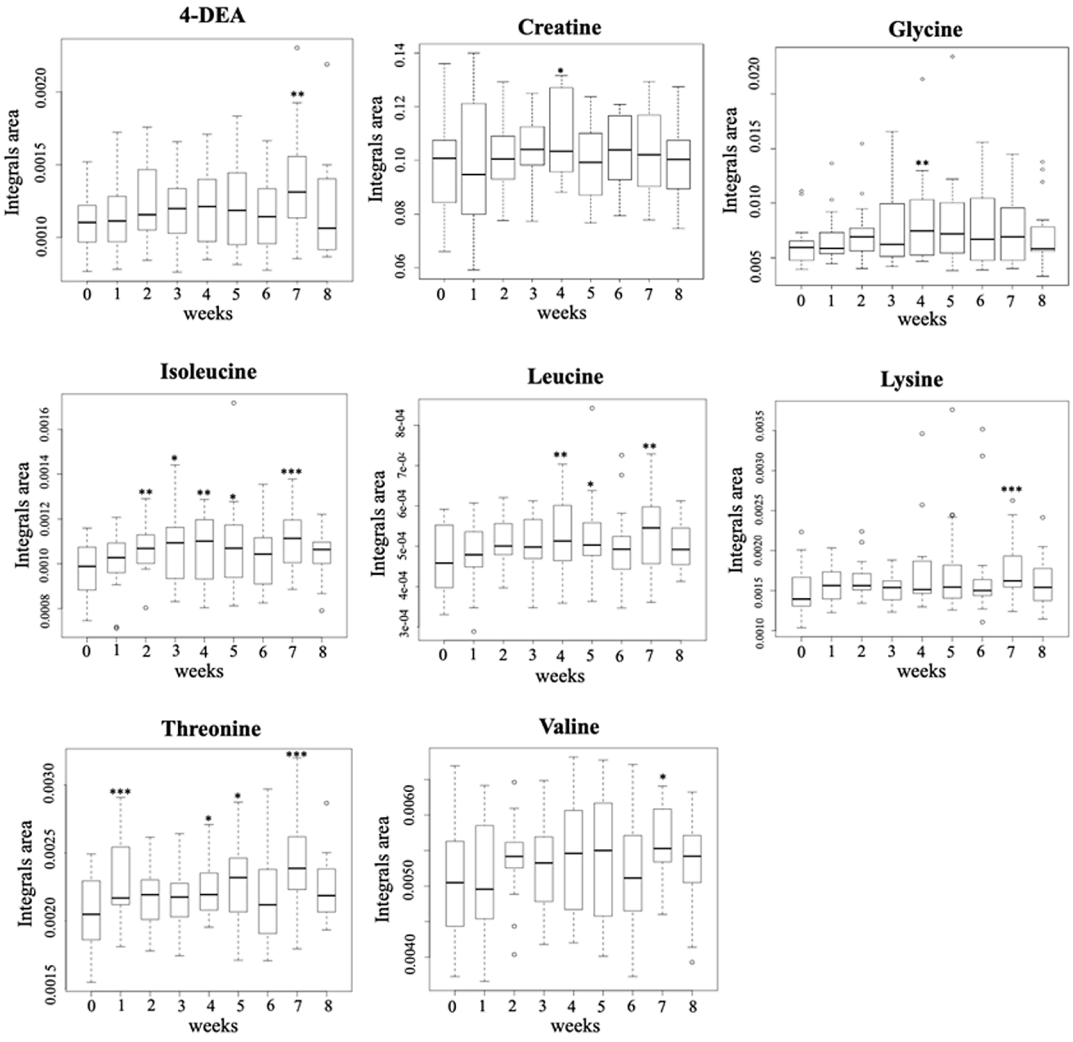


Figure 4. Boxplots of changes in amino acids and derivatives (ordered in alphabetical order). The statistical significance indicated by asterisks corresponds to the comparison of each week's samples against those obtained at baseline (week 0). 4-DEA: 4-deoxyerythronic acid; *$p < 0.05$; ** $p < 0.01$; ***$p < 0.001$.

pometric measures that could impact the urine excretory profile and act as potential confounders.

Regarding the potential new information on markers of legume intake (either metabolism mediated or directly reflecting legume composition), the NMR results suggest that the individual response may be nonlinear, with a possible tendency for amino acid, energy metabolism, and gut microbiota to approach basal levels after 8 weeks (with the exception of a persisting deviation expressed by increased pseudouridine levels). Indeed, weeks 2, 4, and 7 seemed to exhibit a higher number of excreted metabolite variations, with weeks 6 and 8 corresponding to profiles close to baseline. This observation demonstrates the importance of longer-term interventions in identifying time-dependent responses, although confirmation of these results in larger cohorts is of paramount importance. One of the main observed changes reflected on amino acid metabolism, in particular leucine, isoleucine and threonine, among others (Figure 6). All the increased amino acids may be found in legumes' proteins and can, therefore, be excreted in higher amounts due to the increased intake of legumes.[62] However, some of the major legumes' nonessential amino acids,[62] asparagine, aspartic acid, glutamine and glutamate, were not identified in urine, which suggests that they are probably serving other metabolic fates, *e.g.*, acting as anaplerotic substrates to the TCA cycle.[63] In fact, TCA cycle activation upon legume intake has been previously suggested, and only glutamine, leucine, and lysine, as well as trigonelline (observed here to increase in week 3), have been reported as potential urinary biomarkers of legumes' dietary exposure.[13,19] Our results (Figure 6) suggest that anaplerotic amino acids threonine, isoleucine, and, to a lesser extent, leucine may be increased in urine, thus apparently not being completely used up in TCA cycle enhancement, as also noted for glycine (increased in week 4), lysine, or valine (increased in week 7). Glycine appears to increase together with creatine at week 4, whereas 4-DEA accompanies lysine and valine at week 7. Glycine is a precursor of creatine, through guanidinoacetate, and we putatively hypothesize that both glycine and creatine increases may indicate a response of the creatine kinase (CK)/creatine/phosphocreatine system at week 4, with relates to ATP/ADP pool regulation.[64] An enhanced TCA cycle may also be suggested with basis on the accumulation of pyruvate and succinate, the former proposed to arise from the conversion of threonine into α-ketobutyrate, propionyl-CoA, and then to succinil-CoA (Figure 6).[65] Moreover, the increases found for 4-DEA may be associated with threonine variations

