Universitat Autònoma de Barcelona

DESIGN AND VALIDATION OF A SHORT FOOD FREQUENCY QUESTIONNAIRE TO RELATE DIET WITH THE GUT MICROBIOTA

FRANCISCA YÁÑEZ ARAYA

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: http://cat.creativecommons.org/?page_id=184

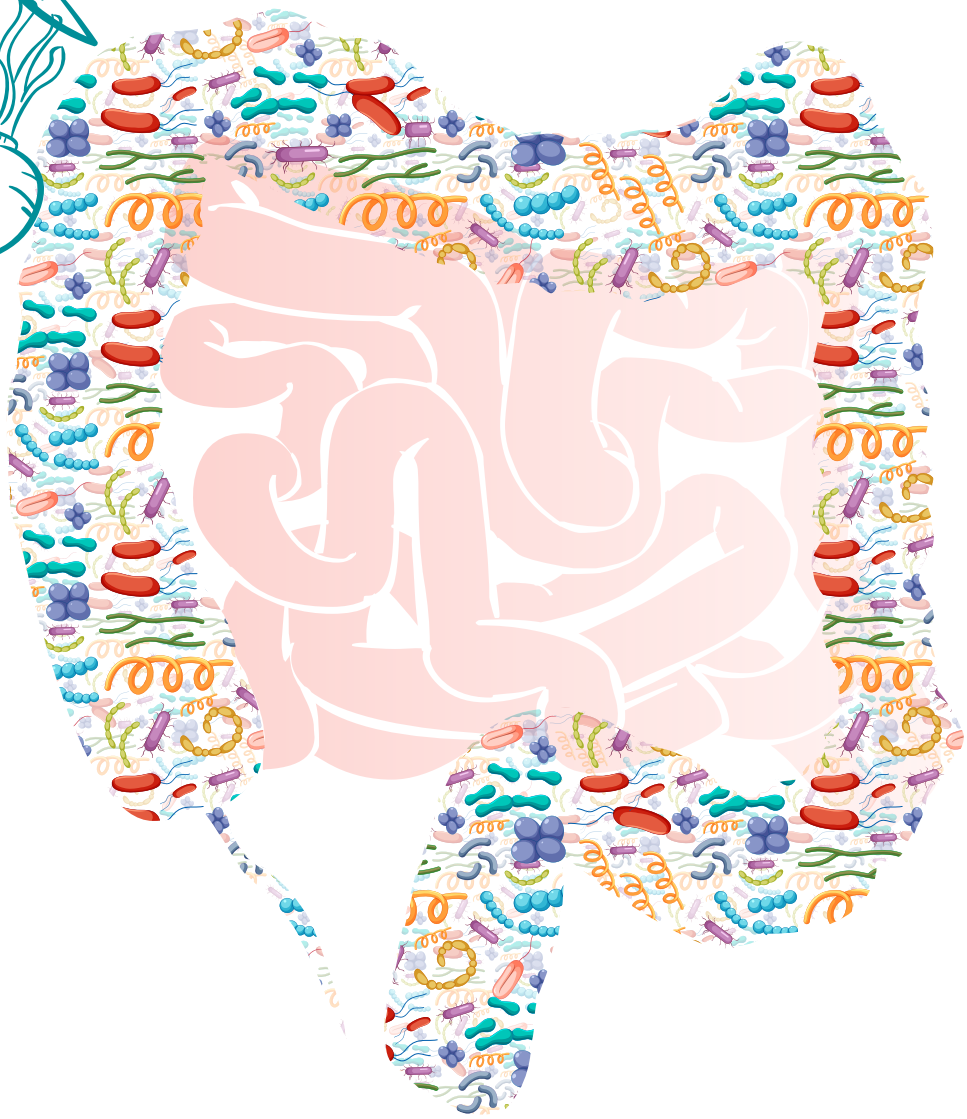
ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: http://es.creativecommons.org/blog/licencias/

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

Design and Validation of a Short Food Frequency Questionnaire to relate Diet with the Gut Microbiota

Francisca Yáñez Araya

Doctoral Thesis



Design and validation of a short food frequency questionnaire to relate diet with the gut microbiota

A doctoral thesis presented by

Francisca Yáñez Araya to aim for the degree of Doctor

Supervisor: Dra. Chaysavanh Manichanh

Tutor: Dr. Carlos Guarner Aguilar

Doctoral Program in Medicine

Department of Medicina

Barcelona, 2021

Universitat Autònoma de Barcelona

AGRADECIMIENTOS

Gracias a Dios y a la Máter… por fin llegó el momento de escribir estas palabras. Han sido casi 5 años que dejé mi tierra, Chile. Dejé a mi familia, amigos, conocidos, trabajo…. casi una vida entera, para emprender este nuevo rumbo y desafío. Se suponía que este viaje sería solo de un año, pero aquí estoy.

Doy mil gracias de haber conocido a la Dra. Chaysavanh Manichanh, directora de esta tesis, que me dio la oportunidad y confió en mí para iniciar esta nueva subárea de dieta en su grupo de investigación. Sin su apoyo incondicional hubiera sido bastante difícil haber continuado los primeros años de esta tesis. Sus conocimientos y consejos me han servido mucho para crecer y seguir creciendo en este mundo de la investigación.

Gracias a Becas Chile ANID, por haber financiado este doctorado y al Dr. Carlos Guarner por haber sido tutor de esta tesis.

Gratitud a todos los colaboradores clínicos Dra. Dopazo, Dra. Goya y Dra. Redecilla entre otros, que me permitieron aprender de otras áreas y gestionar el trabajo en equipo. Muchas gracias al Dr. Casellas que ha estado siempre al tanto de este proyecto. Claudia, gracias por animarme a continuar en esta última etapa. Fina, siempre estaré eternamente agradecida por lo que hiciste por mí cuando inicié este proceso.

Agradezco inmensamente a todas las personas que he conocido en el VHIR que prácticamente han sido mi familia durante estos años. De los que ya no están; Alba y Marta, muchas gracias por enseñarme y haberme dado su apoyo cuando inicié este desafío. Manon francesa loca me ayudaste un montón cuando inicié este proyecto. Carlos y Ferrán, Dios… no sé qué hubiera hecho sin su ayuda, aún recuerdo cuando pasamos todo un verano discutiendo cómo analizar los datos de esta tesis. Cris, gracias por estar animándome, apoyándome en todo y ser mi confidente. Guillaume, gracias por toda tu ayuda, apoyo, y por permitirme conocer otras facetas de ti… tu humor.

De lo que están; Elba, Mercé y Laura gracias por hacerme reír y dar vida al laboratorio, sin Uds. no sería lo mismo. Ana y Bea, muchas gracias por haberme permitido compartir un poco más con Uds.

De mi actual grupo dar las gracias a Blanca, Gerard, Iñigo y Zixuan por haberme ayudado en estos últimos meses. Y sobre todo agradecer a Zaida que me ayudo en esta etapa final. Encarna, estos últimos años has sido mi confidente, paño de lágrimas, guía, etc. no sabes cuánto te agradezco todo el apoyo y sabiduría que me has dado.

Finalmente agradecer a todo el personal del VHIR; coordinación, biobanco, celadores, personal de limpieza etc. Especialmente a Helena, Ana, Isabel, Ray, Margarita y Flor.

Y en último lugar, quiero dar las gracias a mi guapito y familia. Edgar, lo mejor que me ha pasado desde que llegué a Barcelona, fue conocerte. Me has acompañado incondicionalmente, hemos estado apoyándonos en las muy buenas y en las muy malas. Has sido mi paño de lágrimas, has tenido la capacidad de soportarme estos últimos meses, incluso me montaste una fiesta en casa para darme más ánimo. Me has dado fuerzas, me has enseñado y ayudado mucho a lo largo de estos años… has y seguirás siendo mi compañero de vida.

Bamba y Vicho, mis hermanos, gracias por estar siempre apoyándome y darme las fuerzas para seguir adelante desde tan lejos. Papo y mamá, nada de esto hubiera sido posible sin Uds. Les agradezco infinitamente por apoyarme incondicionalmente desde siempre, con todo su amor. Todos sus esfuerzos para sacarnos adelante a mis hermanos y a mí, han valido la pena. Les doy infinitas gracias por hacer de mí, la mujer que soy. Y no sé si este es el medio, pero les quiero pedir perdón por no haber estado con Uds. en los momentos buenos y malos de estos años fuera, sobre todo a ti, hermana.

Os quiero mucho a todos, cada uno se lleva un pedacito de mi corazón.

Y ahora a comenzar una nueva etapa y emprender un nuevo desafío….

Para ustedes,
papo y mamá

LIST OF ABBREVIATIONS

16S rRNA, 16S ribosomal RNA

24HR, 24 hours dietary recall/records

ASVs, amplicon sequence variants

AUSNUT, Australian Food, Supplement and Nutrient Database

BMI, body mass index

BMRest, basal metabolic rate estimated

CCC, concordance correlation coefficient

CHOL, cholesterol

C-section, caesarean section

EIrep, energy intake reported

EFSA, European Food Safety Authority

ENALIA2, Encuesta Nacional de Alimentación en la Población Adulta, Mayores y Embarazadas 2.

FCDB/FCT, databases or food composition table

FFQ, food frequency questionnaire

FODMAP, fermentable oligo-, di-, mono-saccharides, and polyols

GFD, gluten-free diet

HEI, Healthy Eating Index

HF-diet, hight fat diet

HFD-index, Healthy Food Diversity Index

IASE, Indice de Alimentación-Saludable (Healthy Feeding Index)

IBD, inflammatory bowel disease

ICC, intraclass correlation coefficients

KD, ketogenic diet

MAC, microbiota accessible carbohydrates

MaAsLin2, Microbiome Multivariable Association with Linear Models

MD, Mediterranean diet

MEDAS, Mediterranean Diet Adherence Screeners

MUFA, monounsaturated fatty acid

NRI, nutritional reference intake

inFCDB, in-house Food Composition Database

OTUs, operational taxonomic units

PAL, physical activity level

PCR, polymerase chain reaction

PUFA, polyunsaturated fatty acid

QIIME, Quantitative Insights Into Microbial Ecology

SCFA, short-chain fatty acid

SFA, saturated fatty acid

sFFQ, short food frequency questionnaire

sFFQp, short food frequency questionnaire pilot

SOP, standard operating procedure

TMAO, trimethylamine N-oxide

VHIR, Vall d´Hebron Research Institute

UK, United Kingdom

USDA, United States Department of Agriculture

UWM, unweighted mean

WM, weighted mean

INDEX

SUMMARY

Diet is recognised as the main driver of changes in the gut microbiota. However, linking habitual dietary intake with the composition and activity of the microbiome remains a challenge. Indeed, the collection of dietary data in conjunction with microbiome data has been under-evaluated, leading to most microbiome studies containing little or no information about diet.

This thesis aims to design and validate a short food frequency questionnaire (sFFQ) to relate diet to the microbiome composition in a healthy adults population.

For this purpose, we conducted two consecutive studies (n = 84). A first pilot study (n= 40) was designed to build and validate a simplified semi-quantitative 46-item sFFQ against a reference method (mean of three 24-h dietary recall (3-24HR)) to evaluate the dietary habits of our population. Furthermore, a second study (n = 44) was implemented to redesign and validate a web-based 58-item sFFQ from the data provided by the pilot study, using the 3-24HR as a reference method. Faecal samples were collected to relate the gut microbiome profile (based on the 16S rRNA gene) with the dietary data extracted from the sFFQ.

Relative validation analysis provided acceptable classification and agreement for 13 out of 24 (54%) food groups and 20 out of 29 nutrients (69%) based on at least three of the following statistical methods: Wilcoxon test, Bland & Altman analysis, Spearman´s correlation, intraclass correlation coefficient and cross-classification. Microbiome analysis showed that higher diversity was positively associated with age, vaginal delivery, and fruit intake. In contrast, diversity was negatively associated with BMI classification, processed meats, ready-to-eat meals, sodium, and saturated fat. Our analysis also revealed a correlation between food or nutrient groups and microbial composition.

Overall, the present work provided the first validated dietary assessment tool to correlate diet intake with microbiome data for a healthy adult population. This tool could be adapted and used in population studies and pave the way to better understand the contribution of diet to health.

RESUMEN

Se reconoce que la dieta es el principal impulsor de los cambios en la microbiota intestinal. Sin embargo, vincular la ingesta dietética habitual con la composición y la actividad del microbioma sigue siendo un desafío. De hecho, la recopilación de datos dietéticos junto con los datos del microbioma se han subestimado, lo que ha llevado a que la mayoría de los estudios de microbioma contengan poca o ninguna información sobre la dieta.

El objetivo de esta tesis es diseñar y validar un cuestionario de frecuencia de consumo alimentario corto (CFCAc) para relacionar la dieta con la composición de microbioma en una población de adultos sanos.

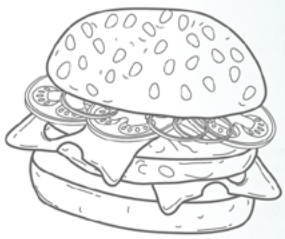
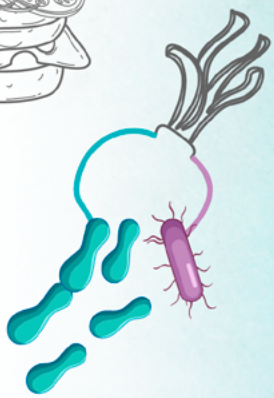
Para ello, realizamos dos estudios consecutivos (n = 84). Se diseñó un primer estudio piloto (n= 40) para construir y validar un CFCAc semicuantitativo de 46 items contra un método de referencia (media de tres recordatorios alimentarios de 24 horas (24HR)), para evaluar los hábitos alimentarios de nuestra población. Además, se implementó un segundo estudio (n = 44) para rediseñar y validar el CFCAc de 58 ítems realizado vía web a partir de los datos proporcionados por el estudio piloto, utilizando tres 24HR como método de referencia. Además, se recolectaron muestras fecales para relacionar el perfil del microbioma intestinal (basado en el gen 16S rRNA) con los datos dietéticos extraídos del CFCAc.

El análisis de validación relativa proporcionó una clasificación y concordancia aceptables para 13 de 24 (54%) grupos de alimentos y 20 de 29 nutrientes (69%) basado en al menos tres de los siguientes métodos estadísticos: Prueba de Wilcoxon, análisis de Bland & Altman, correlación de Spearman, coeficiente de correlación intraclase y clasificación cruzada. El análisis del microbioma mostró que una mayor diversidad se asoció positivamente con la edad, el parto vaginal y la ingesta de fruta. Por el contrario, la diversidad se asoció negativamente con la clasificación del IMC, las carnes procesadas, las comidas listas para consumo, el sodio y las grasas saturadas. Nuestro análisis también reveló una correlación entre los grupos de alimentos, nutrientes y la composición microbiana.

En resumen, el presente trabajo proporcionó la primera herramienta de evaluación dietética validada para correlacionar la ingesta dietaría con los datos del microbioma en una población adulta sana. Esta herramienta podría adaptarse y utilizarse en estudios de población y allanar el camino para comprender mejor la contribución de la dieta a la salud.






1. INTRODUCTION

1.1 DIET

Feeding is the set of actions during which we provide food to our body (1). It is a complex process as the selection and availability of meals depend on geography, culture, religion, socioeconomic situation and psychological aspects (2). Our daily consumed foods contain thousands of specific chemicals components digested and absorbed by our bodies to maintain a proper function. These food components can be classified (3,4) as detailed in Figure 1.

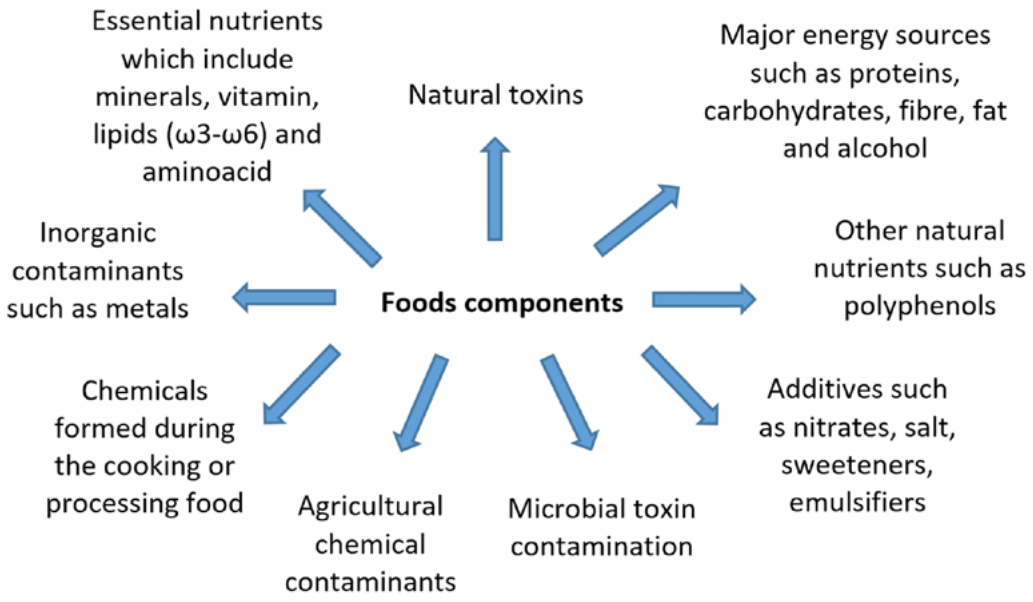


Figure 1. Chemical food component categories (that are not mutually exclusive). Adapted from (1,4).

Food intake is difficult to measure as it varies from day to day. Therefore, in most dietary assessment studies, habitual intake is measured, which corresponds to the average dietary intake of an individual over a relatively long period of time (5).

Although there are broad interests in measuring human food consumption (6), dietary intake is a complex human exposure. It is multidimensional since people consume multiple

foods, macronutrients, micronutrients, phytochemicals, among others, in different combinations. Moreover, it is dynamic since it varies over time depending on the life cycle stage we find ourselves (7).

Assessment of nutritional status is important at individual as well as group (population) level (8,9).

The individual-level assessment reflects the habitual intake, allows classifying individuals according to their intake level within the population. This information may show great dietary adequacy or insufficiency, which could translate into dietary advice, prescription of a therapeutic diet or evaluation of the nutritional intervention.

The group-level assessment reflects the frequency and distribution of dietary intake and estimates the average consumption of foods and nutrients of a population. The results provide the nutritional status of the group. The assessment can explore the relationship of a diet with various health indices or can attempt to identify the intake of desirable nutrients for health status improvement.

To evaluate the habitual diet of a population, several aspects should be taken into considerations (5):

- The dimensions or "layers" of food that are of interest (e.g. the dimensions of food could be based on its content on vitamins, minerals, macronutrients, chemicals (such as types of polyphenols), contaminants, or its grouping in different food categories, chemical (such as types of polyphenols), etc.) (7).

- Assess how close these estimated values are to the true habitual intake of the study population.

- Evaluate the degree to which the data provided by the dietary assessment tool correspond to the food dietary exposure of interest. (e.g. if the objective is to measure the consumption of vitamin A in the population. Does this instrument measure it? To what degree?).

- The objectives for the use of a specific dietary assessment method (e.g. to obtain an accurate estimate of a total absolute intake of a single dietary component, such as fat or iron, a method with properties fitted of capturing the frequency of consumption and amounts consumed of all foods contributing to the dietary component of interest) (5).

1.1.1 DIETARY ASSESSMENT METHODS

The dietary assessment evaluates food and nutrients intake and dietary patterns. It is one of the four nutritional evaluation approaches alongside anthropometric, biochemical parameters, and clinical examination (9).

The direct dietary assessment collects information at the personal level, classified according to different criteria, including the time frame of the data collected (retrospective or prospective) and duration of recorded time (one year, one month, etc.) and others (10,11). There are currently several methods to evaluate dietary intake. However, the choice of one of them may depend on many factors such as environment, population, age, literacy, interviewer training and experience, cost, choice of nutrients and the study's objective. In the following section, we will describe the most used assessment methods.

1.1.1.1 Food diary or food record methods

The food diary is an open method where the participants self-report daily the food and beverages consumed during a specific time to minimise memory dependence. Each respondent must be trained in the level of detail required to describe the food adequately; the portion sizes consumed, the name of the food (brand name), culinary techniques (cooking type) and food mix recipes to complete the food record (12).

The portion size of foods can be assessed by estimating volumes (visual atlas, food models or food packages) or weighing foods before and after consumption. The Weighed Food Record is the most accurate method at the individual level. It has been one of the most widely used methods in the United Kingdom (UK) and Europe to evaluate habitual food and nutrient intake (9,10).

1.1.1.2 Dietary history

Burke developed the dietary story in 1947, it evaluates in detail the habitual intake and eating patterns over a long or short time period. This approach includes a 24-hour dietary recall (24HR), a 3-day dietary recall, and a checklist of foods consumed during a specific period. Highly trained personnel are usually needed for its administration, being impractical for population studies. The checklist of food was the forerunner of the more structured dietary questionnaires in use today (e.g. food frequency questionnaires (FFQs)) (4,10,12).

1.1.1.3 24-hour dietary recall

The 24HR method is an open method in which personnel trained in interview techniques asks respondents to recall and report all foods and beverages consumed during the previous

24 hours. The 24-hour period starts with the first thing eaten by the respondent in the morning until the last food item intake before he/she get up the following day (10,12). It measures habitual intake; however, a single record is not enough to measure individuals' food and nutrient intake. Therefore, a minimum of 2-3 non-consecutive days, including a weekend day, is usually required to capture the day's variability (13,14). Additionally, multi-day data collection is often used to validate an FFQ (10).

When there is more than one interviewer, it is necessary to carry out a standardised procedure and structured interviews to help collect the dietary intake information, such as the automated multi-step method to lower interviewer bias (14–16). The interview can be administered face-to-face, by telephone or via self-administered computer-assisted. To help describe serving sizes, photo atlases, food models, or standard household measures could be used. The food data can be recorded in an open or structured 24HR (9–11).

Once the information has been collected, the interviewer code the raw or cooked food (depending on the aim of the study and the nutritional composition table used) and convert the data into grams for further analysis (11).

1.1.1.4 Food frequency questionnaire

FFQ emerged from the dietary history food intake checklist. It measures the frequency of consuming certain foods or food groups at a specific time (week, month, year, etc.). It evaluates the diet in the long term and can be self-administered, reducing the costs and burden of the participants compared to other methods. Therefore, it is often considered the most suitable method for nutritional epidemiological studies (10,12,17).

The FFQ can be adapted from an existing questionnaire or developed from some basic principles. Once created, the FFQ requires validation.

The questionnaire includes a list of closed foods and a category of consumption frequency according to the study's objectives. Sometimes, it can contain information on the size of the portions, allowing to estimate the amounts of food and nutrients intake (9).

- *Food list*

An FFQ is constituted by a food list. To set up a food list, preliminary considerations should be taken into account (17–19) as follows:

- The food must be used reasonably often by an appreciable number of individuals.
- The selected foods must contain the nutrient(s) of interest for the study.
- To be discriminatory, the use of the foods must vary from person to person.

Generally, the food list of an FFQ may include between 5 and 350 foods items; there are several approaches to create the list: simple selection, selection by exclusion, selection by open data and food selection through a bibliographic review (10,17):

- *Simple selection:* examining the nutritional composition tables and identifying the foods with the nutrients of interest, with the risk of selecting foods with low consumption.
- *Selection by exclusion:* starting a list of potentially essential foods for the study, but some of them could be, later on, eliminated based on the results of a pilot study. However, before removing a food item, its nutritional value to the study's objectives should be considered. For example, beef liver should not be eliminated if the aim is to study vitamin A intake.
- *Selection by open data*: identifying and selecting foods that contribute significantly to the nutrient intake of the study population using a 24HR.
- *Food selection through a bibliographic review:* searching for foods of interest and using data from national food consumption surveys carried out in the population of interest.

● *Frequency categories*

Frequency categories should always be continuous, with no gaps, as the sensitivity of the questionnaire could be reduced, and respondents could be frustrated if they do not find their response. The number of choices are between 1 and 12 but may mainly depend on the intended use of the questionnaire (19).

The range of frequency choices should reflect the time frame of interest. The frequency categories should emphasise the extreme consumption frequency of the distribution for most foods (e.g. number of times per week o per day). However, from foods that are eaten infrequently but make a significant contribution to nutrient intake (e.g. liver), it is important to include a less frequent option, say less than once a month. If there are options of more than once a day, this tends to lead to gross overestimates for some people (19).

● *Portion size*

The estimation of the portion size of each food item of the questionnaire allows us to classify the FFQ into three categories:

- Qualitative: the amount consumed is not determined.
- Semi-quantitative: the standard portion size is given to each item of the FFQ, or an option is provided to choose between different sizes.
- Quantitative: the usual portion size is determined by employing homemade measurements or photographic food albums (11,12).

The FFQ must be associated with a food composition database to estimate the energy and nutrient intake from the semi-quantitative and quantitative questionnaires (10). Currently, several food composition tables are available and will be explained in section 1.2

1.1.1.5 Short dietary assessment methods or screeners

Short questionnaires (or screeners) are often used to measure the intake of certain food groups (vegetables, fruits) or dietary patterns. They are generally used as a last resource if the burden on participants and investigators is too high compared to other methods such as the 24HR. Extracted nutritional data are limited, mainly due to the incapacity to assess energy intake. The food list of the questionnaires varies between 5 and 40 foods, and it can either be administered face-to-face or as telephone interviews, or it can be self-administered (11,12,14). One example is the Mediterranean Diet Adherence Screeners (MEDAS) that contains 14 questions, validated in Spain and recently in the UK (20,21). The Healthy Eating Index (HEI) of 13 items assesses the quality of diet based on adherence to the Dietary Guidelines recommendations for the American population (22). The "Indice de alimentación saludable" (IASE) for the Spanish population is based on the HEI methodology and the recommendations of the Dietary Guidelines (23) and the Healthy Food Diversity Index (HFD-index) that captures dietary diversity (24).

1.1.1.6 Integration of new technologies to improve dietary assessment methods

Technological development has enabled the creation of new tools to collect and process dietary intake. There are currently six main groups of innovative technologies that are promising to improve, complement or replace traditional dietary assessment methods (12,25):

- Personal digital assistant technologies (McClung *et al.* (26), Fukuo *et al.* (27)).
- Mobile phone-based technologies (Shang *et al.* (28), Ruiz *et al.* (29)).
- Camera- and tape recorder-based technologies (Higgins *et al.* (30), Dahl Lassen *et al.* (31)).
- Interactive computer-based technologies (Baranowski *et al.* (32), Vereecken *et al.* (33)).
- Web-based technologies (Subar *et al.* (34)).
- Scan-and sensor-based technologies (Makhsous *et al.* (35), Jia *et al.* (36)).

1.1.2 FOOD COMPOSITION DATABASE

Databases or food composition tables (FCDB/FCT) are used to convert the data obtained from the dietary assessment into energy and nutrients.

The nutritional composition of foods could be highly variable since it may be affected by environmental factors (insolation, type of soil, water regime to grow fruits and vegetables), genetics and culinary processes (8). Therefore, the inclusion of novel foods into the FCDB/FCT is essential to account for changes in the population dietary behaviour (10).

At an international level, there are databases with a large number of foods and nutrients, such as those from the US (USDA National Nutrient Database for Standard Reference) (37), France (Table of the nutritional composition of foods (CIQUAL)) (38) and Australia (Australian Food, Supplement and Nutrient Database (AUSNUT)) (39).

In Spain, there are 19 nutritional programs based on European or Spanish FCDB (40). The Spanish FCDB have several limitations:

- They include a relatively limited number of foods:

Food composition database	Number of foods
International FCDB	
USDA (37)	8,618
CIQUAL (38)	3,185
AUSNUT (39)	5,740
Spanish FCDB	
Farrán et al. 2003 (41)	698
Mataix-Verdú et al. 2009 (42)	1,100
Moreiras et al. 2019 (43)	± 900
Martín Peña et al. 1997 (44)	700
Ortega et al. 2004 (45)	700
BDCA et al. 2010 (46)	950

- They do not include the nutrients or food components necessary for diet and microbiome association studies, such as separating soluble and insoluble fibre, starch, lactose, polyols, polyphenols, etc.
- They do not include post-cooking nutritional values.
- They majority are not updated. New foods such as vegetable drinks, kombucha, kimchi, among others, have not been included.
- Some are not free access.

The choice of an FCDB/FCT may depend on the geolocation of the population of interest as the food components (fibre, vitamins, polyphenols) could vary greatly depending on each country's climatic conditions and production practices (47–49). Therefore, the use of an FCDB/FCT from one country may not be appropriate for another.

1.1.3 METHODOLOGICAL ASPECTS OF THE VALIDATION OF A FOOD FREQUENCY QUESTIONNAIRE (FFQ)

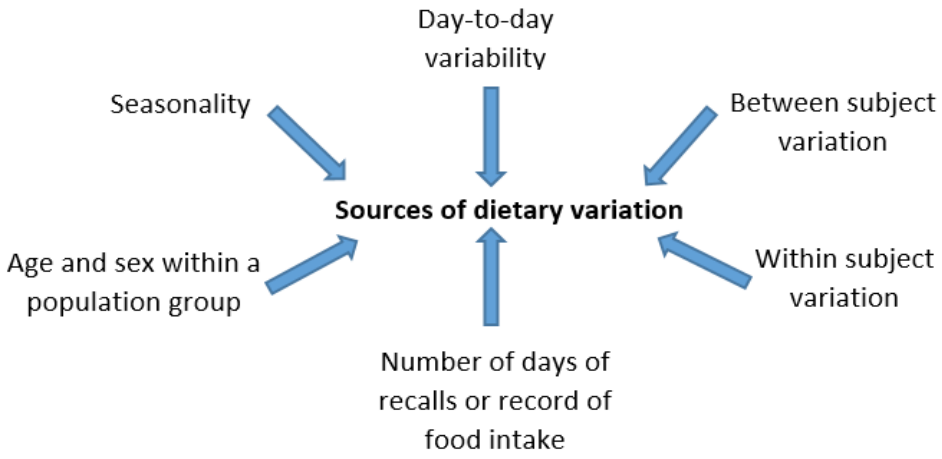


Figure 2. Potential sources of confounding factors in dietary assessment methods.

The usual dietary intake could be assessed through different methods (as we commented previously). However, a series of factors could significantly influence the design, planning, data collection and analysis and should be taken into consideration (Figure 2).

The use of these instruments are facing two key challenges (9,12):

- We need to know how to deal with measurement error (frequent in the data collected with dietary assessing tools).
- We must determine the extent to which a given tool accurately measures true intake. For this, the reproducibility and validity of the dietary assessment instruments should be evaluated.

1.1.3.1 Errors in dietary assessments methods

The food and nutrient intake measurement presents a certain degree of error. The direction and extent of error may vary depending on the method used, the population of interest and the nutrient under investigation. Two types of errors, random error and systematic error (bias), are often encountered (5).

Random error is bi-directional, leads to inaccurate measurements resulting in unreliable conclusions, and can affect the reproducibility of a method. It arises from three sources: individual biological variation, sampling error and measurement error.

Bias is a condition where the results deviate from the true value in a consistent direction. It can reduce the accuracy and validity of a method. It arises from three sources: selection bias, measurement bias, and confounding bias (9).

Several sources of error can significantly affect the estimation of habitual intake and the validity of the dietary assessment methods, but measurement errors and bias can be reduced and corrected (5).

Measurement errors in dietary intake

Measurement errors and biases can be broadly defined as deviations from the true value (true mean, true variation, or both). We can evaluate them by calculating the sample mean and the variation around the mean, expressed by the variance (or standard deviation) (12).

In epidemiological studies, these errors can occur at two different levels, at the individual level (within-person) and the group or population level (between-person). Several measurement errors can be minimised by incorporating different strategies and quality control procedures during the dietary data collection, as detailed in Table 1.

In addition, several statistical techniques have been developed to deal with the errors and improve the performance of diet assessment tools at the data analysis stage. To correct systematic and random errors, linear regression calibration, analysis of variance, and energy adjustment could be used. The energy adjustment improves the correlation coefficient values of dietary intake obtained from the test method (FFQ) and the reference method of the validation studies (24HR). One possible explanation for this adjustment is that energy and nutrient intake errors are correlated and tend to cancel out each other in energy-adjusted nutrient intake. The nutrient density method and the residual nutrients are the most used methods (12,17).

Table 1. Major measurement errors in dietary assessment methods and strategies to reduce them. Adapted from (9).

