Clinical Nutrition ESPEN 63 (2024) 332–345

Contents lists available at ScienceDirect

Clinical Nutrition ESPEN

journal homepage: http://www.clinicalnutritionespen.com

Original article

Impact of a legumes diet on the human gut microbiome articulated with fecal and plasma metabolomes: A pilot study

Helena Ferreira [a, b], Daniela Duarte [b], Tatiana J. Carneiro [b], Célia Costa [a], Joana C. Barbosa [a], João E. Rodrigues [b], Paulo Alves [c], Marta Vasconcelos [a], Elisabete Pinto [a, d], Ana Gomes [a], Ana M. Gil [b, *]

[a] Universidade Católica Portuguesa, CBQF - Centro de Biotecnologia e Química Fina – Laboratório Associado, Escola Superior de Biotecnologia, Porto, Portugal
[b] Department of Chemistry and CICECO-Aveiro Institute of Materials, University of Aveiro, Aveiro, Portugal
[c] Universidade Católica Portuguesa, CIIS - Centro de Investigação Interdisciplinar em Saúde, Escola Enfermagem (Porto), Portugal
[d] EPIUnit - Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal

ARTICLE INFO

Article history:
Received 18 April 2024
Accepted 27 June 2024

Keywords:
Legumes
Vegetarian diet
Metabolomics
NMR
Plasma
Fecal extracts
Microbiome

SUMMARY

Background & aims: Legumes intake is known to be associated with several health benefits the origins of which is still a matter of debate. This paper addresses a pilot small cohort to probe for metabolic aspects of the interplay between legumes intake, human metabolism and gut microbiota.
Methods: Untargeted nuclear magnetic resonance (NMR) metabolomics of blood plasma and fecal extracts was carried out, in tandem with qPCR analysis of feces, to assess the impact of an 8-week pilot legumes diet intervention on the fecal and plasma metabolomes and gut microbiota of 19 subjects.
Results: While the high inter-individual variability hindered the detection of statistically significant changes in the gut microbiome, increased fecal glucose and decreased threonine levels were noted. Correlation analysis between the microbiome and fecal metabolome lead to putative hypotheses regarding the metabolic activities of prevalent bacteria groups (*Clostridium leptum* subgroup, *Roseburia* spp., and *Faecalibacterium prausnitzii*). These included elevated fecal glucose as a preferential energy source, the involvement of valerate/isovalerate and reduced protein degradation in gut microbiota. Plasma metabolomics advanced mannose and betaine as potential markers of legume intake and unveiled a decrease in formate and ketone bodies, the latter suggesting improved energy utilization through legume carbohydrates. Amino acid metabolism was also apparently affected, as suggested by lowered urea, histidine and threonine levels.
Conclusions: Despite the high inter-individual gut microbiome variability characterizing the small cohort addressed, combination of microbiological measurements and untargeted metabolomics unveiled several metabolic effects putatively related to legumes intake. If confirmed in larger cohorts, our findings will support the inclusion of legumes in diets and contribute valuable new insight into the origins of associated health benefits.

1. Introduction

A dietary shift towards more plant-based diets is keystone in a sustainability-driven path [1], in tandem with exploiting the beneficial health impacts of reducing excessive intake of animal-based foods, particularly red meat [2], and increasing legumes intake [3]. Legumes are nutrient-dense dry seeds from the *Leguminosae* family, the use of which has however become shadowed by less nutritious but more profitable crops [4]. Legumes include different bean, pea, chickpea, lentil and lupin species, also known as pulses [5]. These provide relatively high amounts of protein [6], fiber [7] and micronutrients [8], along with a wide range of bioactive compounds (*e.g.*, phenolic acids, tannins, and flavonoids) [9]. Recognized major health benefits of legumes include reduction

* Corresponding author.
E-mail address: agil@ua.pt (A.M. Gil).

https://doi.org/10.1016/j.clnesp.2024.06.051

in cardiovascular disease (CVD) risk factors (*e.g.* obesity, dyslipidemia, hyperglycemia, inflammation, oxidative stress) [10] and beneficial modulation of gut microbiota [2,11,12].

In particular, legumes fermentable carbohydrates, *e.g.,* soluble fiber and resistant starch, are believed to enhance intestinal proliferation of beneficial microbial populations (namely, *Lactobacillus* and *Bifidobacterium*) [13] and stimulate the production of short-chain fatty acids (SCFAs) (*e.g.,* acetate, propionate, butyrate) [14], associated with lower risk of metabolic disorders including CVD [15]. SCFAs are also believed to help maintain the immune intestinal barrier integrity [16], while a portion can enter blood circulation and further influence cell function in peripheral tissues [17]. Researchers have attempted to characterize these mechanisms through human randomized control trials. For instance, analysis of feces of healthy adults given lupin fiber showed an increase in *Bifidobacterium* spp abundance and a decrease in the *Clostridium ramosum, Clostridium cocleatum*, and *Clostridium spiroforme* group abundance, despite no changes in total bacteria abundance (the authors also suggested that gut microbiota responses may be highly individual-specific) [18]. Moreover, fecal samples from healthy and pre-metabolic syndrome subjects showed lower abundance of *Eubacterium limosum,* after consumption of pinto beans [19], whereas a chickpea-enriched diet led to a lower number of individuals testing positive for pathogenic and putrefactive gut bacterial species (*Clostridum lituseburense* groups) [20]. Furthermore, analysis of fecal samples from colorectal cancer (CRC) survivors (to seek possible relationships between navy beans intake and cancer risk) revealed that navy bean powder increased microbial richness in *Lachnobacterium* spp., and in *Clostridium*, *Lachnospira*, and *Coprococcus* [21]. The authors suggested a set of fecal metabolites as markers of glutathione regulation and antioxidant defense, detoxification of xenobiotics, cancer cell proliferation and apoptosis. In addition, metabolic pathways involving lysine and phytochemicals were modulated by navy bean intake [21]. Interestingly, another study directed at overweight and obese adults revealed that daily intake of yellow pea hulls induced no distinction from the placebo group, in terms of the fecal abundance of several microbial groups [22].

Therefore, additional data is required to demonstrate solid links between bacteria composition, SCFAs or other active metabolites and health outcomes [12,23]. Metabolomics of feces and/or biofluids either using mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy [24], in tandem with microbiome characterization, may contribute to such mechanistic insight. Metabolomic studies of plant-based diets and pulses intake have mainly focused on urine [25–32] and blood serum or plasma [33–39]. As a result, several gut bacterial-derived metabolites have been identified [29,31,32,40–42], namely, SCFAs, trimethylamine *N*-oxide (TMAO), benzoate, hippurate, 3-indoxyl sulfate, dimethylsulfone, di/trimethylglycine, and trigonelline. For instance, low meat consumers excrete less taurine, carnitine, acetylcarnitine, 1- and 3-methylhistidine and TMAO [31], a fingerprint which may underly health benefits associated with vegetarian diets, as is the case for TMAO which possesses proatherogenic [43] and procarcinogenic properties [44]. Regarding pulses, habitual consumers exhibited different excretory patterns related to TMAO and dimethylamine (DMA) [29] and multiple blood/urine metabolites have been associated with gut microbiota modulation by navy beans intake, and possible associated cancer prevention roles [45]. Stool metabolomics has been increasingly recognized in the context of food-gut-organ axis research [46,47]. In particular, the effect of combined rice bran and navy beans on human plasma, urine and stool metabolic profiles was investigated to study their impact on CRC risk [47], although no significant differences were observed possibly due to insufficient bran/bean dose or length of intervention. Therefore, further nutritional metabolomic studies are needed for a wider variety of dry legumes (preferably whole, as more accessible than bran extracts or powders [48] and more holistic data collection (*e.g.*, nutritional status, food intake data, gut microbiota dynamics). Such data will aid in the use of legumes as alternative protein sources in sustainable and healthier diets.

This paper addresses the effects of a pilot longitudinal 8-week legume-based food intervention on the gut microbiota composition and fecal/plasma metabolic profiles of 19 omnivorous healthy young adults (predominantly female). A combination of real-time quantitative polymerase chain reaction (qPCR) assessment of gut bacterial composition specifically targeting eight bacterial subpopulations commonly affected by diet [2] and untargeted NMR metabolomics is employed. Despite the small cohort and simple setup considered, the results obtained pave the way for translation to larger cohorts for confirmation and support of the inclusion of legumes in diets.

2. Experimental

2.1. Subjects and study design

Methodological details about the recruitment of subjects recruitment and the study design can be found elsewhere [49]. The subjects' cohort comprised 19 volunteers (18 women and 1 man); in relation to the unbalanced 1/18 M/F ratio, the authors opted for not excluding the male subject as this would impact negatively on an already small cohort (thus subsequently analyzed as predominantly female). This was also supported by the absence of significant differences between the male subject's data and that of the remainder subjects. The median age of the participants was 29.0 years (P25 = 24.0; P75 = 37.0). The median BMI was 21.9 kg/m^2 (P25 = 21.3; P75 = 26.0). Most participants were healthy, though some had specific health conditions such as depressive disorders (n = 1), allergic conditions (n = 3), and thyroid disorders (n = 2). In terms of regular medication, 12 participants used oral contraceptives, 1 used antidepressant drugs and 2 used thyroid medications. Moreover, 4 participants reported using vitamin and/or mineral supplements. Most participants were non-smokers (n = 16; 84.2%), and nine (47.4%) engaged in regular physical activity. The study followed a quasi-experimental design with a one-group comparative approach, involving the replacement of one lunch meal with a vegetarian legume-based meal on five weekdays, for 8 consecutive weeks, and using as control the baseline information on each subject (Fig. 1). Ethical compliance with the Declaration of Helsinki was ensured, supported by approval from the Institute of Bioethics at the Catholic University of Portugal (Ethics Screening Report 11/2017). Participants were required to meet specific eligibility criteria, including daily intake of animal-protein foods, limited legume consumption before undergoing the intervention, and a commitment to follow the intervention diet. Exclusion criteria comprised vegetarianism, severe allergies, chronic illnesses, recent antibiotic use, and pregnancy or breastfeeding. Sociodemographic and lifestyle data were collected at the study's outset, along with dietary information and anthropometric assessment at weeks 0 and 8, following standard [50–52].

2.2. Blood and stool collection

Participants were asked to provide blood samples at weeks 0 (baseline), 4 and 8 (Fig. 1). Blood samples were collected after overnight fasting (12 h) and were obtained from peripheral veins by a skilled nurse, following standard safety procedures. Samples were collected into sodium-heparin tubes, centrifuged within 30 min (1500 rpm, 10 min, 20 °C), and then split into 2 mL aliquots

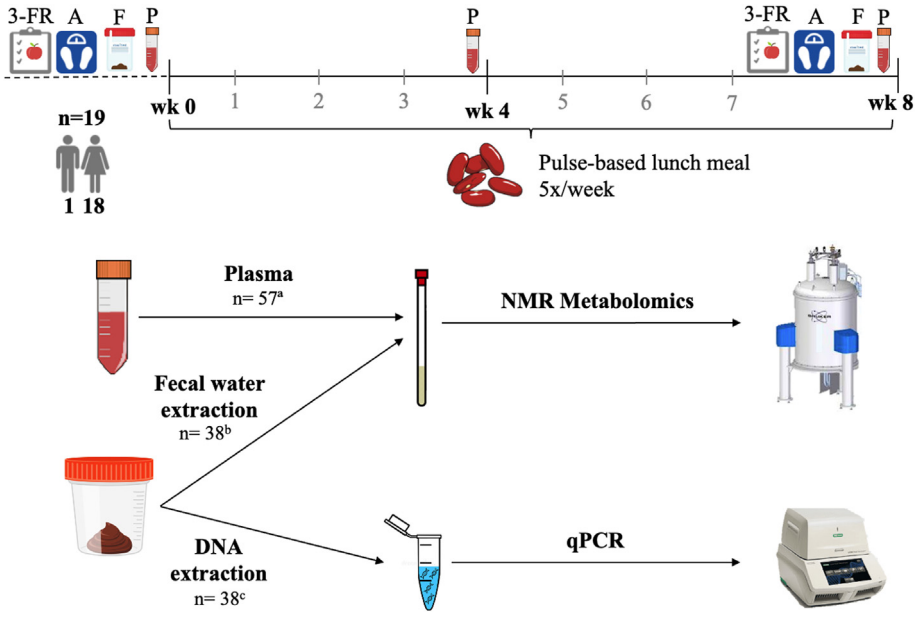

Fig. 1. Schematic representation of the study design. Abbreviations: 3-FR – 3-day food records; A – Anthropometry assessment; F – Feces collection; P – Plasma collection; NMR – Nuclear Magnetic Resonance (NMR) spectroscopy; n – number of subjects or samples; qPCR - Quantitative polymerase chain reaction; wk – week. Some samples were later excluded due, either to technical problems during NMR spectral acquisition or failure of sample delivery by the volunteers, so that final analyzed sample numbers are as follows: [a] 54 and 55 plasma samples were considered for CPMG and diffusion-edited NMR analysis; [b] 37 fecal samples were considered for NMR analysis and [c] 34 fecal samples for qPCR fecal analysis.