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

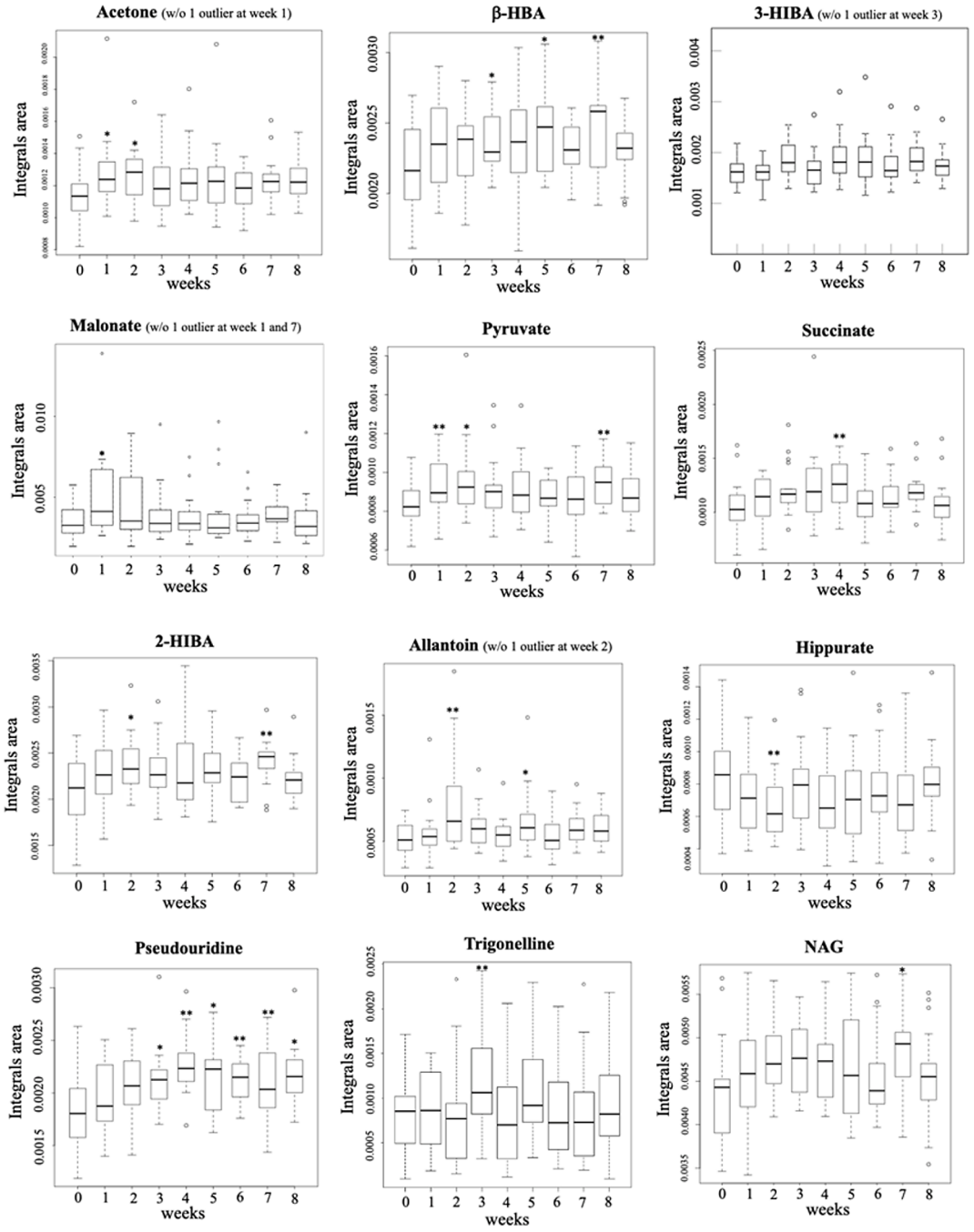


Figure 5. Boxplots of changes in all identified metabolites (except amino acids and derivatives). The statistical significance indicated by asterisks corresponds to the comparison of each week's samples against those obtained at baseline (week 0). 2-HIBA: 2-hydroxyisobutyrate; 3-HBA: 3-hydroxybutyrate; 3-HIBA: 3-hydroxyisobutirate; NAG: *N*-acetyl glycoproteins; w/o: without; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

since the former is a byproduct of threonine degradation.[66] The early (weeks 1 and 2) increase in excreted ketone body acetone, followed by the ketone body 3-HBA (weeks 3, 5, and 7), may be indicative that the organism is responding with an increased energetic status to the legume diet, using hepatic lipid storages to increase ketone bodies in circulation to promote TCA cycle activation in extrahepatic tissues. Such a tentative hypothesis would be consistent with the correlations found between saturated fat and TCA cycle intermediate α-KG (week 4) and polyunsaturated fat and citrate (week 7), which would suggest an important contribution of fat reserves for energy production. Interestingly, total carbohydrates appear as the expected dominant energy source at week 8, an idea putatively advanced with basis on their correlation with *cis*-aconitate.

Ketone bodies have been previously noted to vary in an observational legume intake study, with acetoacetate and isobutyrate decreasing,[13] instead of increasing, as noted here for 3-HBA (probably due to high interindividual variability in different cohorts). We also hypothesize that the increase in 3-HIBA, a product of valine degradation,[67] may relate to leucine and isoleucine and, indirectly, to 3-HBA (Figure 6), although such possibility needs further investigation. Indeed, 3-HIBA is one of the hydroxy acids found elevated in the urine and serum