Measurement errors	Strategies to reduce these errors
Non-response bias The reason why people may refuse to participate in the survey or drop out of the intervention study could be due to a lack of understanding of the study objectives.	Simplify the dietary assessment methods. Carrying out mailed or telephoned reminders in surveillance studies. Training interviewers to convey warmth, understanding and trust.
Respondent's bias This error corresponds to one of the most significant sources of error in dietary assessments and results from over-reporting or under-reporting the food consumed.	Identify misreporting through biomarkers-doubly labelled water, urinary biomarkers, Goldberg cut-off method, among others.
Interviewer bias It may occur if different interviewers probe for information to differing degrees, intentionally omit specific questions, or record subjects' responses incorrectly.	Use standardized 24HR.
Memory lapses of the respondent This may result from unintentional omission or addition of food.	Use multi-pass interview techniques, probing questions, standardised prompts, or memory aids such as food models. Minimising the period of time between actual food intake and recall.
An incorrect estimate of the size of the portion consumed This error can arise from respondents failing to accurately quantify the amount of food consumed or the misconception of the average portion size.	Use household measures, food photographs or food models as a memory aid.
Omission of information on the use of nutritional supplements This omission could lead to errors in the calculation of nutrient intake.	Include close-ended questions about the specific brand taken, the amount per pill, frequency, and duration of use.
Encoding error This error may happen when portion size estimates are converted from household measures into grams and when food items are assigned to codes.	Establish a standardised coding system or use bar-code scanners.
Mistake in the handling of mixed dishes This error may result in incorrect estimates of their nutrient content, per se, as well as errors in their assignment to a specific food group.	Mixed dishes should be broken down into simple ingredients to be classified into their appropriate food groups.

1.1.3.2 Reproducibility of an FFQ

The reproducibility or reliability analysis evaluates the consistency of the repeated measurements of a specific assessment method. The reproducibility can be determined through a test-retest design; for this, an identical FFQ is repeated on the same person under a similar time period, after a defined time interval between repetitions.

The consistency of consumption, not the precision of measurement, is evaluated as dietary intake is rarely totally replicated. Only estimates are obtained as values can be affected by a series of factors such as (9):

- Sample size.
- The time interval between administration.
- Mode of administration (24HR or FFQ).
- The method used to estimate the serving size.
- Seasonality may influence changes in eating habits.
- Nutrient intake that can vary between measurements.
- Random measurement errors.
- Intra and inter-individual variability.

Random measurement errors can be minimised, for example, by increasing the number of repetitions of 24HR. However, reducing intra- and inter-individual variability is not recommended, as variability may be a characteristic of a group of people's true habitual food intake (10).

Statistical methods to evaluate the reproducibility of an FFQ

The reproducibility of the FFQ can be evaluated at the individual and group level, based on the test-retest method, through different statistical tests.

At the individual level, correlation analysis or intraclass correlation coefficient can be used to assess the degree to which two administrations are related. The percentage of misclassification is calculated by comparing the number of pairs (repetition of the same FFQ) of FFQ with a similar agreement.

A paired t-test or a Wilcoxon test can be used at the group level to compare the mean or median intake, respectively. Bland & Altman's analyses can be used to test the degree of correspondence between the mean and the standard deviation of two replicates through limits of agreement, which will be detailed in section 1.3.3.

If there are more than two days of feeding measurements (e.g. 24HR), the variance could be analysed to estimate the variability between- and within- individuals (9,12,19).

1.1.3.3 Validation of an FFQ

Validation determines the degree to which a questionnaire measures what it intends to measure. Methods for assessing dietary intake are usually difficult to validate since there is not a gold standard method has been yet reported; therefore, "relative validation" is often performed. Relative validity is a process in which the dietary values estimated by a test method are compared to those from a reference method (FFQ v/s 24HR). The reference methods must have a greater degree of demonstrated validity and errors independent of the test method; however, they rarely occur (10,12). The results of validation can be affected by systematic errors.

One way to avoid these errors is to use an external marker independent of the intake, such as biomarkers. Nevertheless, they have limited use since they are only available for a limited number of nutrients (50).

When designing validation methods, a series of factors may be taken into account (10):

- The selection of subjects for the validation study should be representative of the population under investigation. Validation can be done on a small sample but must be large enough to estimate the correlation between the reference method and the test with reasonable precision.
- The reference method should have equivalent objectives and measure similar parameters to the test method.
- The test methods should be administered before the reference method in a validation study to simulate the situation chosen as the objective of the study.
- The spacing between the administration of both instruments should be considered so that the test method does not influence the responses of the reference method.
- Physiological characteristics such as sex, age, socioeconomic status, ethnicity, and health status can affect the validation results.

Several combinations are recommended between the test method and the reference method to perform a relative validation of dietary assessment methods (e.g. single 24HR v/s single-day weighted record, dietary history v/s single day weighted record). However, a good agreement between these combinations does not necessarily indicate validity.

Indeed, the agreement could reflect similar errors between both methods. Moreover, the limited validity of a test method can be attributed to errors associated with the reference method (9).

Statistical methods to perform a relative validation of an FFQ

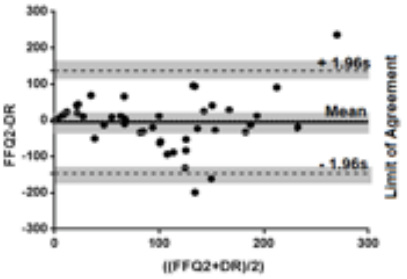
Currently, there are several statistical approaches to the relative validation of an FFQ. However, there is no consensus on the ideal method and the number of statistical tests required to validate a diet assessment tool as it depends on the objectives of the study. The results of the different statistical tests should be interpreted with caution because they reflect different facets of validity, such as concordance, association or bias at the group and individual levels of a particular food or nutrient. In addition, an FFQ could be valid for some foods and nutrients but not for all (17,51).

Assessing the distribution of the data should be a preliminary step before the validity evaluation. A logarithmic transformation could be applied to data that present an abnormal distribution, as well as the use of non-parametric statistical tests. The most commonly used statistical methods to perform a relative validation of an FFQ and interpretation criteria applied to the results are shown in Table 2.

In summary, dietary assessment methods, specifically the FFQ, are considered a relevant tool for obtaining data on habitual intake patterns since its principal objective is to assess diet over long periods of time in large populations (17,52). Indeed, the FFQ is a relatively inexpensive, quick and easy-to-implement method that can provide in-depth insights into food, nutrient intake and dietary patterns (19). However, the FFQ must be validated in the population for which it has been designed before being applied. To validate an FFQ, the reproducibility and validity of its measurements must be verified using different statistical methods (17,51).

Various studies in nutritional epidemiology have shown the association between diet and the risk of suffering from certain chronic non-communicable disorders or cancers (6,53). However, a new actor appears to influence this relationship between diet and disease risk, the microbiome (54).

Table 2. Statistical tests, interpretation criteria and example of the Bland & Altman plot of the relative validation of an FFQ (9,13,51,55).

Stadistical test	Facet of validity reflected	Interpretation criteria		
		Good outcome	Acceptable outcome	Poor Outcome
At group level				
Paired t-test/ Wicoxon test	Agreement To examine if the mean or median of each food or nutrient estimated by the FFQ and reference method are different	P>0.05		P≤ 0.05
Bland Altman analysis: Correlation between mean and mean difference	Presence, direction and extent of bias (proportional bias)	P>0.05		P≤ 0.05
Bland Altman plot	Illustrates the magnitude of disagreement, identifies outliers and trends in bias between the FFQ and 24-DR. It also reflects the overestimation and underestimation of the estimates.			
At individual level				
Correlation coeffcient Pearson/Spearman	Strenght and direction of association bethween two methods	≥ 0.50	0.20-0.49	<0.20
Intraclasss correlation	Takes into account the degree of correlation and the extent of the disagreement within pairs	≥ 0.4		
Deatennuation coeffiicient	It is used to adjust the correlation coefficients when there is a variation from day to day (intra-individual variation). For example, when we use multiple 24h-DR			
Percent difference	Agreement (size and direction of error		0.0 - 10.0%	> 10%
Croos-classification (tertiles/quartiles or quintiles)	Agreement (including chance) - In same tertile - In opposite tertile	≥ 50% in same tertile/quartile ≤ 10% in opposite tertile/quartile		< 50% in same tertile/quartile > 10% in opposite tertile/quartile
Weighted Kappa statistics (coefficients)	Agreement (excluding chance)	≥ 0.61	0.20 - 0.60	< 0.20

*Bland & Altman plot: The Bland & Altman plot is showing the difference between the measurement (test - reference measure) (y-axis) against the mean of the two measures (test measure + reference measure/2) (x-axis) for each subject. The limit of agreement (95% confidence limits of the normal distribution) is calculated as the mean difference ±1.96 SD and reflect over and underestimation of estimates.

1.2 MICROBIOME

Humans have co-evolved with millions of microorganisms (bacteria, fungi, viruses, archaea and protozoa) to establish a complex and mutually beneficial relationship. In the gastrointestinal tract (56), almost 10 million non-redundant bacterial genes have been identified (57). The microbiota is estimated to outnumber host cells by a ratio of 10:1 (58). However, they have recently estimated that the ratio is much closer to 1:1 (59).

1.2.1 TECHNICAL APPROACHES FOR MICROBIOME ANALYSIS

Recent advances in high-throughput sequencing have expanded our understanding of the gut microbiome, particularly prokaryotes belonging to bacterial and archaeal domains. The most common methods of microbiome characterisation are targeted sequencing and shotgun sequencing on microbial DNA or RNA. Targeted sequencing approaches focus on a specific microbial gene used to describe the taxonomic composition of the samples. In contrast, shotgun sequencing focuses on the whole extracted genomic DNA. This approach allows identifying gut microbial potential functions; however, not all these genes are active at sampling. Other omics approaches have been developed to overcome these limitations. Metatranscriptomic, metaproteomics and metabolomics are approaches based on RNA shotgun sequencing to identify active genes, overall proteins and metabolites seeking to detect metabolic activities, respectively (60).

Targeted-sequencing methods involve the amplification of marker genes by Polymerase Chain Reaction (PCR). PCR is performed at an orthologous locus (marker gene) to differentiate microbial community members depending on the analysed microorganisms. The most commonly analysed locus is the 16S ribosomal RNA (16S rRNA) gene containing several conserved regions, each flanked by nine hyper-variable areas V1 to V9 that potentially differ between species. Conserved areas allow the use of universal primers to detect all bacteria present in a sample, whereas hypervariable regions determine which microorganism is being detected (61,62).

Overall, V4 is the most analysed hypervariable region that allows microbial identification down to the genus level in biological samples. Besides, their short length (254 base pairs approximately) is suitable for small-length read technologies such as the HiSeq and the MiniSeq Illumina platforms, which uses paired-end reads that can be joined into ~300 nucleotides (63,64).

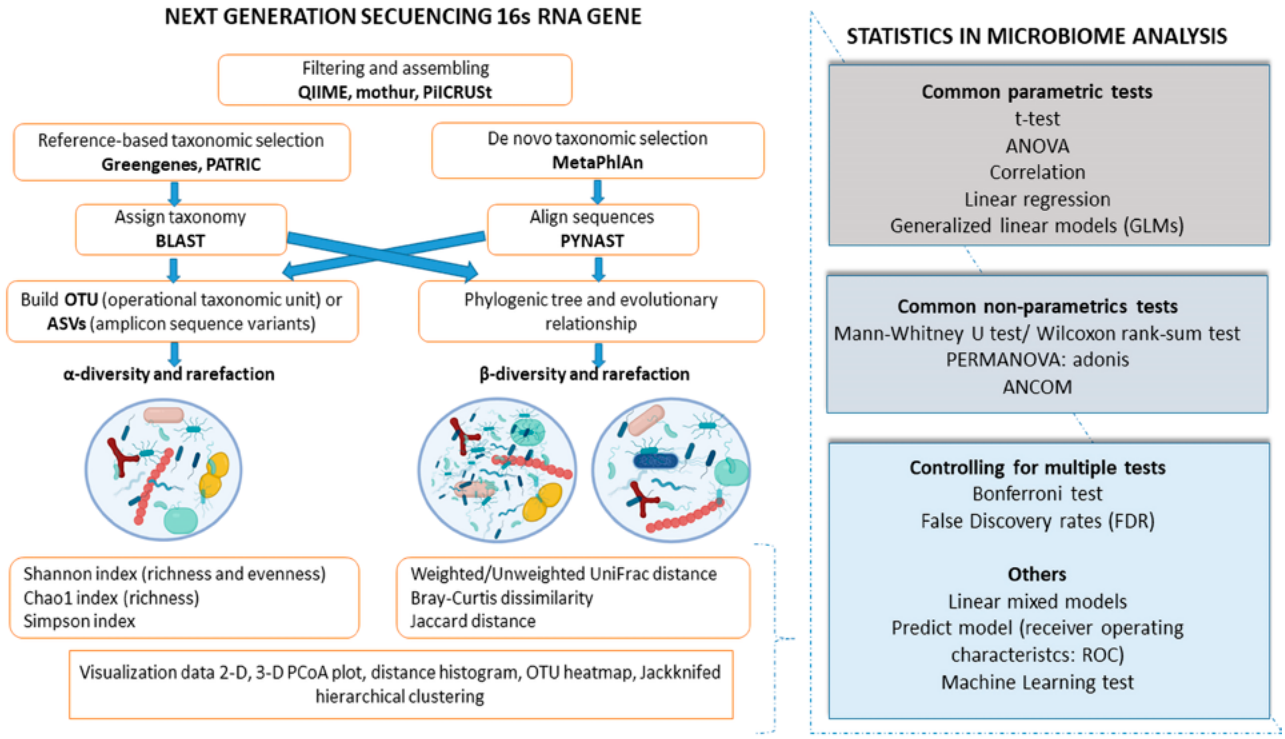


Figure 3. Summary of bioinformatics workflow and statistical analysis of the microbiome studies. Adapted from (65,66).

High-throughput targeted sequencing results in millions of sequences that correspond to bacteria and archaea in microbial communities. The 16S rDNA sequences could be processed using several available bioinformatic pipelines and public databases to recover taxonomic profiles.

The raw sequence reads obtained by the Illumina platform are filtered and processed using bioinformatics pipelines such as QIIME (67,68) (Quantitative Insights Into Microbial Ecology), mothur (69), UPARSE (70), DADA2 (71), etc. Most of these tools use public databases such as Greengenes (72), Silva (73), EzBioCloud (74) and PATRIC (75) and are based on clustering raw sequences into operational taxonomic units (OTUs) following different clustering algorithms. They classify each OTU into a bacterial taxon down to the species level. Recently developed tools such as DADA2, implemented in QIIME2, denoise, dereplicate and filter out chimaeras sequences into amplicon sequence variants (ASVs).

The OTU and ASVs data generate bacteria abundances tables used to perform taxonomic classification, alpha- and beta-diversity, and differential abundance analyses. The summary of bioinformatics workflow and statistical analysis of the microbiome studies by 16S RNA is shown in Figure 3.

1.2.2 GUT MICROBIOME PROFILING

Prokaryotes can be classified into different phylogenetic levels: phyla, class, order, family, genus and species. 90% of the gut microbiota is composed of the phyla Firmicutes and Bacteroidetes followed in minor proportions by Proteobacteria, Actinobacteria, Fusobacteria and Verrucomicrobia phyla. (76,77).

The human gut harbours a diverse community of microbes that exhibit significant between-person variations (56,78). 1200 species of bacteria have been detected in most human samples, of which 160 to 400 species are estimated to present in each individual (76,79). The intestinal microbiome has relatively slow temporal dynamics compared to the oral or cutaneous microbiome. Possibly reflecting trends in response to long-term factors such as dietary patterns (80).

Also, the gut microbiome has been shown to have high temporal stability. It conforms to a power-law function where significant differences in community composition occur on shorter time scales while a stable core set of strains persists at longer scales. Higher stability is typically observed in Actinobacteria and Bacteroidetes phyla compared to Proteobacteria and Firmicutes phyla (81).

In the gastrointestinal tract, the density and composition of the microbiota vary along the longitudinal and transverse axes. The difference in the chemical-nutritional gradient and the compartmentalised immunological activity of the host increases the number and diversity of species from the stomach to the colon (82,83). The colon is the most colonised area since it contains 70% of the microorganisms that make up our body and has the highest metabolic activity (59,84). The close symbiotic relationship between the host and the gut microbiota leads to various functions that help support and maintain the host's health. Among the main tasks performed by the intestinal microbiota is modulation of the immune system and interaction with the enteric nervous system, protection against colonisation by pathogenic organisms, and metabolic activities (84,85).

One of the most important metabolic functions of the gut microbiota is to ferment the components of the diet not digested by the small intestine. In addition, its proteolytic potential converts ingested dietary and endogenous proteins (e.g. host enzymes, mucin and sloughed off intestinal cell content) into several metabolic substrates. As a result, critical metabolic substrates, such as short-chain fatty acids (SCFA), amino acids and derivatives,

phenolic compounds, and gases including H_2, CO_2 and H_2S are produced that contribute to the host´s health. (86,87).

SCFA have different functions (88):

- Butyrate is the primary energy source for the colonocytes as it provides 70% of the ATP production in the colon. It has anticancer activity by inducing apoptosis of malignant cells and regulating the expression of genes involved in the inhibition of histone deacetylation. It also activates intestinal gluconeogenesis.
- Propionate inhibits endogenous cholesterol synthesis and regulates gluconeogenesis in the liver.
- Acetate is the most abundant SCFA; it is essential for the growth of other bacteria and is used as a substrate for the synthesise of lipids and cholesterol at the peripheral level (87,89).

Other functions of butyrate and propionate are to serve as ligands regulators of the excretion of enteroendocrine hormones, that control or modulate appetite and food intake, as shown in mice (90).

Likewise, the intestinal microbiota participates in other metabolic functions, supports the absorption of Ca^{+2}, Mg^{+2} and Fe^{+2} ions, synthesises vitamin K and other B-complex vitamins, and participates in the bile acids catabolism (86,91).

The close functional relationship between humans and their microbiome is essential to maintain proper homeostasis. However, several intrinsic and extrinsic factors can negatively affect microbial composition and activity, triggering alterations in the microbial community, namely dysbiosis. A decrease in microbial diversity has been reported in low and high-grade inflamed mucosa that has been associated with a vast number of diseases (92), as shown in Figure 4.

1.2.3 FACTORS AFFECTING THE GUT MICROBIOME

Many studies have shown that host genetics can impact the structure and composition of the gut microbiome (93,94). For instance, the association of nucleotide-binding oligomerisation domain-containing protein 2 (NOD2) mutations with some inflammatory pathologies, including Crohn's disease, have been shown to be associated with immune response and microbiome alterations (95,96). Nevertheless, the combination of multiple extrinsic factors such as mode of birth, frequency of drug use such as antibiotics, lifestyle, age, environment and especially diet have a significant impact on the microbial community during our lives (97–99).

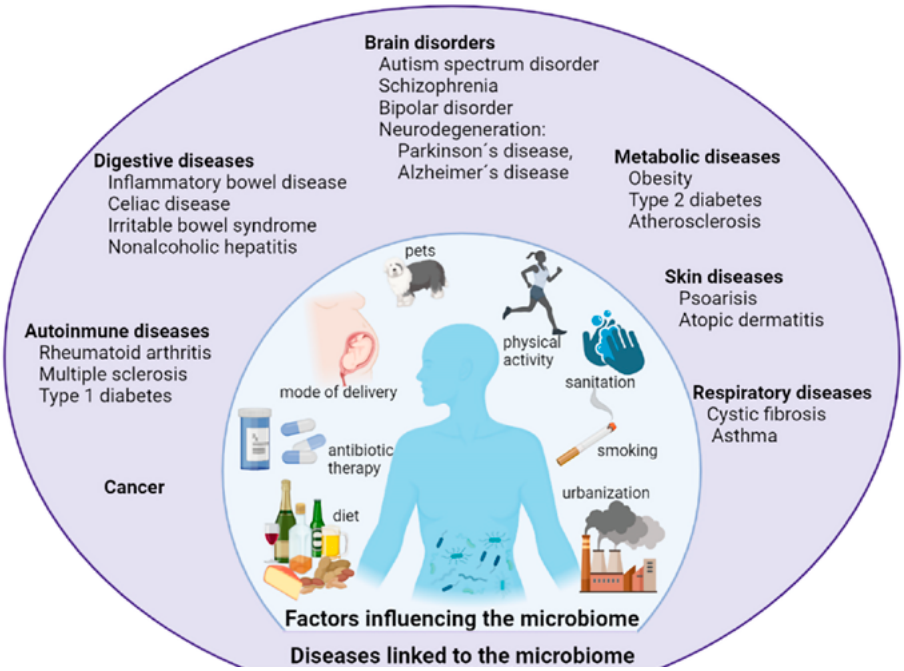


Figure 4. Factors that influence the gut microbiome composition and the association of dysbiosis in developing diseases in the host. Adapted from (100). Created with BioRender.com

1.2.3.1 Age, delivery and type of feeding

The age of the host has been shown to have a significant effect on the gut microbiota. Colonisation begins when we are born and varies depending on the mode of delivery. Infants born vaginally are shown to be mostly colonised by the mother's vaginal and faecal microbiota. In contrast, those babies born by caesarean section (C-section) are often inhabited by the mother's skin microbiota. The lack of inoculation of bacteria from the birth canal in children born by C-section could negatively affect the development of their intestinal microbiota since it increases the susceptibility to specific pathogens and the risk of developing certain diseases such as atopy allergy and asthma (101,102). Colonisation may be chaotic during the first months of life since the microbiome composition could be affected by several environmental factors such as diseases, antibiotic intake, and feeding (103,104). Various studies have observed that infants fed with breast milk have a different microbiota composition than those fed formula-fed infants. Breast milk provides several substrates for bacterial function and offers certain bioactive compounds that help digestion and absorption of nutrients, immune protection, and antimicrobial defence, absent in formula milk (104,105). When solid feeding begins, the most significant compositional change in the infant's gut microbiome occurs.

1.2.3.2 Lifestyle

New evidence suggests that physical activity can modify the composition and functional activity of the microbiota in mice (106–108) and humans (109). An increase in the richness, microbial diversity and more significant presence of *Faecalibacterium prausnitzii* has been observed in rugby players (110). Individuals who perform high-intensity cardiorespiratory exercises present greater microbial diversity than those who do them at a lower intensity (111). However, it is unclear what type of physical activity is the most appropriate since most studies have been done on mice. Human studies have been inconsistent; the modality, intensity, and duration of exercise necessary to produce a significant effect on the gut microbiota have yet to be determined (112).

Smoking also has shown an impact on the intestinal microbiota. Various studies indicated an increase in the relative abundance of Firmicutes and Actinobacteria and a reduction of Bacteroidetes and Proteobacteria in stool samples from individuals undergoing smoking cessation (113,114). However, more evidence is necessary as there are few studies.

1.2.3.3 Antibiotic therapy

The use of antibiotics (115,116) and non-antibiotic drugs (117) may have harmful effects on the microbiome. Several studies showed that antibiotics alter the microbiome composition by increasing, decreasing or eliminating certain species and producing an imbalance between Firmicutes and Bacteroidetes. This alteration depends on the class of antibiotic, the dose, the period of exposure, the pharmacological action and the bacterial target (118). After antibiotics intake, the recovery of microbial diversity can take months or years, as it varies between individuals (119).

In a recent study, researchers screened more than 1,000 marketed non-antibiotic drugs against 40 representative gut bacterial strains. They found that 24% of the drugs with human targets, including members of all therapeutic classes, inhibited the growth of at least one strain *in vitro*. Therefore, these pharmaceuticals could contribute to a decrease in microbiome diversity in modern Western societies (120).

1.2.3.4 Environmental factors

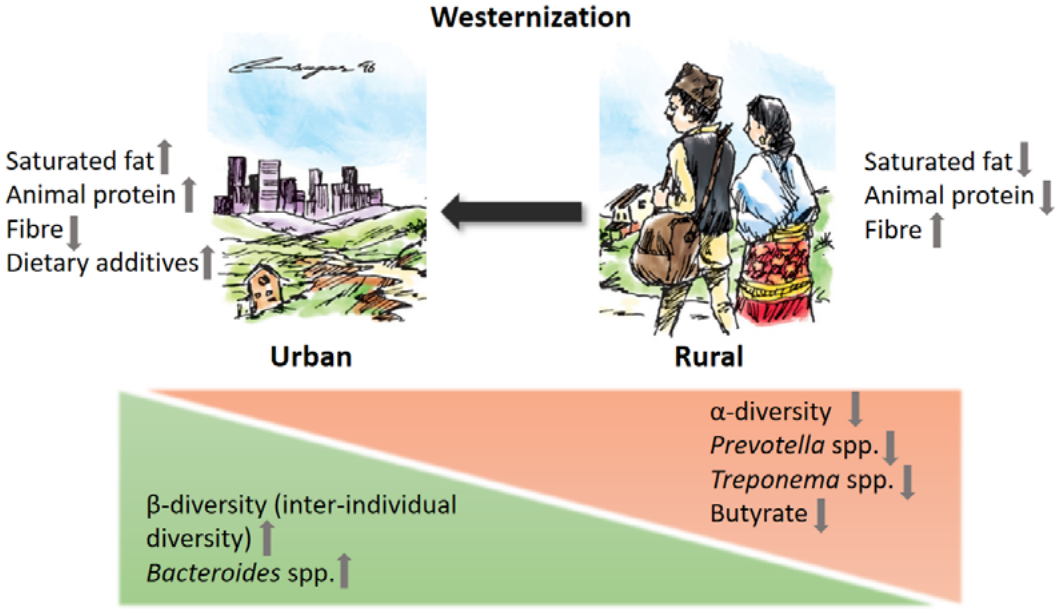


Figure 5. Changes in dietary patterns following westernisation, accompanied by alterations in dietary components, result in remarkable changes in the gut microbiome composition and function. Adapted from (121).

Several studies have reported changes in the taxonomic composition of faecal samples from different ethnicities and geographic areas (Figure 5). De Filippo and colleagues (99) compared children's intestinal microbiome from a rural village in Burkina Faso and European children. They observed a higher relative abundance of Bacteroidetes (associated with a rural diet) in African children and a higher proportion of Firmicutes (associated with a diet high in protein and low in fibre) in European children. Yatzunenko and colleagues (98) compared the bacterial communities between the American population and two rural villages in Venezuela and Malawi, and they observed a clear difference in faecal microbiota composition. One of the possible causes is that the American diet (a diet rich in protein) differs significantly from the diet of Venezuela and Malawi (a diet rich in corn and cassava).

Additionally, differences in cultural tradition, exposure to pets and livestock, sanitation, hygiene, and climate could also influence how and where the gut microbiome is acquired. However, the diet has a dominant role over these variables (99). Notably, urban areas offer a wide variety of foods, which could lead to more significant inter-individual variability of the gut community than the rural areas' more homogeneous and simple diets (121).

1.3 DIET AND MICROBIOME

Diet constitutes one of the principal determinants in establishing the cross-talk between the intestinal microbiome and host genes (122). Multiple studies have suggested that diet is the central modulator of microbial communities in humans and other species (123). Indeed, diet would explain 57% of the microbial variation in mice and around 20% in humans (along with drugs and anthropometric measurements), which makes the diet a potential therapeutic target in the management of diseases through the modulation of the gut microbiome (124,125).

Short- and long-term dietary intervention studies have demonstrated the ability to alter microbial diversity in humans rapidly; however, these changes may be transitory since the microbial composition returns to a basal state (126–128). Even after extensive dietary changes, an individual's microbiota maintains its unique personalised composition, suggesting that the forces that control the ecological homeostasis of the microbiome extend far beyond diet (129,130).

Although the intestinal microbiome has high plasticity in terms of composition, its function in an individual shows high stability against external and internal disturbances, suggesting that microbial communities are resilient and resistant to changes (92,131). However, more permanent modulation of the microbiome community has been shown possible in several cases. For instance, post-infectious gastroenteritis (induced by bacteria, parasites or viruses), which stimulated aberrant immune responses and was associated with an alteration of the microbiome composition, could increase the risk of developing irritable bowel syndrome (IBS) (132,133). A study carried out in germ-free mice, which underwent a faecal transplant of twins discordant for obesity, showed the transmission of the lean and obese phenotype (134). Cohabitation experiments showed that "lean microbiota" could be successfully transmitted into mice with "obese microbiota" and could partially prevent obesity. However, this was only the case while the mice were consuming low-fat / high-fibre diets. Once the animals were exposed to high-fat, low-fibre diets, all mice experienced an increase in body mass and fat mass, and co-housing of the lean and obese mice failed to attenuate or block the development of obesity (134). These findings emphasised the importance of diet in developing obesity and the close relationship between diet and microbiota composition (135).

Habitual dietary patterns that acts as continuous external disturbances with a constant level maintained over a period of time (136), could drive the establishment of stable, dominant microbial networks in the gut microbiome of each individual (129, 130, 137). Therefore, we expect that a permanent change of diet could induce the proliferation of very low previously undetectable species, increasing microbial diversity and potentially beneficial taxa (130).

Before evaluating the degree to which a dietary intervention could modify the microbial communities, the effect of the different dietary habits (Western-type, vegan and vegetarian, Mediterranean diet, etc.) of the population in the gut microbiome should be investigated. The lesson learned from the impact of diet on the gut microbiota will provide support to carry out nutritional interventions that could be individualised and sustainable over time.

Population-level studies have observed that the habitual diet represents only a tiny proportion of microbiome variation (125,138,139). Insignificant and modest differences have been found between groups of people who consume very different dietary patterns, such as omnivores v/s vegetarians (140) and omnivores, vegetarians and vegans with three different levels of adherence to the Mediterranean diet (141).

However, Johnson *et al.* evaluated the habitual diet and found that the interactions between food and microbes are highly personalised. This personalised response might be one of the reasons how diet has been shown to have a small effect in shaping the gut microbiome in population-level studies; therefore, the impact of diet could be much more significant than these studies suggest (142). Indeed, if the same foods impact different bacterial populations in an individualised manner, such effects may not be detectable in cross-sectional investigations. In addition, Johnson and colleagues observed that the microbial composition was more strongly associated with food choice than with conventional nutrient profiles. Therefore, evaluating diet using the conventional nutrient approach could be insufficient to connect dietary intake with microbiome variation. Unfortunately, the limited available information on food composition (as discussed above, section 1.2) relevant to the microbiota (additives, preservatives, polyphenols, etc.) makes it challenging to detect correlations between microbial functions and nutrient intake (142,143).

Another factor recently explored is the impact of seasonality on the habitual diet and the microbiome. In animal models, it has been reported that seasonality probably affects the composition and function of the gut microbiome in response to seasonal changes in diet (144,145). Smits *et al.* collected faecal samples from the Hadza tribe of Tanzania in a longitudinal manner during one year (146). They showed that the microbiota of these hunter-gatherers reflects the seasonality of their diet, where the dry and wet seasons affect the availability of certain types of food. Also, they observed a considerable shift in the intestinal microbiota composition with cyclical characteristics between seasons, with several taxa undetectable in one season and reappeared in the next.

Other studies have revealed that the gut microbiota can be affected by circadian rhythms, as they fluctuate depending on the availability of nutrients (147). The feeding regimen could also influence gut microbial composition and function. Kaczmarek *et al.* observed in humans that several bacteria were related to the timing of eating and overnight-fast duration (148). At the same time, Collado *et al.* discovered that the timing of the main meal could affect the diurnal rhythms of the microbial profile in ten individuals. They found that eating late (17:30hr) increases diversity in saliva, specifically pro-inflammatory microbial taxa, but has no effect on the faecal composition (149).

We currently understand that diet widely affects the microbial community composition within an individual (150). However, identifying the exact importance of specific foods and nutrients in shaping the microbiome composition across populations remains a challenge (142). Therefore, building new diet assessment tools should improve our understanding of these close relationships.

Next, the current state of knowledge will describe some food components and dietary habits that may impact the gut microbiome signature.

1.3.1 FOOD COMPONENTS AND GUT MICROBIOME

1.3.1.1 Carbohydrates and gut microbiome

Carbohydrates (CHOs) are separated into two classes, digestible or indigestible. Digestible carbohydrates such as glucose, fructose, galactose (present in sugar, honey, fruit, pastries) and starch (present in cereals and derivatives and potatoes) are enzymatically degraded in the small intestine. In contrast, the indigestible carbohydrates are resistant to enzymatic degradation and are released into the large intestine (151).

Recent human studies indicate that dietary mono and disaccharides affect intestinal microbiota and increase *Bifidobacteria* while decreasing *Bacteroides* (152,153). The intake of polyols, another type of digestible carbohydrates, produces a similar effect (154).