[53]. Participants self-collected two stool samples in coded specimen containers provided by the research team, together with collecting instructions. Feces samples were collected at two time points: week 0 (baseline) and week 8 (Fig. 1). After collection, samples were immediately stored at 4 °C by the participants and delivered at Faculty of Biotechnology - Catholic University of Portugal (FB-CUP) in an ice chest, as soon as possible (and no later than 18 h post-collection). Sample collection date and time were recorded upon delivery. Both plasma and fecal aliquots were stored at −80 °C, until analysis.

2.3. Fecal bacterial quantification by qPCR

2.3.1. DNA extraction and quantification

The NZY Tissue gDNA Isolation kit (NZYTech, Lisbon, Portugal) was utilized to perform DNA extraction [54,55]. Briefly, 250 mg of fecal sample was added to 1 mL TE buffer (10 mM Tris/HCl; 0.1 mM EDTA, pH 8.0) and the mixture was resuspended by vigorous vortex homogenization. Then samples were centrifuged (4000 *g*, 15 min) and the supernatant discarded. This process was repeated until obtention of a colorless supernatant. Next, the pellet was resuspended in 1 mL of buffer NT1. From this, 200 μL of the resuspended sample were transferred to a new tube, which was then vortexed and incubated at 95 °C, for 10 min. After the first incubation, 25 μL of Proteinase K were added and the samples were vortexed and incubated again at 56 °C, for 1 h 45 min, vortexing occasionally during this period. After the second incubation time, the samples were vortexed again and 200 μL of buffer NL was added, followed by vortex agitation for 10 s and centrifugation (11,500 *g*, 10 min). The supernatant was transferred to a new tube and the pellet discarded. Then, 210 μL of 100% ethanol were added and samples immediately vortexed. The mixture was transferred into a NZYSpin Tissue Column, placed in a 2 ml collection tube and centrifuged (11,500 *g*, 1 min). The flow-through was discarded and the column was placed back in a new collection tube. Afterward, 500 mL of buffer NW1 were added to the column and centrifuged again (11,500 *g*, 1 min). The flow-through was discarded and the column was placed back in the collection tube. Then, 600 mL of buffer NW2 was added to the column, centrifuged (11,500 *g*, 1 min), the flow-through was discarded and the column was placed back in the collection tube. Another centrifugation was carried out for 2 min at 11,500 *g*. Finally, the column was placed in a clean tube and 100 μL of buffer NE were added directly into the membrane column. The mixture was incubated at room temperature for 1 min followed by centrifugation (11,500 *g*, 2 min) to elute DNA. Once extracted, the concentration and purity of genomic DNA (gDNA) was determined through the analysis of Abs_{260nm}/Abs_{230nm} and Abs_{260nm}/Abs_{280nm} ratios, using a Thermo Scientific™ μDrop™ Plate coupled with a Thermo Scientific™ Multiskan™ FC Microplate Photometer (Thermo Fisher Scientific, Waltham, USA) [56]. The obtained concentration of gDNA was standardized to 20 ng/μL. The stock and diluted gDNA samples were stored at − 20 °C until further analysis.

2.3.2. Standard calibration curves with gDNA

Standard calibration curves with gDNA were obtained according to Applied Biosystems instructions [57], comprising the following steps: i) identification of the genome size of the organism of interest; ii) identification of the number of copies of the gene of interest (16 S rRNA gene) per genome of the organism of interest; iii) calculation of DNA mass per genome and per copy of 16 S rRNA gene; iv) calculation of gDNA mass needed to contain the number of copies of interest; v) calculation of gDNA concentration needed to achieve the number of copies of interest; vi) preparation of gDNA serial dilutions. Genomic DNA standards from *Akkermansia muciniphila*, *Enterococcus gilvus* (*Enterococcus* spp.), *Bifidobacterium longum* subsp. *Infantis* (*Bifidobacterium* spp.), *Clostridium leptum* (*C. leptum* subgroup), *Roseburia hominis* (*Roseburia* spp.), and *Phocaeicola vulgatus* (former *Bacteroides vulgatus* - *Bacteroides* spp.) were obtained from DSMZ (DSMZ, Braunschweig, Germany). *Lactocaseibacillus rhamnosus* Lcr35 strain culture (*Lactobacillus* spp.) obtained from Laboratórios Azevedos – Indústria Farmacêutica, S.A., and *Faecalibacterium prausnitzii* (*F. prausnitzii*) A2-165 (=DSM

17677) culture purchased from DSMZ, were stored in the internal collection of CBQF [Centre for Biotechnology and Fine Chemistry], cultured according to the manufacturer's instructions, and used to extract the corresponding genomic DNA. Information about genome size and number of 16 S rRNA gene copies was gathered from public databases, *e.g.,* GenomeNet database (Kyoto University Bioinformatics Center), the ribosomal RNA database (Center for Microbial Systems, University of Michigan) or the NCBI (National Center for Biotechnology Information) genome database (Table S1). To build the standard calibration curves, the log of the number of 16 S rRNA gene copies of each bacterial strain was plotted against the quantification cycle (Cq).

2.3.3. Real-time quantitative-PCR (or qPCR)

Real-time quantitative-PCR was performed using a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories, Inc., Hercules, USA) on the fecal DNA [58]. The cycling program of qPCR included the following steps: i) initial denaturation/enzyme activation, for 5 min, at 95 °C; ii) denaturation, for 10 s, at 95 °C; iii) annealing, for 1 min, at primer specific temperature; iv) extension, for 30 s at 72 °C; v) melt curve generation, for 5 s, at 60–97 °C with an increment of 0.5 °C. The amplification reaction integrated steps 2, 3, and 4 which were repeated 45 times (45 cycles). The annealing temperature of each primer (STABvida, Lisbon, Portugal) was optimized by testing a range of annealing temperatures above and below the respective calculated T_m. Table S1 lists the primer sequences, and respective optimal annealing temperature, the genome size, and the number of copies of the 16 S rRNA gene, for each used bacterial strain. For each qPCR reaction a final volume of 10 μL was used containing: i) 5 μL of 2x iQTM SYBR® Green Supermix (Bio-Rad Laboratories, Inc., Hercules, USA); ii) 2 μL of ultrapure water; iii) 1 μL of sample diluted DNA (20 ng/μL); and iv) 1 μL of forward and reverse primers (100 nM) targeting the 16 S rRNA gene. A calibration curve was constructed in every run and assays were performed in quadruplicate.

2.4. Sample preparation and NMR analysis

2.4.1. Stool samples

A fecal extract protocol was optimized for these samples based on previous reports [59–64]. The extraction method employed a stool:solvent proportion of 1 mg of sample to 10 μL of solvent. Hence, *ca.* 70 mg of feces, thawed at room temperature, were mixed by vortex (30 s) with 700 μL of phosphate buffer (0.1 M K_2HPO_4/NaH_2PO_4, pH 7.4), containing 10% D_2O, 0.01% (w/v) of NaN_3 (microbial protectant) and 0.5 mM of sodium 3-([2,2,3-d_4]trimethylsilyl) propionate (TSP, for chemical shift referencing). Then, samples were submitted to three rapid freeze (instant)-thaw (10 min, 4 °C) cycles using liquid nitrogen. Next, homogenization was performed for 10 cycles as follows: ultrasonication 20 s >> vortex 10 s >> break 30 s, at room temperature. Slurries were then centrifuged at 16,000 *g*, for 1 h, at 4 °C, and the supernatants collected. The remaining residues were once again subjected to the extraction process, freeze–thaw cycles, homogenization, and a 30 min centrifugation. Supernatants from the first and second extractions were combined and centrifuged again for 10 min before storing at -80 °C until NMR analysis.

Before NMR analysis, fecal extracts were thawed at room temperature and centrifuged at 16,000 *g*, for 10 min, at 4 °C. Next, 540 μL of supernatants were mixed with 70 μL of 99% D_2O and 0.002% TSP. When necessary, pH adjustment to 7.40 ± 0.02 was performed using KOD (4 M) or DCl (4 M). Then 600 μL were transferred into 5 mm NMR tubes. NMR spectra were acquired using a Bruker Avance III spectrometer, operating at 500.13 MHz for proton and equipped with a 5 mm TXI probe. For each sample, a standard one-dimensional (1D) ^{1}H NMR spectrum was acquired with water peak suppression using a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence ("cpmgpr1d"; Bruker library). Acquisition parameters were as follows: 298 K, 256 scans, 32 k complex data points, 10 kHz spectral width, 1.59 s acquisition time, and 4 s relaxation delay. CPMG spectra were acquired with water presaturation, with 80 loops (n) and a total spin–spin relaxation time (2nτ) of 64 ms (with $\tau = 400$ μs). All free induction decays (FID) were multiplied by a 0.3 Hz line-broadening factor prior to Fourier Transformation (FT). Spectra were manually phased, baseline corrected, and referenced internally to and internally referenced to the TSP singlet at δ 0.00 (TopSpin 4.1.3). Peak assignments were performed using 2D NMR experiments, Chenomx NMR suite 9.0 Profiler (Chenomx Inc., Edmonton, Canada), consultation of the Human Metabolome Database [65], the Bruker Biorefcode spectral database (BBIORFCODE-2–0–0 database, Bruker Biospin, Rheinstetten, Germany), as well as already published scientific literature [40,60,62,64,66–68]. To aid peak assignment, Statistical Total Correlation Spectroscopy (STOCSY) [69] was also used in some instances.

2.4.2. Plasma samples

Plasma samples were thawed at room temperature and homogenized in a vortex mixer. Then, 400 μl of saline solution (NaCl 0.9% in 10% D_2O) were added to 200 μl of sample. After centrifugation (4500 g, 5 min, 25 °C), 550 μl of each sample was transferred to 5 mm NMR tubes. NMR spectra were recorded, at 300 K, using a Bruker® Avance DRX 500 spectrometer operating at 500.13 MHz for proton, with a 5 mm TXI probe. For each sample, three one-dimensional (1D) ^{1}H NMR spectra were obtained: a standard spectrum ("noesypr1d"), a CPMG spectrum ("cpmgpr") and a diffusion-edited spectrum ("ledbpgp2s1dpr"), all pulse programs retrieved from the Bruker library. Standard spectra were acquired with a 100 ms mixing time and water suppression during the relaxation delay (4 s) and mixing time. CPMG spectra were acquired with water presaturation, with 80 loops (n) and a total spin–spin relaxation time (2nτ) of 64 ms (with $\tau = 400$ μs). Diffusion-edited spectra were recorded with a diffusion time of 100 ms, a pulsed-field gradient (G1) of 1 ms, a spoil gradient (G2) of 0.6 ms, an eddy current recovery time (τ) of 5 ms, and 90% of the maximum gradient strength (48.15 G/cm). All 1D spectra were acquired into 32 k complex data points with 10330.58 Hz spectral width. Each FID was zero-filled to 64 k points and multiplied by a 0.3 Hz exponential line-broadening function prior to FT. Spectra were manually phased, and baseline corrected, and chemical shifts referenced internally to the α-glucose H1 resonance (5.23 ppm). Peak assignments were performed as described for fecal samples and based on specific literature [53,70–72].