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

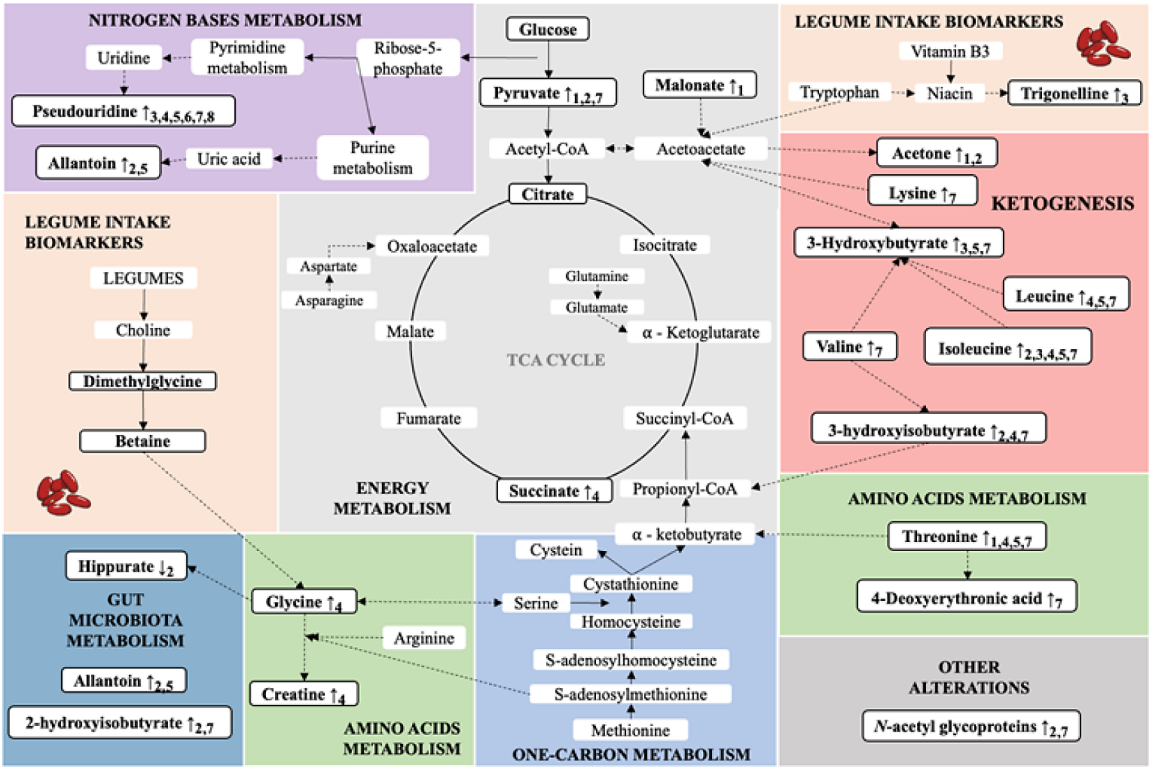


Figure 6. Schematic representation of putative biochemical explanations for the metabolite changes observed during the 8-week legume-based intervention. Metabolites detected in urine samples are boxed and shown in bold, although it should be noted that the complete list of identified compounds may be found in Table S2. Metabolites with statistically significant variations are boxed and shown in bold and underlined, with arrows and numbers to indicate the direction of significant variations ($p < 0.05$) and corresponding intervention weeks, respectively; dashed pathway arrows indicate several intermediate pathway steps. Background color code: dark blue: gut microbiota metabolism; green: amino acid metabolism; light blue: one-carbon metabolism; light gray: energy metabolism; peach: legume intake biomarker; pink: ketogenesis; purple: nitrogen base metabolism; dark gray: other alterations. The diet intervention involved a set of legumes and not just one type, as schematically depicted in this figure (please check the Experimental Section for more specific information).

samples of patients affected by ketoacidosis, possibly as a result of deviant branched chain amino acid (BCAA) metabolism.[68] Here, ketoacidosis does not clinically apply; however, the apparent increase in ketone bodies may be an indication of enhanced use of fat reserves, as suggested above, along with a deviation in BCAA metabolism (all leucine, isoleucine, and valine seem to increase at some stage in the intervention), possibly giving rise to 3-HIBA, as well as feeding into the TCA cycle.

The apparent slight increase in *N*-acetylated glycoproteins (NAG) metabolism early in the intervention (weeks 1 and 2) is consistent with previous reports related to plant-based diets,[29] similarly to changes in glycine[26,28,29] (here, seen increased at week 4), succinate[29] (increased at week 4), and hippurate[27,29] (decreased at week 2). We putatively hypothesize that NAG may arise directly from legumes, namely from plant lectins.[69,70] The decrease in hippurate noted here (week 2) is in apparent contradiction of earlier reports of increased urinary levels,[27,29] particularly related to ingestion of fruits and vegetables[71] and gut bacteria fermentation.[72] However, the fact that a consistent hippurate variation is not observed along the 8 weeks suggests that its variation may reflect a dynamic metabolic response or, most probably, high interindividual variability. In addition, the increase in allantoin may relate to gut bacteria with the ability to synthesize uric acid-metabolizing enzymes (uricase, allantoinase, and allantoicase)[73] to degrade uric acid into 5-hydroxyisourate, allantoin, allantoate and urea. However, allantoin may also be related to a potentially beneficial antioxidant protective mechanism,[74] consistently with the expected role of legumes in reducing oxidative stress.[75] Pseudouridine is a nonclassical nucleoside detected in human urine[76] as a degradation product of RNA,[77] and believed not to be metabolized by the human organism.[78] Its persisting increase in relation to legume intake is reported here for the first time, to the best of our knowledge. As pseudouridine has also been found in beans (in higher amounts than in lentils or chickpeas,[79] having recently been studied in relation with plant metabolism,[80] we suggest that increased pseudouridine may directly reflect legume intake. In addition, trigonelline, a recognized legume consumption marker,[13,14,17,19] was here hinted to increase but only early in the intervention, which suggests a possible limitation of its use in longer interventions and/or when interindividual variability is high. Malonate has also been detected in chickpeas and soybean plant tissues and root nodules[81−84] and recognized as a flavonoid precursor in fruits, grains, and legumes.[71] It is possible, therefore, that malonate may also serve as a legume marker, although it is again subject to confirmation in larger cohorts.

Finally, it is interesting to note the highlighted correlations between vitamin B_6, potassium, and iron with different metabolites. Regarding vitamin B_6, its apparent relationship with energy-related compounds is consistent with this vitamin being a cofactor of numerous enzymatic reactions for energy production, mainly from protein sources.[85] It is possible, therefore, that legume ingestion activates protein catabolism,

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

largely due to the higher protein intake. If confirmed in the future, the potassium/creatinine negative correlation may also result from a higher relative intake of legumes, generally richer in potassium compared to meat,[86] for which creatinine is a marker.[87] Finally, the apparent negative relationship between iron intake and lysine may relate to a previous suggestion that L-lysine may play a role in improving iron absorption,[88,89] although this has yet to be further demonstrated.

5. CONCLUDING REMARKS

We report on a preliminary longitudinal NMR metabolomics study of the urinary metabolic fingerprint of a small pilot (n = 18) legume-based 8-week intervention on a group of traditionally omnivorous adults. Despite high interindividual variability and subsequent relatively low statistical robustness of PLS-DA models, the results suggest that the human organism response to legume intake may be highly active in the first weeks, subsequently tending toward baseline. Anaplerotic activation of the TCA cycle is suggested together with ketogenesis and enhanced catabolism of fat resources, although concomitant protein and carbohydrate catabolism may also occur, at least until week 7. Increased amino acid excretion is part of an apparent signature of response, along with trigonelline, malonate, and possibly pseudouridine as markers of legume intake. Although at a preliminary level, the abovementioned results suggest that legume intake may cause unique changes in the human excretory metabolome, justifying the need for further research in expanded cohorts.