Indigestible carbohydrates, known as dietary fibre, are fermented by gut microbes in the large intestine. However, not all dietary fibre is fermentable (Figure 6). Several clinical trials showed that a high fibre dietary intervention increases the relative abundance of beneficial bacteria such as *Bifidobacterium* spp., *Lactobacillus* spp., *Akkermansia* spp., *Faecalibacterium* spp., *Roseburia* spp., *Bacteroides* spp. and *Prevotella* spp. (155). In addition to reducing the Firmicutes/Bacteroidetes ratio (156) and improving microbial diversity (99,157,158). Compositional differences between various types of resistant starch have been found to have differential effects on the host microbiota (159). The consumption of 100g/day of resistant starch type 2 and type 4 for three weeks in healthy adults is associated with an increase in *Bifidobacterium* spp., *Faecalibacterium* spp., *Eubacterium* spp. and *Ruminococcus* spp. (160).

Fermentable non-digestible carbohydrates are considered prebiotic since they are non-digestible dietary components with beneficial effects on the host's health. They selectively stimulate the growth and activity of certain microorganisms (161), in addition to being part of the "microbiota accessible carbohydrates", or MAC (162), a term used to describe the CHOs that are metabolically available to gut microbes.

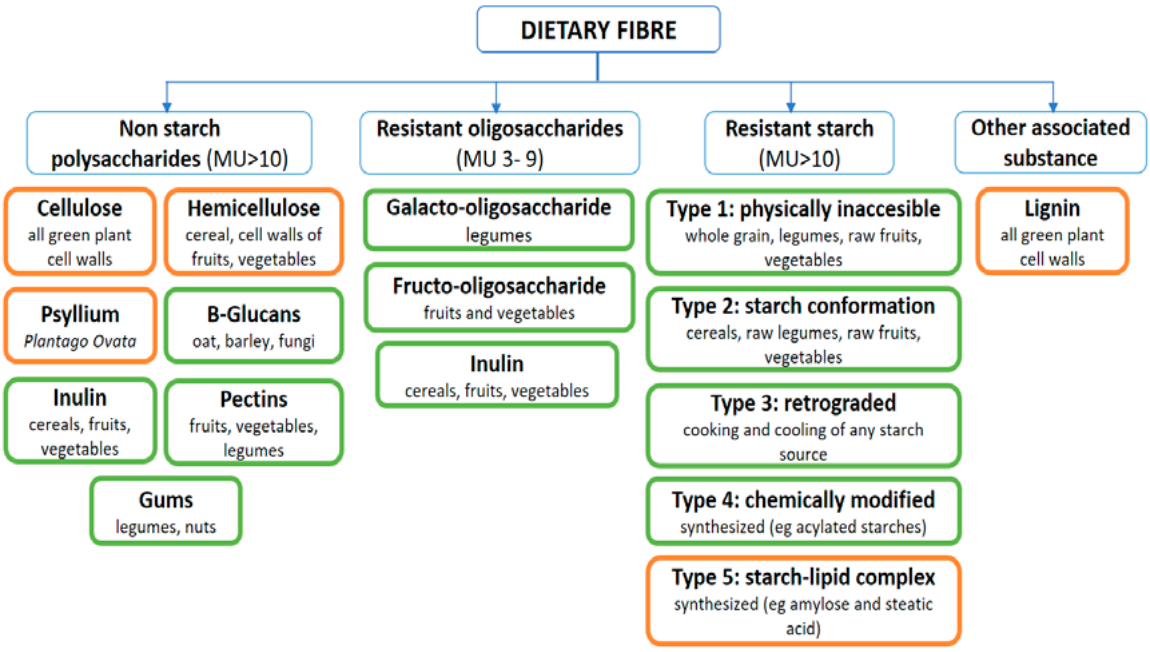


Figure 6. Scheme of fibre subtypes and their food sources. The red rectangular represents fibre with low fermentability, and the green rectangular represents fibre with high fermentability. Abbreviations: MU, monomeric units. Adapted from (159,163).

1.3.1.2 Fat and gut microbiome

The quantity and quality of fats in the diet may severely affect the composition and function of the gut microbiome, which in turn influences the host's metabolism. High-fat diets (HF-diet) in animal models decrease Bacteroidetes and increase Firmicutes and Proteobacteria (164,165). In particular, saturated fatty acid (SFAs) could induce dysbiosis and the consequent alterations of the intestinal barrier (151). In humans, an HF-diet, mainly SFAs, is associated with a reduction in microbiome richness and diversity (166). An interventional study showed that an HF-diet in healthy adults is related to increases in *Alistipides* and *Bacteriodes* species and decreases in *Faecalibacterium* species (167).

The modulation of the gut microbiome by dietary fats is still unclear since findings are inconsistent. However, a recent systematic review in human studies showed that HF-diet and SFA could exert unfavourable effects on the intestinal microbiota and are associated with an unhealthy metabolic state. The authors of this review reported that diets rich in monounsaturated fatty acids (MUFA) could negatively affect the gut microbiota or metabolic health outcomes, but not polyunsaturated fatty acids (PUFA) (166).

Another study performed in individuals with metabolic syndrome suggests that a low-fat diet led to an increased faecal abundance of *Bifidobacterium*. Inversely, a high total fat/ high SFA diet increased the relative abundance of *Faecalibacterium prausnitzii* compared

to baseline. Finally, subjects with high MUFA intake did not encounter shifts in the relative abundance of any bacterial genera. However, they did have overall reduced bacterial load and plasma total- and LDL-cholesterol (168). Supplementation with omega 3-PUFA (in fatty fish) in healthy subjects reversibly increases the abundance of butyrate-producing bacteria such as *Bifidobacterium*, *Roseburia* and *Lactobacillus*; however, it does not produce significant changes in α or β diversity or the composition at the phylum level (169).

1.3.1.3 Protein and gut microbiome

The effect of proteins on the microbiome composition varies according to the type (vegetable or animal) between individuals (170). A rich animal protein intake (red meat and dairy products) can increase the abundance of bile-tolerant anaerobic bacteria (*Alistipes, Bilophila* and *Bacteroides* species). On the contrary, it decreases levels of Firmicutes (*Roseburia* species, *Eubacterium rectale, Ruminococcoccus bromii*) that metabolise dietary plant polysaccharides (127,171,172). The increased risk of inflammatory bowel disease (IBD) has been associated with prolonged consumption of animal protein through the cumulative production of hydrogen sulfide (H_2S) by sulfate-reducing bacteria (*Desulfovibrio* and *Bilophila* species) (173). However, a significant impact on the abundance of these reducing bacteria has not been observed when consuming a diet rich in sulfur amino acids (methionine, cysteine, taurine) (174). Additionally, increased trimethylamine N-oxide (TMAO) (from L-carnitine (red meat) metabolism by bacteria and the liver) was shown to be associated with the risk of atherosclerosis and could be increasing cardiovascular risk (173).

On the other hand, a high vegetal protein diet intake (based on glycated pea proteins) was shown to increase commensal beneficial bacteria (*Bifidobacterium* and *Lactobacillus)*, which increases the production of SCFA in humans. On the contrary, a decrease in pathogens such as *Bacteroides fragilis* and *Clostridium perfringers* was observed (161,175).

1.3.1.4 Micronutrients and gut microbiome

Various micronutrients such as vitamins taken as dietary supplements has been shown to shape microbial composition and provide critical functions within bacteria (150). Vitamin D could affect intestinal microbiota composition. Vitamin D3 supplementation leads to a shift in intestinal bacterial composition in Crohn´s disease patients but not in healthy controls (176).

Metals can drastically alter the microbiome, as they appear to be necessary cofactors for numerous physiological processes in mammals and bacteria (150). Iron is an essential micronutrient for the growth of pathobionts that could therefore influence the microbiota structure. In fact, in mice, a diet rich in heme was shown to decrease microbial diversity, increase the abundance of Proteobacteria, and reduce the abundance of Firmicutes (177). In developing countries, zinc deficiency could be a risk factor for life-threatening childhood diarrhoea as it increases populations of pathobionts (178).

Salt intake has also been associated with microbiome alteration; recent data suggest that diets high in salt are mediated by reduced levels of *Lactobacillus* and subsequent increases in pro-inflammatory T-helper 17 cells in experimental animals and humans (150,179).

1.3.1.5 Polyphenols and gut microbiome

Dietary polyphenols are bioactive compounds present in various foods (Figure 7). In the last decades, they have been extensively studied for their antioxidant properties. Recently, they have been shown to have a significant impact on the gut microbiome. (180,181). However, it is difficult to determine the true effects of these compounds on human health since absorption and availability remain unclear and controversial. Furthermore, there is a wide interindividual variation in their responses, probably due to differences in the gut microbiome composition (182).

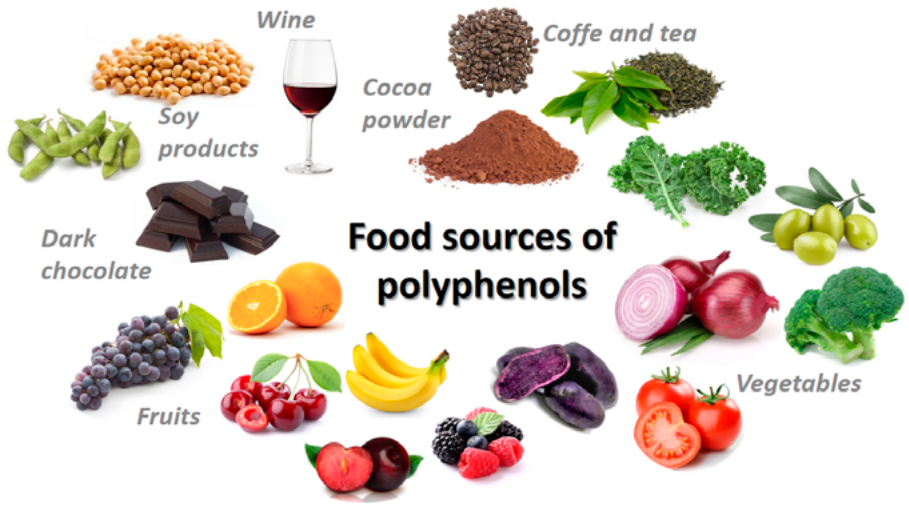


Figure 7. Primary sources of dietary polyphenols (180).

Polyphenols and the gut microbiome have a two-way relationship. Polyphenols are biotransformed into metabolites by the intestinal microbiome and could modulate the microbial community's composition by inhibiting pathobionts and stimulating beneficial bacteria (181). *In vitro*, an enrichment of the *Bifidobacterium* and *Lactobacillus* genera has been observed after using different extracts rich in polyphenols (153,183). In a human intervention study of 22 people, cocoa-flavanols were reported to promote the growth and proliferation of *Bifidobacterium spp*. and *Lactobacillus spp*. (184). As an opposite effect, the use of anthocyanins would also stimulate the growth of *Enterococcus* spp. (185).

1.3.1.6 Dietary additives and gut microbiome

In the last decades, the consumption of processed foods, characteristic of the Western dietary pattern, has increased considerably. Many studies suggest that the harmful effect of this type of diet may be due to food additives (such as sweeteners, emulsifiers, colourants, preservatives and fortifying agents) (150) since they significantly impact the gut microbiota and its gut homeostasis (186,187).

There is evidence that non-caloric artificial sweeteners such as saccharin, sucralose, aspartame, cyclamate, neotame, and acesulfame-potassium cause dysbiosis and disrupt metabolic homeostasis in rodents (122). In humans, only the effect of saccharin and sucralose (154) has been observed. In a small-scale intervention trial in humans, Suez *et al.* (188) showed that saccharin consumption is associated with the induction of glucose intolerance through compositional (relative increase in the Bacteroidales order, Lactobacillales order and decrease in the Clostridial order) and functional alterations of the intestinal microbiota. These metabolic effects are transferable to germ-free mice by faecal microbiota transplantation. However, the response to artificial sweeteners in healthy individuals is personalised, possibly linked to differences in the composition and function of the microbiota (150,188).

On the contrary, investigations on natural sweeteners such as steviol glycosides (extracted from the stevia leaf) have not reported consistent microbial changes in anaerobic faecal cultures taken from healthy human subjects (151).

The supplementation of dietary emulsifiers such as polysorbate-80 and carboxymethyl cellulose (present in chewing gum, soft drinks, pastry, among others) in mice produces a reduction in Bacteroidales an increase in *Ruminococcus gnavus* and other mucolytic bacteria. These alterations could increase bacterial translocation through the intestinal epithelium *in vitro*, promoting chronic gut inflammation and developing colitis and metabolic syndrome (189).

A recent review showed that exposure to TiO_2 (colourant present in cakes, sweets and sauces) might produce variations in the abundance of specific bacterial species and lead to intestinal dysfunctions in animal models (190).

1.3.2 EFFECTS OF DIETARY HABITS ON GUT MICROBIOME

Previously we have discussed the critical association between diet and gut microbiota. Nevertheless, the approaches discussed previously could be reductionist since most research studies were implemented on isolated nutrients. Our diet is accompanied by a mixture of nutrients, and bioactive compounds continuously and indefinitely supplied to our intestinal ecosystem. Therefore, it is considered relevant to evaluate the effect of habitual diets or dietary patterns that comprises a combination of nutrients and non-nutrients on the gut microbiome rather than a specific nutrient (151).

1.3.2.1. Gut microbiome in the vegetarian diet

The vegetarian diet is characterised by excluding the consumption of meat. However, within vegetarianism, there are distinctions depend on whether they include animal-based foods such as fish, chicken, eggs, or dairy products, or if the diet is 100% plant-based (vegans or strict vegetarians). The vegans are a subgroup of vegetarians (191).

Some researchers reported higher proportions of Bacteroides/Prevotella, *Bacteroides thetaiotaomicron*, *Clostridium clostridioforme*, *Klebsiella pneumoniae* and *Faecalibacterium prausnitzzi* and lower ratios of *Clostridium* cluster XIVa and *Bilophila wadsworthia* in vegetarians and vegans compared to omnivores (192,193). However, intervention and cross-sectional studies have observed modest differences in the gut microbiota between omnivores and vegetarians (140,194,195). These studies suggested that the effects of dietary patterns on the microbiota are more significant at the genus and species level and relatively minimal on more general compositional characteristics, such as diversity and richness (150).

1.3.2.2. Gut microbiome in the Western diet

It has been reported that the Western diet (characterised by being rich in total fat, animal protein and refined sugars) shows a marked underrepresentation of the *Prevotella* genera (98,99) and a decrease in total bacteria and beneficial species such as *Bifidobacterium* and *Eubacterium* (129,172,196). Simultaneously, a diet based on animal protein intake increases the abundance of bile-tolerant microorganisms such as *Alistipes, Bilophila* and *Bacteroides*, and decreases the levels of Firmicutes that metabolise dietary polysaccharides (127).

1.3.2.3. Gut microbiome in the Mediterranean diet

The Mediterranean diet (MD) is rich in fruits, vegetables, olive oil, nuts, legumes, whole grains and wine. Several studies have shown that greater adherence to the MD has a beneficial impact on the gut microbiota (141,197,198). However, a recent MD intervention

study (N=343) only showed minor changes in the *Lachnospiraceae* and *Ruminococcaceae* families and did not found differences in the alpha, beta diversity and *Prevotella / Bacteroides* ratio one year post-intervention (199).

1.3.2.4. Gut microbiome and other types of diet

Many studies have highlighted a pronounced change in the composition of the gut microbiota due to the Gluten-free diet (GFD), low-FODMAP (fermentable oligo-,di-, monosaccharides, and polyols) diet and the ketogenic diet (KD), employed as a treatment for different diseases. However, the modification of the gut microbiota seems to be more pronounced in patients who followed a GFD and low-FODMAP diet, while the effect was not clear for KD. More studies are needed to confirm and expand on these findings (200).

1.3.3 COLLECTING DIETARY DATA IN GUT MICROBIOME STUDIES

Dietary assessment in microbiome studies is based on different nutritional assessing methods (See section 1.1). Nevertheless, studies integrating diet and microbiome published to date contain limited information on dietary intake. Most of them collected dietary information by one-day food records, 24HR or FFQs. However, these methods are associated with advantages and disadvantages relevant to the investigation of diet-microbiome (Table 3).

One of the main dietary tools used is the FFQ. The advantage of FFQs is that they are convenient, easy to administer, and require less participant time than other methods. A study participant indicates the frequency and amount consumed of certain foods, and these answers are used to estimate the total caloric and primary nutrient intake. Although FFQs cannot capture diet as accurately as other methods, global dietary parameters through dietary patterns or a healthy eating index could be evaluated (201). The dietary patterns estimated by FFQ provides an idea of how the habitual diet contributes to the microbiome's profile modulation. These findings are supported by research showing that changes in dietary patterns or habitual diet affect the makeup of the microbiome (98,99,139). However, the currently used FFQs are not designed and validated to provide data that can link foods with changes in microbial composition. In addition, the extensive list of foods and the measurement of food consumption from the previous year -of the majority of FFQs used- increases the burden on the participants and does not allow to capture the changes in diet.

Therefore, creating a tool that allows evaluating diet, not from a nutrition-centric perspective but a microbiologist-centric perspective, shall allow a more adapted way to correlate dietary patterns with microbial community composition and activity and by extrapolation to associate food with health status (143).

Table 3. Dietary assessment methods, main advantages and disadvantages, and diet-microbiome research where these methods have been used. Adapted from (202,203).

Dietary assessment method	Description	Output	Study design
Weighed food record	Participants weigh and record all food and beverages in real-time Usual duration 3–7 days	Energy, fibre, macronutrient, micronutrients, additional food constituents (depending on food composition data) Food groups Number of meals	Experimental 'whole diet' interventions; supplementation studies to check background intake Cohort Case-control Cross-sectional Longitudinal
Unweighed food record	Participants record estimated quantities of food and beverages in real-time Usual duration 3–7 days	As for weighed food record	As for weighed food record
24-h recall Examples of online versions: ASA24, myfood24, Intake24	Food and drink intake between midnight to midnight the day prior to data collection through a structured interview with a trained interviewer	As for weighed food record	Cohort Case-control Cross-sectional Longitudinal
Food frequency questionnaire Examples: Harvard FFQ, EPIC FFQ, AES, Food4Me	Questionnaire that assesses the frequency of consumption of individual foods over a defined period (e.g.,1 year) Most include 80–120 items	Energy, fibre, macronutrient, micronutrients and additional food constituents (dependent on food composition data) Food groups	Experimental studies requiring long-term diet data Cohort Case-control Cross-sectional Longitudinal
Diet quality and dietary patterns			
Diet quality Examples: 'HDI, HDS, HEI, AHEI, MDS	*A priori* score measuring overall healthfulness of the diet based on current evidence	Components aggregated to obtain a final score. Higher score indicates better diet quality.	Cohort Case-control Cross-sectional Longitudinal
Diet pattern analysis	*A posteriori* approach Derives patterns using principal components/ exploratory factor analysis or cluster analysis Patterns such as 'prudent' or 'Western' can be derived	Identifies foods consumed together (principal components analysis) or clusters individuals with differing dietary intakes (cluster analysis)	Cohort Case-control Cross-sectional Longitudinal

Note: Blue: **(Number item FFQ, time frame recall)** * The American Gut Project cohort used VioScreenFFQ. ¥ TwinUK cohort where they used EPIC-STUDY NORFOLK FFQ. ⍰ FFQ Human Nutrition of Wageningen University. ⍰ NutritionQuest by Block98.2. ⍰ National Cancer Institute's Diet History Questionnaire II2 (DHQ). Abbreviations: AES, Australian Eating

Advantages in diet-microbiome research	Disadvantages in diet-microbiome research	Microbiota-dietary studies that used this tool method
Most precise measurement of actual dietary intake Useful for associating very recent dietary intake with microbiota profile Good agreement with biological dietary biomarkers	Very burdensome for participant Burdensome for researcher May influence eating behaviour May require assessment of interobserver agreement between coders	De Filippis 2016 (141)
Less burdensome for the participant than weighed food record Useful for associating very recent dietary intake with microbiota profile	Burdensome for researcher May influence eating behaviour Participants may under-or over-estimate quantities	Johnson 2019 (142); Jenning 2018 (204); Dao 2016 (205); David 2014 (126); David 2014 (127); Cotillard 2013 (171)
Low burden for participant Online versions available Multiple 24-h recalls demonstrating good agreement with dietary biomarkers Single 24-h recall acceptable for large cross-sectional studies	Risk of recall error Requires trained interviewer Interviewer bias (data accuracy dependent on interviewer expertise, consistency between interviewers) Single 24-h recall is usually not appropriate due to day-to-day variation in dietary intake	Bellikci-Koyu 2019 (206); Vangay 2018 (139); Fujio 2017 (207); Wu 2016 (140); Zeevi 2015 (208); Yatzunenko 2012 (98); Wu 2011 (129).
Accounts for weekly/seasonal variation in intake Useful for assessment of habitual diet–microbiome associations Low burden for participant Simple to administer Practical for large scale studies Validated tools available for specific populations, specific nutrients	Time consuming for the participant (up to 60 min) Requires mathematical skill to calculate intake using frequency categories Infrequently consumed foods may be missed due to fixed food lists Greater risk of under-reporting and error compared with other methods	Bolte 2021 **(183 item, 1m)** (209); Asnicar 2021¥ **(156 item, 1y)** (210); Gacesa 2021 **(183 item,1m)** (211); Taylor 2020* **(155 item, 3m)** (212); Le Roy 2019¥ **(156 item, 1y)** (213); Skikany 2019 **(127 item, 1y)** (214); McDonald 2018* **(155 item, 3m)** (215); Rothschild 2018 (create by dietitian) (125); Zeevi 2015 (create by dietitian) (208); Suez 2014 **(126 item, 1y)** (188); David 2014 **(134 item, 1y)** (127); Claesson 2012 **(147 item, 1y)** (216); Wu 2011 (**- ,1y**) (129).
Accounts for complexity of the diet and interactive effects of dietary factors Many indices validated by relating index score against health outcomes	Majority require nutrient intake assessment (i.e. food record, 24-h recall or FFQ) for calculating final score	Taylor 2020 (212); Rothschild 2018 (125); De Fillipis 2016 (141).
Accounts for complexity of the diet and interactive effects of dietary factors Can be used as a covariate to determine if the effect of a nutrient is independent of the overall dietary pattern	Requires nutrient intake assessment Patterns empirically derived from data not from diet-health evidence Arbitrary decisions required (e.g., food groups, number of factors/clusters to be retained)	Bolte 2021 (209); Johnson 2019 (142); Shikany 2019 (214).

Survey; AHEI, Alternate Healthy Eating Index; ASA-24, Automated Self-administered 24-hour Dietary Assessment Tool; FFQ, food frequency questionnaire; HDI, Healthy Diet Indicator; HDS, Healthy Diet Score; HEI, Healthy Eating Index; MDS, Mediterranean Diet Score. Acronyms: m: month y: years

2. HYPOTHESIS

HYPOTHESIS

While collecting and analysing microbiome data methods have improved over the past decade, there has been little change in the analysis and collection of dietary data. The current relationship results between diet and intestinal microbiome have been based on epidemiological studies that capture the usual diet through different evaluation methods, as shown in section 1. However, these tools were designed and validated for association studies of cancer or chronic non-communicable diseases and not to establish associations with the gut microbiome.

First, we believe that developing a new short FFQ (sFFQ) would provide a satisfactory relative validity and reproducibility, and thus, properly assessing the dietary intake of a healthy population between 18 and 65 years.

Second, we hypothesise that the newly developed sFFQ based on the foods with higher potential importance for the gut microbiome and based on the foods most consumed by the population could establish a relationship between diet and intestinal microbiota composition for future population studies. Additionally, this sFFQ could be easily adapted to a population with intestinal diseases or metabolic syndromes and a population from another region of the world.

3. OBJECTIVES

OBJECTIVES

3.1 MAIN OBJECTIVES

Design and validate a short semi-quantitative food frequency questionnaire to relate diet with microbiome composition in healthy adults.

3.2 SPECIFIC OBJECTIVES

1. Design and validate a simplified semi-quantitative food frequency questionnaire self-administered to be used in studies of gut microbiota.

2. To characterise the diet of our study population and compare the results with population studies carried out on the Spanish population.

3. To characterise the composition of the intestinal microbiota in relation to participant´s characteristics, the consumption of certain food groups, energy and nutrients.

4. METHODS

4.1 ETHICAL COMMITTEE´S APPROVAL FOR HUMAN STUDIES

The protocols were presented and approved by the Vall d'Hebron University Hospital ethics committee code PR (AG) 84-2020 (Barcelona, Spain). The subjects were informed about the study during the admission interview and gave their consent before inclusion.

4.2 STUDY DESIGN

We carried out an observational, monocentric and longitudinal study between April 2017 and August 2020 at the Vall d'Hebron Research Institute (VHIR), Barcelona, Spain. We designed the research to assess the association between food intake and the intestinal microbial community composition by developing and validating a sFFQ. We used a 24HR as a reference method for the relative validation of the sFFQ. For this purpose, between April 2017 and June 2018, a pilot study was conducted. Healthy volunteers underwent two self-administered sFFQs to evaluate the reproducibility and three 24HRs, including one non-working day. Participants also provided two consecutive stool samples, one just after the first sFFQ and 24HR and one after the second sFFQ and last 24HR. The samples were processed to determine the microbial composition.

After detecting some problems in designing and validating the pilot sFFQ (sFFQp), we integrated some modifications into the next version of the sFFQ. Finally, between February and August 2020, we re-validated the sFFQ with 44 new participants. The study design is shown in Figure 8.

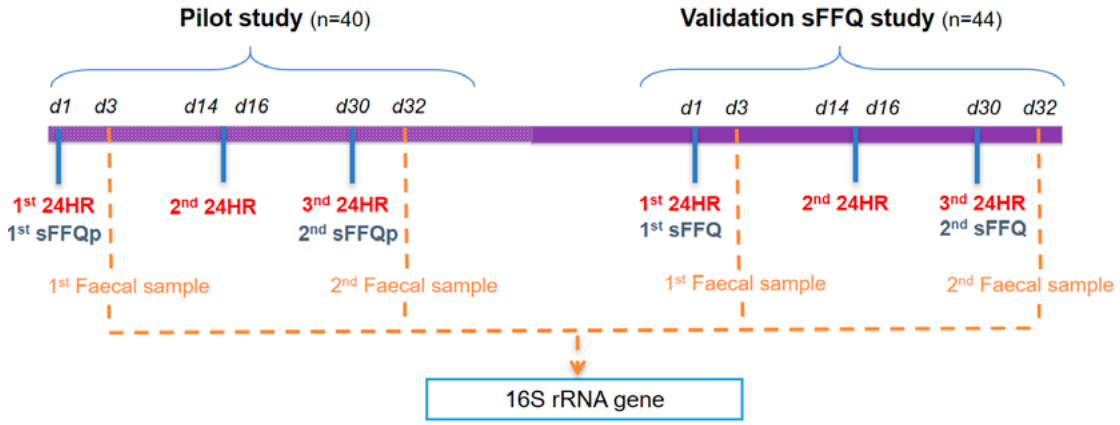


Figure 8. Study design for the validation of the sFFQ. In the pilot and final study, participants were interviewed three times by trained staff over one month in the form of three dietary recalls (3-24HRs) (baseline, day 15 (d15) and day 30 (d30)) and complete two sFFQs (baseline and day 30). Participants in the pilot and validation study provided two frozen stool samples (day2 or d3 and d32 or d33) for microbiome analysis. d, day; 24HR, 24-hour dietary record; sFFQp, pilot short food frequency questionnaire; sFFQ, short food frequency questionnaire.

4.3 STUDY POPULATION

A total of 98 healthy volunteers from Barcelona, Spain (45 people in the pilot study and 53 in the sFFQ validation) were recruited between May 2017 and August 2020 by disseminating an announcement (approved by the committee of ethics) in the Vall d'Hebron Hospital Campus of the University of Barcelona, Spain.

Power calculation showed that to validate the FFQ (by mean of three 24HR), a minimum of 40 subjects would be needed to give 85% power to detect a significant correlation at the 5% level.

Exclusion criteria were:

- Under 18 years and over 70 years.
- Intake of antibiotics in the three months before study entry.
- Intake of proton pump inhibitor medications.
- Any disorders potentially associated with an altered gut microbiome, such as diabetes, digestive pathology, inflammatory bowel disease, or autoimmune disease.

4.4 24HR: REFERENCE METHOD FOR THE RELATIVE VALIDATION THE SHORT FFQ

In the pilot and validation study, we administered three 24HRs to healthy subjects to assess the habitual diet and validate the sFFQ. Trained staff performed the three interviews, two of them during week-days and one at the weekend, through which they collected food consumption of the previous day from the first intake in the morning to the last meal and liquid consumed during the night. In the pilot, we conducted face-to-face interviews. While in the validation study, we carried out the interviews face-to-face or online due to COVID-19 norms.

To avoid the introduction of potential biases during the 24HR recall response, we conducted the interviews as randomly as possible in time, taking into account the participants' availability. We assigned an alphanumeric code to each person to maintain anonymity and registered the food in a 24HR open-form designed by the Nutrition Department of the University of Chile (Annex 1: 24HR collection form). The document includes questions about meal-type and meal-times, specific ingredients of each preparation, amount eaten in homemade serving size measures, net grams and food description (detail of the particular characteristics of the food such as fresh, frozen, canned, type of cooking, brand, etc.). Also, we included a question about the consumption of dietary supplements, vitamins or minerals and drugs.

As memory aids, we used photographic albums: "Guide for dietary studies" from the University of Granada (217) and the "SU.VI.MAX Portions Alimentaires" (218) to evaluate the serving size of each food and liquid objectively.

To estimate the serving weight provided by the participant, we created a "Table of standardised household measures" based on the publication of Moreira's *et al.* from the Complutense University of Madrid (43) and on the measurement of homemade foods (food for which information on their weights was not available) realised in the lab. For the latter, different laboratory staff members weighed the food (with household measures such as a level or stacked teaspoon) to recover an average weight for each food.

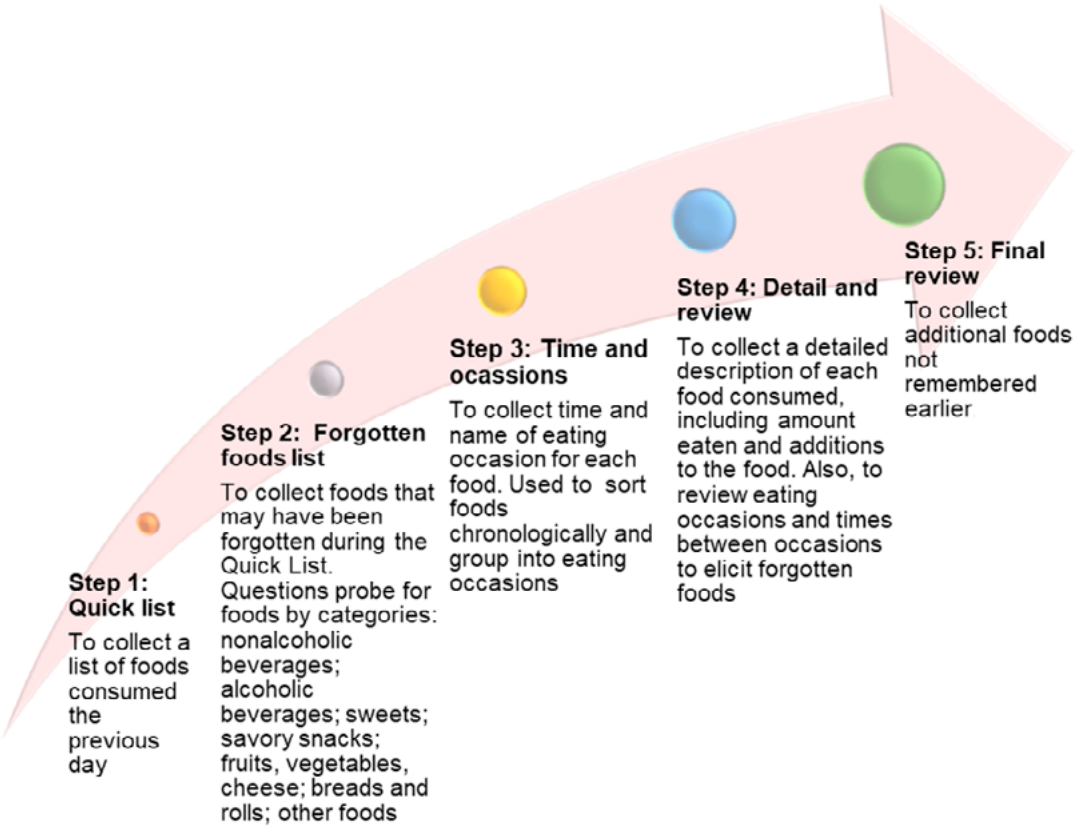


Figure 9. Automated multiple pass method. Adapted from (16).