2.5. Statistical analysis

Descriptive statistics, performed using the IBM® SPSS® software version 27 (SPSS Inc., Chicago, IL, USA), was used to describe the sociodemographic data, health-related features, as well as legumes intake of the participants. The Shannon index [73] was used to assess variations in bacterial alpha diversity between baseline and the end of the intervention. Shapiro-Wilk test was used to assess variable normality and paired Student's t-test or Wilcoxon test were applied accordingly to all the above variables. For plasma NMR data analysis, the spectral regions corresponding to water (4.45–5.05 ppm) and ethanol resonances (1.15–1.20 and 3.62–3.68 ppm, contamination at collection) were excluded from the data matrix. For NMR fecal data, only the spectral region corresponding to water (4.64–4.92 ppm) was excluded. A final spectral region between 0.50 and 9.50 ppm was considered for both

spectra (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.) [74]. Normalization to total spectral area was carried out and spectra were scaled to unit variance (UV) (SIMCA-P 11.5, Umetrics, Umea, Sweden). Multivariate analysis (MVA) comprised both principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA). PLS-DA model robustness was assessed by Monte Carlo Cross Validation (MCCV) using 500 runs [75,76]. For each of the 500 randomly generated classification models, Q^2 values (predictive power), number of Latent Variables (LV), and confusion matrices of original and randomly permuted classes were retrieved. PLS-DA models were considered robust for $Q^2 > 0.5$. Sensitivity, specificity and classification rates (CR) were also computed (but found not too hold statistical robustness due to high inter-subject variability). Peak integration was carried out and followed by normalization to total spectral area. Effect size (ES) values were calculated [77], and integral areas presenting simultaneously ES > 0.5 and ES error > ES were selected for statistical comparison. The Shapiro-Wilk test was used to assess variable normality and paired Student's t-test or Wilcoxon test were applied accordingly. The resulting p values were corrected for false discovery rate (FDR). Finally, the (non-parametric) Spearman rank correlation was used to calculate the correlation between NMR data and bacterial quantification through qPCR, using the RStudio® package [78]. Only correlations values > |0.5| and p value < 0.05 were considered, consistently with reported recommendations [79]. Correlation data was displayed using RStudio® package [80]. For all data statistical analysis, a p value lower than 0.05 was considered as the limit for statistical significance.

3. Results

3.1. Food and nutrient intake

Over the 8-week trial the participants performed a median of 38 legume-based meals ($P_{25} = 34$; $P_{75} = 39$), resulting in a median intake of 78.8 g ($P_{25} = 74.45$; $P_{75} = 90.91$) of cooked legumes per meal. The daily macronutrient intake calculated from the 3-day food records collected at baseline and week 8 is summarized in Table S2. Data revealed that, except for the amount of total protein, which was noted to increase, there were no further statistically significant ($p < 0.05$) variations in the daily macronutrient intake between baseline and the final week of the trial.

3.2. Bacterial quantification using qPCR

Fig. S1 represents the variation per participant [17 (16 F/1 M) out of 19] of the 8 bacterial sub-populations studied, between baseline and the end of the 8-week food intervention. The results are expressed as changes in the number of copies of the gene 16 S rRNA/ng of DNA. Results revealed a high interindividual variability after the 8-week legume-based food intervention, with no clear net tendency on the modulation of the analyzed gut bacterial sub-groups. However, compared to other sub-populations, *F. prausnitzii* displayed a tendency to increase in abundance for most individuals (n = 11/17). In turn, *Enterococcus* spp. did not vary throughout the 8 weeks, for several subjects (n = 10/17). Notably, *C. leptum* subgroup, *F. prausnitzii*, *Lactobacillus* spp. and *Roseburia* spp., which belong to the same Phylum (Firmicutes) [81], exhibited a similar overall behavior in five participants in response to the dietary intervention.

Fig. S2 represents the relative abundance of the 8 bacterial sub-populations per participant, comparing baseline against week 8. Although a large subject-dependence was visible, *Roseburia* spp. frequently displayed a higher relative abundance (>50% in n = 9/17), compared to other bacterial sub-populations. However, a higher relative abundance (>50%) of *Bifidobacterium* and *C. leptum* subgroup was found in some cases (n = 2/17). In turn, *Enterococcus* spp. had the lowest overall relative abundance, along with by *A. muciniphila* and *Lactobacillus* spp. In terms of bacterial alpha diversity, as a group, the 8-week legume-based dietary intervention did not significantly affect the Shannon index considering the bacterial sub-populations under study (p = 0.379). However, a more detailed analysis revealed that bacterial diversity could be either increased or decreased by the present dietary challenge. For instance, five individuals (P1, P2, P6, P11, P17) expressed an increase in bacterial diversity, whereas eleven (P3, P4, P5, P7, P8, P10, P12, P13, P14, P15, P16) expressed a decrease, as evidenced by the variation in the Shannon index value at the end of the two months.

3.3. Fecal metabolome analysis

Fig 2a and b shows the average NMR spectra of the samples collected at baseline (week 0) and week 8. A total of 63 metabolites have been identified (Table S3), confirming previous reports of feces NMR assignment [40,60,62,64,66–68]. Visual comparison of the spectra in Fig. 2 suggests apparent increases in succinate (peak 15, δ 2.41), histidine (peak(s) 23, δ 3.16, δ 3.23), glucose (peak(s) 28, δ 5.23) and hypoxanthine (peak(s) 37, δ 8.21, δ 8.19) (Fig. 2b, red arrows). In addition, apparent decreases are noted in malonate (peak 22, δ 3.11) (Fig. 2b, blue arrows). However, these apparent visual changes may not be representative and require validation through multivariate and univariate statistical analysis. Fig. S3 shows that both unsupervised (PCA) and supervised (PLS-DA) analysis do not reveal any group separation, thus indicating that no consistent strong compositional differences are detected in the metabolic profile of feces, upon the legumes diet intervention on the small cohort considered. However, peak integration and univariate analysis reveals small variations in glucose (increase, p value 0.005) (Fig. 2c, left) and in threonine (decrease, p value 0.025) (Fig. 2c, right), none of which remained significant upon FDR correction. The former had indeed been suggested through visual inspection of the spectra, while other visual changes in the average spectra (Fig. 2a) were revealed to be statistically non-significant. Notably, no significant changes in SCFAs were observed, particularly in those detected by NMR, namely acetate, formate, *n*-butyrate, propionate, succinate and valerate.

3.4. Correlation between NMR results and qPCR

The results of correlation analysis between the NMR metabolic profile of feces and qPCR data (relative abundance of the selected 8 bacterial families) are summarized in Fig. 3a and b, with thresholds r>|0.5| and r>|0.6|, respectively. For this correlation analysis, data was available for 16 (15 F/1 M) of the 19 participants and, overall, most correlations fell within the r = |0.50–0.70| interval, with no correlations found for r>|0.75|. Results revealed correlations with the 3 most abundant bacterial sub-populations from the Firmicutes Phylum [81], namely *C. leptum* subgroup, *Roseburia* spp. And (in higher abundance) *F. prausnitzii*. The correlations found involved metabolites related to organic acid metabolism, amino acid metabolism and glycolysis. Only *F. prausnitzii* showed a significant positive correlation with glucose, along with *n*-butyrate and nicotinate (Fig. 3a). Additional positive correlations (r > 0.5–0.6) were found for acetate with both *F. prausnitzii* and the *C. leptum* subgroup. In addition, negative correlations (Fig. 3a and b) included inverse variations between the *C. leptum* subgroup and valerate, isovalerate, alanine, valine and *p*-cresol, whereas *F. prausnitzii* varied inversely with acetoacetate, malate, valerate, isovalerate, 2-hydroxyvalerate, 2-hydroxy-3-methylvalerate, 3-methyl-2-oxovalerate, 5-aminovalerate, methionine, p-

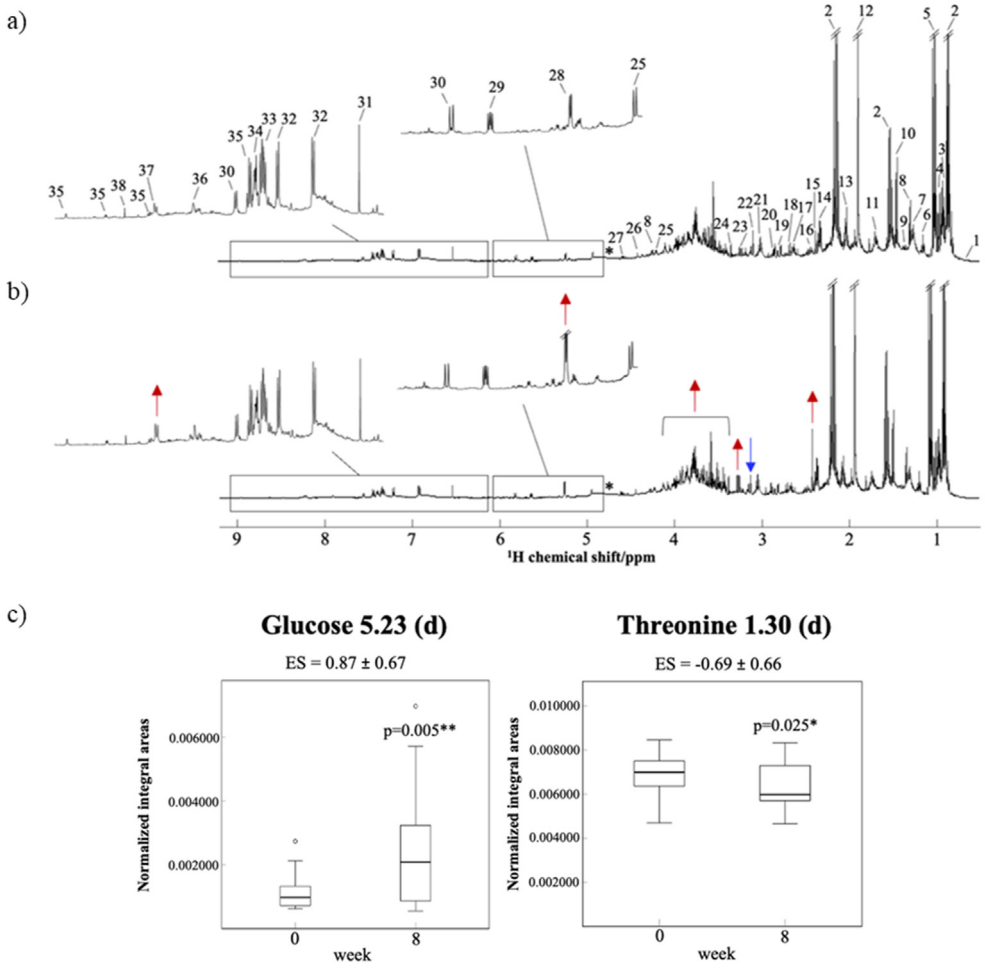

Fig. 2. a) and b) Average ^{1}H NMR spectra of fecal samples from (a) baseline (n = 19) and (b) week 8 (n = 18) and (c) Boxplots of spectral regions with statistically significant variations during the 8-week legume trial (none of the variations remained significant upon FDR correction). Arrows indicate visible spectral alterations: blue arrows indicate a visible decrease and red arrows indicate a visible increase. Abbreviations: 1 – bile acids; 2 – n-butyrate; 3 – isoleucine; 4 – leucine; 5 – propionate; 6 – ethanol (contamination); 7 – valerate; 8 – threonine; 9 – 2-hydroxy-valerate; 10 – alanine; 11 – leucine; 12 – acetate; 13 – proline; 14 – valine; 15 – succinate; 16 – glutamine; 17 – methionine; 18 – malate; 19 – aspartate; 20 – trimethylamine (TMA); 21 – lysine; 22 – malonate; 23 – histidine; 24 – methanol; 25 – ribose; 26 – dihydroxyacetone; 27– xylose; 28 – glucose; 29 – uracil-diphosphate glucose (UDP-glucose); 30 – uracil; 31 – fumarate; 32 – tyrosine; 33 – phenylalanine; 34 – 3-phenyl-propionate; 35 – nicotinate; 36 – xanthine; 37 – hypoxanthine; 38 – formate. *Excluded regions – water (4.64–4.92 ppm). d – doublet; ES – Effect size; *$p < 0.05$; ** $p < 0.01$.