This study is not without limitations, namely, the small size of the cohort, which was also predominantly female, so that results may be somewhat biased toward female metabolism. This adds to the absence of a separate control group (hindered by practical stumbling blocks), the need to rely on the individuals for urine sample collection, and the subjectivity of individual food records. It is also important to note that although the study benefitted from its longitudinal nature (several collection points per subject, taking week 0 samples as personalized controls), the high interindividual variability may mask additional metabolic features, thus justifying the use of more specific multivariate analysis methods to extract meaningful information (*e.g.,* linear mixed models). The results obtained will necessarily require future validation in larger cohorts and, possibly, enhanced control over meal components (although the aim here was precisely to mirror free-living conditions) in tandem with knowledge on legume origin/composition. Furthermore, the advanced putative explanatory hypotheses and potential new legume biomarkers will require further investigation, for instance, through isotopic tracing studies in animal models and a more complete assignment strategy to identify presently unknown compounds varying as a result of legume intake.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c04795.

PCA and PLS-DA scores scatter plots obtained from the ^{1}H NMR spectra of urine representing baseline samples versus week 7, before (a) and after (b) variable selection (Figure S1); PCA and PLS-DA score scatterplots obtained from the ^{1}H NMR spectra of urine representing baseline samples (week 0, yellow circles) versus each of the 8 weeks of the intervention (1−8; a−h) (Figure S2); boxplots of all varying unassigned metabolites (Ui) (Figure S3); statistical correlation (Spearman) plots p<0.05 and r>|0.6|) between varying urinary metabolites (as detected by NMR, x-axis) with levels of macro- and micronutrients intake and daily legumes intake (y-axis) at (a) week 4 vs baseline; n =13; (b) week 7 vs baseline; n = 13, and (c) week 8 vs baseline; n = 13 (Figure S4); anthropometric assessments (Table S1); metabolites assigned in the ^{1}H NMR urine spectra of the subjects included in the intervention (Table S2); quality predictive power (Q2) values for PLS-DA models before and after variable selection (Table S3); list of correlations (|r| > 0.6) registered between intake parameters and urine varying metabolites, for baseline alone and weeks 4, 7 and 8, compared to baseline (week 0), corresponding to the maps shown in Figure S4 (Table S4) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Elisabete Pinto − *CBQF - Centro de Biotecnologia e Química Fina − Laboratório Associado, Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Porto 4200-072, Portugal; EPIUnit - Instituto de Saúde Pública, Universidade do Porto, Porto 4050-600, Portugal*; Email: ecbpinto@ucp.pt

Ana M. Gil − *Department of Chemistry and CICECO-Aveiro Institute of Materials, Universidade de Aveiro, Aveiro 3810-193, Portugal*; orcid.org/0000-0003-3766-4364; Email: agil@ua.pt

Authors

Helena Ferreira − *CBQF - Centro de Biotecnologia e Química Fina − Laboratório Associado, Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Porto 4200-072, Portugal*

Daniela Duarte − *Department of Chemistry and CICECO-Aveiro Institute of Materials, Universidade de Aveiro, Aveiro 3810-193, Portugal;* Present Address: Instituto de Investigação e Inovação em Saúde (i3S), University of Porto, Porto, Portugal; orcid.org/0000-0001-9460-1162

João A. Rodrigues − *Department of Chemistry and CICECO-Aveiro Institute of Materials, Universidade de Aveiro, Aveiro 3810-193, Portugal*

Marta W. Vasconcelos − *CBQF - Centro de Biotecnologia e Química Fina − Laboratório Associado, Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Porto 4200-072, Portugal*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c04795

Author Contributions

#E.P. and AMG authors contributed equally to this work. The authors' responsibilities are as follows: H.F., M.W.V., E.P., and A.M.G. designed the dietary intervention study; H.F. implemented the food trial and performed all data collection, including the biological samples for NMR analysis; H.F., D.D., and J.R. conducted NMR metabolomics analysis as well as data processing and statistical analysis; H.F. and A.M.G. drafted the original manuscript and have primary responsibility for the

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

final content; M.W.V., E.P., and A.M.G. supervised the study execution, validated data outputs, and provided essential intellectual inputs; and all authors revised and approved the final manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by National Funds from FCT - Fundação para a Ciência e a Tecnologia through project UIDB/50016/2020 and by the project "Transition paths to sustainable legume-based systems in Europe" (TRUE), which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 727973. HF would like to acknowledge FCT for doctoral grant ref SFRH/BDE/132240/2017. This work was developed within the scope of the project CICECO-Aveiro Institute of Materials, UIDB/50011/2020, UIDP/50011/2020 & LA/P/0006/2020, financed by national funds through the FCT/MCTES (PIDDAC). The NMR spectrometer is part of the National NMR Network (PTNMR) and partially supported by Infrastructure Project N° 022161 (cofinanced by FEDER through COMPETE 2020, POCI and PORL and FCT through PIDDAC). All the authors acknowledge the help of Joana Ferrão Silveira, MSc, and Kimhoung Say, MSc, during the food intervention and data collection procedures, as well as the food company Eurest (https://www.eurest.pt/) for providing the trial meals. The INCREASE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 862862.

■ ABBREVIATIONS

2-HIBA: 2-hydroxyisobutyrate; 3-HBA: 3-hydroxybutyrate; 3-HIBA: 3-hydroxyisobutyrate; 3-HIVA: 3-hydroxyisovalerate; 4-DEA: 4-deoxyerythronic acid; BCAA: branched-chain amino acid; br: broad NMR resonance; CK: creatine kinase; d: doublet; dd: doublet of doublets; DMA: dimethylamine; DRVs: dietary reference values; E: total daily energy intake; EFSA: European Food Safety Authority; ES: effect size; m: multiplet; MCCV: Monte Carlo cross-validation; MVA: multivariate analysis; NAG: *N*-acetylated glycoproteins; NE: niacin equivalents; P_{25}: 25th percentile; P_{75}: 75th percentile; PCA: principal component analysis; PLS-DA: partial least-squares discriminant analysis; PREDIMED: Primary Prevention of Cardiovascular Disease with a Mediterranean Diet; RE: retinol equivalents; reg.: complex spectral region; s: singlet; t: triplet; TCA: tricarboxylic acid cycle; TMA: trimethylamine; TMAO: trimethylamine-*N*-oxide; TSP: 3-(trimethylsilyl)-propionic acid-d4; Ui: unassigned resonance i, numbered by increasing chemical shift; UV: unit variance; VIP: variable importance to projection; α-KG: α-ketoglutarate

■ REFERENCES

(1) Chen, C.; Chaudhary, A.; Mathys, A. Dietary Change and Global Sustainable Development Goals. *Front. Sustain. Food Syst.* **2022**, *6*, 6.