To reduce possible bias introduced by interviewers, we applied a standard operating procedure (SOP) (Annex 2: Standard operating procedure to data collection of the 24HRs) for data collection of the 24HRs, based on the five-step interview proposed by the USDA (Figure 9) (15,16).

4.4.1 EXTRACTION OF ENERGY AND NUTRIENT FROM THE 24HR

4.4.1.1 Pilot study

The calculation of energy and nutrients of the 24HR from the pilot study was performed using the PCN-GRAMS1.1 program. This program uses the nutritional composition database from the "Centro de Ensayos de Nutrición y Dietética Superior" (CESNID, Spain) (41), which includes a list of 698 foods and 33 nutrients.

We obtained the nutritional data of foods not available in this program from the FCT of the French Agency for Food Safety, Environment and Occupational Health (CIQUAL, France) (38) and the nutrition labelling of food.

Nevertheless, we encountered some issues quantifying energy and nutrients using this methodology (see section 3.1 in results); therefore, we made a series of modifications in the validation study.

4.4.1.2 Validation study

Initially, we transformed the 24HR food data into energy and nutrients using the EvalFinut® software (219). We selected this software because it uses a unified database of food composition BEDCA developed by the government of Spain, which includes a list of 950 foods and 31 nutrients (46). In addition, it contains the USDA food database, consisting of a list of 8,618 foods and 150 nutrients (37).

However, we encountered many issues during its use:

- It only allowed the connection of a single user.
- The long food list could lead to possible food selection biases (for instance, the variety of bread made it challenging to classify different types of cereal bread).
- The food search system did not allow working with food codes.
- Lack of updated nutritional information on new and processed foods.
- Difficulty to include new recipes and foods.
- Therefore, we decided to create a new database of the nutritional composition of foods.

Development of in-house food composition database (ihFCDB)

We created an updated food database based on the food code to facilitate the search and decrease encoding error in food selection to nutritional quantification of the 24HRs. For this purpose, we combined several food composition databases, including the BEDCA Spanish database (46), the Moreiras' table (43) and the USDA nutritional database (37). The mixed dishes and recipes were broken down into simple ingredients; the participants were asked to describe the recipes and cooking procedures in detail. We then calculated their nutritional composition following the recommendations of the "Workbook to develop nutritional food labels" based on the Chilean Health Food Regulation (220).

Our in-house food composition database (ihFCDB) currently contains 1104 foods and mixed dishes, grouped into 13 food groups and 31 nutrients plus energy for 100g of food.

4.4.2 ANALYSIS OF DATA EXTRACTED FROM THE 24HR

We verified the correct estimation of the homemade measurements in grams of food recorded by the 24HRs, according to the "Standardised table of homemade measurements" and SOP for data collection of the 24HRs (Annex 2: Standard operating procedure to data collection of the 24HRs) in both studies.

Furthermore, in the validation study, we checked the correct estimation of grams of the homemade measurements and corroborated the proper selection of the food code (from ihFCDB) by the staff for the nutritional quantification of the 24HRs.

4.5 SHORT FOOD FREQUENCY QUESTIONNAIRE (sFFQ)

We conducted a brainstorming group among researchers with knowledge of the microbiome in our centre before developing the questionnaire. The objective was to acknowledge what type of food groups or other variables could be related to microbiome modulation and, therefore, could be considered and used to generate questions to be integrating into the sFFQ (Figure 10). To develop the sFFQp and then the sFFQ, we followed the recommendations from previous works by Cade (18,19), Willett (17,221) and Lombard (51), as we will describe in the next sections.

4.5.1 THE SEMI-QUANTITATIVE AND SHORT FOOD FREQUENCY QUESTIONNAIRE: PILOT STUDY (sFFQp)

4.5.1.1 Design and development

We developed the food list of the sFFQp based on:

- The food consumption data published by the ENALIA2 survey (National Food Survey on adults, the elderly and pregnant women) (222).
- FFQs applied in the Spanish population (223,224) and other countries (225–227).
- According to current evidence of the foods that have been linked to the gut microbiome profile (127,129,171,215,228–233).

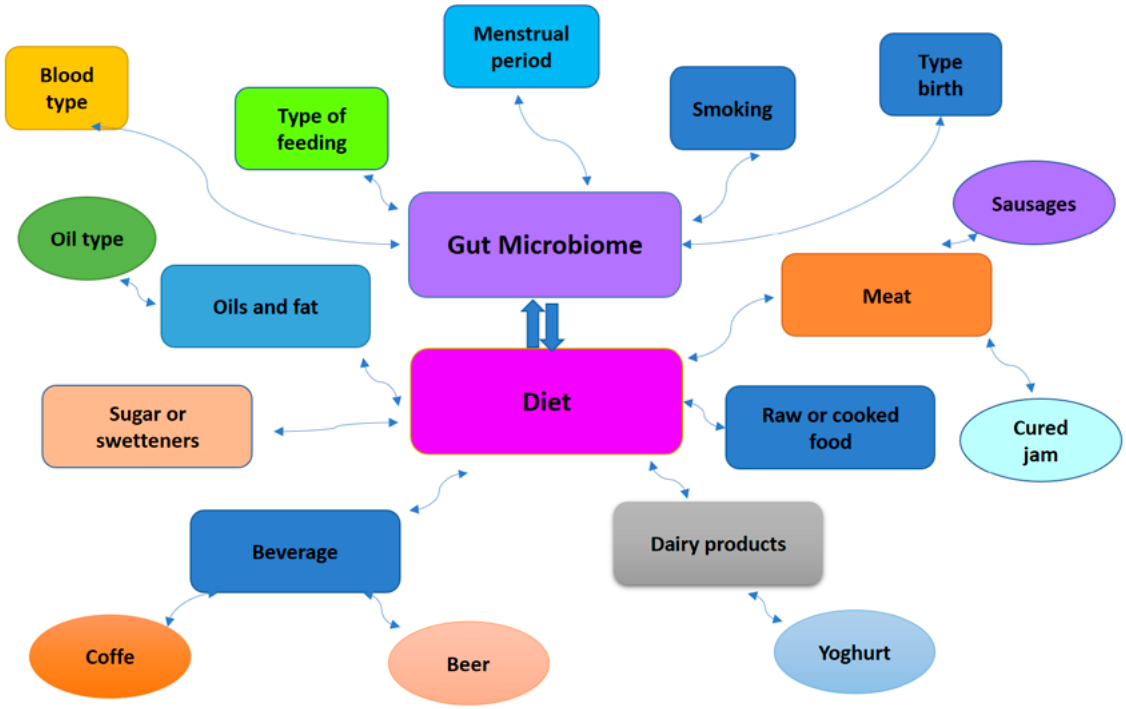


Figure 10. Brainstorming group. Variables of interest in gut microbiota were indicated by our researcher and including in the FFQ.

In total, we selected 216 foods to include in the list according to their higher intake within the population and their relationship with the gut microbiome. Then, we grouped them into 46 items according to their nutritional composition and group of food, for instance, vegetables, fruits, bread, legumes, etc.

We categorised the frequency of consumption for the previous month into six possible responses for each item:

Never	*1 or 2 per month*	*1 or 2 per week*	*3 or more per week*	*Once per day*	*2 or more times per day*

We estimated the standard food serving size in grams and household measures for each item of the questionnaire, based on the food portion on the FFQ of the PREDIMED study and the results of different survey and guidelines, including the ENALIA2 (222), the Guidelines of the Community Nutrition Spanish Society (234), the Guideline of the Scientific Committee "5 a day " (235,236) and the size of the portions assigned by the food industry.

We also added questions that could pinpoint to relevant factors with a potential effect on microbiome composition changes, as previously demonstrated, such as demographic data (age, sex) (216,237,238),

anthropometric data (usual weight and height) (239), other variables (blood group (240), type of delivery (241), follow a special diet (151,242), consumption of ready-to-eat meals (243)) and supplements or drugs intake (117,120).

Finally, we applied a pre-questionnaire to the participants and research staff to help identify any difficulties in the pilot questionnaire completion.

4.5.1.2 Method of administration of the sFFQp

The two sFFQp, in paper format, was self-administered on the same day of the first and third 24HR interviews (one day and d30) (Figure 8). Once we got the responses from both diet questionnaires, we checked the missing data with the participants.

4.5.1.3 Construction of the energy and nutritional composition table data extracted from sFFQp items

Using the PCN-GRAMS1.1 program (commented in section 4.1.1), we calculated the unweighted mean of each of the 46 items. For example, for the vegetable item formed by lettuce, endive, sprouts, and mixed salad, we calculated the mean content of energy and nutrients per 100 grams of each of these foods.

4.5.1.4 Analysis of responses of the sFFQp

The frequencies and quantities of food consumed by the participants were collected monthly and weekly by the sFFQp. They were divided by 30 and 7, respectively, to get the mean daily intake of food, energy, and nutrients. For example, for the legume item with a consumption frequency of 1 to 2 times per week (1.5 / 7 days = 0.21) and a serving size of 150 grams, the final value would be 0.21 x 150g = 31.5g/d. These values will be compared later with the average daily intake of the 24HR to determine the relative validity of the questionnaire.

Once the sFFQp and 24HRs were obtained, we classified and consolidated the foods and liquids into 15 food groups for statistical analysis.

1..Vegetable	*9. Sausages*
2. Fruits	*10. Oils and fats*
3. Nuts and seeds	*11. Cakes and pastries*
4. Cereals or grains	*12. Sauces and condiments*
5. Milk and dairy products	*13. non-beverage alcohol*
6. Meat, meat products	*14. Beverages alcohol*
7. Fish and seafood	*15. Sugar and sweet*
8. Legumes	

We then evaluated the data and performed statistical analyses to validate the sFFQp, and decided to make a few modifications to the design of the questionnaire based on the observations and results obtained in order to improve. The reasons for these modifications are indicated in the subsection 3.3 of the results section.

4.5.2 THE SEMI-QUANTITATIVE AND SHORT FOOD FREQUENCY QUESTIONNAIRE: VALIDATION STUDY (sFFQ)

The validation study was performed using the modified sFFQ. We kept most of the components; however, we implemented some changes that are detailed below.

4.5.2.1 Design and development of the sFFQ

We used the food selection criteria from the previous sFFQ and the food list based on the information provided by the prior 24HRs (n=40). This design allowed us to identify the foods most consumed by our study population, several of which were absent in the pilot questionnaire. In total, we selected 310 foods on the basis of the above-mentioned criteria and on to higher intra- and interindividual variability of consumption. Subsequently, these foods were grouped into 58 items based on their possible impact on the gut microbiome and were subsequently disaggregated according to their nutritional composition.

The frequency of consumption “1 or 2 per month” was changed to “1 or 3 per month”. Previously, we did not consider the frequency of consumption of three times per month.

Serving size was kept for most food items of the sFFQ. However, we considered the values of the food portions obtained from the 24HR of the pilot study. From the analysis of the dietary intake using the sFFQp, we also realised that participants could have considerable difficulty perceiving the serving size in grams of certain items, such as 30g of nuts. Therefore, to facilitate estimation of the serving size in grams of the 58 items, size indication in gram was replaced by household measurements (tablespoon, a dessert teaspoon, etc.) or portions assigned by the food industry (such as yoghurt container, canned beer, etc.) in the sFFQ.

To further improve the estimation of the amount of food consumed in the sFFQ, we included memory aids such as food photographs and added three consumption alternatives for the serving size: “1/2 of the standard serving size”, “standard serving size”, and “double the standard serving size” based on the recommendations of Subar and colleagues (244).

In addition, we added some questions that could be related to the microbiome, such as menstruation period (245), smoking (113), use of sweeteners (188) and amount of liquids consumed per day based on previous studies.

Finally, the final version of the sFFQ (Annex 3: Short food frequency questionnaire) was evaluated and tested on non-health professionals volunteers, researchers, and medical doctors from the VHIR, Barcelona, Spain.

4.5.2.2 Creation of food photography

According to previous studies, we created a support document based on food photographs to estimate the amounts of food consumed in the sFFQ (246–250).

Methodology

a. Food selection and serving size

We selected foods with the highest consumption rate and nutritional interest and with the greatest variability in amount inside the sFFQ.

We generated two types of food photographs:

- A unique photo with three different meal timing; breakfast, lunch and dinner and meal snack (Figure 11); for which we applied the standard food serving size.
- A series of three food photographs of the meals with the highest interpersonal and intrapersonal consumption variability (was observed in the 24HR of the pilot study), Figure 12. We used the 1: 2: 3 ratios for the portion size, where "2" corresponded to the standard serving size of the sFFQ.

b. Preparation of the food photographs

To take the photos, we used a Canon DS126071 camera and an electronic kitchen scale (EKS 8250) with a 5kg capacity. In addition, we used different eating utensils, such as white plates with different diameters (19cm and 26cm diameter), cutlery (21cm long knife, 20cm long fork, 18,5cm long spoon, 13cm long teaspoon), bowls and cups with different capacity, glasses (400, 300 and 100cc) and wine glass (150cc) to show portion size of the several foods.

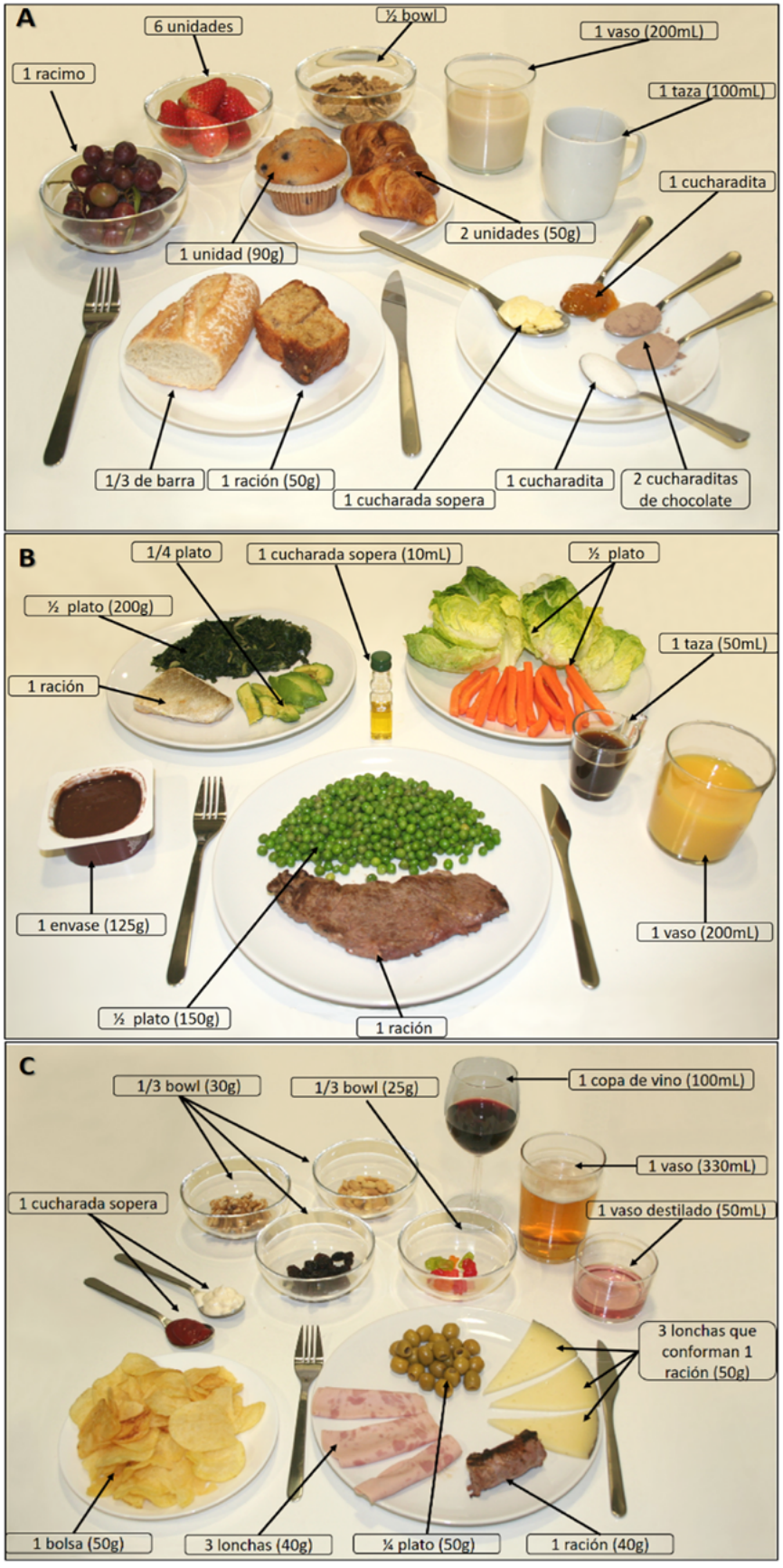


Figure 11. Photographs of the three meal times used in the sFFQ. A: Breakfast; B: Lunch and dinner; C: Meal snack.

c.- Procedure for producing the food photographs

First, we cooked some of the food. We weighed food before and after cooking. Second, we served the food and drinks on/in the plates, bowls and glasses with the characteristics mentioned above. Third, we added silverware to the table to get a perspective of the proportions and portion sizes. Fourth, we took the photographs on a white background, 42° above the horizontal and light standard conditions. Finally, we edited the photographs and added them to the online version of the sFFQ. The pictures, printed in colour and size 140x95mm, were used as memory aids in the 24HRs.

4.5.2.3 Method of administration of the sFFQ

The sFFQ was self-reported using SurveyMonkey Inc. (San Mateo, California, USA) online platform. On the first 24HR interview, we provided the participants with a web link or QR code to complete the online sFFQ. Once we obtained the responses, we verified the missing data. We contacted participants if there was a lack of response to any of the items on the sFFQ.

4.5.2.4 Energy and nutrient composition of the sFFQ items

To estimate the energy and nutrient intake of the sFFQ, we used our ihFCDB. We calculated and evaluated the differences in nutritional composition between the unweighted and weighted mean of energy and nutrient for each item.

For the unweighted mean of energy and nutrients, we added the nutritional values per 100g of food that made up a given item, and we divided by the total values.

For the weighted mean, we based our calculation on the data obtained from the 24HR of pilot study and the amount consumed in the ENALIA2 survey (222). The foods collected from the 24HR and ENALIA2 were classified and cross-check into 58 items the sFFQ. We then recovered the proportions contributed by each food to each of the 58 items from the 24HR of the pilot study and used them as weighted factors to calculate the energy and nutrient intake for each item in the sFFQ through the ihFCDB.

4.5.2.5 Analysis of the sFFQ responses

We created an excel sheet linked to an external database to calculate the nutritional composition of each food item (with unweighted and weighted values of energy and nutrients). This database is a nutritional composition table of the sFFQ. We applied an SOP (Annex 4: Standard operating procedure to sFFQ nutritional composition quantification) to facilitate the analysis and eliminate possible errors in the transformation from weekly or monthly consumption to daily consumption frequency and in the calculations of g/day of the sFFQ to compare with the values obtained from the 24HR for its validation.

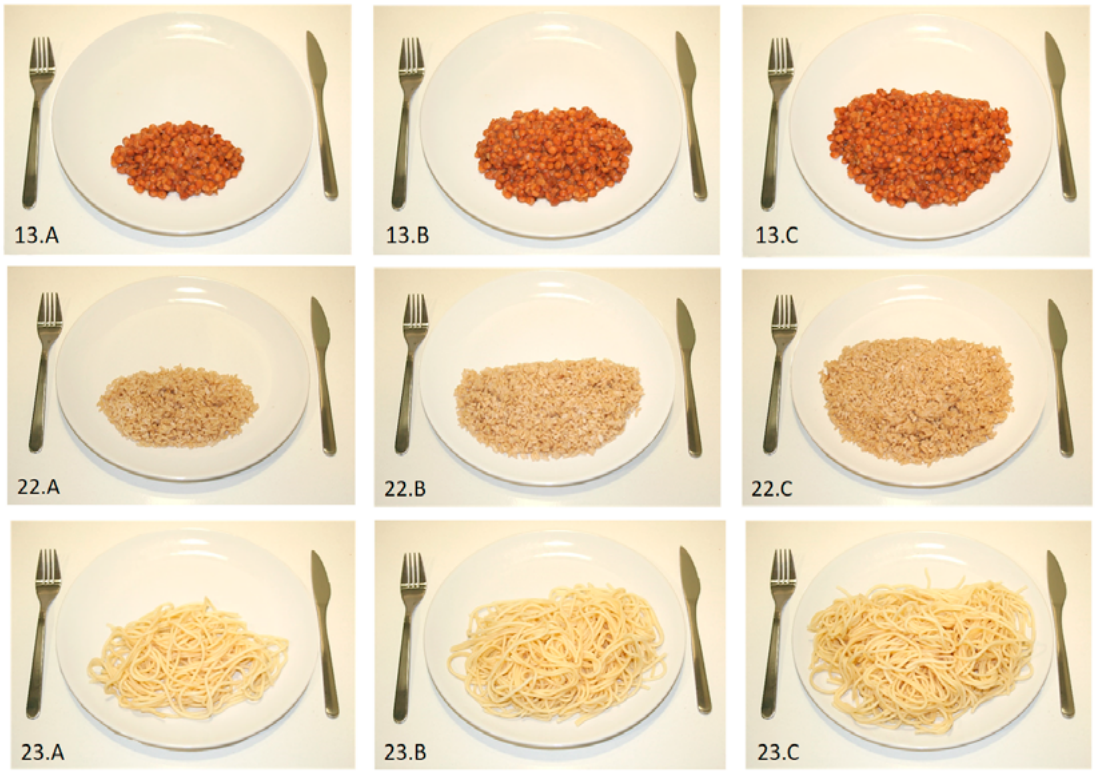


Figure 12. Photographic series of food used in the sFFQ in the 1:2:3 ratios. Letter B or ratio 2, corresponds to the standard serving size of the sFFQ.

Finally, we consolidated the information obtained from the sFFQs and 24HRs. We classified foods and drinks into 13 groups based on the EUROCODE 2 Classification (251). Nevertheless, we regrouped and subdivided them into 24 subgroups foods based on their potential and similar role on the gut microbiota modulation:

1. Alcoholic beverage	*13. Pastries and sweetbreads*
2. Appetisers	*14. Potatoes and other tubers*
3. Biscuits breakfast cereals and cereals bars	*15. Ready to eat meals*
4. Chocolates and derivatives	*16. Sauces and condiments*
5. Fats and oils	*17. Sausages and other meat products*
6. Fish and shellfish	*18. Sugar and other sweets*
7. Fruit and fruit products	*19. Vegetables and vegetable products*
8. Legumes	*20. White bread*
9. Meats and eggs	*21. White grains and white pasta*
10. Milk and dairy products except for fermented milk	*22. Wholegrain or wholemeal bread*
11. Non Alcoholic beverage	*23. Wholemeal grains and wholemeal pasta*
12. Nuts and seeds	*24.Yoghurt and fermented milk*

4.6 IDENTIFICATION OF UNRELIABLE RESPONSES TO QUESTIONS OF sFFQs AND 24HRS

We verified each value obtained from the quantification of the sFFQs and 24HRs. When a possible outlier was detected, we examined each data entry involved in leading to this value. For instance, we excluded participants with calorie intake values lower than 800 kcal/day or higher than 4,200 kcal/day in men and lower than 600 kcal/day or higher than 3,500 kcal/day in women, in both the sFFQs and mean of 3-24HRs (225,252).

4.7 STATISTICAL ANALYSES OF THE RELATIVE VALIDATION AND REPRODUCIBILITY OF THE sFFQ IN THE PILOT AND VALIDATION STUDIES

In the pilot and validation studies, we calculated median and 25-75 percentile of food, energy and nutrient consumption from the mean 3-24HRs and the two sFFQs. Nutrient intake values were log-transformed to transform skewed data to conform to normality. The nutrients were energy-adjusted using the residual and density method in the pilot and validation study, respectively (253) to control the confounding effect of calories.

We evaluated the relative validity (sFFQ2 versus mean 3-24HRs) and the reproducibility (sFFQ1 versus sFFQ2) of the sFFQ using a series of statistical tests.

In the pilot study, we used Wilcoxon signed-rank test, the Bland & Altman plots, Spearman's correlation coefficient, the intraclass correlation coefficient and the concordance correlation coefficient (CCC) in the pilot study. While in the final study, we evaluated the concordance between the two measures. First, at the group level, we applied the Wilcoxon signed-rank test to determine the differences in food, energy and nutrient consumption between the two sFFQs and the mean 3-24HR. Second, we used the Bland & Altman analysis to check the degree of agreement between sFFQ2 and 24HRs. Third, we plotted the differences between the two methods (sFFQ2 – mean 3-24HRs) against the mean intake of the measures ((sFFQ2 + 24HR)/2) and evaluated the limits of the agreement defined as the mean ± 1.96 SD of the mean between both methods (254). In addition, to illustrates the magnitude of the possible systematic difference, we calculated the 95% CI of the means differences. Finally, we calculated the Spearman correlation between the mean and the mean difference of the two methods to reflect proportional bias (55).

To evaluate associations between the two different measurements at the individual level, we used Spearman's correlation coefficient to estimate the strength and direction of the association (51). To control inter-and intra-individual variation, we calculated the

ICC (255,256) between sFFQ and mean 3-24HR. Moreover, we used cross-classification to categorised individuals into equal third or opposite third for food group and energy-adjusted nutrient intake extracted from sFFQ2 and 24HR (257).

We decided not to use the CCC in the final validation as it provided similar information to the ICC (258). Instead, we prefered to add statistical tests that allowed us to evaluate other facets of the validation (see section 1.3 of Introduction) not examined in the pilot, such as a cross-classification analysis. We performed the statistical analysis on GraphPad Prism (v8) and the R_Studio (Version 1.4.1106) package.

4.8 COMPARISON OF THE SELF-REPORTED INTAKE OUR POPULATION WITH THE NUTRITIONAL REFERENCE INTAKES OF THE SPANISH POPULATION

We compared the macro and micronutrient intake data obtained from the sFFQ of this study with the Nutritional Reference Intakes (NRI) of the Spanish population according to age and sex (259). The disparity between reported consumption and the level needed for adequacy was calculated comparing with 80% of the Spanish NRI (260), as suggested by the ANIBES study (288).

4.9 HEALTHY EATING INDEX OF OUR POPULATION

To evaluate the quality of the diet of our population, we calculated the Healthy Eating Index (IASE) for the Spanish population from the consumption frequency data obtained by the sFFQ (23). The IASE comprises 10 variables as listed below. We grouped the sFFQ items into the ten variables.

IASE variables	sFFQ items
Daily consumption	
Grains and derivatives	18-19-20-21-22-23
Vegetables	1-2-3-4-5-6-7-8-9-10
Fruits	14-15
Dairy products	24-25-26-28-29-30-31
Weekly consumption	
Meats	32-33-34
Legumes	13
Occasional consumption	
Cold-processed meats	35
Sweets	44-46
Beverage	39
Diet variety	Depending on whether you meet each of the daily and weekly recommendations

For each variable, a score was assigned according to the frequency of consumption obtained by the sFFQ. For the "daily consumption" variables, the scoring criteria were 10 (once per day), 7.5 (3 or more times per week), 5 (1-2 times per week), 2.5 (<1 time per week) and 0 (never or rarely) points. For the "weekly consumption" variables, the scores were 10 (1-2 times per week), 7.5 (3 or more times per week), 5 (<1 time per week), 2.5 (once per day), and 0 (never or rarely) points, and for the "occasional consumption" variables, the scores were 10 (never or rarely), 7.5 (<1 time per week), 5 (1-2 time per week), 2.5 (3 or more times per week) and 0 (once per day), points.

We calculated the IASE by adding the score obtained for each variable (maximum score being 100 points). Then, we divided the food classification into three categories according to the total score; healthy eating (> 80 points), need-for-change (>50 to 80 points), and little healthy (<50 points).

4.10 IDENTIFICATION OF MISREPORTING IN THE sFFQ

To evaluate the level of reporting bias on energy intake collected by the sFFQ, we used the EFSA guidance on "Protocol for identification of under- and over-reporting" and "Guidance on the Menu methodology" (261), which is based on the cut-off method developed by Goldberg and revised by Black *et al.* (262,263).

The procedure proposed by EFSA evaluates the energy intake reported (EIrep) of the sFFQ against the supposed energy requirements. EIrep is expressed as a multiple of the mean basal metabolic rate estimated (BMRest), and it is contrasted with the supposed energy expenditure of the studied population. Subsequently, the ratio EIrep: BMRest is referred to as the physical activity levels (PAL).

Table 4. Calculated misreporting cut-off at the group and individual levels for the sFFQ2.

Confidence limit	Group level	Individual level
Lower	1.510	1.122
Upper	1.687	2.280

Note: We used the value of the physical activity level (PAL) of 1.6 for all our population.

The protocol indicates that the analyses should be performed at the group and individual levels. The group level determines the general bias towards reported EIrep, and the individual level shows the rate of under and over reporters. The lower and upper cut-off values for misreporting of our population at both levels (Table 4) were calculated based on the values proposed by Black. The BMRest was estimated using the Schoefield equations (264), and we used the PAL 1.6 value recommended by the EFSA to the adult population between 18-69 years.

4.11 FAECAL MICROBIOME ANALYSIS

4.11.1 SAMPLE COLLECTION AND GENOMIC DNA EXTRACTION

On the first and third food interviews, each participant received a freezer package and a collection kit with instructions to collect their stool samples with a 2-day window (d2 and d32), or if not possible, the first stool after the food interviews (Figure 13). Each participant kept the faecal sample in their domestic freezer at -20°C, then took it to the laboratory (maintaining the cold chain), where we stored it at -80°C until further processing. We made 250mg aliquots of faeces in solid CO_2 (using a dry ice environment), maintaining the sample's frozen state to avoid possible nucleic acid degradation. Before starting the extraction procedure, we added 800mg of 0.1mm previously sterilised zirconia/silica beads to the tubes. We extracted DNA following the recommendations of the International Standards for the Human Microbiome (265).

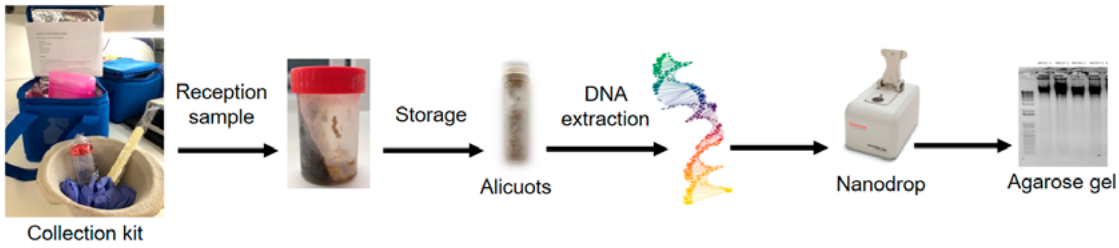


Figure 13. Sample collection and genomic DNA extraction.

To perform chemical lysis, we added 250µl of guanidine thiocyanate to each sample, 40µl of 10% N-lauroyl sarcosine and 500µl of 5% N-lauroyl sarcosine, and we incubated at 70°C for 1 hour. To ensure disruption of most of gram-positive bacteria's cell walls, we performed the additional mechanical disruption using a Beadbeater (Biospec Products ©) (266). We added Polyvinylpolypyrrolidone (PVPP) in multiple wash steps to precipitate and discard aromatic molecules. We performed enzymatic digestion of the RNA of the sample. Then, we precipitated with ethanol the extracted DNA and re-suspended in 200µl of Tris-EDTA buffer.

We quantified the DNA of each sample using a NanoDrop ND-1000 spectrophotometer (Nucliber).

4.11.2 LIBRARY PREPARATION FOR ILLUMINA SEQUENCING

16S rRNA gene amplification

We used the V4 hypervariable region of the 16S bacterial and archaeological RNA gene by PCR for microbiome composition profiling. For this, we used the 5 'ends of the forward (V4F_515_19) and reversed (V4R_806_20) primers that target the V4 region of the 16S gene and were labelled with specific sequences for the Illumina® MiSeq technology (Figure 14). In addition, we specified twelve base-pair Golay codes downstream of the reverse primer sequence (VV4R_806_20) to allow individual sample multiple identifications (48).