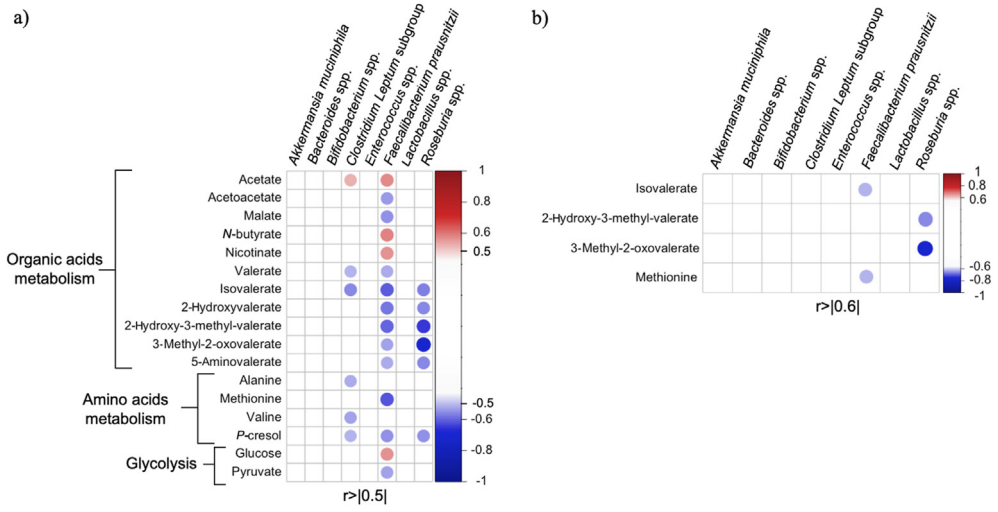

Fig. 3. Spearman rank correlation maps showing significant ($p < 0.05$) correlations between NMR spectra of fecal samples and qPCR results (n = 16): a) correlation values r > |0.5|; b) correlation values r > |0.6|. Blue and red dots: negative and positive correlations, respectively.

cresol and pyruvate. Finally, *Roseburia* spp. seems to correlate negatively with isovalerate and the same 4 valerate derivatives (2-hydroxyvalerate, 2-hydroxy-3-methyl-valerate, 3-methyl-2-oxovalerate, 5-aminovalerate), in addition to *p*-cresol. Of the above correlations, the strongest correlations surviving a threshold of r>|0.6| (Fig. 3b) refer to *F. prausnitzii* inverse variations with isovalerate and methionine and *Roseburia* spp. Inverse correlations with 2-hydroxy-3-methyl-valerate and 3-methyl-2-oxovalerate ($r = 0.704$ for the latter metabolite). Hence, it is interesting to note that, although no changes were observed in fecal SCFAs throughout the intervention, correlation studies unveil that a set of SCFAs are indeed involved in specific pathways, related at least with the three most abundant bacteria. In addition, correlations were sought between bacterial abundance or fecal metabolites, and food and nutrient intake, namely, legumes, fiber, and protein (data not shown), but no meaningful correlations were found. Similar results were obtained for correlations between fecal and plasma metabolic profiles.

3.5. Plasma metabolome analysis

For plasma analysis, both CPMG and diffusion-edited ^{1}H NMR spectra were recorded for the 19-subject cohort (Fig. 4 and S4, respectively). This enabled the assessment of eventual changes both in the pool of low M_w compounds and in that of macromolecules (mainly lipoproteins and glycoproteins, the latter viewed through the *N*-acetyl resonance at 2.02–2.07 ppm), respectively.

Fig. 4 illustrates the high resolution and complexity of plasma CPMG spectra and, specifically, of the average ^{1}H NMR spectrum of plasma samples collected at baseline (Fig. 4a), in which a total of 28 metabolites were identified, with basis on previous reports of plasma NMR assignments [53,70–72]. The average ^{1}H NMR CPMG spectra of plasma samples collected at weeks 4 and 8 are represented in Fig. 4b and c. Visual comparison suggests apparent decreases maintained in both time points (week 4 and 8) in acetate (peak 8, δ 1.90), methanol (peak 19, δ 3.4), glucose (peak 20, δ 3.24), urea (peak 24, δ 5.75), and formate (peak 28, δ 8.45) (Fig. 4b and c, blue arrows). Also, apparent increases are noted in a broader lipid region corresponding to fatty acids (mainly expected to arise from lipoproteins) (peak 6, δ 1.55) (Fig. 4b and c, red arrow).

Again, multivariate and univariate statistics was applied to confirm/discard the above visual observations and Fig. 5 shows that some PCA separation is observed when the CPMG spectra for weeks 0 and 8 are considered (Fig. 5c, left), with the corresponding PLS-DA model exhibiting a high predictive power ($Q^2 = 0.731$), compared to comparison of weeks 0 and 4 ($Q^2 = 0.481$) and weeks 4 and 8 ($Q^2 = 0.366$) (Fig. 5c, right). This suggests that a gradual change occurs in the metabolic profile of plasma, with maximum differences noted between the first and final weeks of the intervention. In relation to diffusion-edited NMR spectra of plasma, no visual differences were observed (Fig. S4). Indeed, no group separation was observed in PCA (Fig. S5) and, despite the apparently satisfactory PLS-DA predictive power ($Q^2 = 0.565$) between weeks 0 and

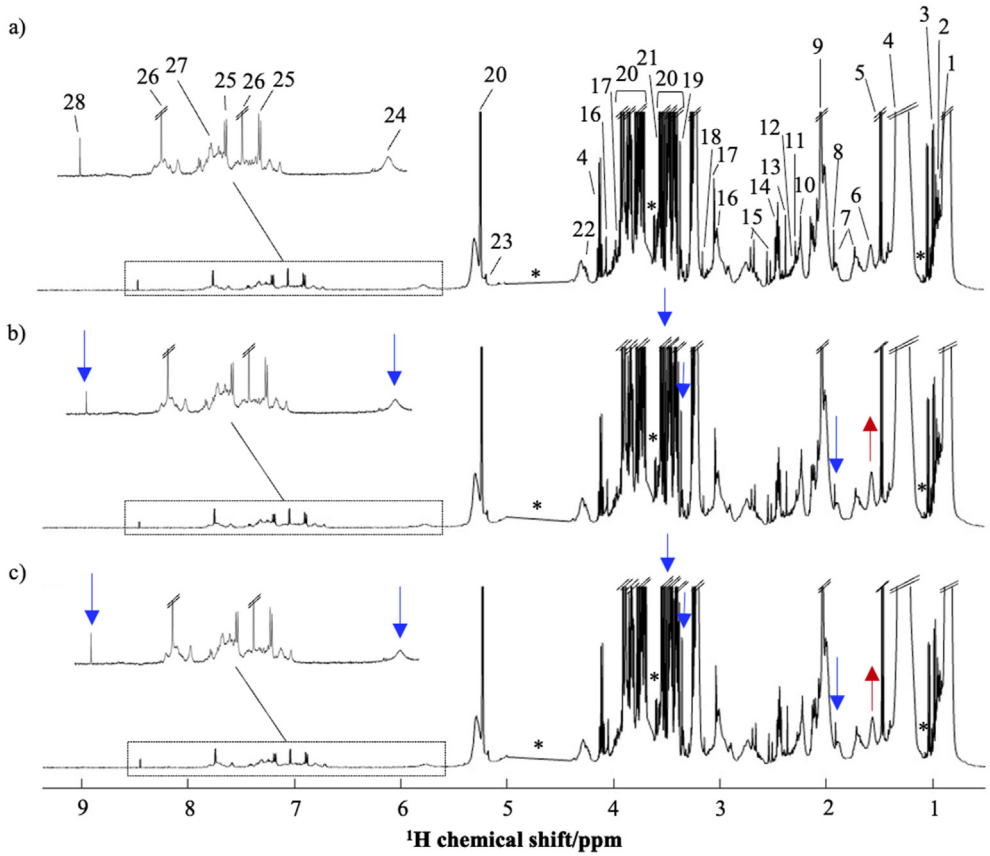

Fig. 4. Average ^{1}H NMR CPMG spectra of plasma samples from a) baseline (week 0), b) week 4 and c) week 8. Arrows indicate visible spectral alterations: blue and red arrows indicate apparent visual decreases and increases, respectively (even for the stronger peaks cut-off due to vertical expansion of the spectra). Abbreviations: 1 – leucine; 2 – valine; 3 – isoleucine; 4 – lactate; 5 – alanine; 6 – lipids (fatty acids -C<u>H</u>$_2$CH$_2$CO- methylene groups); 7– lysine; 8 – acetate; 9 – *N*-acetylated glycoproteins (NAG); 10 – acetone; 11 – acetoacetate; 12 – β–hydroxybutyrate (or β/3-HBA); 13 – pyruvate; 14 – glutamine; 15 – citrate; 16 – creatinine; 17 - creatine; 18 – dimethylsulfone; 19 – methanol; 20 – glucose; 21 – glycine; 22 – threonine; 23 – mannose; 24 – urea; 25 – tyrosine; 26 – histidine; 27 – phenylalanine; 28 – formate; *Excluded regions: water resonance, at 4.4–5.0 ppm; ethanol resonances, at 1.15–1.20 and 3.62–3.68 ppm.

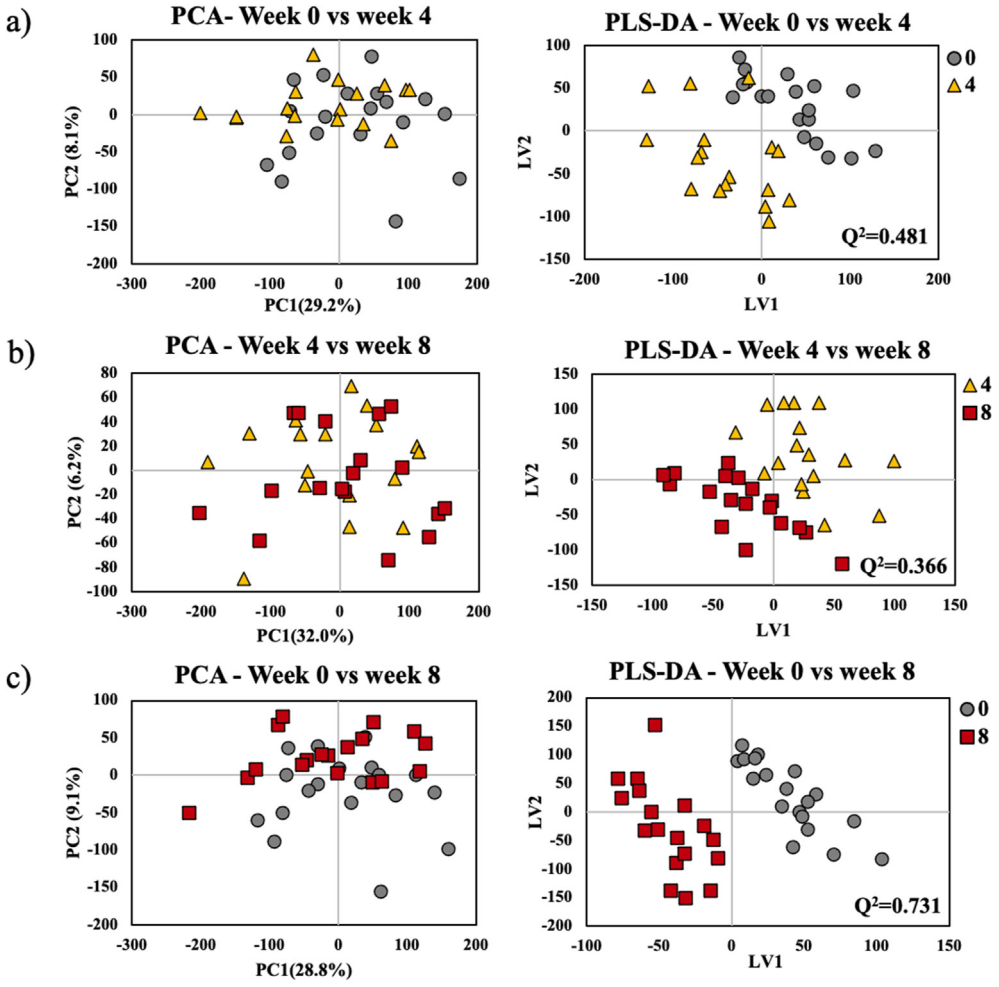

Fig. 5. PCA (left) and PLS-DA (right) scores scatter plots obtained with the ^{1}H NMR spectra (CPMG) of plasma for a) baseline (week 0) *versus* week 4, b) week 4 *versus* week 8, and c) baseline (week 0) *versus* week 8. All models were obtained with 2 principal components (PCA) and 2 latent variables (PLS-DA). Abbreviations and symbols: Q^2 - predictive power of PLS-DA models; PCi - principal components; LVi - latent variables.