(2) Fanzo, J.; Bellows, A. L.; Spiker, M. L.; Thorne-Lyman, A. L.; Bloem, M. W. The Importance of Food Systems and the Environment for Nutrition. *Am. J. Clin. Nutr.* **2021**, *113*, 7.

(3) Willett, W.; Rockström, J.; Loken, B.; Springmann, M.; Lang, T.; Vermeulen, S.; Garnett, T.; Tilman, D.; DeClerck, F.; Wood, A.; et al. Food in the Anthropocene: The EAT-Lancet Commission on Healthy Diets from Sustainable Food Systems. *Lancet* **2019**, *393* (10170), 447−492.

(4) Alae-Carew, C.; Green, R.; Stewart, C.; Cook, B.; Dangour, A. D.; Scheelbeek, P. F. D. The Role of Plant-Based Alternative Foods in Sustainable and Healthy Food Systems: Consumption Trends in the UK. *Sci. Total Environ.* **2022**, *807*, 151041.

(5) Schmidt, C. V.; Mouritsen, O. G. The Solution to Sustainable Eating Is Not a One-Way Street. *Front. Psychol.* **2020**, *11*, 531.

(6) Vasconcelos, M. W.; Grusak, M. A.; Pinto, E.; Gomes, A.; Ferreira, H.; Balázs, B.; Centofanti, T.; Ntatsi, G.; Savvas, D.; Karkanis, A., et al. The Biology of Legumes and Their Agronomic, Economic, and Social Impact. In *The Plant Family Fabaceae*, Hasanuzzaman, M.; Araújo, S.; Gill, S. S., Eds.; Springer: Singapore, 2020; pp. 325.

(7) Foyer, C. H.; Lam, H.-M.; Nguyen, H. T.; Siddique, K. H. M.; Varshney, R. K.; Colmer, T. D.; Cowling, W.; Bramley, H.; Mori, T. A.; Hodgson, J. M.; et al. Neglecting Legumes Has Compromised Human Health and Sustainable Food Production. *Nat. Plants* **2016**, *2* (8), 16112.

(8) Ferreira, H.; Vasconcelos, M. W.; Gil, A. M.; Pinto, E. Benefits of Pulse Consumption on Metabolism and Health: A Systematic Review of Randomized Controlled Trials. *Crit. Rev. Food Sci. Nutr.* **2021**, *61*, 85.

(9) Mullins, A. P.; Arjmandi, B. H. Health Benefits of Plant-Based Nutrition: Focus on Beans in Cardiometabolic Diseases. *Nutrients* **2021**, *13*, 519.

(10) Aiking, H.; De Boer, J. The next Protein Transition. *Trends Food Sci. Technol.* **2020**, *105*, 515.

(11) González-Peña, D.; Brennan, L. Recent Advances in the Application of Metabolomics for Nutrition and Health. *Annu. Rev. Food Sci. Technol.* **2019**, *10* (1), 479.

(12) Wishart, D. S. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol. Rev.* **2019**, *99* (4), 1819−1875.

(13) Madrid-Gambin, F.; Llorach, R.; Vázquez-Fresno, R.; Urpi-Sarda, M.; Almanza-Aguilera, E.; Garcia-Aloy, M.; Estruch, R.; Corella, D.; Andres-Lacueva, C. Urinary 1H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED Study. *J. Proteome Res.* **2017**, *16* (4), 1483.

(14) Baxter, B. A.; Li, K. J.; Zarei, I.; Yao, L.; Rao, S.; Ryan, E. P. Nontargeted and Targeted Metabolomics Identifies Dietary Exposure Biomarkers for Navy Bean and Rice Bran Consumption in Children and Adults. *J. Agric. Food Chem.* **2022**, *70* (45), 14531−14543.

(15) Zarei, I.; Baxter, B. A.; Oppel, R. C.; Borresen, E. C.; Brown, R. J.; Ryan, E. P. Plasma and Urine Metabolite Profiles Impacted by Increased Dietary Navy Bean Intake in Colorectal Cancer Survivors: A Randomized-Controlled Trial. *Cancer Prev. Res.* **2021**, *14* (4), 497−508.

(16) Garcia-Aloy, M.; Ulaszewska, M.; Franceschi, P.; Estruel-Amades, S.; Weinert, C. H.; Tor-Roca, A.; Urpi-Sarda, M.; Mattivi, F.; Andres-Lacueva, C. Discovery of Intake Biomarkers of Lentils, Chickpeas, and White Beans by Untargeted LC − MS Metabolomics in Serum and Urine. *Mol. Nutr. Food Res.* **2020**, *64* (13), 1901137.

(17) Perera, T.; Young, M. R.; Zhang, Z.; Murphy, G.; Colburn, N. H.; Lanza, E.; Hartman, T. J.; Cross, A. J.; Bobe, G. Identification and Monitoring of Metabolite Markers of Dry Bean Consumption in Parallel Human and Mouse Studies. *Mol. Nutr. Food Res.* **2015**, *59*, 795−806.

(18) Harsha, P. S. C. S.; Wahab, R. A.; Cuparencu, C.; Dragsted, L. O.; Brennan, L. A Metabolomics Approach to the Identification of Urinary Biomarkers of Pea Intake. *Nutrients* **2018**, *10* (12), 1911.

(19) Madrid-Gambin, F.; Brunius, C.; Garcia-Aloy, M.; Estruel-Amandes, S.; Landberg, R.; Andres-Lacueva, C. Untargeted 1 H NMR-Based Metabolomics Analysis of Urine and Serum Profiles after Consumption of Lentils, Chickpeas, and Beans: An Extended Meal Study To Discover Dietary Biomarkers of Pulses. *J. Agric. Food Chem.* **2018**, *66* (27), 6997−7005.