Subsequently, we performed a standard PCR using 0.75units of Taq polymerase (AmpliTaq Gold, Life Technologies®) and 20pmol/µl of the forward and reverse primers (IDT Technologies®). A final volume of 50µl was run on a Mastercycler gradient (Eppendorf®) at 94°C for 3 minutes, followed by 35 cycles of 94°C for 45 seconds, 50°C for 60 seconds, 72°C for 90 seconds, and finally a cycle of 72°C for 10 minutes.

Agarose gel electrophoresis and amplicons purification

To visualise the amplicons from the PCR amplification, we first made a 1% agarose gel stained with RedSafe® and ran in 1X Tris Acetate EDTA (TAE) buffer. We then mixed 5µl of each amplicon with 6x loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water). We loaded the amplicon into the solidified agarose gel wells and used a 100bp of DNA XIV molecular weight marker (Roche ©).

The current of the electrophoresis was run at 100V for 30 minutes. Subsequently, we visualised the amplicon bands in a Gel Doc XR + system (Bio-Rad©). The appearance of bands confirmed the generation of amplicons during PCR amplification. In contrast, his absence could be explained by the low presence of bacterial DNA in the sample (especially for samples with low biomass content) or added PCR inhibitors throughout the procedure.

Once the amplification was confirmed, we purified the samples using the QIAquick PCR purification kit (Qiagen, Barcelona, Spain) according to the manufacturer's instructions. Subsequently, we quantified the purified DNA with a NanoDrop ND-1000 spectrophotometer (Nucliber©). Moreover, we examined the amplicons pooled in 2nM by microcapillary electrophoresis. Finally, we used an Agilent 2100 bioanalyser with the DNA 12,000 kit, which solves the distribution of double-stranded DNA fragments up to 12,000bp in length.

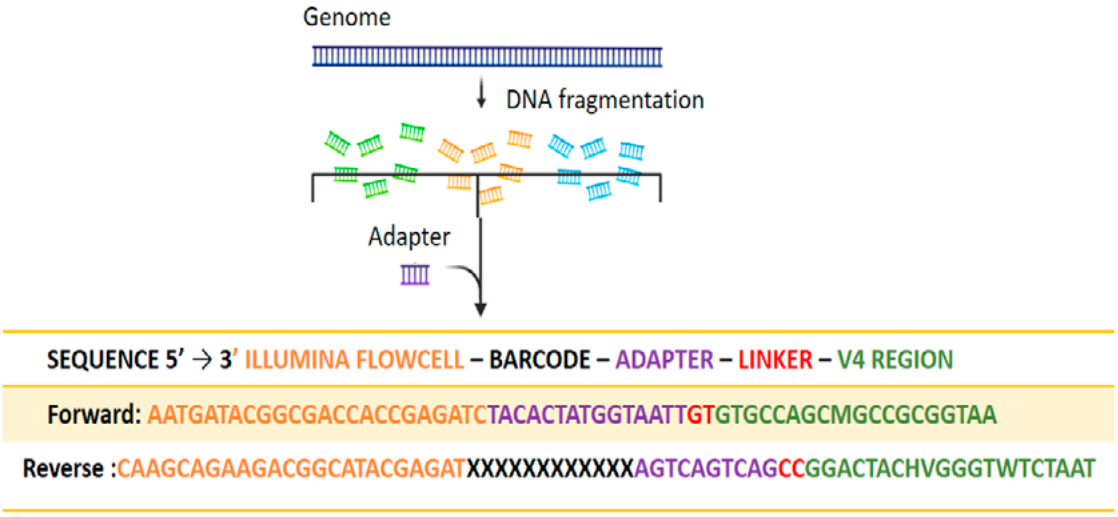


Figure 14. Library preparation. We detailed the primers used to amplify the 16S rRNA gene.

The pooled amplicons (240ng of DNA per sample) were sequenced using the Illumina MiSeq technology at the Autonomous University of Barcelona (UAB, Spain), following the standard Illumina platform protocols.

4.11.3 ILLUMINA SEQUENCING

The following steps of Illumina, Next Generation sequence, are described by the Illumina website (267) (https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf.) Which consists of:

"**Cluster Generation**: For cluster generation, the library is loaded into a flow cell where fragments are captured on a lawn of surface-bound oligos complementary to the library adapters. Each fragment is then amplified into distinct, clonal clusters through bridge amplification (Figure 15). When cluster generation is complete, the templates are ready for sequencing.

Sequencing: Illumina SBS technology uses a proprietary reversible terminator–based method to detect single bases as they are incorporated into DNA template strands (Figure 15). Sequencing reagents, including fluorescently labelled nucleotides, are added, and the first base is incorporated. The flow cell is imaged, and emission from each cluster is recorded. As all four reversible terminator–bound dNTPs are present during each sequencing cycle, natural competition minimises incorporation bias and greatly reduces raw error rates compared to other technologies. The emission wavelength and intensity are used to identify the base. This cycle is repeated "n" times to create a read length of "n" bases.

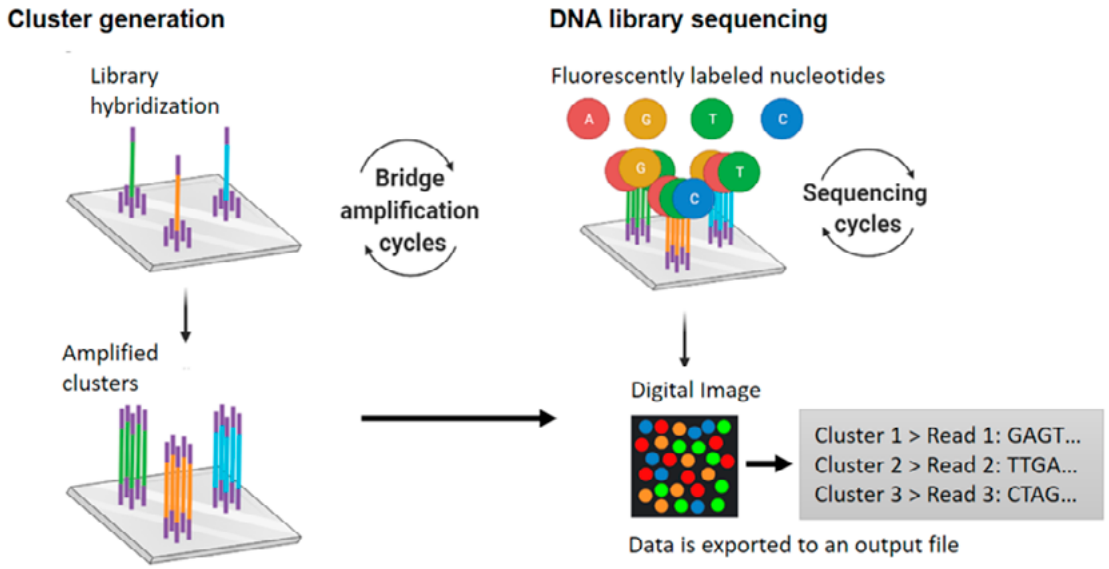


Figure 15. Overview of cluster generation and sequencing of the Illumina Platform. Created by BioRender.com

16S rRNA sequencing

The 16S rRNA amplicons pooled equal concentration (240ng of DNA per sample) were diluted to 2nM and spiked with 15-30% denatured PhiX. Then, denatured templates were further diluted to 5ρM and subsequently combined to give an 85% 16S rRNA gene amplicon library and 15% PhiX control pool.

With the use of a MiSeq (Illumina Technologies©) Reagent Cartridge (300-cycle PE kit), we performed sequencing in the technical support unit of the Autonomous University of Barcelona (UAB Spain), following the standard protocols of the Illumina platform.

4.11.4 16S rRNA SEQUENCES ANALYSIS AND STATISTICAL ANALYSES

Sequence data were analysed using the QIIME 2™, which is a bioinformatics platform that stands for Quantitative Insights Into Microbial Ecology. The sequences were demultiplexed to attribute sequence reads to the appropriate samples and were then denoised and dereplicated into Amplicon Sequence Variants (ASVs) using the dada2 tool, which also filtered out chimaeras. Each sequence read was trimmed to a length of 298 bp. A total of 3.1 million sequences of the 16S rRNA gene were generated from the 166 samples, with a mean of 19,000 sequences per sample. A feature table was generated for all samples with a minimum of 9,159 sequences per sample. One sample with a very low number of reads was removed for further analysis. The feature table of the 165 remaining samples was then used to perform taxonomic classification, alpha- and beta-diversity analyses, and differential abundance measurements in different experimental groups. Taxonomy was assigned to each ASV using a database that combined the Greengenes (version 13.8) and PATRIC (version 2016) databases. To study the association between the microbiome data and clinical or dietary variables, we then used linear mixed models as implemented in the Microbiome Multivariable Association with Linear Models (MaAsLin2) package MaAsLin2 (268), which was setup with the following parameters: normalization="TMM", transform="LOG", correction="BH", analysis_method="LM", max_significance=0.25 (default significance threshold), min_abundance= 0.0001, min_prevalence=0.1. Age, gender, other characteristics of the participants, and dietary data were added as fixed effects. All models were adjusted for gender, and as participant samples from two-time points were included, the participant identification number was added as a random effect. Results with a False Discovery Rate (FDR) lower than 0.25 were considered significant.

5. RESULTS

5.1 STUDY POPULATION

We recruited a total of ninety-eight participants. Five individuals were excluded from the pilot study because they took antibiotics or did not provide stool samples. While nine individuals did not continue during the validation study as a consequence of the COVID-19 pandemic situation. Forty individuals completed the two sFFQp, three 24HR, and provided two stool samples in the pilot study. While 44 participants completed the two sFFQs, three 24HRs, and 42 participants provided the two stool samples in the validation study.

5.1.1 DESCRIPTION OF THE POPULATION IN THE PILOT STUDY

The mean age of the participants in the pilot population was 32.6 years, 65% of the population was female, and the mean BMI was 22 ± 2.6 kg/m^2. 72.5% of this cohort was of Spanish nationality, and the remaining were of other nationalities such as French (n=5), German (n=1), Brazilian (n=2), Chilean (n=1), Colombian (n=1) and Peruvian (n=1) (Table 5). Most of the participants indicated being born via vaginal birth (87.5%), being non-smokers (80%) and consuming a conventional diet (77.5%). Around half of the population in the pilot study reported ready-to-eat-meal consumption (57.5%).

5.1.2 DESCRIPTION OF THE POPULATION IN THE VALIDATION STUDY

The mean age of the population in the validation study was 35.7 years, 47.7% of the population was female, and the mean BMI was 23.1 ± 3.3 kg/m^2. 88.6% of this cohort was of Spanish nationality, and the remaining were of other nationalities such as French (n=2), Honduran (n=1), Brazilian (n=1) and Chinese (n=1). Most of the participants indicated being born via vaginal

birth (84.1%), being non-smokers (65.9%) and consuming a conventional diet (84.1%). Also, 63.6% and 25% reported consuming ready-to-eat meals and sweeteners, respectively (Table 5).

There were no significant differences between the pilot and validation population for age ($p=0.600$; $p<0.05$); gender ($p=1.120$; $p<0.05$) and BMI ($p=0.200$; $p<0.05$).

Table 5. Description of the total population, pilot study and validation study of the sFFQ (mean ± SD).

	Total	Pilot study	Validation study
n	**84**	**40**	**44**
Age (years)	34.2 ± 12.7	32.6 ± 11.1	35.7 ± 14.0
18-29 years, n (%)	44 (52.4)	21 (52.5)	23 (52.3)
30-39 years, n (%)	19 (22.6)	13 (32.5)	6 (13.6)
40-49 years, n (%)	7 (8.3)	2 (5.0)	5 (11.4)
50-59 years, n (%)	9 (10.7)	3 (7.5)	6 (13.6)
> 60 years, n (%)	5 (4.2)	1 (2.5)	4 (9.1)
Female gender, n (%)	47 (55.9)	26 (65.0)	21 (47.7)
BMI (kg/m²)	22.5 ± 3.0	22.0 ± 2.6	23.1 ± 3.3
Weight status, n (%)			
Underweight (< 18.5 kg/m²)	5 (4.2)	3 (7.5)	2 (4.5)
Normal (18.5-24.9 kg/m²)	67 (79.8)	35 (87.5)	32 (72.2)
Overweight (25-29.9 kg/m²)	10 (11.9)	2 (5.0)	8 (18.1)
Obese (> 30 kg/m²)	2 (2.4)	0	2 (4.5)
Nationality, n (%)			
Spanish	68 (81.0)	29 (72.5)	39 (88.6)
European - non-Spanish	8 (9.5)	6 (15)	2 (4.5)
Others	8 (9.5)	5 (12.5)	3 (6.8)
Birth type, n (%)			
Vaginal birth	72 (85.7)	35 (87.5)	37 (84.1)
C-section	12 (14.3)	5 (12.5)	7 (16.0)
Blood type, n (%)			
A	26 (30.9)	14 (16.7)	12 (14.3)
B	4 (4.7)	2 (2.4)	2 (2.4)
AB	1 (1.2)	0	1 (1.2)
O	33 (39.3)	15 (17.9)	18 (21.4)
Unknown	20 (23.8)	9 (10.7)	11 (13.1)

	Total	Pilot study	Validation study
n	**84**	**40**	**44**
Smoking status, n (%)			
Non-smoker	61 (72.6)	32 (80)	29 (65.9)
Smoker	9 (10.7)	2 (5)	7 (15.9)
Former smoker	8 (9.5)	0	8 (18.2)
Unknown	6 (7.1)	6 (15)	0
Diet type, n (%)			
Conventional	67 (79.8)	31 (77.5)	37 (84.1)
Vegetarian diet	6 (7.1)	4 (10.0)	2 (4.5)
Vegan diet	2 (2.4)	1 (2.5)	1 (2.3)
Organic diet	2 (2.4)	2 (5.0)	0
Others diet	7 (8.3)	2 (5.0)	4 (9.1)
Intake of ready-to-eat meals, n (%)			
Yes	51 (60.7)	23 (57.5)	28 (63.6)
No	33 (39.3)	17 (42.5)	16 (36.4)
Intake of sweeteners, n (%)			
Yes	-	-	11 (25.0)
No	-	-	33 (75.0)
Intake of supplements or drugs, n (%)			
Dietary supplements	20 (23.8)	8 (20.0)	12 (27.3)
Probiotics	1 (1.2)	1 (2.5)	0
Oral contraceptive	6 (7.1)	4 (10.0)	2 (4.5)
ACE inhibidors	3 (3.6)	2 (5.0)	1 (2.3)
Fibrate	1 (1.2)	1 (2.5)	0
Statin	1 (1.2)	1 (2.5)	0
Levothyroxine	2 (2.4)	1 (2.5)	1 (2.3)
Other drugs	8 (9.5)	4 (10.0)	4 (9.1)

5.2 RELATIVE VALIDITY AND REPRODUCIBILITY OF THE sFFQp

Daily intake of foods groups, energy and nutrients of sFFQ1p, sFFQ2p and mean 3-24HR are shown in the 25th, 50th and 75th percentiles in Table 6. We selected the sFFQ2p to cover the same time period as the 3-24HR (as discussed in section 1.3.3) to assess its relative validity. For the food group, the sFFQ2p tended to report higher consumption of fish and shellfish, legumes, oil and fat and alcoholic beverages than the 24HR. While for energy and nutrients, sFFQ2p showed an underestimation of energy, water, total protein, animal protein, SFA, cholesterol (CHOL), carbohydrates, sugar digestible, sodium, vitamin A, carotenoids, vitamin D, vitamin B_6, vitamin B_{12} and vitamin C intake than the 24HR (Table 6).

Table 6. Daily consumption of food, energy and nutrients estimated by two sFFQp (sFFQ1p and sFFQ2p) and a mean of 3-24HR.

Food group (g/d)	sFFQ1p P25	sFFQ1p Median	sFFQ1p P75	
Vegetables	103.25	171.00	250.00	
Fruits	200.00	400.00	400.00	
Nuts	6.30	6.30	30.00	
Cereals or grains	123.43	164.73	217.11	
Milk and dairy products	72.00	120.00	177.60	
Meat meat products	59.70	74.70	108.45	
Fish and seefood	16.58	54.60	54.60	
Legumes	31.50	31.50	31.50	
Sausage	1.50	8.80	16.80	
Oils and fat	15.86	30.00	31.20	
Pastries and cake	6.45	12.55	26.33	
Sauces and condiments	1.28	2.60	4.20	
Non-alcoholic beverage	88.75	122.75	151.50	
Alcoholic beverage	21.50	73.05	90.30	
Sugar and sweet	0.47	2.10	5.00	
Energy and nutrients				
Energy (kcal)	1475.80	1677.81	2101.98	
Water (g)	663.07	808.73	988.46	
Total Protein (g)	61.93	76.46	91.50	
Protein vegetal (g)	25.03	32.74	42.55	
Protein animal (g)	29.76	42.65	56.77	
Total Fat (g)	63.30	78.83	100.45	
SFA (g)	17.61	23.93	33.77	
MUFA (g)	24.59	34.11	43.83	
PUFA (g)	11.12	13.79	16.39	
Cholesterol (mg)	206.19	241.42	323.04	
Total Carbohydrates (g)	144.80	167.13	191.28	
Sugar digeribles (g)	63.46	72.53	84.33	
Starch digeribles (g)	67.90	84.15	104.74	
Fibre (g)	21.54	32.19	43.03	
Ethanol (g)	1.28	3.97	6.54	
Sodium (mg)	1219.28	1415.53	2280.02	
Potassium (mg)	3063.52	3889.96	5041.34	
Calcium (mg)	598.84	821.05	1170.91	
Magnesium (mg)	294.40	398.96	499.31	
Phosphorus (mg)	1043.55	1318.31	1527.30	

	sFFQ2p			Mean 3-24HR		
	P25	Median	P75	P25	Median	P75
	103.25	171.00	200.00	169.13	243.87	322.42
	200.00	200.00	400.00	115.17	249.67	426.58
	1.50	6.30	15.00	0.00	8.17	21.92
	103.51	147.30	204.78	167.42	210.42	326.54
	56.40	101.30	145.20	94.58	159.33	300.83
	47.70	69.30	103.01	57.08	103.83	132.00
	18.20	44.20	54.60	4.00	36.04	71.67
	31.50	31.50	31.50	0.00	20.83	45.00
	1.50	12.80	23.50	2.96	19.42	40.00
	15.75	30.00	30.75	15.17	19.25	23.13
	6.10	12.55	20.10	5.17	17.17	40.92
	2.10	2.60	5.18	0.00	6.83	22.50
	73.75	118.75	157.50	92.50	169.00	312.92
	21.50	60.75	90.30	0.00	10.00	95.67
	0.10	2.10	5.00	2.42	6.88	16.58
	1398.76	1584.75	1818.14	1408.98	1755.57	2085.33
	633.36	769.23	891.99	726.31	966.06	1142.25
	59.21	73.12	83.70	63.02	76.10	96.86
	24.76	28.77	37.73	17.78	21.83	26.02
	30.37	41.45	52.14	33.07	46.19	53.60
	66.57	74.56	90.48	57.26	74.00	85.75
	16.95	23.39	28.60	17.66	22.24	33.76
	26.21	34.32	40.10	19.18	25.17	33.14
	10.76	13.28	17.32	7.77	10.41	12.59
	191.16	241.46	307.08	185.73	238.17	295.73
	131.34	154.29	181.50	144.15	181.29	246.60
	49.44	63.94	85.57	61.74	78.62	107.77
	59.27	80.80	103.62	56.33	84.51	102.89
	20.83	31.90	42.38	18.35	21.34	24.48
	1.28	3.79	6.14	0.00	1.44	4.86
	1200.29	1459.85	1773.56	1514.50	1938.01	2659.35
	3041.59	3956.77	4574.09	2363.45	2790.57	3303.51
	531.45	748.98	1067.36	513.81	654.92	922.47
	304.18	380.38	458.41	215.69	255.98	299.33
	1071.50	1161.23	1444.37	836.60	1055.43	1243.34

Food group (g/d)	sFFQ1p P25	sFFQ1p Median	sFFQ1p P75	
Energy and nutrients				
Iron (mg)	12.86	15.02	18.27	
Zinc (mg)	10.76	12.97	17.17	
Vitamin A (mcg.e.r)	516.35	695.63	964.28	
Total retinoids (mcg)	206.54	279.74	410.10	
Total carotenoids (mcg)	1763.43	2359.07	3631.77	
Vitamin D (mcg)	2.11	3.40	3.90	
Vitamin E (mg.e.t)	9.40	10.80	13.96	
Vitamin B1 (mg)	1.12	1.35	1.51	
Vitamin B2 (mg)	1.14	1.39	1.76	
Niacin (mg)	18.55	24.26	29.32	
Vitamin B6 (mg)	1.53	1.94	2.23	
Folic (mcg)	300.81	386.68	442.36	
Vitamin B12 (mcg)	5.10	6.31	8.31	
Vitamin C (mg)	105.21	172.18	197.86	

Note. sFFQp, short food frequency questionnaire pilot; 24HR, 24-h dietary recall.

The relative validation of the sFFQ2 using various statistical tests and the means of 3-24HR as a reference method is shown in Table 7. For the food group, the Spearman correlation values were between 0.33 (nuts and fish and shellfish) and 0.78 (meats and products meat) with a median of 0.45. Correlations were ≥ 0.5 for meats and meat products, sausages, milk and dairy products, pastries and cakes, and sauces and condiments. The ICC presented values between 0.06 (sauces and condiments) and 0.91 (milk and dairy products) with a median of 0.43. Milk and dairy products, sausages, meat and meat products, fruits, cereals and grains, oils and fats, vegetables, cakes and cake, fish and shellfish presented values higher than 0.4. The CCC values were between 0.06 (sugar and sweet) and 0.88 (sausages), with a median of 0.33.

For energy and nutrients, Spearman correlation values were between -0.07 (vitamin D) to 0.78 (animal protein) with a median of 0.36 and between -0.02 (vitamin B_6) and 0.77 (animal protein) with a median of 0.37 for energy-unadjusted and -adjusted values, respectively. Correlations adjusted for energy were ≥ 0.5 for animal protein, cholesterol, and MUFA. ICC values not energy-adjusted ranged between 0.00 (vitamin D) and 0.68 (animal protein) with a median of 0.34. The energy-adjusted values were between 0.00 (vitamin D) and 0.80 (animal protein), with a median of 0.35. Energy-adjusted ICC values greater than 0.4 were for animal protein, cholesterol, MUFA, total carbohydrates,

	sFFQ2p			Mean 3-24HR		
	P25	Median	P75	P25	Median	P75
	12.28	13.80	16.83	8.05	10.23	11.46
	10.16	12.50	15.14	5.80	7.10	8.95
	529.16	666.03	806.98	562.06	775.22	1109.84
	186.79	308.76	381.16	133.26	220.01	293.05
	1662.91	2278.04	2858.95	2078.59	3588.28	5216.17
	1.87	3.36	4.00	0.87	1.47	4.13
	8.91	10.65	12.92	6.55	7.72	11.01
	1.06	1.26	1.49	0.81	1.12	1.38
	1.04	1.27	1.64	0.92	1.17	1.63
	16.91	23.65	30.72	12.94	15.47	22.24
	1.40	1.78	2.00	1.58	1.84	2.14
	263.67	340.17	403.26	254.24	283.62	371.69
	4.43	6.04	7.20	2.40	3.36	5.08
	94.64	142.25	187.48	111.33	179.07	235.08

water, zinc, digestible sugar, ethanol, calcium, folic acid, and total protein. Using CCC, energy-unadjusted values were from 0.15 to 0.67 (median of 0.29) and adjusted values were from 0.19 to 0.70 (median of 0.33).

The reproducibility evaluation of the two sFFQp (sFFQ1p vs sFFQ2p) using Spearman's rank correlation, ICC and CCC, is shown in Annex 5: Table of reproducibility pilot sFFQ. The Spearman correlation for the food group ranged from 0.56 (milk and dairy products) to 0.93 (alcoholic beverage) with a median of 0.78. The ICC values were between 0.53 (sauces and condiments) and 0.98 (meats and meat products), with a median of 0.75. Meanwhile, the CCC values presented a range from 0.53 (sauces and condiments) to 0.98 (meats, meats and products) with a median of 0.75. The Spearman correlation values for energy and nutrients not adjusted for energy ranged from 0.51 for Vitamin B_1 to 0.93 for Vitamin B_{12}, with a median of 0.70. Meanwhile, the adjusted values for energy ranged between 0.59 (digestible sugar) and 0.95 (Vitamin B_{12}), with a median of 0.80. The ICC for values not adjusted for energy was between 0.51 to 0.93, with a median of 0.70. While, for energy-adjusted nutrients, the values were between 0.57 and 0.95, with a median of 0.80. The CCC for unadjusted values were between 0.56 to 0.93 (median of 0.68), and for adjusted energy values, CCC ranged from 0.54 to 0.95 (median of 0.75).

Table 7. Relative validity of sFFQp: correlation between food groups, energy and nutrients intake between sFFQ2p and mean 3-24HRs.

Food groups (g/d)	Wilcoxon signed-rank test	Spearman's rank coefficient	
		Un-adjusted	Energy-adjusted
Alcoholic beverage	0.02	0.35	-
Cereals or grains	0.00	0.46	-
Fish and seafood	0.38	0.33	-
Fruits	0.95	0.41	-
Legumes	0.00	0.39	-
Meat and meat products	0.04	0.78	-
Milk and dairy products	0.00	0.66	-
Non alcoholic beverage	0.38	0.37	-
Nuts	0.38	0.33	-
Oils and fat	0.00	0.47	-
Pastries and cake	0.88	0.55	-
Sauces and condiments	0.50	0.50	-
Sausage	0.11	0.77	-
Sugar and sweet	0.00	0.45	-
Vegetables	0.00	0.43	-
Energy and nutrient			
Energy (kcal)	0.23	0.39	-
Water (g)	0.00	0.45	0.47
Total Protein (g)	0.34	0.35	0.43
Vegetal protein (g)	0.00	0.32	0.17
Animal protein (g)	0.38	0.78	0.82
Total Fat (g)	0.18	0.39	0.37
SFA (g)	0.99	0.41	0.23
MUFA (g)	0.00	0.56	0.59
PUFA (g)	0.00	0.35	0.38
Cholesterol (mg)	0.80	0.68	0.71
Total Carbohydrates (g)	0.00	0.45	0.47
Sugar digeribles (g)	0.02	0.31	0.43
Starch digeribles (g)	0.80	0.35	0.06
Fibre (g)	0.00	0.37	0.30
Ethanol (g)	0.01	0.46	0.44

Intraclass Correlation Coeficient (ICC)		Concordance Correlation Coefficient (CCC)	
Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
0.38	-	0.33	-
0.58	-	0.27	-
0.41	-	0.48	-
0.65	-	0.52	-
0.19	-	0.15	-
0.77	-	0.80	-
0.91	-	0.62	-
0.28	-	0.26	-
0.17	-	0.23	-
0.46	-	0.45	-
0.43	-	0.43	-
0.06	-	0.07	-
0.87	-	0.88	-
0.12	-	0.06	-
0.43	-	0.26	-
0.39	-	0.38	-
0.45	0.47	0.35	0.46
0.35	0.42	0.34	0.40
0.32	0.17	0.21	0.16
0.77	0.80	0.76	0.80
0.39	0.37	0.39	0.33
0.42	0.22	0.41	0.22
0.56	0.59	0.46	0.57
0.34	0.38	0.26	0.36
0.68	0.71	0.67	0.70
0.45	0.47	0.38	0.44
0.30	0.43	0.27	0.42
0.34	0.06	0.33	0.05
0.36	0.29	0.26	0.28
0.44	0.43	0.37	0.42

Food groups (g/d)	Wilcoxon signed-rank test	Spearman's rank coefficient	
		Un-adjusted	Energy-adjusted
Energy and nutrient			
Sodium (mg)	0.00	0.54	0.37
Potassium (mg)	0.00	0.36	0.20
Calcium (mg)	0.24	0.45	0.44
Magnesium (mg)	0.00	0.32	0.29
Phosphorus (mg)	0.00	0.26	0.05
Iron (mg)	0.00	0.13	0.09
Zinc (mg)	0.00	0.34	0.45
Vitamin A (mcg.e.r)	0.05	0.47	0.47
Total retinoids (mcg)	0.00	0.44	0.35
Total carotenoids (mcg)	0.00	0.38	0.37
Vitamin D (mcg)	0.02	-0.07	0.00
Vitamin E (mg.e.t)	0.00	0.17	0.27
Vitamin B1 (mg)	0.02	0.26	0.10
Vitamin B2 (mg)	0.32	0.26	0.32
Niacin (mg)	0.00	0.31	0.28
Vitamin B6 (mg)	0.38	0.09	-0.02
Folic (mcg)	0.03	0.33	0.45
Vitamin B12 (mcg)	0.00	0.38	0.38
Vitamin C (mg)	0.14	0.27	0.30

Green colour code: good or acceptable outcome based on at least three statistical tests. Red colour code: poor outcome, did not pass more than two statistical tests. (Lombard *et al.* (51)). Energy adjustment by residual methods. Student paired t-test or Wilcoxon signed-rank test at $p<0.05$.

The Bland & Altman analysis confirmed the overestimation of sFFQ2p as compared to the mean 3-24HR for fish and shellfish, legumes, oil and fats, and alcoholic beverages. While for energy and nutrients, sFFQ2p tended to overestimate its values by 58% compared with the 24HRs. For most food groups, we observed that the mean error in the intake values by the sFFQ2p changed the dietary intake increase. For fruits, nuts, fish and shellfish, legumes, oils and fats, cakes and cakes, sauces and condiments, sugar and sweets, the low intakes were overestimated, and the sFFQp underestimated the high intake. While for nutrients, this proportion bias was observed only for vegetable protein, fibre,

Intraclass Correlation Coeficient (ICC)		Concordance Correlation Coefficient (CCC)	
Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
0.54	0.35	0.46	0.34
0.34	0.19	0.24	0.19
0.45	0.43	0.44	0.43
0.31	0.27	0.17	0.26
0.26	0.05	0.21	0.05
0.13	0.09	0.06	0.08
0.34	0.45	0.15	0.43
0.42	0.35	0.38	0.35
0.43	0.36	0.36	0.35
0.34	0.32	0.28	0.31
0.00	0.01	-0.06	0.00
0.16	0.26	0.12	0.26
0.26	0.08	0.23	0.08
0.26	0.31	0.26	0.30
0.31	0.27	0.22	0.27
0.06	0.00	0.05	-0.01
0.33	0.43	0.30	0.42
0.38	0.39	0.34	0.38
0.27	0.30	0.25	0.29

sodium, potassium, calcium, magnesium, zinc, vitamin A and total carotenoids (Annex 6: Differences in daily intake between pilot sFFQ2 and mean 3-24HRs based on Bland & Altman plots).

5.3 DESIGN AND DATA TREATMENT OF THE FINAL sFFQ FROM THE RESULTS OF THE PILOT STUDY

In this section, we will explain the re-design of the FFQ from the results of the pilot questionnaire. Also, we will present the considerations that we took into consideration in the treatment of the data for the final validation of the sFFQ.

5.3.1 IDENTIFICATION OF ISSUES RELATED TO THE DESIGN OF THE PILOT QUESTIONNAIRE

From the results obtained from the sFFQp, we observe several errors in its design that we tried to correct to improve the final sFFQ, as detailed below.

1. When using the sFFQp, the participants reported difficulties estimating the questionnaire's standard food serving size since it was described in grams and not in household measures. To reduce this problem;
    - We incorporated food photographs for several items.
    - We defined the food standard serving sizes in household measures.
    - We added three consumption alternatives to the standard serving sizes.
2. Participants did not report consuming two items of processed milk (item 18) and viscera (item 22) using the sFFQp and during the 24HR interview. We compared these findings with the consumption data from the ENALIA2 survey, where the values were also relatively low. Therefore, we decided to remove these two items from the final sFFQ.
3. The consumption of sweeteners between the sFFQp (item 46) and 24HR were not concordant. The participants indicated not to consume them in the sFFQp but reported eating them in the 24HRs. For example, we realised that the participants

misreported food containing sweeteners during the 24HR interviews. Indeed, the presence of sweeteners could be detected only after we checked the nutritional labels and ingredients of the products from the commercial brands.

Consequently, we decided to remove this item from the final sFFQ list, as it provided measurement errors. Still, we replaced it with the following question: "Do you eat foods containing sweeteners like..... yes or no?" at the beginning of the questionnaire.