8, no significant changes were retrieved, as viewed by diffusion-edited NMR spectra (i.e. regarding macromolecules global profile).

Peak integration in the CPMG spectra (Fig. 4) revealed changes that are represented in boxplot form in Fig. 6, along with changes in a set of still unassigned resonances (Fig. S6, mainly involving singlet resonances, of particularly difficult identification). In particular, plasma changes comprised decreases in the two ketone bodies acetoacetate and 3-hydroxybutyrate (3-HBA, although only qualitatively), along with decreases in citrate, formate (from weeks 0–4), histidine, threonine and urea. Conversely, increases in betaine, mannose, and a slight increase in formate from weeks 4–8 were observed.

Considering the above changes distributed by intervention week (Table 1), it becomes clear that citrate (dec.), mannose (inc.), histidine (dec.) and urea (dec.) are earlier significant changes in the circulatory metabolome (although none remain significant upon FDR correction), in tandem with a more marked decrease in formate. This is followed by faint changes between weeks 4 and 8, namely in the levels of betaine (inc.), formate (inc.) threonine (dec.). Notably, betaine and threonine changes are confirmed with higher statistical relevance by comparison of weeks 0 and 8. This comparison also confirms the initial (week 4 *vs* week 0) changes in mannose (inc.) and urea (dec.), revealing an additional decrease in acetoacetate (which exhibits a similar behavior to that of a second ketone body, 3-HBA) (Fig. 6).

4. Discussion

The present dietary intervention did not exert a significant effect over the overall diversity score of the bacterial sub-populations under study (Fig. S3). Still, individual analysis showed that the majority (n = 11/17) of participants expressed a reduction in alpha diversity at the end of the 8-week intervention. Indeed, a previous study revealed that navy bean powder intake increased microbial richness compared to baseline, but caused no effect on diversity either [21]. Such results somehow contradict other evidence that suggests that diets rich in legumes fiber are associated with greater diversity of fecal microbiota [82]. Noteworthy, food intake data (Table S2) showed that daily total fiber intake was not significantly different during the 2-months legume-based dietary intervention, which may help justify the absence of effects on gut bacterial diversity. The microbiological assessment carried out here evaluated the eight bacterial sub-populations which have been reported to be more commonly affected by diet and to have a strong impact in relation to disease states [2], although it is recognized that the whole gut microbiome was not described, nor the interactions between the different bacterial sub-populations were assessed.

The high subject-dependent response to intervention, in terms of gut bacterial quantification, is consistent with previous reports [18,83,84]. The human large intestine is a very complex ecosystem that is still not fully understood [84], and the inter-individual

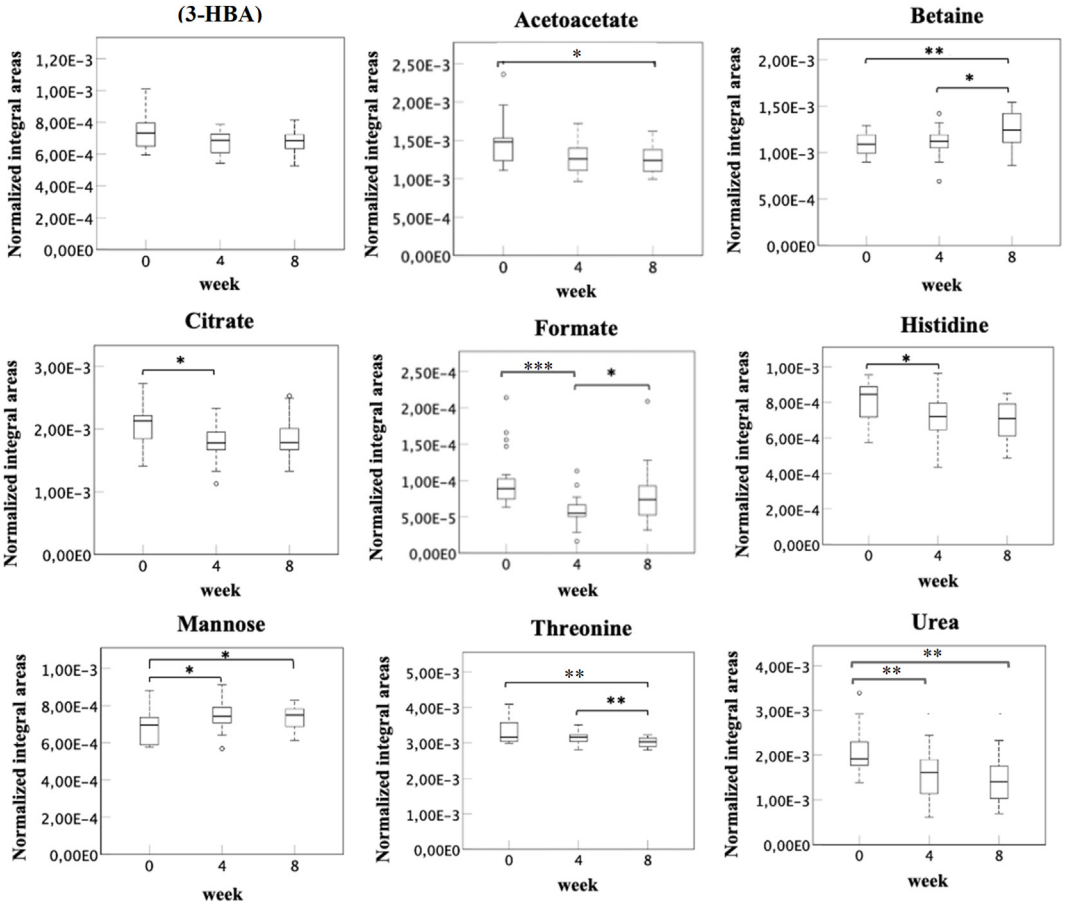

Fig. 6. Boxplots of statistically significant changes in identified plasma metabolites, observed throughout the intervention. Abbreviations: 3-HBA − 3/β-hydroxybutyrate; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 1
Univariate statistical analysis of plasma metabolite variations separated by weeks (statistical significance for p value < 0.05).

Metabolites	ppm (multiplicity)	ES ± ES_{error}	p $value_{uncorrect.}$	p $value_{FDR\text{-}correct.}$
Week 4 *vs* week 0				
Citrate	2.52 (d)	−0.891 ± 0.686	0.022	>0.05
Formate	8.45 (s)	−1.218 ± 0.699	0.00004	0.0006
Mannose	5.18 (d)	0.705 ± 0.674	0.049	>0.05
Histidine	7.04 (s)	−0.781 ± 0.679	0.035	>0.05
Urea	5.75 (br)	−1.019 ± 0.682	0.009	>0.05
U1	4.14 (s)	−0.924 ± 0.688	0.025	>0.05
U2	7.40 (d)	−0.734 ± 0.676	0.038	>0.05
Week 8 *vs* week 4				
Formate	8.45 (s)	0.715 ± 0.684	0.049	>0.05
Betaine	3.27 (s)	0.830 ± 0.691	0.037	>0.05
Threonine	4.23 (dd)	−0.823 ± 0.690	0.033	>0.05
Week 8 *vs* week 0				
Acetoacetate	2.27 (s)	−0.884 ± 0.675	0.013	0.041
Betaine	3.27 (s)	1.119 ± 0.693	0.006	0.033
Mannose	5.18 (d)	0.687 ± 0.663	0.045	>0.05
Threonine	4.23 (dd)	−1.049 ± 0.688	0.010	0.035
Urea	5.75 (br)	−1.348 ± 0.714	0.0008	0.014
U4	1.07 (s)	0.702 ± 0.664	0.014	0.041
U5	1.05 (s)	1.081 ± 0.690	0.003	0.029
U6	3.56 (s)	−0.787 ± 0.669	0.023	0.048
U3	4.28 (br)	−0.914 ± 0.678	0.012	0.041
U2	7.40 (d)	−0.859 ± 0.674	0.018	0.043

p-$value_{uncorrect.}$: p value without false discovery rate (FDR) correction, p-$value_{FDR\text{-}correct.}$: p value obtained upon FDR correction. Unassigned resonances are identified with "Ui", according to chronological order of detection (i.e., from week 0 to week 8). s: singlet; d: doublet; t: triplet; dd: doublet of doublets; m: multiplet; br: broad signal.

variability of gut microbiota composition has been found to be quite high [83]. This agrees with previous studies that show that bacterial communities can display qualitative and quantitative variations in response to different factors related to host (*e.g.*, pH, bile acids, transit time and mucus), environment (*e.g.*, diet and use of antibiotics) and microbial characteristics (*e.g.*, adhesion capability and bacterial enzymes) [84].

Additionally, we had hypothesized that the eight bacterial subpopulations tested were part of the phyla most affected by diet [85] and in our study, four genera from the *Firmicutes* phylum, namely, *C. leptum* subgroup, *F. prausnitzii*, *Lactobacillus* spp. and *Roseburia* spp., exhibited similar responses in different participants. Previous publications had already reported the modulation of these bacterial genera by diets containing legumes [18–22]. Narrowing down to species level, within the group of the eight bacterial subpopulations studied, a high relative predominance of *Roseburia* spp. in several (n = 9) individuals (Fig. S3) was observed. In fact, literature reports that *Roseburia* spp. are among the group of dominant butyrate-producing *Firmicutes*, which together with *F. prausnitzii* and *Eubacterium rectale*, make up 7–24% of the total bacteria in the healthy human colon [86].

The above observations are, however, consistent with the weak changes noted in the metabolic profile of feces, with only a slightly enhanced excretion of glucose, together with a decrease in threonine. Higher fecal glucose content may arise from the complex carbohydrates found in legumes, namely, fibers, resistant starch and oligosaccharides (mainly in the form of α-galactosides or raffinose family oligosaccharides) [87], which although not hydrolyzed by gastric or intestinal enzymes, may be fermented by gut microbiota [11,88]. Gut bacteria can breakdown these compounds and use their subunits, namely glucose, as energy sources [89,90]. Indeed, the prebiotic potential of legumes' fermentable carbohydrates has been described previously [91]. We, therefore, propose that glucose is here present in amounts largely sufficient to meet the energy requirements of gut bacteria. In relation to the small decrease in fecal threonine levels, it is known that this amino acid can be degraded by the gut microbiota to generate SCFAs, especially propionate [92]. Thus, the lower threonine content in feces could reflect overall gut microbiota activity, although it is not possible to relate this metabolite to the specific bacterial groups tested in the present work. Regarding the absence of statistically significant variations in fecal SCFAs (namely those detected here by NMR: acetate, formate, *n*-butyrate, propionate, succinate, and valerate), this may be explained by different reasons. Firstly, literature reports that fecal SCFA concentrations may not reflect their concentration and production rate in the intestine as the host takes up most SCFAs [93]. *In vivo* measurements of fecal SCFA concentrations reflect just 5% of the total production since more than 95% are rapidly absorbed by the colonocytes [94]. So, fecal samples alone may not be enough to realistically reflect production and potential absorption of SCFAs [94]. For this reason, because SCFA production is difficult to measure *in vivo*, most experiments have been done through *in vitro* experiments with intestinal or fecal microbiota as inoculum [93]. For instance, a study found that propionate production from fecal material fermented *in vitro* with bean flour was higher in volunteers consuming beans (compared with those who consumed soup) [19]. Still, *in vitro* fermentation differs from the *in vivo* situation because, during the bacteria isolation steps, the microbiota diversity appears to alter considerably, and later products accumulate during fermentation [93]. Secondly, it is possible that the duration of the food challenge and the legumes dose used were not sufficiently high to cause measurable changes in fecal SCFAs. Indeed, the total daily fiber intake did not change over the eight weeks, even though the legume-based meals provided greater amounts of fiber (data not shown) [49] and, as fiber is the main substrate of SCFA-producing bacteria, significant changes in these compounds are not expected. Furthermore, a recent study that analyzed fecal microbiota activity markers in vegans and omnivorous found no significant differences in stool SCFA contents between the two groups, thus suggesting that gut microbiota is highly adaptable to different nutrient availability, so that SCFA levels are maintained to guarantee energy supply of colonocytes and overall intestinal homeostasis [95].