(20) Posma, J. M.; Garcia-Perez, I.; Heaton, J. C.; Burdisso, P.; Mathers, J. C.; Draper, J.; Lewis, M.; Lindon, J. C.; Frost, G.; Holmes, E.; Nicholson, J. K. Integrated Analytical and Statistical Two-

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

Dimensional Spectroscopy Strategy for Metabolite Identification: Application to Dietary Biomarkers. *Anal. Chem.* **2017**, *89* (6), 3300−3309.
(21) Levatte, M.; Keshteli, A. H.; Zarei, P.; Wishart, D. S. Applications of Metabolomics to Precision Nutrition. *Lifestyle Genomics* **2022**, *15* (1), 1−9.
(22) Van Duynhoven, J. P. M.; Jacobs, D. M. Assessment of Dietary Exposure and Effect in Humans: The Role of NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 58−72.
(23) Brennan, L. NMR-Based Metabolomics: From Sample Preparation to Applications in Nutrition Research. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 42−49.
(24) Emwas, A.-H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Gowda, G. A. N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; Wishart, D. S. NMR Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9*, 123.
(25) Shibutami, E.; Takebayashi, T. A Scoping Review of the Application of Metabolomics in Nutrition Research: The Literature Survey 2000−2019. *Nutrients* **2021**, *13*, 3760.
(26) Lindqvist, H. M.; Rådjursöga, M.; Torstensson, T.; Jansson, L.; Ellegård, L.; Winkvist, A. Urine Metabolite Profiles and Nutrient Intake Based on 4-Day Weighed Food Diary in Habitual Vegans, Vegetarians, and Omnivores. *J. Nutr.* **2021**, *151* (1), 30.
(27) Garcia-Perez, I.; Posma, J. M.; Gibson, R.; Chambers, E. S.; Hansen, T. H.; Vestergaard, H.; Hansen, T.; Beckmann, M.; Pedersen, O.; Elliott, P.; Stamler, J.; Nicholson, J. K.; Draper, J.; Mathers, J. C.; Holmes, E.; Frost, G. Objective Assessment of Dietary Patterns by Use of Metabolic Phenotyping1424 Col: 374: A Randomised, Controlled, Crossover Trial. *Lancet Diabetes Endocrinol.* **2017**, *5* (3), 184.
(28) O'Sullivan, A.; Gibney, M. J.; Brennan, L. Dietary Intake Patterns Are Reflected in Metabolomic Profiles: Potential Role in Dietary Assessment Studies. *Am. J. Clin. Nutr.* **2011**, *93* (2), 314−321.
(29) Xu, J.; Yang, S.; Cai, S.; Dong, J.; Li, X.; Chen, Z. Identification of Biochemical Changes in Lactovegetarian Urine Using 1 H NMR Spectroscopy and Pattern Recognition. *Anal. Bioanal. Chem.* **2010**, *396* (4), 1451−1463.
(30) Stella, C.; Beckwith-Hall, B.; Cloarec, O.; Holmes, E.; Lindon, J. C.; Powell, J.; Van Der Ouderaa, F.; Bingham, S.; Cross, A. J.; Nicholson, J. K. Susceptibility Of Human Metabolic Phenotypes To Dietary Modulation Research Articles. *J. Proteome Res.* **2006**, *5*, 2780−2788.
(31) Schmidt, J. A.; Fensom, G. K.; Rinaldi, S.; Scalbert, A.; Gunter, M. J.; Holmes, M. V.; Key, T. J.; Travis, R. C. NMR Metabolite Profiles in Male Meat-Eaters, Fish-Eaters, Vegetarians and Vegans, and Comparison with MS Metabolite Profiles. *Metabolites* **2021**, *11*, 121.
(32) Prochazkova, M.; Budinska, E.; Kuzma, M.; Pelantova, H.; Hradecky, J.; Heczkova, M.; Daskova, N.; Bratova, M.; Modos, I.; Videnska, P.; et al. Vegan Diet Is Associated With Favorable Effects on the Metabolic Performance of Intestinal Microbiota: A Cross-Sectional Multi-Omics Study. *Front. Nutr.* **2022**, *8*, 783302.
(33) Buzzard, M. 24-h Dietary Recall and Food Record Methods. In *Nutritional Epidemiology*; Oxford University Press: New York, 1998; pp. 5073.
(34) WHO Part 3: Training & Practical Guides - Section 4: Guide to Physical Measurements. In*The WHO STEPS Surveillance Manual*; WHO, 2008; pp. 3−4−13−4−15.
(35) Stewart, A.; Marfell-Jones, M.; Olds, T.; Ridder, H. *International Standards for Anthropometric Assessment*; Lower Hutt: New Zealand, 2011.
(36) Rodrigues, S.; Franchini, B.; Graça, P.; de Almeida, M. D. V. A New Food Guide for the Portuguese Population: Development and Technical Considerations. *J. Nutr. Educ. Behav.* **2006**, *38* (3), 189.
(37) EFSA*Dietary Reference Values for Nutrients Summary Report*EFSA Supporting Publications201714e15121E
(38) Emwas, A. H.; Roy, R.; McKay, R. T.; Ryan, D.; Brennan, L.; Tenori, L.; Luchinat, C.; Gao, X.; Zeri, A. C.; Gowda, G. A. N.; Raftery, D.; Steinbeck, C.; Salek, R. M.; Wishart, D. S. Recommendations and Standardization of Biomarker Quantification Using NMR-Based Metabolomics with Particular Focus on Urinary Analysis. *J. Proteome Res.* **2016**, *15* (2), 360.
(39) Emwas, A. H.; Luchinat, C.; Turano, P.; Tenori, L.; Roy, R.; Salek, R. M.; Ryan, D.; Merzaban, J. S.; Kaddurah-Daouk, R.; Zeri, A. C.; Nagana Gowda, G. A.; Raftery, D.; Wang, Y.; Brennan, L.; Wishart, D. S. Standardizing the Experimental Conditions for Using Urine in NMR-Based Metabolomic Studies with a Particular Focus on Diagnostic Studies: A Review. *Metabolomics* **2015**, *11* (4), 872.
(40) Takis, P. G.; Schäfer, H.; Spraul, M.; Luchinat, C. Deconvoluting Interrelationships between Provides a Powerful Analysis Tool. *Nat. Commun.* **2017**, *8*, 1662.
(41) Wishart, D. S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B. L.; Berjanskii, M.; Mah, R.; Yamamoto, M.; Jovel, J.; Torres-Calzada, C.; Hiebert-Giesbrecht, M.; Lui, V. W.; Varshavi, D.; Varshavi, D.; Allen, D.; Arndt, D.; Khetarpal, N.; Sivakumaran, A.; Harford, K.; Sanford, S.; Yee, K.; Cao, X.; Budinski, Z.; Liigand, J.; Zhang, L.; Zheng, J.; Mandal, R.; Karu, N.; Dambrova, M.; Schiöth, H. B.; Greiner, R.; Gautam, V. HMDB 5.0: The Human Metabolome Database for 2022. *Nucleic Acids Res.* **2022**, *50* (D1), D622−D631.
(42) Laíns, I.; Duarte, D.; Barros, A. S.; Martins, A. S.; Carneiro, T. J.; Gil, J. Q.; Miller, J. B.; Marques, M.; Mesquita, T. S.; Barreto, P.; et al. Urine Nuclear Magnetic Resonance (NMR) Metabolomics in Age- Related Macular Degeneration. *J. Proteome Res.* **2019**, *18* (3), 1278−1288.
(43) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. The Human Urine Metabolome. *PLoS One* **2013**, *8* (9), No. e73076.
(44) Cloarec, O.; Dumas, M. E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical Total Correlation Spectroscopy: An exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal. Chem.* **2005**, *77* (5), 1282.
(45) Veselkov, K. A.; Lindon, J. C.; Ebbels, T. M. D.; Crockford, D.; Volynkin, V. V.; Holmes, E.; Davies, D. B.; Nicholson, J. K. Recursive Segment-Wise Peak Alignment of Biological 1H NMR Spectra for Improved Metabolic Biomarker Recovery. *Anal. Chem.* **2009**, *81* (1), 56.
(46) Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLSDA Cross Validation. *Metabolomics* **2008**, *4* (1), 81.
(47) Wiklund, S.; Nilsson, D.; Eriksson, L.; Sjöström, M.; Wold, S.; Faber, K. A Randomization Test for PLS Component Selection. *J. Chemom.* **2007**, *21* (10−11), 427.
(48) Kaur, P.; Rizk, N.; Ibrahim, S.; Luo, Y.; Younes, N.; Perry, B.; Dennis, K.; Zirie, M.; Luta, G.; Cheema, A. K. Quantitative Metabolomicand Lipidomic Profiling Reveals Aberrant Amino Acid Metabolism inType 2 Diabetes. *Mol. BioSyst.* **2013**, *9*, 307−317.
(49) Diaz, S. O.; Barros, A. S.; Goodfellow, B. J.; Duarte, I. F.; Galhano, E.; Pita, C.; Almeida, M. D. C.; Carreira, I. M.; Gil, A. M. Second Trimester Maternal Urine for the Diagnosis of Trisomy 21 and Prediction of Poor Pregnancy Outcomes. *J. Proteome Res.* **2013**, *12* (6), 2946−2957.
(50) Berben, L.; Sereika, S. M.; Engberg, S. Effect Size Estimation: Methods and Examples. *Int. J. Nurs. Stud.* **2012**, *49* (8), 1039.
(51) Pinto, J.; Diaz, S. O.; Aguiar, E.; Duarte, D.; Barros, A. S.; Galhano, E.; Pita, C.; Do Céu Almeida, M.; Carreira, I. M.; Spraul, M.; Gil, A. M. Metabolic Profiling of Maternal Urine Can Aid Clinical Management of Gestational Diabetes Mellitus. *Metabolomics* **2016**, *12* (6), 105.
(52) Wei, T.; Simko, V. *R Package "Corrplot": visualization of a Correlation Matrix (Version 0.84)*, 2017.
(53) Schober, P.; Boer, C.; Schwarte, L. A. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Analg.* **2018**, *126* (5), 1763−1768.