4. Many participants reported consuming plant-based drinks not included in the sFFQp. Therefore, we add this item to the new sFFQ.

5. The high difference in energy and fat content of milk, cheese and yoghurt did not allow us to convert these foods reported by the 24HR into items 15 (high-fat dairy products, such as whole milk and cheese) and 16 (low-fat dairy products, such as skim milk and low-fat cheese) of the sFFQp. Accordingly, we split them into six different items in the final sFFQ.

6. The low agreement between sFFQp and 24HR in fibre intake led us to regroup items 1 and 2 (of the pilot questionnaire) of vegetables into 12 different food groups since their fibre content varied considerably.

7. The wide variety of new foods and the few available nutritional information (micronutrients) make it challenging to choose a food with a similar nutritional profile in the FCDB. Therefore, we decided to replace the CESNID nutritional composition table used in the pilot study with our in-house database (see section 5.1.3 of the method). The CESNID table was insufficient to provide the nutritional composition data of the 491 foods registered by the 24HRs.

The results of the relative validation of the sFFQp were not satisfactory to validate the questionnaire since the various statistical tests used did not evaluate the different facets of validity of the questionnaire. In addition to the issues in the design of the sFFQp, commented previously, the statistical tests showed low correlation and agreement between sFFQp and 24HRs based on the most used interpretation criteria for validation (51). However, these data helped us improve our questionnaire. Moreover, we used the information provided by the 24HR to know the usual intake of our study population and thus enhance the design of this study.

To adapt and replace possible deficiencies in our food list and increase the external validity of the final sFFQ, the 491 foods reported by the 24HR from the pilot study were compared and cross-checked with the data from the ENALIA2 survey. We selected 310 foods based on their highest intake within the population (Figure 16), greater intra- and inter-individual consumption variability, and potential relationship with the gut microbiome as commented in sections 5.1.1 and 5.1.2 of methods.

In summary, the final sFFQ food list comprises 310 foods, classified into 58 items and grouped into 24 food groups.

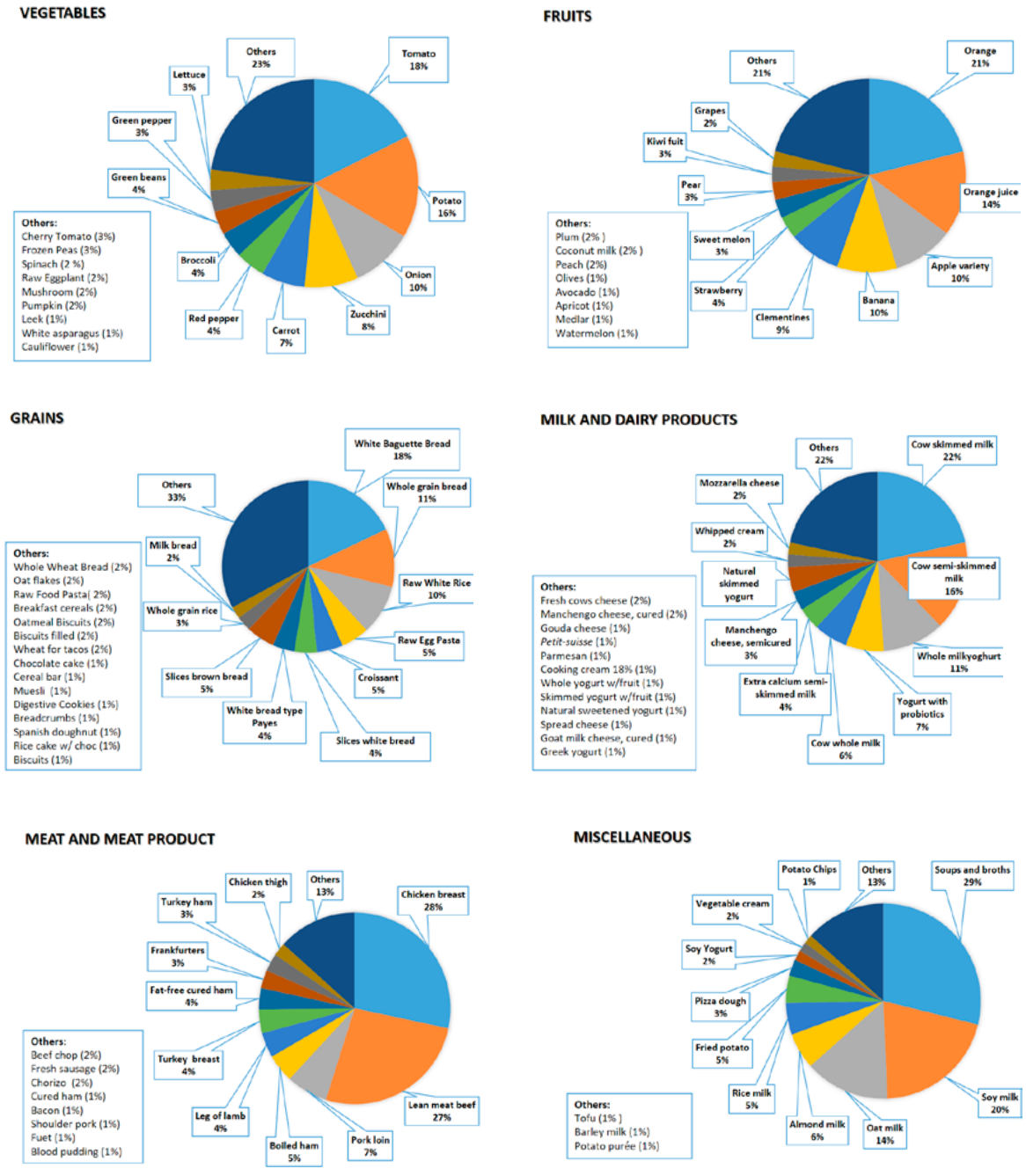


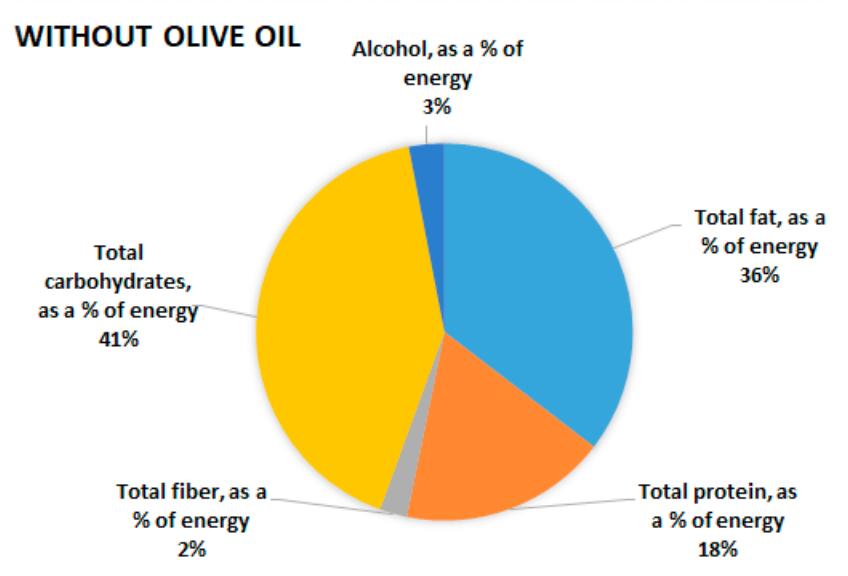
Figure 16. Foods (%) most consumed by the population from the pilot study were grouped into vegetables, fruits, cereals, meat and meat products, milk and dairy products and miscellaneous. Values obtained from 24HRs. Foods that contributed to <1% of the food group are not shown.

5.3.2 DATA TREATMENT OF THE RELATIVE VALIDATION OF THE FINAL sFFQ

While interpreting the results from the pilot study, two questions arose about analysing the data obtained from the sFFQp and the 24HR before performing the statistical analyses of the relative validation of the final sFFQ.

Will it be necessary to include olive oil consumption according to the intake of each individual, or do we assign an average intake value for all participants?

Dietary data obtained from the pilot study showed that 60% of the individuals exceeded the fat intake recommendations (by 25 to 35%) (259). Since olive oil contributes significantly to lipid intake, we wonder if these values were due to an erroneous quantification of olive oil intake during the 24HRs. To test this hypothesis, we evaluated the contribution of olive oil to total fat intake as a percentage of total kcal. In addition, we calculated the nutritional composition of the data obtained from the mean 3-24HRs of the final study, with and without olive oil.



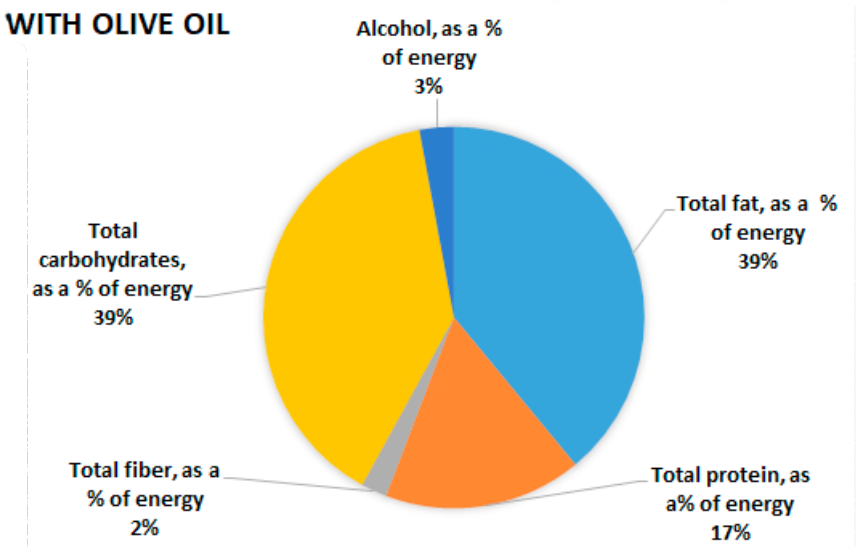


Figure 17. Contribution of olive oil to total fat intake as a percentage of total calories from 24HRs of the final study.

As a result, we observed that the olive oil intake, when it is included or not modified significantly total energy ($p<0.0001$; $p=0.05$) and total fat ($p<0.0001$; $p=0.05$). However, it would only explain 3% of the total fat consumption of the population in the final study (Figure 17). Therefore, we decided to quantify the olive oil according to each participant's intake and not assign an average consumption value as was performed in another study (224).

How do we calculate the nutritional composition of each item of the sFFQ?

Each sFFQ requires a database to estimate energy and nutrient intake. However, to our knowledge, there are two possible approaches to calculate nutritional composition (223,225). We calculated the unweighted (UWM) and weighted mean (WM) of energy and nutrients for each item of the sFFQ (see section 5.2.4 of methods) to evaluate differences between the nutritional value of sFFQ-UNW, sFFQ-WM and 24HR (reference methods).

As a result, we observed significant differences ($p < 0.05$) between sFFQ-UWN and sFFQ-WN nutritional composition values for all nutrients except phosphorus. While comparing the energy and nutrient values of the 24HRs with the values obtained from sFFQ-UWN and sFFQ-WN, the differences were significantly different for 60% and 43% of the nutrients, respectively (Table 8). Therefore, we decided to use the sFFQ-WM nutritional composition since fewer differences with the 24HRs values. In addition, when calculating the nutritional composition, we gave more or less strength to the food depending on its intake in our population.

Table 8. Means and standard deviations of three 24HRs compared with nutritional composition (weighted and unweighted) of the sFFQ2.

	Mean 3-24HR		Nutritional Composition sFFQ2 Weighted (FFQ-WM)			Nutritional Composition sFFQ2 Unweighted (FFQ-UWM)		
	Mean	SD	Mean	SD	p*	Mean	SD	p*
Energy (kcal)	1832.70	417.52	1621.41	534.72	0.09	1689.25	550.51	0.30
Total fat (g)	81.13	31.05	60.64	28.04	< 0.0001	67.42	29.42	0.02
Total protein (g)	77.58	30.63	77.60	37.46	0.41	77.19	36.92	0.47
Total water (g)	1211.89	371.13	1274.95	594.00	0.04	1261.78	587.25	0.06
Total fiber (g)	19.76	7.04	23.02	10.86	0.01	24.63	11.26	0.00
Total carbohydrates (g)	179.08	48.74	171.20	67.55	0.83	173.06	66.80	0.70
MUFA (g)	34.40	12.02	25.29	11.63	< 0.0001	28.57	12.44	0.01
PUFA (g)	13.09	5.57	9.88	4.51	0.00	10.07	3.94	0.00
SFA (g)	27.69	13.79	20.38	13.38	0.00	22.99	14.50	0.05
Alcohol (g)	7.65	10.67	7.54	8.66	0.50	8.95	10.30	0.04
Cholesterol (mg)	323.48	208.27	261.43	152.67	0.01	268.91	154.85	0.03
Vitamin A (µg retinol eq)	843.78	546.46	895.53	532.99	0.61	1293.77	748.38	< 0.0001
Vitamin D (µg)	5.71	8.41	8.49	9.57	< 0.0001	11.72	11.63	< 0.0001
Vitamin E (mg α-tocoferol)	10.76	5.05	11.03	5.36	0.36	12.55	5.65	0.01
Biotine (µg)	3.02	5.60	1.38	1.25	0.60	1.90	1.26	0.33
Folate (µg)	295.84	108.74	397.50	175.87	< 0.0001	427.72	183.12	< 0.0001
Eq Niacin (mg)	31.99	12.79	71.31	59.32	< 0.0001	54.24	37.29	< 0.0001
Vitamin B5 (mg)	0.66	1.06	1.16	1.10	< 0.0001	1.29	1.07	< 0.0001
Riboflavin (mg)	1.89	0.88	1.85	0.82	0.81	1.96	0.87	0.27
Tiamin (mg)	1.27	0.43	1.29	0.51	0.20	1.32	0.49	0.12
Vitamin B12 (µg)	24.96	71.93	24.24	71.61	0.98	25.81	71.31	0.02
Vitamin B6 (mg)	2.35	1.60	2.24	0.90	0.08	2.15	0.86	0.28
Vitamin C (mg)	129.99	74.95	180.93	94.98	< 0.0001	167.85	84.40	0.00
Calcium (mg)	766.14	348.38	869.87	616.52	0.04	883.01	594.08	0.02
Iron (mg)	14.05	3.96	13.85	5.59	0.90	15.22	6.35	0.16
Potassium (mg)	3064.32	852.25	3244.75	1202.74	0.10	3347.08	1255.50	0.03
Magnesium (mg)	387.41	215.31	403.96	161.95	0.05	413.50	164.04	0.03
Sodium (mg)	2309.00	854.52	1441.17	718.07	< 0.0001	1532.68	758.31	< 0.0001
Phosphorus (mg)	1301.88	437.24	1366.22	590.62	0.07	1370.80	583.28	0.06
Iodine (µg)	114.08	128.45	105.59	56.61	0.40	104.20	55.92	0.42
Selenium (µg)	82.01	28.74	67.23	28.58	0.00	77.89	34.23	0.54
Zinc (mg)	8.51	2.82	8.29	3.25	0.89	8.20	3.20	0.80

p<0.05 between sFFQ2 (unweighted and weighted) and mean 24HRs using paired t-test and Wilcoxon test.

5.4 RELATIVE VALIDITY AND REPRODUCIBILITY ANALYSIS OF THE FINAL sFFQ

Dietary data extracted from the 24HRs were classified into the 58 food items, as they are presented in the sFFQ to compare and validate the questionnaire. The 58 food items were also grouped into 24 food groups, total energy and 29 nutrients. The participants spent an average of 22 min answering the sFFQ (SD=16.2 min, max=86 min, min=4 min).

The daily intake of food groups, energy and macro- and micronutrients from sFFQ1, sFFQ2, and mean 3-24HR are shown in the 25th, 50th and 75th percentiles in Table 9.

For the food group, the sFFQ2 tended to report higher consumption of appetiser and vegetables than the mean 3-24HR. While for energy and nutrients, sFFQ2 showed an underestimation of energy intake, total fat, PUFA, SFA, cholesterol, sodium and selenium than the mean 3-24HR.

Table 9. Daily consumption of food, energy and nutrients was estimated by two sFFQ (sFFQ1 and sFFQ2) and a mean of 3-24HR.

	sFFQ1		
Food group (g/d)	P25	Median	P75
Alcoholic beverage	18.56	84.97	168.20
Appetizers	3.30	3.30	10.50
Biscuits breakfast cereals and cereal bars	1.73	10.50	25.00
Chocolates and derivatives	0.79	2.52	7.26
Fats and oils	8.12	10.92	13.75
Fish and shellfish	27.39	51.06	67.49
Fruit and fruit products	47.57	184.30	216.50
Legumes	19.80	43.88	57.71
Meats and eggs	42.32	69.30	136.20
Milk and dairy products except fermented milk	30.71	79.48	224.40
Non Alcoholic beverage	63.83	150.00	260.60
Nuts and seeds	1.24	1.98	15.00
Pastries and sweets breads	0.00	3.30	3.30
Potatoes and other tubers	9.90	25.65	31.50
Ready to eat meals	0.00	13.20	19.80
Sauces and condiments	2.64	5.52	12.21
Sausages and other meat products	2.97	5.33	9.45
Sugar and other sweets	0.00	2.37	8.40
Vegetables and vegetable products	204.20	295.30	436.60
White bread	5.30	32.20	44.80
White grains and white pastas	16.85	25.31	45.24
Wholegrain or wholemeal bread	0.00	4.62	31.85
Wholemeal grains and wholemeal pastas	0.00	0.00	7.95
Yogurt and fermented milk	8.25	26.25	80.00
Energy and nutrients			
Energy (kcal/d)	1262.00	1588.00	2014.00
Total fat (g/d)	41.53	56.13	71.78
Total protein (g/d)	56.60	72.71	93.31

	sFFQ2			Mean 3-24HRs		
	P25	Median	P75	P25	Median	P75
	22.02	101.30	208.40	0.00	110.00	305.80
	2.06	4.28	10.50	0.00	0.00	8.33
	2.06	10.50	18.60	0.00	15.42	40.00
	0.79	2.52	7.68	0.00	3.83	16.17
	6.50	10.53	20.48	7.54	12.00	16.67
	27.39	43.86	69.18	6.75	20.28	76.33
	53.80	139.60	259.90	32.50	158.70	397.00
	20.03	31.50	54.68	4.83	27.50	49.58
	40.88	69.30	151.50	24.58	99.17	156.70
	34.68	83.73	218.60	45.83	130.80	191.50
	65.48	156.60	335.20	111.70	231.70	366.90
	0.99	5.13	11.85	0.00	8.00	20.00
	0.00	3.30	10.50	0.00	8.92	35.64
	9.90	31.50	43.88	0.00	16.67	66.67
	1.65	13.20	20.70	0.00	47.50	130.10
	2.64	4.08	12.69	5.17	13.70	33..82
	2.97	5.33	9.45	0.00	28.33	58.83
	0.17	2.37	9.60	0.00	3.33	12.50
	193.90	278.50	427.40	107.40	174.50	254.00
	7.35	14.70	44.80	7.50	35.33	65.04
	16.85	25.31	51.36	10.67	28.33	65.50
	0.58	5.99	20.48	0.00	0.00	18.42
	0.00	1.32	5.48	0.00	0.00	0.00
	8.25	26.25	80.00	0.00	41.67	83.33
	1285.00	1571.00	1877.00	1546.00	1761.00	2059.00
	39.89	52.14	74.85	60.21	74.08	95.80
	53.11	71.31	86.65	58.00	72.02	85.13

	sFFQ1		
Food group (g/d)	P25	Median	P75
Energy and nutrients			
Total water (g/d)	861.20	1094.00	1593.00
Total fiber (g/d)	17.07	21.63	31.59
Total carbohydrates (g/d)	120.80	155.60	222.00
Alcohol (g/d)	1.75	4.64	8.43
MUFA (g/d)	17.61	21.47	31.94
PUFA (g/d)	6.89	9.60	11.71
SFA (g/d)	12.97	18.72	22.85
Cholesterol (mg/d)	157.00	242.20	335.40
Vitamin A µg retinol (eq/d)	552.50	738.50	1066.00
Vitamin D (µg/d)	3.12	5.15	9.42
Vitamin E (mg α tocoferol/d)	8.23	9.50	12.92
Folate total (µg/d)	284.90	390.20	520.70
Total niacin equivalent (mg/d)	35.14	50.10	99.30
Riboflavin (mg/d)	1.17	1.86	2.32
Tiamin (mg/d)	1.00	1.22	1.69
Vitamin B12 (µg/d)	3.59	4.74	6.55
Vitamin B6 (mg/d)	1.72	2.18	2.88
Vitamin C ascorbic acid (mg/d)	119.00	163.90	245.30
Calcium (mg/d)	554.40	742.20	1066.00
Iron (mg/d)	9.99	13.74	17.81
Potassium (mg/d)	2477.00	3244.00	3986.00
Magnesium (mg/d)	273.10	399.50	521.50
Sodium (mg/d)	951.30	1388.00	1652.00
Phosphorus (mg/d)	1015.00	1300.00	1665.00
Iodine (µg/d)	67.33	96.10	124.50
Selenium (µg/d)	45.60	62.48	84.33
Zinc (mg/d)	6.48	7.98	9.80

sFFQ, food frequency questionnaire; 24HR, 24-h dietary recall

	sFFQ2			Mean 3-24HRs		
	P25	Median	P75	P25	Median	P75
	802.40	1153.00	1584.00	937.20	1184.00	1492.00
	14.25	18.08	27.24	13.47	18.80	24.86
	115.70	172.70	201.00	152.10	176.40	212.30
	1.56	5.38	10.40	0.00	3.52	10.56
	16.98	21.56	36.02	26.77	32.33	38.20
	6.68	8.94	11.79	9.21	11.46	16.98
	12.98	17.52	25.09	19.65	24.80	33.36
	175.30	233.10	279.50	175.30	300.90	437.50
	580.40	724.40	1091.00	542.90	750.30	1010.00
	3.26	5.72	8.75	1.81	3.27	5.4
	7.45	9.17	13.98	7.29	9.63	11.90
	265.00	366.10	454.80	217.40	274.10	366.90
	35.77	50.04	91.34	23.09	28.89	39.53
	1.29	1.59	2.28	1.34	1.70	2.08
	0.88	1.16	1.56	0.91	1.15	1.52
	3.56	4.49	6.29	2.85	4.10	7.49
	1.72	2.09	2.45	1.52	1.96	2.52
	115.10	153.80	241.20	69.64	125.40	166.20
	488.60	688.30	1120.00	521.00	702.50	920.10
	9.49	12.30	15.27	11.35	13.46	16.32
	2405.00	2976.00	3788.00	2379.00	2967.00	3679.00
	257.50	379.20	466.00	280.40	348.90	432.80
	965.20	1331.00	1704.00	1825.00	2053.00	2610.00
	1010.00	1248.00	1587.00	1056.00	1236.00	1384.00
	74.65	93.11	124.10	70.93	94.12	113.50
	49.62	61.61	76.66	65.20	82.62	89.66
	5.59	7.69	9.51	6.52	7.90	10.04

The relative validity of the sFFQ2 compared with the 24HRs is shown in Table 10. The means intake of fifteen groups of food (out of 24) did not show significant differences between both methods. The Spearman correlation values were between 0.18 (appetisers) and 0.78 (non-alcoholic beverage) with a median of 0.42. 45.8% of the food groups had correlation values higher than 0.5. The ICC presented values between 0.05 (ready to eat meals) and 0.83 (alcoholic beverage) with a median of 0.36. 37.5% of the food groups had correlation values higher than 0.4.

For energy and nutrients, we observed significant differences only in 10 of 30 nutrients evaluated by 24HR and sFFQ2. The Spearman correlation values for unadjusted and energy-adjusted dietary data were from 0.01 (sodium) to 0.66 (alcohol) with a median of 0.42 and between 0.10 (sodium) and 0.71 (total carbohydrates) with a median of 0.46, respectively. 44.8% of the nutrient values adjusted for energy had correlation values ≥ of 0.5. ICC values for not adjusted data ranged from 0.05 (iodine) to 0.99 (B_{12} vitamin) with a median of 0.44. Meanwhile, the ICC values energy-adjusted data were between 0.08 (selenium) and 0.98 (B_{12} vitamin), with a median of 0.55, with 72% of the ICC values being higher than 0.4

The Bland & Altman analysis confirmed the underestimation and overestimation of some foods and nutrients data collected from the sFFQ2 compared with the 24HR. The 95% confidence intervals of the mean difference (sFFQ2 – mean 3-24HR) did not reach the line of equality (y=0) (Table 10); see the value of the mean difference of the 95% CI in the Bland & Altman analysis). The underestimation was observed for biscuits, breakfast cereals and cereals bars, chocolates and derivatives, pastries and sweets bread, ready-to-eat meals, sauces and condiments, sausages, energy, total fat, PUFA, SFA, cholesterol, sodium and selenium. The overestimation was detected for appetisers, vegetables, vitamin D, total folate, total niacin and vitamin C.

Moreover, we observed a systematic variation in agreement (proportional bias) between the two methods for 9 of the 24 food groups. The sFFQ overestimated low intakes and underestimated high intakes of biscuits, breakfast cereals and cereals bars, chocolates and derivatives, pastries and sweet bread, potatoes and other tuber, sauces and condiments, sausages and other meat products, sugar and other sweets, vegetables and vegetable products, white bread and wholegrain bread (Annex 7: Differences in daily intake between final sFFQ2 and mean 3-24HRs based on Bland & Altman plots).

While, for energy and nutrients, the sFFQ2 underestimated low intakes and overestimated high consumption of water, carbohydrates, fibre, folic acid, niacin, vitamin C, calcium, potassium and phosphorus. At the same time, it overestimated low intakes and underestimated high cholesterol intakes (Figure 18).

Table 10. Relative validation of sFFQ2 for food group, energy and nutrients, with the mean 3-24HRs as the reference method.

Food group (g/d)	Wilcoxon signed-rank test	Bland & Altman analysis		
	p value	Mean Difference	[95% CI]	Lower LOA[a]
Alcoholic beverage	0.15	-10.87	[-46.08, 24.34]	-237.90
Appetizers	<0.0001	3.62	[0.74, 6.51]	-14.99
Biscuits breakfast cereals and cereals bars	0.59	-8.85	[-15.03, -2.68]	-48.65
Chocolates and derivatives	0.21	-2.90	[-5.59, -0.21]	-20.25
Fats and oils	0.61	1.72	[-1.39, 4.83]	-18.35
Fish and shellfish	0.01	14.21	[-8.56, 36.98]	-132.5
Fruit and fruit products	0.06	3.98	[-62.58, 70.54]	-425.10
Legumes	0.04	2.59	[-12.93, 18.11]	-97.46
Meats and eggs	0.30	-4.41	[-26.35, 17.53]	-145.80
Milk and dairy products except fermented milk	1.00	29.64	[-17.16, 76.44]	-272.00
Non Alcoholic beverage	0.46	32.27	[-77.41, 141.95]	-674.80
Nuts and seeds	0.10	-3.99	[-8.76, 0.79]	-34.75
Pastries and sweets breads	0.67	-17.67	[-28.68, -6.66]	-88.63
Potatoes and other tubers	0.02	-7.02	[-23.98, 9.95]	-116.40
Ready to eat meals	0.04	-68.08	[-102.41, -33.75	-289.30
Sauces and condiments	<0.0001	-15.03	[-22.10, -7,96]	-60.57
Sausages and other meat products	0.22	-26.90	[-36.55, -17.25]	-89.09
Sugar and other sweets	0.97	-4.05	[-9.73, 1.64]	-40.65
Vegetables and vegetable products	<0.0001	134.30	[68.56, 200.04]	-289.50
White bread	0.22	-6.87	[-16.86, 3.12]	-71.25
White grains and white pastas	0.95	-8.13	[-18.75, 2.49]	-76.55
Wholegrain or wholemeal bread	0.01	3.75	[-4.38, 11.89]	-48.68
Wholegrain cereals	0.00	-1.99	[-7.04, 3.06]	-34.55
Yogurt and fermented milk	0.59	-1.46	[-17.92, 15.00]	-107.60
Energy and nutrients				
Energy (kcal/d)	0.00	-218.10	[-388.59, -47.61]	-1317.00
Total fat (g/d)	<0.0001	-20.69	[-30.03, -11.35]	-80.88
Total protein (g/d)	0.35	-0.01	[-9.05, 9.04]	-58.30
Total water (g/d)	0.78	70.71	[-68.80, 210.22]	-828.60
Total fiber (g/d)	0.50	2.31	[-0.74, 5.36]	-17.34
Total carbohydrates (g/d)	0.16	-10.02	[-28.33, 8.29]	-128.10
Alcohol (g/d)	0.02	0.42	[-1.57, 2.42]	-12.44

Bland & Altman analysis		Spearman's rank coefficient		Intraclass Correlation Coeficient (ICC)		Cross-Classification	
Upper LOA[a]	Spearman correlation coefficient[b]	Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted	Same third	Opposite third
216.20	-0.26	0.76	-	0.83	-	72.73	2.27
22.23	0.20	0.18	-	0.27	-	25.00	9.09
30.94	-0.51	0.65	-	0.58	-	54.55	4.55
14.45	-0.37	0.44	-	0.38	-	54.55	9.09
21.79	0.32	0.35	-	0.22	-	29.55	9.09
161.10	0.26	0.35	-	0.25	-	38.64	18.18
433.11	-0.05	0.63	-	0.60	-	56.82	6.82
102.60	-0.11	0.18	-	0.25	-	36.36	20.45
137.00	-0.13	0.65	-	0.56	-	65.91	6.82
331.30	0.18	0.64	-	0.47	-	59.09	6.82
739.30	0.07	0.78	-	0.34	-	68.18	2.27
26.78	-0.26	0.58	-	0.49	-	63.64	11.36
53.29	-0.76	0.24	-	0.12	-	43.18	18.18
102.40	-0.38	0.28	-	0.12	-	52.27	22.73
153.20	-0.80	0.31	-	0.05	-	47.73	13.64
30.51	-0.72	0.46	-	0.31	-	52.27	9.09
35.29	-0.84	0.36	-	0.26	-	38.64	15.91
32.56	-0.39	0.57	-	0.33	-	63.64	11.36
558.10	0.37	0.30	-	0.40	-	35.09	13.64
57.51	-0.22	0.76	-	0.59	-	61.36	4.55
60.29	-0.40	0.58	-	0.47	-	54.55	6.82
56.18	0.01	0.24	-	0.20	-	29.55	15.91
30.57	0.27	0.45	-	0.47	-	59.09	15.91
104.70	-0.13	0.68	-	0.63	-	59.09	2.27
880.90	0.18	0.30	-	0.37	-	34.09	15.01
39.50	0.03	0.28	0.53	0.43	0.58	54.55	6.82
58.29	0.19	0.46	0.66	0.69	0.73	52.27	2.27
970.00	0.49	0.61	0.62	0.59	0.55	54.55	4.55
21.96	0.40	0.53	0.56	0.49	0.60	50.00	11.36
108.00	0.30	0.50	0.71	0.49	0.75	63.64	2.27
13.29	-0.19	0.66	0.71	0.79	0.75	61.36	4.55

Food group (g/d)	Wilcoxon signed-rank test p value	Bland & Altman analysis Mean Difference [95% CI]	Lower LOA[a]
Energy and nutrients			
MUFA (g/d)	<0.0001	-0.92 [-4.90, 3.05]	-34.86
PUFA (g/d)	<0.0001	-3.33 [-4.99, -1.67]	-14.01
SFA (g/d)	<0.0001	-7.15 [-10.85, -3.45]	-31.02
Cholesterol (mg/d)	0.06	-63.80 [-123.92, -3.68]	-451.40
Vitamin A µg retinol (eq/d)	0.64	31.37 [-160.20, 222.94]	-1203.00
Vitamin D (µg/d)	<0.0001	3.00 [0.77, 5.24]	-11.39
Vitamin E (mg α tocoferol/d)	0.87	0.07 [-1,65, 1.79]	-11.02
Folate total (µg/d)	0.00	88.10 [41.39, 134.81]	-213.10
Total niacin equivalent (mg/d)	<0.0001	37.17 [22.80, 51.54]	-55.47
Riboflavin (mg/d)	0.48	-0.04 [-0.29, 0.20]	-1.63
Tiamin (mg/d)	0.75	0.00 [-0.15, 0.16]	-1.01
Vitamin B12 (µg/d)	0.76	-0.63 [-2.39, 1.12]	-11.94
Vitamin B6 (mg/d)	0.63	-0.11 [-0,60, 0.38]	-3.28
Vitamin C ascorbic acid (mg/d)	0.00	50.48 [21.19,79.77]	-138.30
Calcium (mg/d)	0.48	84.52 [-28.93, 197.97]	-646.70
Iron (mg/d)	0.17	-0.56 [-2.49, 1.38]	-13.03
Potassium (mg/d)	0.95	132.60 [-224.99, 490.19]	-2173.00
Magnesium (mg/d)	0.52	10.58 [-48.14, 69.30]	-367.90
Sodium (mg/d)	<0.0001	-871.80 [-1165.38, -578.22]	-2764.00
Phosphorus (mg/d)	0.99	56.01 [-77.24, 189.26]	-802.80
Iodine (µg/d)	1.00	-8.98 [-51.61, 33.65]	-283.70
Selenium (µg/d)	0.00	-14.19 [-25.95, -2.43]	-89.97
Zinc (mg/d)	0.18	-0.31 [-1.20, 0.57]	-6.03

Green colour code: good or acceptable outcome based on at least three statistical tests. Red colour code: poor outcome, did not pass more than two statistical tests. (Lombard *et al.* (51)). Energy adjustment by residual methods.