Correlations between microbiota and fecal metabolic profile proved useful to pinpoint possible specific metabolite/bacteria relationships, in spite of the observed little or no changes in both domains upon the intervention. However, interpretation of correlation studies should be carried out with care, allowing for the possibility of other bacterial groups (other than the 8 groups studied here) contributing towards the modulation of the same metabolites, especially given the known highly dynamic and cross-feeding effects characteristic of gut bacterial communities [96]. Results revealed correlations with the 3 most abundant bacterial sub-populations from the Firmicutes Phylum [81], namely *C. leptum* subgroup, *Roseburia* spp. and *F. prausnitzii*, these saccharolytic species thus expected to contribute most to the changes measured by NMR.

Regarding the *C. leptum* subgroup, the positive correlation with acetate is consistent with the expected fermentation of dietary fibers by this bacterial group, resulting in the production of SCFAs [97], including acetate and butyrate (although no correlation is noted here with the latter). The concomitant negative correlations with valerate and isovalerate, along with alanine, valine and *p*-cresol may reflect lower extent of protein degradation processes associated with *C. leptum* [97,98]. Indeed, branched-chain fatty acids (BCFAs) (*e.g.,* isovalerate and derived compounds) result from protein degradation namely, involving branched-chain amino acids (BCAA): valine (negatively correlated here), leucine, and isoleucine [99]. Also, phenols such as *p*-cresol has been assigned as a product of tyrosine degradation [98]. Other studies have found that Taxa, which correlated positively with butyrate or acetate, were in general negatively correlated with BCFAs concentrations in plasma samples [100]. This may be consistent with the negative correlations seen here with fecal samples (despite the absence of actual BCFAs changes in feces or plasma). Many of the above considerations also apply to *F. prausnitzii,* which correlates positively with both acetate and butyrate (and nicotinate), and negatively with valerate, isovalerate and four related compounds, as well as with methionine, *p*-cresol and pyruvate (for unclear reasons, at this stage). *F prausnitzi,* is a recognized acetate-consuming and butyrate-producing bacteria [101] and, given this information alone, a negative correlation for acetate might be expected. The fact that *F. prausnitzii* abundance varies directly with acetate (positive correlation) suggests the importance of other variables in determining acetate levels. These may include the possibility of cross-feeding between the *C. leptum* (acetate-producing) and *F. prausnitzii* (acetate-consuming) *via* fiber fermentation [96,102]. Furthermore, *in silico* models have revealed that acetate consumption enables *F. prausnitzii* to harvest more energy from carbon sources, namely glucose [101], which may underlie the positive correlation with fecal glucose. However, the exact origin(s) of the correlation with glucose is unclear, as this metabolite may involve the activity of other bacteria, the metabolism of the host itself and/or result directly from legumes intake. Indeed, *F. prausnitzii* is known to be able to switch between substrates derived from the diet or the host [102]. As with the *C. leptum* subgroup, *F. prausnitzii* seems to be involved in lesser protein degradation, again as expected according to literature [98,99], but now highlighted by a larger number of negatively-correlated isovalerate/valerate compounds, in tandem with methionine (apparently specific of

F. prausnitzii), and, again, *p*-cresol. Interestingly, the set of (iso) valerates negative correlations observed for *F. prausnitzii* is almost identical to those found for *Roseburia* spp., the latter only revealing correlations related with protein degradation (i.e. without evidence of acetate, butyrate or glucose involvement, among other metabolites).

The plasma metabolome seems to primarily be a source of legumes intake markers and of information on host energy metabolism, although a slight relationship with fecal metabolome may also be advanced, as will be shown below. The increase in mannose may arise from the intake of mannose-containing cell wall polysaccharides in legumes fiber, in particular beans [103]. This is in agreement with a similar observation based on the plasma of subjects upon ingesting navy beans [46]. Increased betaine may also arise from ingested legumes, through hepatic degradation of choline found in legumes [29]. Indeed, urinary studies of consumers of chickpea, lentils and/or beans have suggested effects on choline metabolism [29], while betaine has been correlated positively with plant-based diets [39], although a previous MS metabolomics report detected a decrease in plasma betaine levels upon navy beans [45]. It is possible that circulatory betaine levels are dependent on the specific type of legumes incorporated in the diet. We also hypothesize that choline metabolism may not only explain betaine variations but also the increase in formate, which may arise from dimethylglycine formation also from choline [29,104]. However, given that *F. prausnitzii* displayed a tendency to increase in abundance in most individuals ($n = 11/17$), it is also possible that this bacteria will form formate (and butyrate), as part of SCFAs arising from acetate consumption [105,106], although this would most likely give rise to increased fecal formate (not observed), its connection to plasma remaining unclear. Most importantly, evidence is found in plasma of a disturbance of energetic metabolism, namely expressed by a small decrease in citrate (tricarboxylic acid (TCA) cycle intermediate) and of ketone bodies acetoacetate and 3-HBA (the latter showing no statistical relevance but following a similar trend to that of acetoacetate). These are usually used for enhanced energy production in extra-hepatic tissues, when additional energy sources are required by the organism. Indeed, legume consumption is known to affect energy metabolism [29] and a parallel urine metabolomics account of the same subjects studied here (data not published) has indeed suggested increased excretion of acetoacetate and 3-HBA. To our knowledge, no similar observations have been reported before and we suggest that decreased circulatory ketone bodies, in tandem with their enhanced excretion, may reflect that legumes help the organism to meet the necessary energy requirements, without extra contribution of ketone bodies. This is consistent with a previous suggestion of less active gluconeogenesis in legume consumers (precisely also based on urinary reduction of acetoacetate, isobutyrate and some TCA cycle intermediates, including citrate) [29]. A lesser extent of gluconeogenesis is indicative of sufficiently high circulatory glucose levels and, indeed, preliminary clinical and anthropometric data of the same cohort studied here (data not published) has suggested an increase in glucose metabolism-related biomarkers, in tandem with a beneficial cholesterol-lowering effect. Finally, we note that the decreases in plasma levels of threonine (also decreased in feces) and urea suggest some impact of the diet on amino acid metabolism, although this issue requires further investigation.

5. Conclusion

This work describes the effect of a pilot 8-week legumes diet intervention on human gut microbiota and fecal/plasma metabolomes of a small cohort of 19 subjects (predominantly female), to gather information, at the molecular level, on the relationship between legumes intake and gut microbiota response, potentially reflected in related health benefits. To our knowledge, this is the first study that integrates microbiome (qPCR data) and fecal and plasma metabolomics (using untargeted NMR) towards this particular aim. Results showed that no statistically significant changes affected the gut microbiome (due to high inter-individual variability within the small $n = 19$ cohort under study), whereas increased glucose and decreased threonine were observed in feces. However, statistical correlation between the two domains enabled putative hypotheses to be put forward regarding the metabolism of the three most abundant bacteria groups, namely *C. leptum* subgroup, *Roseburia* spp. and *F. prausnitzii,* namely considering i) glucose high levels in feces, at least in part originating from fiber breakdown and well in sufficiency of bacterial requirements, ii) the particular sets of SCFAs produced by fermentation by each group (even though their levels did not change in feces or plasma), and iii) the apparent lesser protein degradation characterizing the three groups, throughout the intervention. Although additional contributions from unmeasured bacteria may also determine metabolite levels, the observed metabolite/bacteria correlations provide valuable insight into gut microbiota response to the legumes diet. Furthermore, plasma metabolomics of the same subjects unveiled mannose and betaine as possible legumes intake markers, as well as evidencing lower levels of formate and ketone bodies throughout the length of the intervention. Taking baseline as each individual's control, the latter observation was consistent with a slowing down of the organism mechanisms to increase circulatory glucose (namely, ketogenesis), thus reflecting that the legumes carbohydrates seem to help the organism meet its energy requirements, with less biochemical effort. An impact on amino acid metabolism was also evident through decreasing urea and some amino acids levels (mainly threonine), although no clear explanation could be advanced at this stage.

We emphasize that this work is not without limitations, namely related to the small cohort size and the lack of a control diet/group, due to practical constraints and difficulty in recruiting healthy volunteers in higher numbers. Also, the cohort considered here was mainly female (one male in a total of 19 subjects) which may bias the results towards a possible female-specific response to the diet. The importance of gender in this context is, indeed, an aspect to be further investigated.

Author contributions

The authors' responsibilities are as follows: HF, MWV, EP and AMG designed the dietary intervention study; HF implemented the food trial and performed all data collection; PA was responsible for defining the methodology for blood collection; HF and TJ optimized the NMR stool sample preparation protocol; HF and DD performed NMR analysis, data processing and statistical analysis; HF, AMG and JER performed correlation analysis; CC optimized DNA extraction and qPCR protocols; HF and JCB conducted qPCR analysis as well as data processing and statistical analysis; AG validated data outputs of microbiota analysis and provided essential intellectual inputs; HF and AMG drafted the original manuscript and have primary responsibility for the final content; MWV, EP and AMG supervised the Dietary fibres in pulse seeds and fractions intervention study execution, validated data outputs, and provided essential intellectual inputs; and all authors revised and approved the final manuscript.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgements

The authors acknowledge the help of Joana Ferrão Silveira, MSc and Kimhoung Say, MSc during both the food intervention and data collection procedures, as well as the food company EUREST for providing the trial meals.

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. AMG acknowledges the project CICECO-Aveiro Institute of Materials and FCT/MCTES (PIDDAC), UIDB/50011/2020 (DOI 10.54499/UIDB/50011/2020), UIDP/50011/2020 (DOI 10.54499/UIDP/50011/2020) & LA/P/0006/2020 (DOI 10.54499/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 is partially supported by PTNMR Infrastructure Project no. 022161 (co-financed by FEDER through COMPETE 2020, POCI and PORL and FCT through PIDDAC).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clnesp.2024.06.051.

References

[1] Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, et al. Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. Lancet 2019;393:447. https://doi.org/10.1016/S0140-6736(18)31788-4.

[2] Singh RK, Chang HW, Yan D, Lee KM, Ucmak D, Wong K, et al. Influence of diet on the gut microbiome and implications for human health. J Transl Med 2017;15:1–17.

[3] Ferreira H, Pinto E, Vasconcelos MW. Legumes as a cornerstone of the transition toward more sustainable agri-food systems and diets in Europe. Front Sustain Food Syst 2021. https://doi.org/10.3389/fsufs.2021.694121.

[4] Vasconcelos MW, Grusak MA, Pinto E, Gomes A, Ferreira H, Balázs B, et al. The biology of legumes and their agronomic, economic, and social impact. In: Hasanuzzaman M, Araújo S, Gill SS, editors. The plant family fabaceae. Singapore: Springer; 2020. p. 3–25.

[5] Sozer N, Holopainen-Mantila U, Poutanen K. Traditional and new food uses of pulses. Cereal Chem 2017;94:66–73. https://doi.org/10.1094/CCHEM-04-16-0082-FI.

[6] Hall C, Hillen C, Robinson JG. Composition, nutritional value, and health benefits of pulses. Cereal Chem 2017;94:11–31. https://doi.org/10.1094/CCHEM-03-16-0069-FI.