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468

(54) Hinkle, D. E.; Wiersma, W.; Jurs, S. G. *Applied Statistics for the Behavioral Sciences*; Houghton Mifflin, 2003.
(55) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D608−D617.
(56) Wallace, M.; Hashim, Y. Z. H.-Y.; Wingfield, M.; Culliton, M.; McAuliffe, F.; Gibney, M. J.; Brennan, L. Effects of Menstrual Cycle Phase on Metabolomic Profiles in Premenopausal Women. *Hum. Reprod.* **2010**, *25* (4), 949−956.
(57) Tosh, S. M.; Yada, S. Dietary Fibres in Pulse Seeds and Fractions: Characterization, Functional Attributes, and Applications. *Food Res. Int.* **2010**, *43* (2), 450−460.
(58) Marinangeli, C.; Curran, J.; Barr, S. I.; Slavin, J.; Puri, S.; Swaminathan, S.; Tapsell, L.; Patterson, C. A. Enhancing Nutrition with Pulses: Defining a Recommended Serving Size for Adults. *Nutr. Rev.* **2017**, *75* (12), 990−1006.
(59) Moreira, R.; Santos, C. S.; Vasconcelos, M. W.; Pinto, E. Qual a Evolução Da Predisposição Dos Portugueses Para Consumir Leguminosas Como Fonte Proteica Principal? *Acta Port. Nutr.* **2022**, No. 28, 26−29.
(60) Duarte, M.; Vasconcelos, M.; Pinto, E. Pulse Consumption among Portuguese Adults: Potential Drivers and Barriers towards a Sustainable Diet. *Nutrients* **2020**, *12* (11), 3336.
(61) Ferreira, H.; Vasconcelos, M.; Gil, A. M.; Oliveira, B.; Varandas, E.; Vilela, E.; Say, K.; Silveira, J.; Pinto, E. Impact of a daily legume-based meal on dietary and nutritional intake in a group of omnivorous adults. *Nutr. Bull.* **2023**, *48* (2), 190−202.
(62) Margier, M.; Georgé, S.; Hafnaoui, N.; Remond, D.; Nowicki, M.; Du Chaffaut, L.; Amiot, M.-J.; Reboul, E. Nutritional Composition and Bioactive Content of Legumes: Characterization of Pulses Frequently Consumed in France and Effect of the Cooking Method. *Nutrients* **2018**, *10*, 1668.
(63) Owen, O. E.; Kalhan, S. C.; Hanson, R. W. The Key Role of Anaplerosis and Cataplerosis for Citric Acid Cycle Function. *J. Biol. Chem.* **2002**, *277* (34), 30409−30412.
(64) Wyss, M.; Kaddurah-Daouk, R. Creatine and Creatinine Metabolism. *Physiol. Rev.* **2000**, *80* (3), 1107−1213.
(65) Hayamizu, K. *Amino Acids and Energy Metabolism: an Overview. An Overview*; Elsevier Inc, 2017. DOI: .
(66) Appiah-Amponsah, E.; Shanaiah, N.; Nagana Gowda, G. A.; Owusu-Sarfo, K.; Ye, T.; Raftery, D. Identification of 4-Deoxythreonic Acid Present in Human Urine Using HPLC and NMR Techniques. *J. Pharm. Biomed. Anal.* **2009**, *50* (5), 878−885.
(67) Nilsen, M. S.; Jersin, R. Å.; Ulvik, A.; Madsen, A.; McCann, A.; Svensson, P.-A.; Svensson, M. K.; Nedrebø, B. G.; Gudbrandsen, O. A.; Tell, G. S.; Kahn, C. R.; Ueland, P. M.; Mellgren, G.; Dankel, S. 3-Hydroxyisobutyrate, A Strong Marker of Insulin Resistance in Type 2 Diabetes and Obesity That Modulates White and Brown Adipocyte Metabolism. *Diabetes* **2020**, *69* (9), 1903−1916.
(68) Landaas, S. Accumulation of 3-Hydroxyisobutyric Acid, 2-Methyl-3-Hydroxybutyric Acid and 3-Hydroxyisovaleric Acid in Ketoacidosis. *Clin. Chim. Acta* **1975**, *64* (2), 143−154.
(69) Lagarda-Diaz, I.; Guzman-Partida, A. M.; Vazquez-Moreno, L. Legume Lectins: Proteins with Diverse Applications. *Int. J. Mol. Sci.* **2017**, *18*, 1242.
(70) Katoch, R.; Tripathi, A. Research Advances and Prospects of Legume Lectins. *J. Biosci.* **2021**, *46* (4), 104.
(71) Krupp, D.; Doberstein, N.; Shi, L.; Remer, T. Hippuric Acid in 24-h Urine Collections Is a Potential Biomarker for Fruit and Vegetable Consumption in Healthy Children and Adolescents. *J. Nutr.* **2012**, *142* (7), 1314−1320.
(72) Vernocchi, P.; Del Chierico, F.; Putignani, L. Gut Microbiota Profiling: Metabolomics Based Approach to Unravel Compounds Affecting Human Health. *Front. Microbiol.* **2016**, *7*, 1144.