[a] Upper and lower limits of agreement (LOA)

[b] Spearman correlation was calculated between the difference and the average values of the sFFQ2 and mean 3-24HRs. Student paired T-test or Wilcoxon signed rank test at ρ< 0.05.

Bland & Altman analysis		Spearman's rank coefficient		Intraclass Correlation Coeficient (ICC)		Cross-Classification	
Upper LOA[a]	Spearman correlation coefficient[b]	Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted	Same third	Opposite third
16.37	0.07	0.34	0.50	0.34	0.48	40.91	4.55
7.35	-0.21	0.37	0.39	0.41	0.40	40.91	13.64
16.72	-0.06	0.23	0.46	0.55	0.62	47.73	11.36
323.80	-0.44	0.50	0.55	0.44	0.60	52.27	6.82
1266	0.00	0.27	0.23	0.27	0.38	38.64	15.91
17.39	0.22	0.39	0.41	0.70	0.83	43.18	11.36
11.15	0.12	0.42	0.45	0.43	0.51	47.73	9.09
389.20	0.41	0.57	0.37	0.43	0.55	47.73	15.91
129.80	0.71	0.34	0.20	0.16	0.13	34.09	25.00
1.54	0.16	0.59	0.53	0.57	0.54	38.64	4.55
1.01	0.19	0.38	0.35	0.47	0.39	40.91	13.64
10.68	-0.24	0.40	0.53	1.00	0.98	50.00	9.09
3.06	0.00	0.55	0.34	0.29	0.16	45.45	9.09
239.30	0.33	0.43	0.35	0.41	0.34	52.27	11.36
815.70	0.41	0.45	0.51	0.62	0.55	43.18	6.82
11.91	0.20	0.35	0.46	0.25	0.49	54.55	11.36
2438.00	0.35	0.48	0.26	0.42	0.42	38.64	13.64
389.00	0.26	0.64	0.70	0.51	0.61	50.00	0.00
1021.00	-0.25	-0.01	0.22	0.25	0.20	38.64	13.64
914.80	0.41	0.43	0.48	0.66	0.65	50.00	9.09
265.70	0.05	0.24	0.28	0.05	0.08	47.73	15.91
61.59	0.07	0.28	0.10	0.21	0.08	29.55	18.18
5.40	0.15	0.51	0.64	0.57	0.68	68.18	2.27

The cross-classification values satisfactorily classified participants based on their consumption obtained from the sFFQ2 and the 24HRs. More than 50% of the participants were classified in the same tertile for 14 food groups and 13 nutrients. Meanwhile, less than 10% were classified in the opposite tertile for 12 food groups and 15 nutrients.

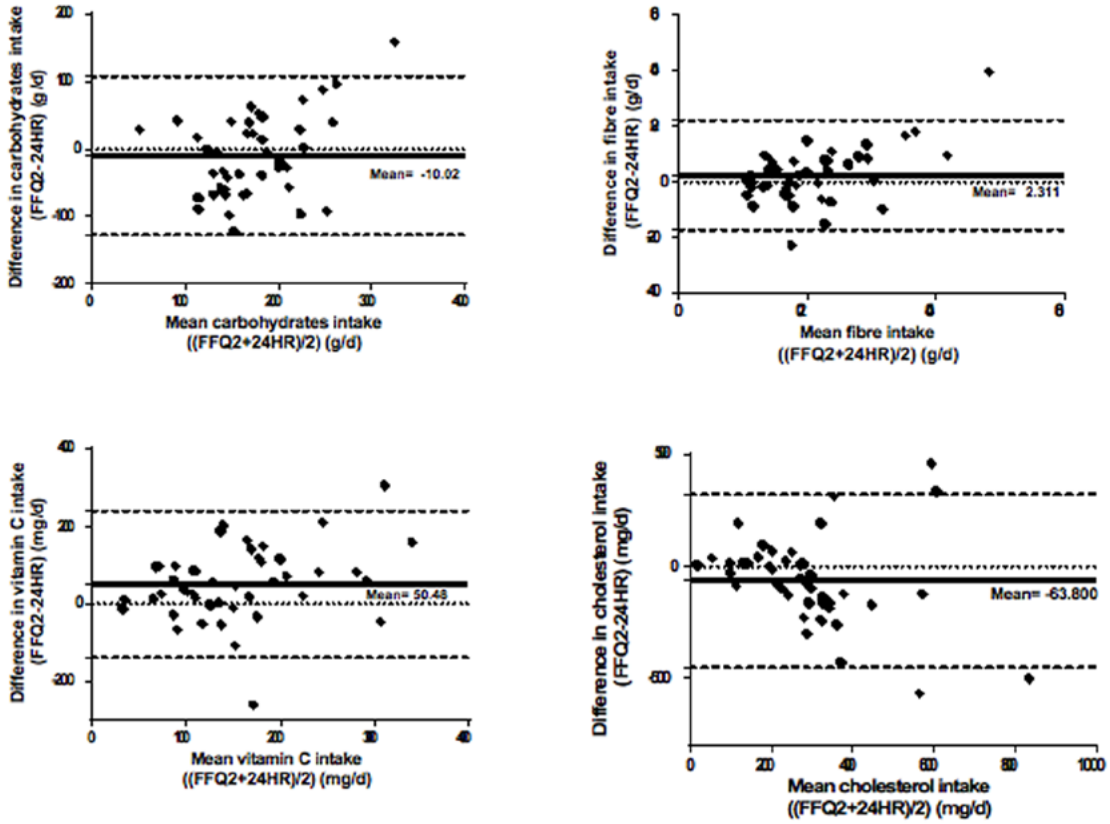


Figure 18. Bland & Altman plots show the relations between the difference in daily intake of carbohydrates, fibre, vitamin C and cholesterol estimated by sFFQ2 and mean 3-24HRs, and the corresponding mean daily intake calculated by the two methods. Plain lines represent the mean difference, and the dashed line represents lower and upper 95% limits of agreement.

Reproducibility was evaluated using the Wilcoxon signed-rank, Spearman's rank correlation, and ICC in Annex 8: Table of reproducibility of the final sFFQ. The mean intake of the 24 food groups did not show significant differences between both sFFQs. The Spearman correlation ranged from 0.33 (appetisers) to 0.90 (wholemeal grains), with 87.50% of the values being above 0.5 (median of 0.72). The ICC values were around 0.21 (potatoes and other tubers) to 0.90 (sugar and other sweets), with 62.50% of the values being above 0.4 (median 0.63).

For energy and nutrients, the mean intake did not show significant differences between both sFFQs. The Spearman correlation ranged from 0.65 to 0.87 (median 0.78) and 0.54 to 0.86 (median of 0.70) for unadjusted and energy-adjusted data, respectively. The ICC ranged between 0.47 to 0.99 (median 0.73) for data unadjusted for energy. While the values were between 0.58 and 0.96 (median 0.74). All Spearman correlation and ICC values were above 0.5 and 0.4, respectively.

5.5 CHARACTERIZATION OF THE STUDY POPULATION DIET

The daily intake values for food groups, energy and nutrient obtained by the sFFQ2 are presented in medians and quartiles in Table 9.

The diet of our study population is made up of the following food groups in descending order: vegetables (21%), non-alcoholic beverages (17%), fruits (14%), milk and dairy products except fermented milk (10%), alcoholic beverages (9%), meat and eggs (6%), fish and shellfish (4%), yoghurt and fermented milk (3%) and legumes (3%) as shown in Figure 19.

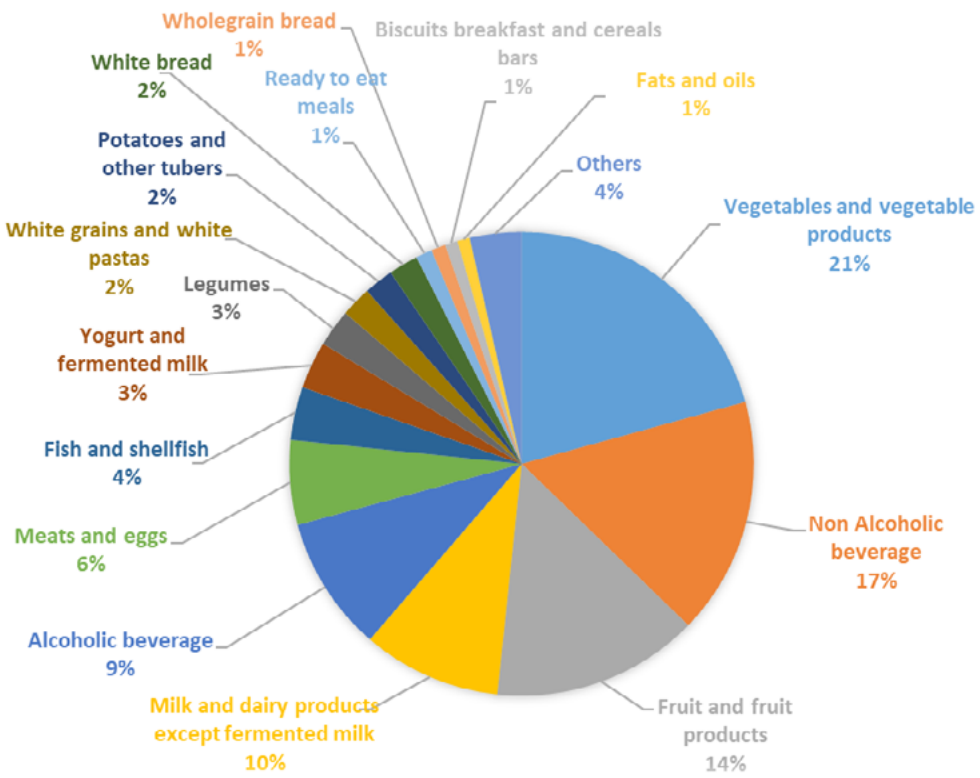


Figure 19. Percentage of food group most consumed by our population as assessed by the sFFQ2. Values less than 1% were grouped in others.

The mean energy intake was 1614.63 ± 568.1 kcal/d. The contribution of macronutrients, expressed as a percentage of energy, was 42% for carbohydrates, 33.4% for total fat, 19% for proteins, 3.5% for alcohol and 2.7% for fibre.

According to the recommendations for nutritional intakes for the Spanish population (259), 79.55% of the participants met 80% of the recommendations for daily protein intake (0.83g/kg per day), 63.6% for total fat (20-35%E), 38.6% for carbohydrates (45-60%E) and 34.1% for fibre (> 25 g/day). Values below the recommendations were observed for protein, carbohydrates and fibre, and those above were observed for total fat.

More than 50% of the people met the recommendations for daily micronutrient intake by sex and age, except for vitamin D (20.45%), calcium (47.73%), iodine (27.27%) and zinc (45.45%), as we showed in Figure 20.

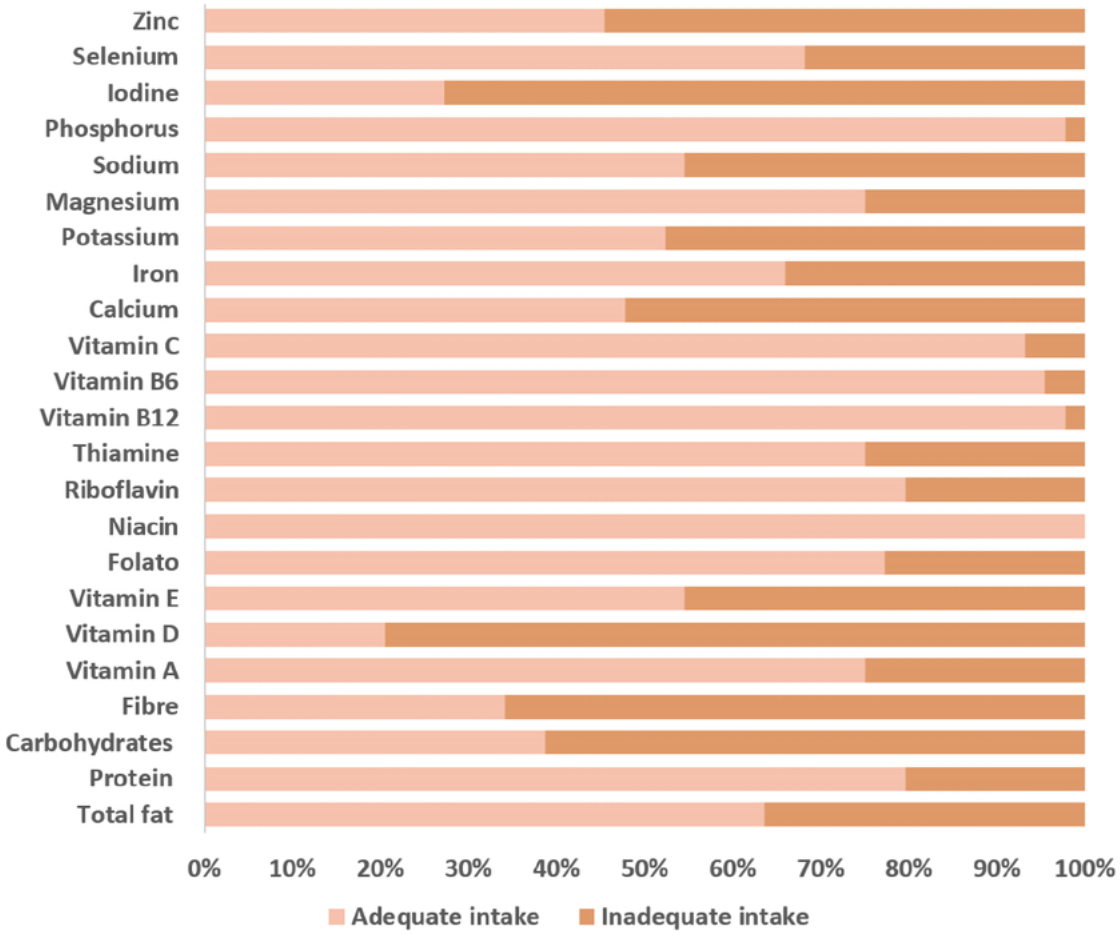


Figure 20. Percentage of participants with adequate and inadequate daily nutrient intake based on the 80% cut-off of the recommendations for the Spanish population.

Based on the frequency of consumption data obtained from the sFFQ2, we classified our population according to the quality of the diet given by the Healthy Eating Index (IASE) for the Spanish people. As a result, 23% of the participants showed healthy feeding, 70% need-for-change, and 7% have a little-healthy diet.

5.5.1 EVALUATION MISREPORTING OF SFFQ

As commented in section 10 of the method, we followed the EFSA recommendations to calculate misreporting information, which is based on the Goldberg and Black works (261-263). The sFFQ2 tended to under-report energy at the group level. At the individual level, 48% and 83% of females and males tended to underreport energy intake using the sFFQ2, respectively.

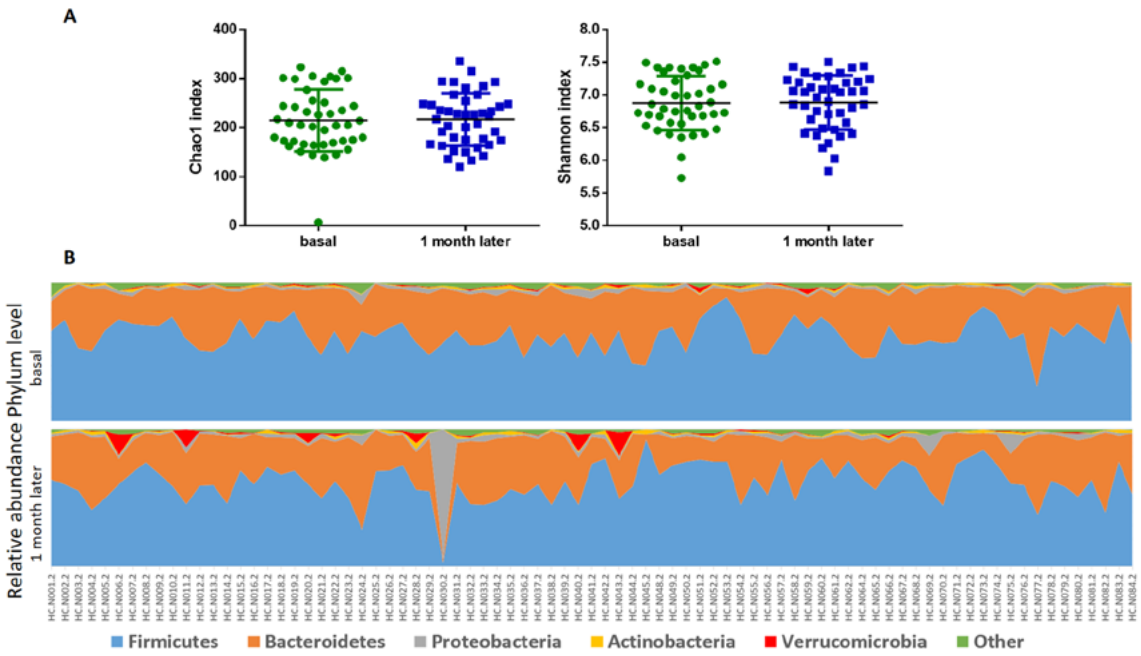


Figure 21: Fecal microbiome composition of the population by sequencing the V4 region of the 16S rRNA gene. A. Microbial richness was calculated based on the Chao1 index (A, left) and microbial richness and evenness on the Shannon index (A, right). Using the Wilcoxon test, the microbiome of the basal faecal sample and one month later did not present significant differences. $p<0.05$. **B.** Relative abundance (%) of phyla at basaline and one month later. We excluded three participants (one sample did not provide sufficient high-quality sequences and two subjects provided only one sample at baseline). (N=84, 165 faecal samples).

5.6 GUT MICROBIOME PROFILES AND ALPHA-DIVERSITY

We collected two faecal samples from participants in the pilot (n=40) and validation (n=44) studies at baseline and one month later (first stool after each sFFQ). Microbiome composition was analysed based on the amplification and sequencing of the V4 region of the 16S rRNA gene.

In total, we identified 14 phyla and 143 genera in 165 stool samples. The alpha diversity estimated by the Shannon index and Chao1 index was 6.72 ± 0.51 and 195.03 ± 61.60, respectively (Figure 21.A). There were no differences in alpha diversity (Shannon index p=0.87; Chao index p=0.82), at the phylum level (p>0.05) and the genus level (p>0.05, for 11 main genera) between the baseline stool (BS) and the stool obtained one month later (LS).

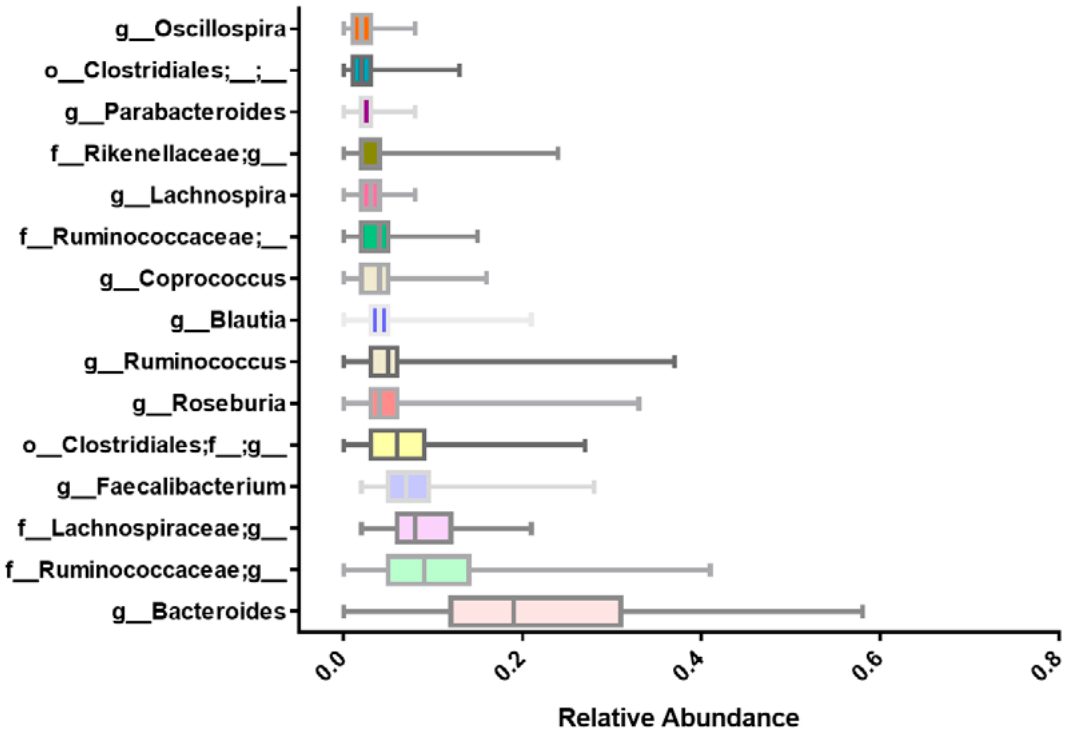


Figure 22. Distribution of the abundances of the genera that are part of the core microbiota. The core microbiota is defined as the list of genera present in at least 95% of the samples in the study (N=84, 165 faecal samples).

The most relative abundant phyla were Firmicutes (61.89 ± 12.89%) and Bacteroidetes (31.53 ± 12.47%), followed by Proteobacteria (2.54 ± 7.41%) and Actinobacteria (1.10 ± 0.84%). We detected Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria in all individuals except in one sample at BS and four samples at LS. We identified a high inter-individual variability in the microbiota composition: the relative abundances ranged from 2.85 to 92.54% for Firmicutes, 2.73 to 72.04% for Bacteroidetes, 0.00 to 94.42% for Proteobacteria, 0.00 to 17.40% for Verrucomicrobia and 0.00 to 4.73% for Actinobacteria (Figure 21B).

The current data set yielded a core microbiota (i.e., the genera shared by 95% of the samples) composed of 15 genera with a median core abundance (MA) of 72.25% (Figure 22).

5.7 CORRELATION BETWEEN PARTICIPANT´S CHARACTERISTICS AND MICROBIAL DIVERSITY AND TAXA

We evaluated the association between microbial alpha-diversity (richness and evenness) or taxonomic profile (relative abundance of microbial genera) and several characteristics of the participants, including age, BMI, gender, smoking habit, blood type and type of birth. To this end, we used linear mixed models implemented in the MaAsLin2 tool. We took into account the longitudinal setting of the study, as the participant identification number (Subject ID) was added as a random effect.

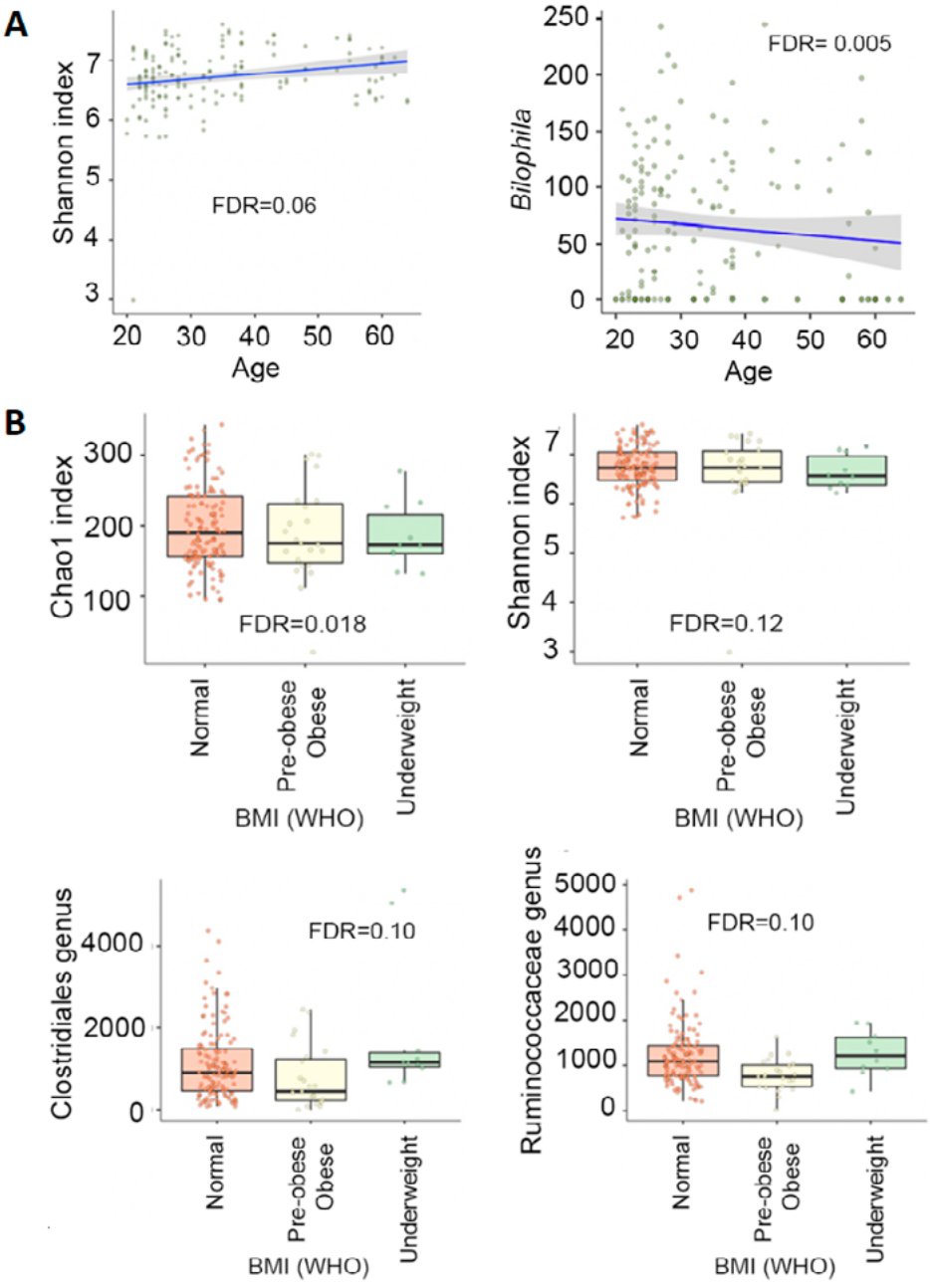


Figure 23. Association between participants' characteristics and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 were used to associate microbiome with demographic and clinical data such as age (A) and BMI (B) in the pilot and validation studies (N=84, 165 faecal samples). BMI, body mass index (BMI) based on the classification of the World Health Organization (WHO).

We found that age was positively correlated with diversity (FDR=0.06 for Shannon index) and negatively correlated with *Bilophila* (FDR=0.005) (Figure 23A). Gender was not found associated with diversity but was associated with a depletion of *Bilophila* (FDR=0.04) in male participants. The pre-obese and obese group (BMI above 25) was associated with a lower microbial diversity (FDR=0.018 for Chao1 and FDR=0.12 for Shannon) and depleted

in members of the Clostridiales order such as *Faecalibacterium* (FDR=0.10) (Figure 23B). Vaginal birth was associated with higher microbial diversity (FDR<0.02 for Chao1 (richness) and Shannon (evenness) indices) and was enriched in several bacterial genera, including an unclassified genus from the Ruminococcaceae family (FDR= 6.37 x 10^{-07}), from the Clostridiales order (FDR= 5.80 x 10^{-05}) and RF39 (FDR= 0.0006) as compared with C-section births (Figure 24.A).

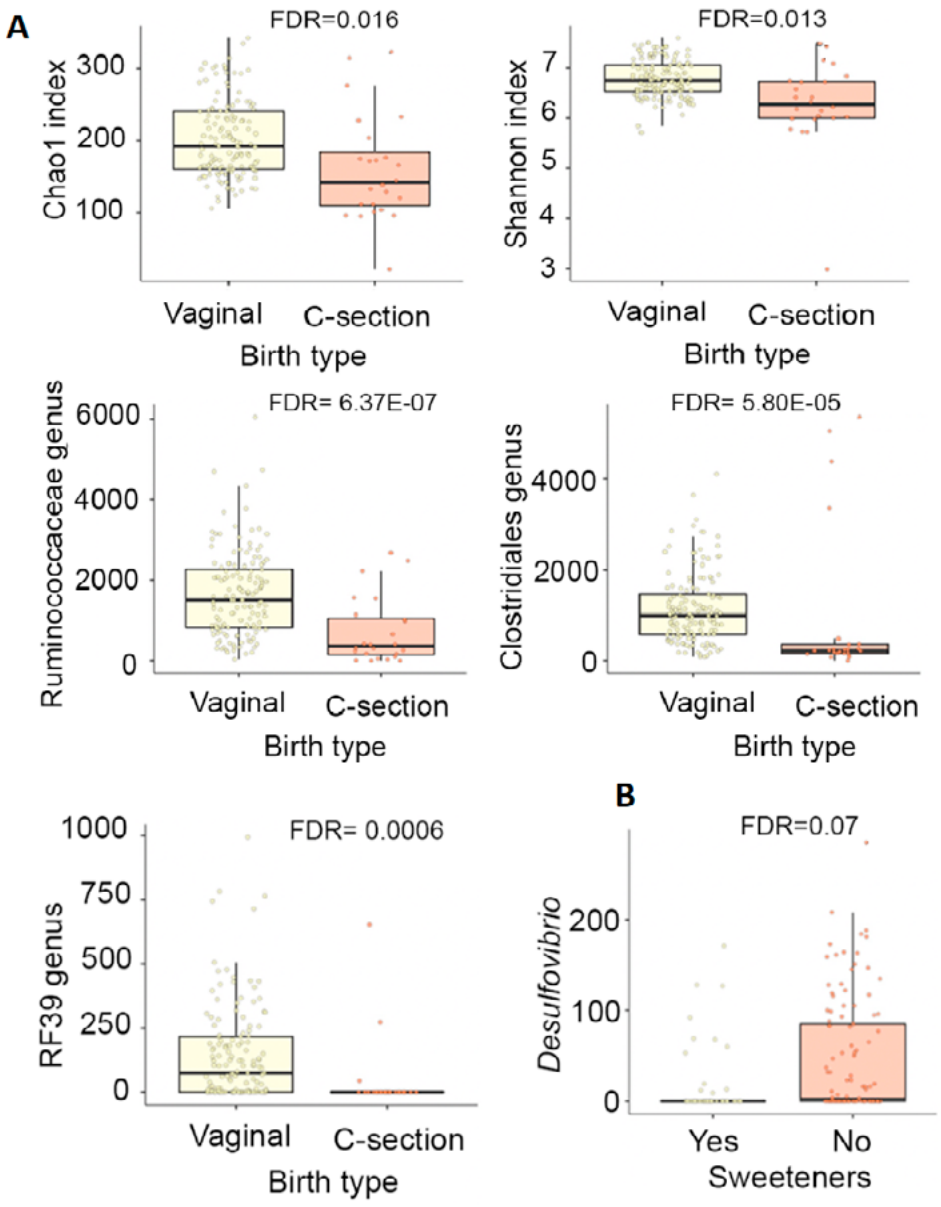


Figure 24. Association between participants' characteristics and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 were used to associate microbiome with demographic and clinical data such as the birth type (A) and intake sweeteners (B) in the pilot and validation studies (N=84, 165 faecal samples).

Smoking habit (smoker, non-smoker or ex-smoker) was not associated with diversity or any microbial taxon. The use of sweeteners was negatively associated with *Desulfovibrio* (FDR=0.07) (24.B).