[7] Tosh SM, Yada S. Dietary fibres in pulse seeds and fractions : characterization, functional attributes, and applications. Food Res Int 2010;43:450–60. https://doi.org/10.1016/j.foodres.2009.09.005.

[8] Marinangeli C, Curran J, Barr SI, Slavin J, Puri S, Swaminathan S, et al. Enhancing nutrition with pulses: defining a recommended serving size for adults. Nutr Rev 2017;75:990. https://doi.org/10.1093/nutrit/nux058.

[9] Singh B, Singh JP, Kaur A, Singh N. Phenolic composition and antioxidant potential of grain legume seeds: a review. Food Res Int 2017;101:1–16. https://doi.org/10.1016/j.foodres.2017.09.026.

[10] Ferreira H, Vasconcelos MW, Gil AM, Pinto E. Benefits of pulse consumption on metabolism and health: a systematic review of randomized controlled trials. Crit Rev Food Sci Nutr 2020;85. https://doi.org/10.1080/10408398.2020.1716680.

[11] Kadyan S, Sharma A, Arjmandi BH, Singh P, Nagpal R. Prebiotic potential of dietary beans and pulses and their resistant starch for aging-associated gut and metabolic health. Nutrients 2022;14. https://doi.org/10.3390/nu14091726.

[12] Marinangeli C, Harding SV, Zafron M, Rideout TC. A systematic review of the effect of dietary pulses on microbial populations inhabiting the human gut. Benef Microbes 2020;11:457–68. https://doi.org/10.3920/BM2020.0028.

[13] Makki K, Deehan EC, Walter J, Bäckhed F. The impact of dietary fiber on gut microbiota in host health and disease. Cell Host Microbe 2018;23:705–15. https://doi.org/10.1016/j.chom.2018.05.012.

[14] Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell 2016;165:1332–45. https://doi.org/10.1016/j.cell.2016.05.041.

[15] Tindall AM, Petersen KS, Kris-Etherton PM. Dietary patterns affect the gut microbiome-the link to risk of cardiometabolic diseases. J Nutr 2018;148:1402–7. https://doi.org/10.1093/jn/nxy141.

[16] Lin L, Zhang J. Role of intestinal microbiota and metabolites on gut homeostasis and human diseases. BMC Immunol 2017;18:1–25. https://doi.org/10.1186/s12865-016-0187-3.

[17] Mcloughlin RF, Berthon BS, Jensen ME, Baines KJ, Wood LG. Short-chain fatty acids, prebiotics, synbiotics, and systemic inflammation: a systematic review and meta-analysis. Am J Clin Nutr 2017:930–45. https://doi.org/10.3945/ajcn.117.156265.

[18] Smith SC, Choy R, Johnson SK, Hall RS, Wildeboer-Veloo AC, Welling GW. Lupin kernel fiber consumption modifies fecal microbiota in healthy men as determined by rRNA gene fluorescent in situ hybridization. Eur J Nutr 2006;45:335–41. https://doi.org/10.1007/s00394-006-0603-1.

[19] Finley John W, Burrell JB, Reeves PG. Pinto bean consumption changes SCFA profiles in fecal fermentations, bacterial populations of the lower bowel, and lipid profiles in blood of humans. J Nutr 2007;137:2391–8. https://doi.org/10.1093/jn/137.11.2391.

[20] Fernando W, Hill JE, Zello GA, Tyler RT, Dahl WJ, Van Kessel AG. Diets supplemented with chickpea or its main oligosaccharide component raffinose modify faecal microbial composition in healthy adults. Benef Microbes 2010;1:197–207. https://doi.org/10.3920/BM2009.0027.

[21] Sheflin AM, Borresen EC, Kirkwood JS, Boot CM, Whitney AK, Lu S, et al. Dietary supplementation with rice bran or navy bean alters gut bacterial metabolism in colorectal cancer survivors. Mol Nutr Food Res 2017;61. https://doi.org/10.1002/mnfr.201500905.

[22] Lambert JE, Parnell JA, Tunnicliffe JM, Han J, Sturzenegger T, Reimer RA. Consuming yellow pea fiber reduces voluntary energy intake and body fat in overweight/obese adults in a 12-week randomized controlled trial. Clin Nutr 2017;36:126–33. https://doi.org/10.1016/j.clnu.2015.12.016.

[23] Losasso C, Eckert EM, Mastrorilli E, Villiger J, Mancin M, Patuzzi I, et al. Assessing the influence of vegan, vegetarian and omnivore oriented westernized dietary styles on human gut microbiota: a cross sectional study. Front Microbiol 2018;9:1–12. https://doi.org/10.3389/fmicb.2018.00317.

[24] Levatte M, Keshteli AH, Zarei P, Wishart DS. Applications of metabolomics to precision nutrition. Lifestyle Genom 2022;15:1–9. https://doi.org/10.1159/000518489.

[25] Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, et al. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. Lancet Diabetes Endocrinol 2017;5:184. https://doi.org/10.1016/S2213-8587(16)30419-3.

[26] Harsha PSCS, Wahab RA, Cuparencu C, Dragsted LO, Brennan L. A metabolomics approach to the identification of urinary biomarkers of pea intake. Nutrients 2018;10. https://doi.org/10.3390/nu10121911.

[27] Lindqvist HM, 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;151:30. https://doi.org/10.1093/jn/nxaa019.

[28] 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. https://doi.org/10.1021/acs.jafc.8b00047.

[29] Madrid-Gambin F, Llorach R, Vázquez-Fresno R, Urpi-Sarda M, Almanza-Aguilera E, Garcia-Aloy M, et al. 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;16:1483. https://doi.org/10.1021/acs.jproteome.6b00860.

[30] Posma JM, Garcia-Perez I, Heaton JC, Burdisso P, Mathers JC, Draper J, et al. Integrated analytical and statistical two-dimensional spectroscopy strategy for metabolite identification: application to dietary biomarkers. Anal Chem 2017;89:3300–9. https://doi.org/10.1021/acs.analchem.6b03324.

[31] Stella C, Beckwith-hall B, Cloarec O, Holmes E, Lindon JC, Powell J, et al. Susceptibility of human metabolic phenotypes to dietary modulation research articles. J Proteome Res 2006:2780–8. https://doi.org/10.1021/pr060265y.

[32] Xu J, Yang S, Cai S, Dong J. Identification of biochemical changes in lacto-vegetarian urine using 1 H NMR spectroscopy and pattern recognition. Anal Bioanal Chem 2010;396:1451. https://doi.org/10.1007/s00216-009-3338-z.

[33] Kim H, Yu B, Li X, Wong KE, Boerwinkle E, Seidelmann SB, et al. Serum metabolomic signatures of plant-based diets and incident chronic kidney disease. Am J Clin Nutr 2022;116:151–64. https://doi.org/10.1093/ajcn/nqac054.

[34] Lindqvist HM, Rådjursöga M, Malmodin D, Winkvist A, Ellegård L. Serum metabolite profiles of habitual diet: evaluation by 1H-nuclear magnetic resonance analysis. Am J Clin Nutr 2019;110:53–62. https://doi.org/10.1093/ajcn/nqz032.

[35] Schmidt JA, Fensom GK, Rinaldi S, Scalbert A, Gunter MJ, Holmes MV, et al. NMR metabolite profiles in male meat-eaters, fish-eaters, vegetarians and vegans, and comparison with MS metabolite profiles. Metabolites 2021;11(2):121. https://doi.org/10.3390/metabo11020121.