(73) Wang, J.; Chen, Y.; Zhong, H.; Chen, F.; Regenstein, J.; Hu, X.; Cai, L.; Feng, F. The Gut Microbiota as a Target to Control Hyperuricemia Pathogenesis: Potential Mechanisms and Therapeutic Strategies. *Crit. Rev. Food Sci. Nutr.* **2022**, *62* (14), 3979−3989.
(74) Il'yasova, D.; Spasojevic, I.; Wang, F.; Tolun, A. A.; Base, K.; Young, S. P.; Marcom, P. K.; Marks, J.; Mixon, G.; DiGiulio, R.; Millington, D. S. Urinary Biomarkers of Oxidative Status in a Clinical Model of Oxidative Assault. *Cancer Epidemiol. Biomarkers Prev.* **2010**, *19* (6), 1506−1510.
(75) Ferreira, H.; Pinto, E.; Gil, A. M.; Vasconcelos, M. W. Chapter 15 - Potential Role of Pulses in the Development of Functional Foods Modulating Inflammation and Oxidative Stress. In *Current Advances for Development of Functional Foods Modulating Inflammation and Oxidative Stress*, Hernández-Ledesma, B.; Martínez-Villaluenga, C. B. T., Eds.; Academic Press, 2022; pp. 287309. DOI: .
(76) Zhu, Y.; Wang, P.; Sha, W.; Sang, S. Urinary Biomarkers of Whole Grain Wheat Intake Identified by Non-Targeted and Targeted Metabolomics Approaches. *Sci. Rep.* **2016**, *6* (1), 36278.
(77) Preumont, A.; Snoussi, K.; Stroobant, V.; Collet, J. F.; Van Schaftingen, E. Molecular Identification of Pseudouridine-Metabolizing Enzymes. *J. Biol. Chem.* **2008**, *283* (37), 25238−25246.
(78) Spenkuch, F.; Motorin, Y.; Helm, M. Pseudouridine: Still Mysterious, but Never a Fake (Uridine)! *RNA Biol.* **2014**, *11* (12), 1540−1554.
(79) Llorach, R.; Favari, C.; Alonso, D.; Garcia-Aloy, M.; Andres-Lacueva, C.; Urpi-Sarda, M. Comparative Metabolite Fingerprinting of Legumes Using LC-MS-Based Untargeted Metabolomics. *Food Res. Int.* **2019**, *126*, 108666.
(80) Ramakrishnan, M.; Rajan, K. S.; Mullasseri, S.; Palakkal, S.; Kalpana, K.; Sharma, A.; Zhou, M.; Vinod, K. K.; Ramasamy, S.; Wei, Q. The Plant Epitranscriptome: Revisiting Pseudouridine and 2′-O-Methyl RNA Modifications. *Plant Biotechnol. J.* **2022**, *20*, 1241−1256.
(81) Streeter, J. G. Carbohydrate, Organic Acid, and Amino Acid Composition of Bacteroids and Cytosol from Soybean Nodules 1. *Plant Physiol.* **1987**, *85* (3), 768−773.
(82) Stumpf, D. K.; Burris, R. H. A Micromethod for the Purification and Quantification of Organic Acids of the Tricarboxylic Acid Cycle in Plant Tissues. *Anal. Biochem.* **1979**, *95* (1), 311−315.
(83) Li, J.; Copeland, L. Role of Malonate in Chickpeas. *Phytochemistry* **2000**, *54* (6), 585−589.
(84) Rudrapal, M.; Khan, J.; Dukhyil, A. A.; Alarousy, R. M.; Attah, E. I.; Sharma, T.; Khairnar, S. J.; Bendale, A. R. Chalcone Scaffolds, \Bioprecursors of Flavonoids: Chemistry, Bioactivities, and Pharmacokinetics. *Molecules* **2021**, *26*, 7177.
(85) Tardy, A.-L.; Pouteau, E.; Marquez, D.; Yilmaz, C.; Scholey, A. Vitamins and Minerals for Energy, Fatigue and Cognition: A Narrative Review of the Biochemical and Clinical Evidence. *Nutrients* **2020**, *12*, 228.
(86) Martínez-Pineda, M.; Yagüe-Ruiz, C.; Caverni-Muñoz, A.; Vercet-Tormo, A. Cooking Legumes: A Way for Their Inclusion in the Renal Patient Diet. *J. Ren. Nutr.* **2019**, *29* (2), 118−125.
(87) Cross, A. J.; Major, J. M.; Sinha, R. Urinary Biomarkers of Meat Consumption. *Cancer Epidemiol. Biomarkers Prev.* **2011**, *20* (6), 1107−1111.
(88) Rushton, D. H. Nutritional Factors and Hair Loss. *Clin. Exp. Dermatol.* **2002**, *27* (5), 396−404.
(89) Bertinato, J.; Lavergne, C.; Vu, N. A.; Plouffe, L. J.; Wood, C.; Griffin, P.; Xiao, C.-W. L-Lysine Supplementation Does Not Affect the Bioavailability of Copper or Iron in Rats. *J. Trace Elem Med. Biol.* **2016**, *38*, 194−200.

https://doi.org/10.1021/acsomega.4c04795
ACS Omega 2024, 9, 43453−43468