[36] Hernández-Alonso P, Becerra-Tomás N, Papandreou C, Bulló M, Guasch-Ferré M, Toledo E, et al. Plasma metabolomics profiles are associated with the amount and source of protein intake: a metabolomics approach within the PREDIMED study. Mol Nutr Food Res 2020;64:2000178. https://doi.org/10.1002/mnfr.202000178.
[37] Miles FL, Orlich MJ, Mashchak A, Chandler PD, Lampe JW, Duerksen-Hughes P, et al. The biology of veganism: plasma metabolomics analysis reveals distinct profiles of vegans and non-vegetarians in the adventist health study-2 cohort. Nutrients 2022;14. https://doi.org/10.3390/nu14030709.
[38] Schmidt JA, Rinaldi S, Ferrari P, Carayol M, Achaintre D, Scalbert A, et al. Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. Am J Clin Nutr 2015;102:1518–26. https://doi.org/10.3945/ajcn.115.111989.
[39] Wang F, Baden MY, Guasch-Ferré M, Wittenbecher C, Li J, Li Y, et al. Plasma metabolite profiles related to plant-based diets and the risk of type 2 diabetes. Diabetologia 2022;65:1119–32. https://doi.org/10.1007/s00125-022-05692-8.
[40] Jacobs DM, Deltimple N, van Velzen E, van Dorsten FA, Bingham M, Vaughan EE, et al. 1H NMR metabolite profiling of feces as a tool to assess the impact of nutrition on the human microbiome. NMR Biomed 2008;21:615–26. https://doi.org/10.1002/nbm.1233.
[41] Mayengbam S, Lambert JE, Parnell JA, Tunnicliffe JM, Nicolucci AC, Han J, et al. Impact of dietary fiber supplementation on modulating microbiota–host–metabolic axes in obesity. J Nutr Biochem 2019;64:228–36. https://doi.org/10.1016/j.jnutbio.2018.11.003.
[42] He W, Bertram HC. NMR-based metabolomics to decipher the molecular mechanisms in the action of gut-modulating foods. Foods 2022;11(17):2707. https://doi.org/10.3390/foods11172707.
[43] Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med 2013;19:576–85. https://doi.org/10.1038/nm.3145.
[44] Chan CWH, Law BMH, Waye MMY, Chan JYW, So WKW, Chow KM. Trimethylamine-N-oxide as one hypothetical link for the relationship between intestinal microbiota and cancer - where we are and where shall we go? J Cancer 2019;10:5874–82. https://doi.org/10.7150/jca.31737.
[45] Zarei I, Baxter BA, Oppel RC, Borresen EC, Brown RJ, Ryan EP. Plasma and urine metabolite profiles impacted by increased dietary navy bean intake in colorectal cancer survivors: a randomized-controlled trial. Cancer Prev Res 2021;14:497–508. https://doi.org/10.1158/1940-6207.CAPR-20-0270.
[46] Baxter BA, Li KJ, Zarei I, Yao L, Rao S, Ryan EP. Nontargeted and targeted metabolomics identifies dietary exposure biomarkers for navy bean and rice bran consumption in children and adults. J Agric Food Chem 2022;70:14531–43. https://doi.org/10.1021/acs.jafc.2c02378.
[47] Hill EB, Baxter BA, Pfluger B, Slaughter CK, Beale M, Smith HV, et al. Plasma, urine, and stool metabolites in response to dietary rice bran and navy bean supplementation in adults at high-risk for colorectal cancer. Frontline Gastroenterol 2023;2. https://doi.org/10.3389/fgstr.2023.1087056.
[48] Marinangeli C, Curran J, Barr SI, Slavin J, Puri S, Swaminathan S, et al. Enhancing nutrition with pulses: defining a recommended serving size for adults. Nutr Rev 2017;75:990–1006. https://doi.org/10.1093/nutrit/nux058.
[49] Ferreira H, Vasconcelos M, Gil AM, Oliveira B, Varandas E, Vilela E, et al. Impact of a daily legume-based meal on dietary and nutritional intake in a group of omnivorous adults. Nutr Bull 2023;48:190–202. https://doi.org/10.1111/nbu.12613.
[50] Buzzard M. 24-Hour dietary recall and food record methods. In: Nutritional epidemiology. New York: Oxford University Press; 1998. p. 50–73.
[51] Stewart A, Marfell-Jones M, Olds T, Ridder H. International standards for anthropometric assessment. Lower Hutt, New Zealand: International Society for the Advancement of Kinanthropometry (ISAK); 2011.
[52] WHO. Part 3: training & practical guides - section 4: guide to physical measurements. In: The WHO STEPS surveillance manual; 2008. 3-4-1–3-4-15.
[53] Laíns I, Duarte D, Barros S, Martins AS, Gil J, Miller JB, et al. Human plasma metabolomics in age-related macular degeneration (AMD) using nuclear magnetic resonance spectroscopy. PLoS One 2017;12:1–18. https://doi.org/10.1371/journal.pone.0177749 [May].
[54] NZYtech. NZY tissue gDNA isolation kit protocol. 2018.
[55] NZYtech. NZY Tissue gDNA Isolation kit - support protocol for isolating genomic DNA from stool. 2016.
[56] Matlock B. Thermo Fisher scientific - technical note 52646: assessment of nucleic acid purity, Thermo Fisher scientific - technical note 52646. USA: Wilmington; 2015.
[57] Applied Biosystems. Creating standard curves with genomic DNA or plasmid DNA templates for use in quantitative PCR. 2003.
[58] Life Technologies. Realtime PCR handbook, Realtime PCR handbook. 2014.
[59] Deda O, Gika HG, Wilson ID, Theodoridis GA. An overview of fecal sample preparation for global metabolic profiling. J Pharm Biomed Anal 2015;113:137–50. https://doi.org/10.1016/j.jpba.2015.02.006.
[60] Lamichhane S, Yde CC, Schmedes MS, Jensen HM, Meier S, Bertram HC. Strategy for nuclear-magnetic-resonance-based metabolomics of human feces. Anal Chem 2015;87:5930–7. https://doi.org/10.1021/acs.analchem.5b00977.
[61] O'Sullivan V, Madrid-Gambin F, Alegra T, Gibbons H, Brennan L. Impact of sample storage on the NMR fecal water metabolome. ACS Omega 2018;3:16585–90. https://doi.org/10.1021/acsomega.8b01761.
[62] Silva JCP, Mota M, Martins FO, Nogueira C, Gonçalves T, Carneiro T, et al. Intestinal microbial and metabolic profiling of mice fed with high-glucose and high-fructose diets. J Proteome Res 2018;17:2880–91. https://doi.org/10.1021/acs.jproteome.8b00354.
[63] Wu J, An Y, Yao J, Wang Y, Tang H. An optimised sample preparation method for NMR-based faecal metabonomic analysis. Analyst 2010;135:1023–30. https://doi.org/10.1039/b927543f.
[64] Yang Y, Wang L, Wang S, Huang R, Zheng L. An integrated metabonomic approach to studying metabolic profiles in rat models with insulin resistance induced by high fructose. Mol Biosyst 2014;10:1803–11. https://doi.org/10.1039/c3mb70618d.
[65] Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Rosa V, et al. Hmdb 4.0: the human metabolome database for 2018. Nucleic Acids Res 2018;46:608–17. https://doi.org/10.1093/nar/gkx1089.
[66] Lamichhane S, Yde CC, Schmedes MS, Jensen HM, Meier S, Bertram HC. Strategy for nuclear-magnetic-resonance-based metabolomics of human feces. Anal Chem 2015;87:5930–7. https://doi.org/10.1021/acs.analchem.5b00977.
[67] Savorani F, Rasmussen MA, Mikkelsen MS, Engelsen SB. A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. Food Res Int 2013;54:1131. https://doi.org/10.1016/j.foodres.2012.12.025.
[68] Zhao Y, Wu J, Li JV, Zhou N, Tang H, Wang Y. Gut microbiota composition modifies fecal metabolic profiles in mice. J Proteome Res 2013;12:2987–99. https://doi.org/10.1021/pr400263n.
[69] Cloarec O, Dumas M, Craig A, Barton RH, Trygg J, Hudson J, et al. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1 H NMR data sets. Anal Chem 2005;77:1282. https://doi.org/10.1021/ac048630x.
[70] Diaz SO, Pinto J, Graça G, Duarte IF, Barros AS, Galhano E, et al. Metabolic biomarkers of prenatal disorders: an exploratory NMR metabonomics study of second trimester maternal urine and blood plasma. J Proteome Res 2011;10:3732–42. https://doi.org/10.1021/pr200352m.
[71] Nicholson JK, Foxall PJD, Spraul M, Farrant RD, Lindon JC. 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Anal Chem 1995;67:793–811. https://doi.org/10.1021/ac00101a004.
[72] Pinto J, Almeida LM, Martins AS, Duarte D, Barros AS, Galhano E, et al. Prediction of gestational diabetes through NMR metabolomics of maternal blood. J Proteome Res 2015;14:2696–706. https://doi.org/10.1021/acs.jproteome.5b00260.
[73] Finotello F, Mastrorilli E, Di Camillo B. Measuring the diversity of the human microbiota with targeted next-generation sequencing. Briefings Bioinf 2018;19:679–92. https://doi.org/10.1093/bib/bbw119.
[74] Veselkov KA, Lindon JC, Ebbels TMD, Crockford D, Volynkin VV, Holmes E, et al. Recursive segment-wise peak alignment of biological 1H NMR spectra for improved metabolic biomarker recovery. Anal Chem 2009;81:56. https://doi.org/10.1021/ac8011544.
[75] Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, et al. Assessment of PLSDA cross validation. Metabolomics 2008;4:81. https://doi.org/10.1007/s11306-007-0099-6.
[76] Wiklund S, Nilsson D, Eriksson L, Sjöström M, Wold S, Faber K. A randomization test for PLS component selection. J Chemom 2007;21:427. https://doi.org/10.1002/cem.1086.
[77] Berben L, Sereika SM, Engberg S. Effect size estimation: methods and examples. Int J Nurs Stud 2012;49:1039. https://doi.org/10.1016/j.ijnurstu.2012.01.015.
[78] Revelle WR. “psych”: procedures for personality and psychological research [WWW document]. 2017. URL, https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research.
[79] Hazra A, Gogtay N. Biostatistics series module 6: correlation and linear regression. Indian J Dermatol 2016;61:593–601. https://doi.org/10.4103/0019-5154.193662.
[80] Wei T, Simko V. R package “corrplot”: visualization of a correlation matrix (version 0.92) [WWW document]. 2021. URL, https://github.com/taiyun/corrplot.
[81] Zhao J, Zhang X, Liu H, Brown MA, Qiao S. Dietary protein and gut microbiota composition and function. Curr Protein Pept Sci 2019;20:145–54. https://doi.org/10.2174/1389203719666180514145437.
[82] Simpson HL, Campbell BJ. Review article: dietary fibre–microbiota interactions. Aliment Pharmacol Ther 2015;42:158–79. https://doi.org/10.1111/apt.13248.
[83] Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. Genome Med 2016;8:51. https://doi.org/10.1186/s13073-016-0307-y.
[84] Ventura M, Turroni F, Strati F, Sinderen van D. Chapter 9 - the gut microbiota in health and disease. In: Advances in molecular and cellular microbiology. Advances in molecular and cellular microbiology; 2014. p. 136–46. https://doi.org/10.1079/9781780640495.0136.
[85] Tomova A, Bukovsky I, Rembert E, Yonas W, Alwarith J, Barnard ND, et al. The effects of vegetarian and vegan diets on gut microbiota. Front Nutr 2019;6. https://doi.org/10.3389/fnut.2019.00047.
[86] La Rosa SL, Leth ML, Michalak L, Hansen ME, Pudlo NA, Glowacki R, et al. The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary β-mannans. Nat Commun 2019;10:905. https://doi.org/10.1038/s41467-019-08812-y.
[87] Ferreira H, Pinto E, Gil AM, Vasconcelos MW. Chapter 15 - potential role of pulses in the development of functional foods modulating inflammation and

oxidative stress. In: Hernández-Ledesma B, Martínez-Villaluenga CBT, editors. Current advances for development of functional foods modulating inflammation and oxidative stress. Academic Press; 2022. p. 287–309. https://doi.org/10.1016/B978-0-12-823482-2.00003-0.
[88] Yang Q, Liang Q, Balakrishnan B, Belobrajdic DP, Feng QJ, Zhang W. Role of dietary nutrients in the modulation of gut microbiota: a narrative review. Nutrients 2020;12:1–57. https://doi.org/10.3390/nu12020381.
[89] Dahl WJ, Alvarez MM. Whole pulses and pulse fiber: modulating gastrointestinal function and the microbiome. In: Health benefits of pulses. Springer; 2019. p. 91–108. https://doi.org/10.1007/978-3-030-12763-3_7.
[90] Singh B, Singh JP, Shevkani K, Singh N, Kaur A. Bioactive constituents in pulses and their health benefits. J Food Sci Technol 2017;54:858–70. https://doi.org/10.1007/s13197-016-2391-9.
[91] Gullón P, Gullón B, Tavaria F, Vasconcelos M, Gomes AM. In vitro fermentation of lupin seeds (Lupinus albus) and broad beans (Vicia faba): dynamic modulation of the intestinal microbiota and metabolomic output. Food Funct 2015;6:3316–22. https://doi.org/10.1039/c5fo00675a.
[92] Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome 2019;7:91. https://doi.org/10.1186/s40168-019-0704-8.
[93] den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. J Lipid Res 2013;54:2325–40. https://doi.org/10.1194/jlr.R036012.
[94] Wang J, Pan J, Chen H, Li Y, Amakye WK, Liang J, et al. Fecal short-chain fatty acids levels were not associated with autism spectrum disorders in Chinese children: a case–control study. Front Neurosci 2019;13:1216. https://doi.org/10.3389/fnins.2019.01216.
[95] Trefflich I, Dietrich S, Braune A, Abraham K, Weikert C. Short- and branched-chain fatty acids as fecal markers for microbiota activity in vegans and omnivores. Nutrients 2021;13. https://doi.org/10.3390/nu13061808.
[96] Saa P, Urrutia A, Silva-Andrade C, Martín AJ, Garrido D. Modeling approaches for probing cross-feeding interactions in the human gut microbiome. Comput Struct Biotechnol J 2022;20:79–89. https://doi.org/10.1016/j.csbj.2021.12.006.
[97] Das B. Chapter One - an introduction to human microbiome. In: Das B, Singh VBT-P in MB, T S, editors. Human microbiome in health and disease - Part A. Academic Press; 2022. p. 1–28. https://doi.org/10.1016/bs.pmbts.2022.06.026.
[98] Saito Y, Sato T, Nomoto K, Tsuji H. Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites. FEMS Microbiol Ecol 2018;94:fiy125. https://doi.org/10.1093/femsec/fiy125.
[99] Huang Z, Boekhorst J, Fogliano V, Capuano E, Wells JM. Distinct effects of fiber and colon segment on microbiota-derived indoles and short-chain fatty acids. Food Chem 2023;398:133801. https://doi.org/10.1016/j.foodchem.2022.133801.
[100] Bourdeau-Julien I, Castonguay-Paradis S, Rochefort G, Perron J, Lamarche B, Flamand N, et al. The diet rapidly and differentially affects the gut microbiota and host lipid mediators in a healthy population. Microbiome 2023;11:26. https://doi.org/10.1186/s40168-023-01469-2.
[101] Almut H, Tanweer KM, Giuseppe P,A,RD, M HHJ, Ines T. Functional metabolic map of Faecalibacterium prausnitzii, a beneficial human gut microbe. J Bacteriol 2014;196:3289–302. https://doi.org/10.1128/JB.01780-14.
[102] Lopez-Siles M, Duncan SH, Garcia-Gil LJ, Martinez-Medina M. Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J 2017;11:841–52. https://doi.org/10.1038/ismej.2016.176.
[103] Shiga TM, Lajolo FM. Cell wall polysaccharides of common beans (Phaseolus vulgaris L.) - composition and structure. Carbohydr Polym 2006;63:1–12. https://doi.org/10.1016/j.carbpol.2005.06.025.
[104] Brosnan JT, Plumptre L, Brosnan ME, Pongnopparat T, Masih SP, Visentin CE, et al. Formate concentrations in maternal plasma during pregnancy and in cord blood in a cohort of pregnant Canadian women: relations to genetic polymorphisms and plasma metabolites. Am J Clin Nutr 2019;110:1131–7. https://doi.org/10.1093/ajcn/nqz152.
[105] He X, Zhao S, Li Y. Faecalibacterium prausnitzii: a next-generation probiotic in gut disease improvement. Can J Infect Dis Med Microbiol 2021;2021:6666114. https://doi.org/10.1155/2021/6666114.
[106] Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol 2017;19:29–41. https://doi.org/10.1111/1462-2920.13589.