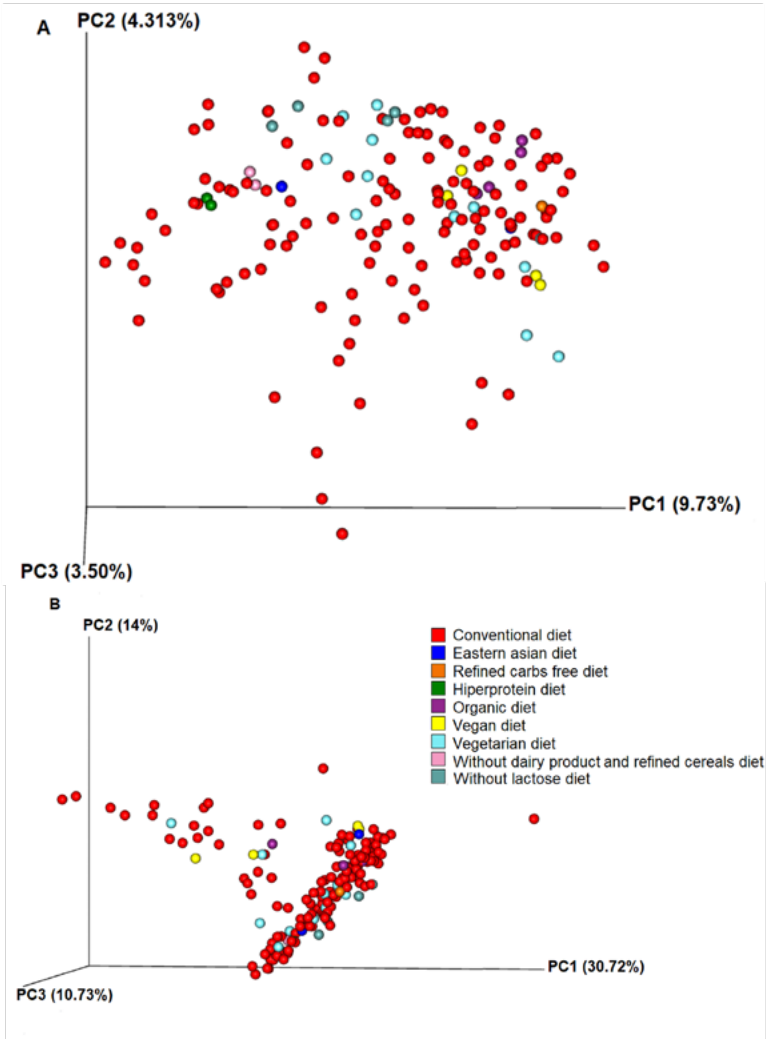


Figure 25. Principal coordinate analysis (PCoA) of weighted (A) and unweighted (B) Unifrac distances of the gut microbiome associated with different types of diet. Unweighted UniFrac distances take into consideration only the microbial composition, and weighted UniFrac distances consider the microbial composition and abundance.

We performed a microbiome clustering analysis based on unweighted and weighted Principal Coordinate Analysis-UniFrac metrics (PCoA) on all cohort samples to visually explore variation between the sample's microbial composition and abundance.

PCoA representations based on unweighted and weighted UniFrac distances showed that different types of diet did not cluster, as shown in Figure 25.

5.8 CORRELATION BETWEEN DIETARY INTAKE AND MICROBIAL DIVERSITY AND TAXA

To correlate microbiome data with dietary intake, we first used the 58 food items from the sFFQs (N=44, 85 faecal samples). One sample did not provide sufficient high-quality sequences, and two subjects provided only one sample baseline). The analysis was performed using MaAsLin2. The results are shown in Figure 26. Item 14, which consisted of fresh fruit, was positively associated with richness and evenness (FDR=0.009 for Chao1 and Shannon) and the relative abundance of a member of the Ruminococcaceae family (FDR=0.1) (Figure 26.A). Item 35, which consists of processed meats, was negatively associated with richness (Chao1, FDR=0.034), evenness (Shannon, FDR=0.03), and an unclassified genus from the Clostridiales order (FDR=0.1) (Figure 26.B). Item 58, which comprised processed foods, was negatively associated with richness and evenness (FDR=0.009 for Chao1 and Shannon) and an unclassified genus from the Clostridiales order (FDR=0.19) (Figure 26.C).

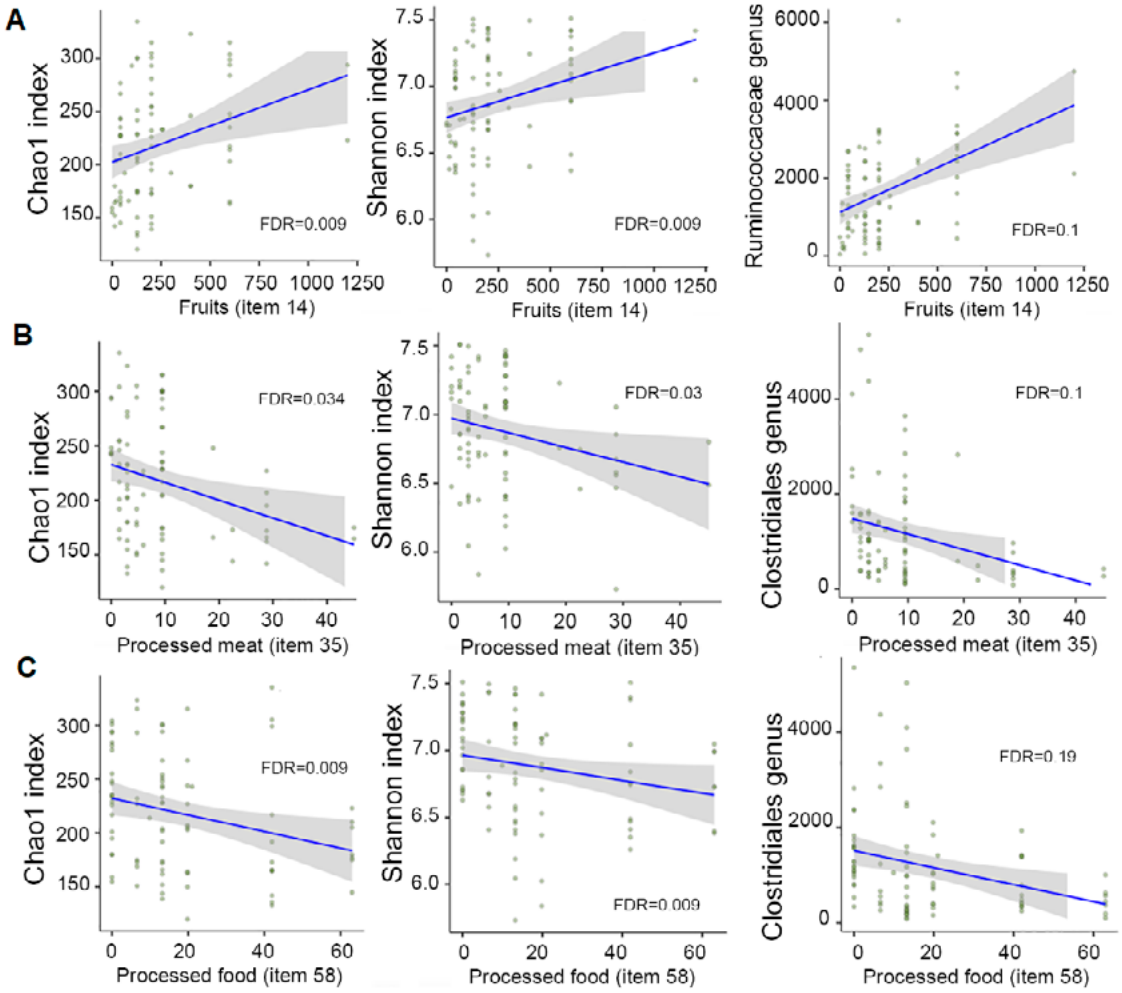


Figure 26. Correlation between food groups and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 were used to analyse the microbiome and dietary data extracted from the validation study. **A**: Item 14, which consisted of fresh fruit such as orange, grapefruit, banana, apple, pear and others. **B**: Item 35, which consisted of processed meats such as salami, pork sausage, blood sausage and others. **C**: Item 58, which comprised processed foods such as pizza, lasagna, cannelloni, chicken nuggets and others. (N=44, 85 faecal samples).

We then correlated microbiome data with the 24 food groups extracted from the sFFQ. Fruits and fruit products, which encompassed food items 14, 15 and 16, were also positively correlated with microbial diversity (FDR=0.005 for Chao1 and Shannon indices). Sausages and other processed meats and ready-to-eat meals, which corresponded to food items 35 and 58, respectively, were found to be negatively correlated with microbial diversity and taxa, as mentioned above. No association was observed between microbiome data and the other items or food groups.

Finally, we did not uncover any association between the 29 nutrients and total energy extracted from the sFFQs and microbial diversity, except for sodium (FDR=0.005) and SFA (FDR=0.04) (Figure 27), whose levels were negatively correlated with both richness and evenness. Sodium was positively correlated with *Holdemania* and negatively correlated with *Ruminococcus* and *Methanobrevicter*, a member of the Ruminococcaceae family (FDR<0.05).

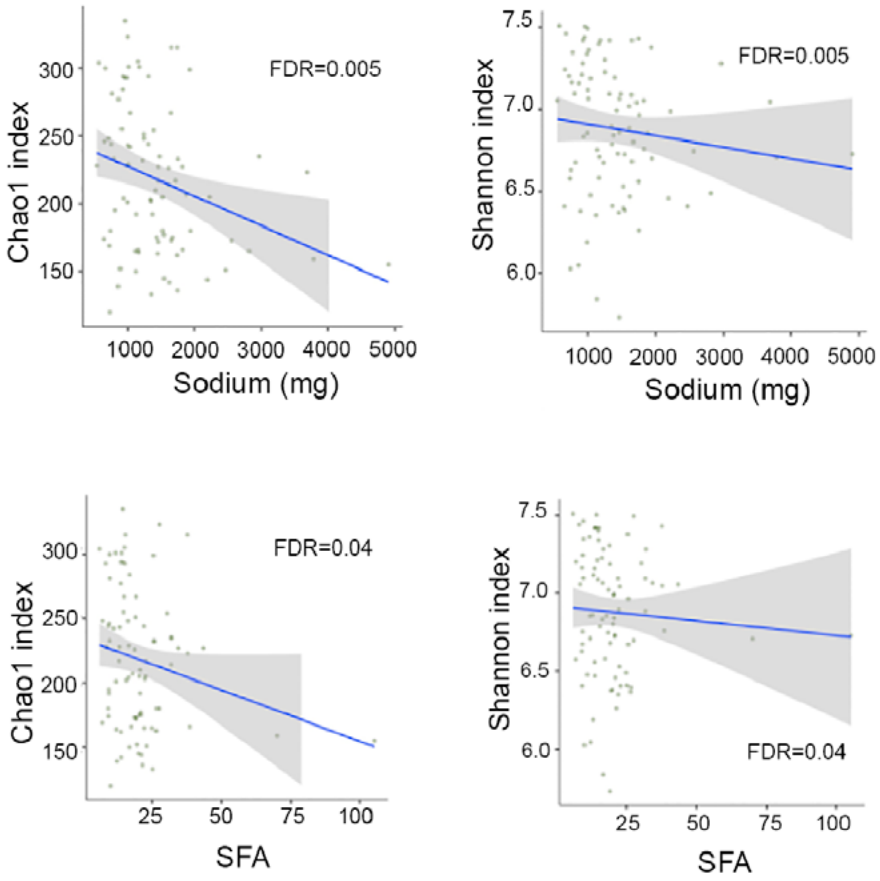


Figure 27. Correlation between nutrients and microbial diversity. Linear mixed models implemented in MaAsLin2 were used to analyse the microbiome diversity and nutrients data extracted from the validation study (N=44, 85 faecal samples). SFA, saturated fatty acids. SFA are displayed in grams and sodium in mg.

Other nutrients such as alcohol, total fat and total fibre were also associated with several microbial genera (Figure 28). Alcohol was positively correlated with two genera from the Coriobacteriaceae family, one of them being *Collinsella* (FDR=0.0004) (Figure 28.A), and negatively correlated with a member of the Peptostreptococcaceae family (FDR=0.03). Total fat and SFA were negatively correlated with *Ruminococcus* (Figure 28.B, D); total fat was also positively correlated with *Clostridium* (Figure 28.C). MUFA were negatively correlated with *Methanobrevibacter* (FDR=0.03) (Figure 28.E). The total fibre was positively correlated with a member of the Clostridiaceae family (FDR=0.009) (Figure 28.F).

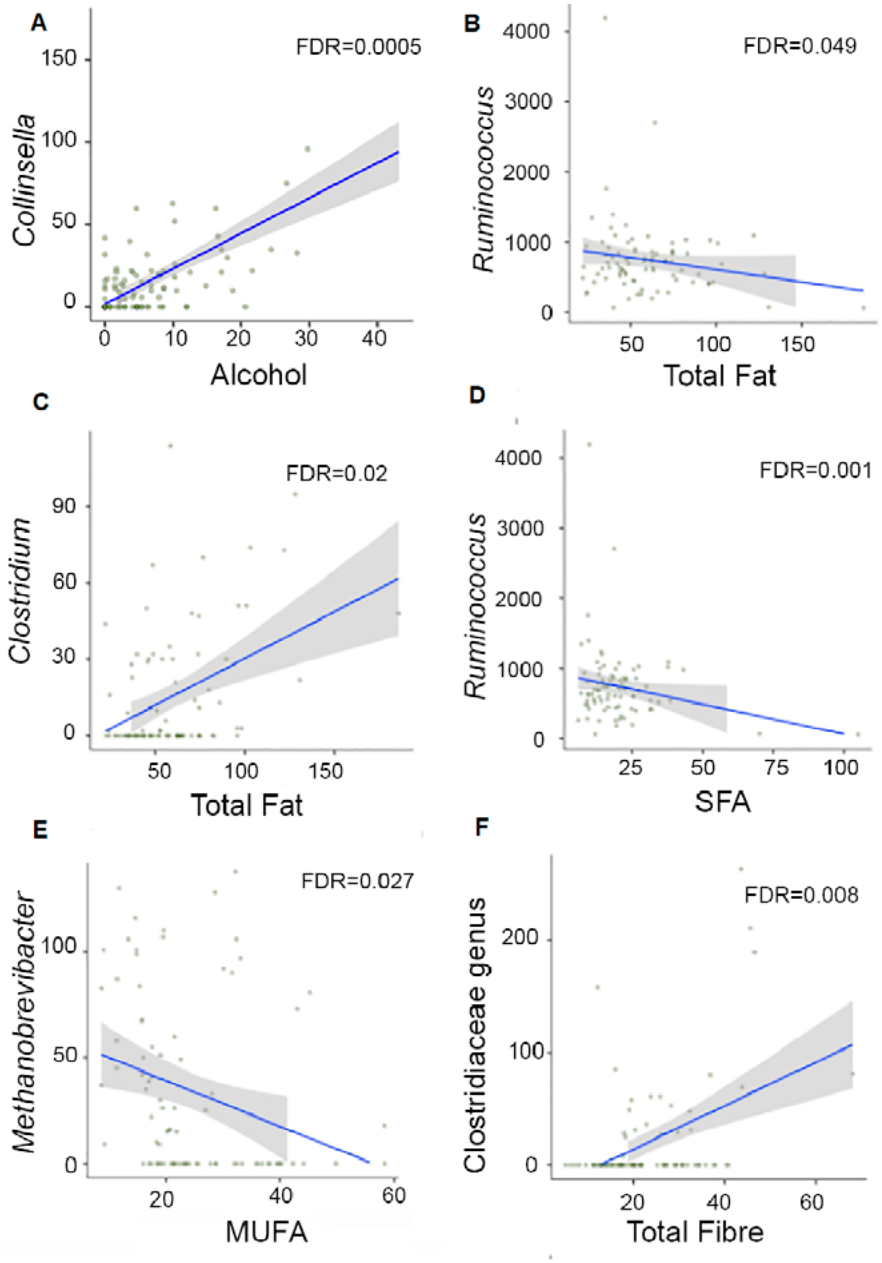
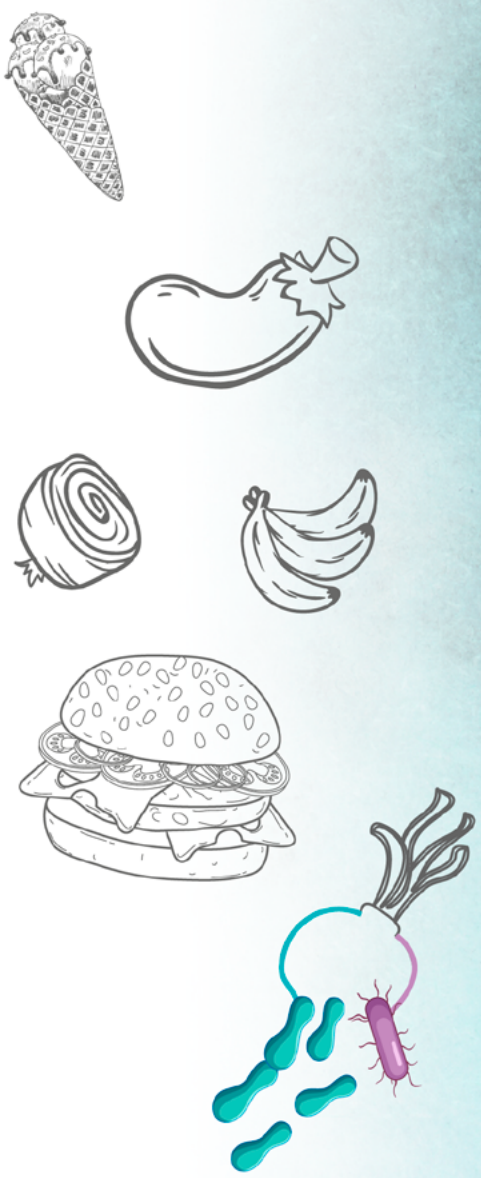


Figure 28. Correlation between nutrients and taxa. Linear mixed models implemented in MaAsLin2 were used to associate the taxonomic profile with nutrients data extracted from the validation study (N=44, 85 faecal samples). SFA, saturated fatty acids; MUFA, monounsaturated fatty acids. All nutrients are displayed in grams.



6. DISCUSSION

DISCUSSION

sFFQs have been developed to establish associations between diet and long-term risk of cancer or chronic non-communicable diseases (269,270). However, their use in association studies with the gut microbiome has exponentially increased, despite a lack of validation of these tools.

To address this knowledge gap, in this thesis, we implemented a pilot study (N=40) to design and improve a new sFFQp and a second study to validate a final version of the sFFQ (N=44) and to establish associations between dietary intake data and the gut microbiome.

6.1 DESIGN AND RELATIVE VALIDATION OF THE sFFQ

The main objective of designing this new sFFQ was to integrate efficiently dietary, demographic and microbiome data. The newly developed online sFFQ, which contains 58 food items converted into 24 food groups and 29 nutrients, reports the dietary intake of an adult population in the last month. The lowest possible number of items was chosen to reduce respondents' burden, reduce omission and response biases, and maximise their full attention. The time frame of one month represented an attempt to match habitual dietary intake with changes in the microbiome community (130). Furthermore, respondents completed the sFFQ in an average of 22 min, which is much less time than performed by participants in studies with the most commonly used FFQs (from 30 to 60 min in FFQs greater than 100 items) (271). We considered that a short FFQ would attract more volunteers, who will be more willing to repeat the experiment several times in a year to cover, for instance, every season or long-term diet fluctuations.

The results showed that the sFFQ had high reproducibility. While for relative validation, the sFFQ reported 20 out of 29 (69%) nutrients and 13 out of 24 (54%) food groups with good or acceptable outcomes compared to the reference method, as evaluated by at least three distinct statistical methods (Table 10).

At the group level, the Bland & Altman analysis reflected good levels of agreement between the sFFQ and the mean 3-24HR, and the plots showed that most of the data fell within the limits of agreement. The sFFQ showed only 24% underestimation and 11% overestimating for food groups, energy and nutrients. The results obtained for several food items (biscuits, breakfast cereals and cereals bars, chocolate and derivatives, pastries and sweetbreads, ready-to-eat meals, sauces-condiments and sausages) for energy and certain nutrients, such as total fat, PUFA, SFA, cholesterol, sodium and selenium, should be interpreted with caution since they were underreported. In contrast, vitamin D, folate, niacin, and vitamin C were overreported in the sFFQ compared to the mean 3-24HR.

For food groups, the substantial overestimation of vegetables observed in our study (+ 134.30g/day) was comparable with other validation studies using FFQs: Verger *et al.* (+ 123g/day) (225), Steinman *et al.* (+ 138g/day) (272) and Shu *et al.* (+ 111g/day) (272).

For energy and nutrients, the underestimation of energy (-218.1 kcal/day, -15%) and total fat (-20.79g/day, 31.7%) is similar to other publications (272–274). However, since the main objective of this dietary assessment was not to measure energy intake, the underestimation of energy should not impact the overall design of our sFFQ. Although there is an underestimation of total fat at the group level, the sFFQ has good validity at the individual level. For fibre, no underestimations or overestimations were observed between the sFFQ and mean 3-24HRs; however, the concordance between both methods decreases as the mean fibre intake increases (275). In addition, for some food groups and nutrient, a considerable dispersion of the data was observed, suggesting a lower concordance when the mean intakes were higher or lower (proportional bias), as it has been observed in other studies (223,276).

The underestimation or overestimation trends of certain foods or nutrients could be explained by a series of characteristics of the participants and the social approval of certain foods (277,278). Indeed, consuming foods considered “good for health”, such as fruits and vegetables, is usually over-reported. In contrast, “bad foods”, such as foods high in fat or sugar, are traditionally under-reported (279–281).

At the individual level, the correlation values for the food groups (Spearman between 0.177 and 0.78 and ICC between 0.049 and 0.83) were within the ranges observed in previous validation studies in adults (226,275,282,283). However, making a direct food group comparison is difficult because of the difference in food classification methods.

The cross-classification values were reasonable to classify the participants according to their intake based on both methods. More than 50% of the participants were ranked in the same tertile for 14 food groups and 13 nutrients. Less than 10% of the participants were classified in the opposite tertile for 12 food groups and 15 nutrients. The most significant discrepancies in the cross-classification analysis were observed for foods eaten sporadically (fish and shellfish, legumes, pastries, ready-to-eat meals, sausages), possibly due to the low probability of encountering these foods during the 3-24HRs interview. A short food list could decrease efficiency in ranking people according to their intakes (284). However, a systematic review on FFQs (281) did not detect an association between the number of items and greater reliability and correlation coefficients compared to our findings.

6.2 COMPARISON OF OUR STUDY WITH OTHER EPIDEMIOLOGICAL STUDIES CARRIED OUT IN SPAIN

The characteristics of diets of our study are similar to results observed in the Spanish epidemiological studies (ANIBES and ENALIA2) (222,285,286). However, for foods groups, the order of distribution of the three main groups consumed was different. For example, in ANIBES and ENALIA2, the most consumed food groups were non-alcoholic beverages and milk dairy products, followed by vegetables in ANIBES and fruits in ENALIA2 (222,287). While in our study, the most consumed food groups were vegetables, followed by non-alcoholic beverages and fruits. These differences may be caused by the demographic characteristics of our population (the majority was healthcare and research workers), who may pay more attention to a healthy diet (higher fruits and vegetable intake). In addition, the emerging consumption of plant-based drinks (in non-alcoholic beverages) in our population may replace the consumption of dairy products observed in these epidemiological studies, as they were carried out between 2013 and 2015. Unfortunately, to the best of our knowledge, there are no updated epidemiological diet studies carried out to date in the healthy Spanish population.

Most of the participants in our study met the dietary reference intakes for the Spanish population for most nutrients (259). However, this was not the case for carbohydrates, fibre, vitamin D, calcium, iodine, and zinc. These findings were similar to those observed in other studies (288–290) except for iodine. This discrepancy could be due to the fact that we did not add any questions on the consumption of added salt, which is fortified with iodine (ionised salt).

The role of fibre as a possible modulator of the gut microbiota is currently recognised. The low intake of fibre (only 34% of the participants comply with the NRI > 25g/day) would suggest a negative microbial diversity. However, no association between the amount of fibre intake and diversity has been found (see section 8 for results), which could be either due to the microbial stability given by a habitual diet or to need to recruit a larger study cohort.

Currently, much importance has been given to vitamin D deficiency, as it has been associated with increased autoimmunity and increased susceptibility to infections (291).

Although levels of vitamin D does not depend only on food intake but also on other factors, such as endogenous synthesis, its deficiency has been demonstrated and has been considered a major public health problem worldwide (292). A study of 2,260 participants showed that, although Spain is a country with many hours of sunshine, a third (33.9%) of the Spanish population could be at risk of vitamin D (25-hydroxyvitamin D) deficiency (293). Moreover, active vitamin D has recently been linked to microbial diversity and several beneficial microbial taxa (294). Therefore, it may be important to consider vitamin D's level and intake in diet-microbiome studies.

A relationship between the HEI index (similar to the IASE but for the US population) and the gut microbiome has been reported (295). Therefore, in our study, we measured the diet quality through the IASE index (based on the compliance to Dietary Guidelines for Spanish people) but did not detect any association (296,297). However, we observed that 70% of our cohort need changes in their diet, similar to reported in another study (23). These low HEI scores were mainly due to the lower weekly consumption of legumes and the higher occasional consumption of could-processed meats, sweets, beverages and low variety of the diet. Moreover, the participants who followed a vegetarian diet had a lower score by not consuming dairy products or meat (10 points for daily consumption of dairy products and 10 points for consumption 1 to 2 times per week of meat). Therefore, although this tool could be helpful to classify the diet of the Spanish population, it may underestimate the quality of the diet of vegetarians.

The agreement of our extracted data with two published epidemiological studies (ANIBES and ENALIA2) investigating the usual diet of the Spanish population suggests that the newly built sFFQ could be applied at the population level. Moreover, the high repeatability of the questionnaire indicates that only one sFFQ would be needed to cover food intake over one month. However, to adapt the questionnaire to another population it would recommended to make additional necessary changes to the sFFQ and validation.

6.3 sFFQ AS A NEW TOOL TO RELATE DIET WITH MICROBIOME

Our study captured the effects of certain demographic data and BMI on the diversity and composition of the gut microbial community, using the data obtained from the pilot and validation study on 84 healthy individuals (165 faecal samples).

We did not found changes in the diversity and microbial composition of the two stool samples collected over one month, which could be a product of the resilience and stability of the intestinal microbiota (92,131). We confirm previous findings of high inter-individual variability in the composition of the gut microbiota (56,78,142). Furthermore, our dataset provided a microbial core composed of 15 genera, of which five (*Bacteroides, Ruminococcacea family, Lachonpiraceae family, Faecalibacterium* and *Blautia*) and seven (*Roseburia* and *Coprococcus*) were previously identified in a cohort of 18,300 (298) and 4,300 (138) individuals, respectively.

Multivariate analysis of variance on distance matrices (weighted and unweighted UniFrac) did not cluster the microbial community by types of diet (reported by the participants) as it has been shown in other studies (140,194,195,199). This observation could be due to the meagre number of participants or to the fact that dietary intake is more variable than the microbial composition both within and between-person (142).

Our findings showed that the depleted bacterial groups in individuals delivered by C-section belonged mainly to Clostridiales (specifically *Ruminococcaeae*) and Mollicutes (RF39), abundant bacterial groups in those born vaginally. The impact of the type of birth on the gut microbial ecosystem has been widely studied during early life (299,300). A persistent effect of mode of delivery on the microbiome composition, the host immune system (301), and natural antibiotics' biosynthesis (302) has been reported during the first years of life. However, to the best of our knowledge, our study is the first to describe this effect in adult subjects. Moreover, our results corroborate previous studies (303,304) regarding the relationship between age and diversity, depletion of Clostridiales and overweight-obesity.

Contradictory findings have been published about the effect of sweeteners such as saccharin on glucose tolerance and dysbiosis in healthy individuals (188,305,306). Our study

observed that sweeteners, which consisted mainly of aspartame and saccharin, decreased *Desulfovibrio*, a sulfate-reducing bacterium.

On the one hand, through the sFFQ, we were able to associate high microbial diversity, which is considered a health-promoting factor (307), with fruits intake and low diversity with processed meat, ready-to-eat meals, total fat, saturated fatty acids and sodium intake. On the other hand, the sFFQ allowed us to correlate food items or nutrients with specific groups of microorganisms. Some members of the Clostridiales order were positively associated with fruits and total fibre, whereas others, including *Ruminococcus*, were negatively related to total fat, saturated fatty acids and sodium intake. The association with the latter should be interpreted with caution, given that the comparison between sFFQ2 and the mean three 24HRs based on Bland & Altman analysis showed that participants underreported sodium.

A low dietary fibre intake has been related to loss of diversity, particularly Clostridiales (308), and the class Clostridia (309). The non-association observed between vegetables and microbial diversity could be explained by the cooking method used. Indeed, fruits and vegetables may contain similar nutrients (bioactive compounds derived from plants, such as polyphenols) that could be lost or could be altered the food matrix during cooking procedures usually used to prepare vegetables and not fruits. Cooking, which transforms fibre and starch, increasing their absorption in the small intestine and thus reducing their fraction in the colon, has been shown to reshape the structure and function of the gut microbiota (310). The significant association between fruits and the gut microbiome observed in our study could be explained by the type of fibre present in the fruits. A recent study compared between-subject variation in fibre sources with the microbiome beta-diversity. They found that subjects who obtained their fruit fibre or grain fibre from similar foods tended to have more similar microbiome profiles than other foods with high fibre content, such as vegetables and legumes. They also concluded that the gut microbiota composition was more related to food choice than the conventional nutrient profile typically used in nutritional research. Probably because of the limited information available on nutrients and substrates relevant to the microbiota as hundreds of additional chemical compounds are present even in a single piece of fruit were ignored (142).

6.4 LIMITATIONS AND STRENGTHS

The present study has several strengths and limitations. One of the limitations of our research is the low heterogeneity of the study population, which can be a significant limitation when generalising our results to other populations, despite finding similar diet results in population studies such as ENALIA2 and ANIBES.

Second, this study relies only on 16S rRNA analysis, revealing only the microbiome composition. In the future, it could be complemented by functional analysis through DNA and RNA shotgun sequencing or metabolomics analysis.

Third, the newly developed sFFQ can capture the habitual diet, as reflected by its validation with a reference method and a comparison of our results with other epidemiological studies. However, the results obtained on several food items, energy and nutrients should be interpreted with caution since they are misreported in the sFFQ compared to three 24HRs. The 24HR, used here as a reference method, is not the gold standard. This reporting method also relies on memory-based and may be biased due to underestimation or overestimation. This limitation could be addressed using metabolomic biomarkers, although only a few validated biomarkers of food intake are available.

While self-report data collection has laid the foundation for numerous studies, the degree of error in measuring energy intake from FFQs has been widely discussed (269,311,312). However, Beaton stated that “there will always be an error in dietary assessments. The challenge is to understand, estimate, and make use of the error structure during analysis” (313). For example, underestimating of energy detected in the sFFQ through the cut-off point estimated by the Goldberg equation will allow us to face these challenges. Assessing the degree of measurement error of our data, correcting these errors using statistical methods or classify individuals according to their level of reporting would allow us to improve the result´s interpretation.

Finally, the database of the nutritional composition of foods used in this study is also a limitation as it is insufficient to measure the wide variety of foods consumed by the population. To date, food databases inadequately capture dietary compounds that may have relevant effects on the microbiota, such as preservatives, additives, organic acids, cooking methods, among others. Currently, 150 key nutritional components have been

documented, representing only a small fraction of more than 26,000 distinct, definable biochemicals present in our foods (314). Even for food components found in our nutritional databases, such as fibre, it remains challenging to relate their intake to microbiome data due to the variety of measurement techniques to dietary fibre and the existence of a large number of different chemical structures that are jointly referred to as “fibre” (315,316).

First, the great strength of this FFQ is that it is a validated questionnaire that measures the habitual intake of food groups and nutrients relevant to the gut microbiota in healthy adults. Previous studies used FFQs to measure specific nutrients relevant to the intestinal microbiota (non-starch polysaccharide (317), fibre (318), inulin and oligosaccharides (319)). However, before developing these specific tools, we believe that we should first understand how the habitual diet interacts with the gut microbiota. This knowledge will further improve our understanding of the microbiome modulation by dietary interventions, in particular, testing specific nutrients.

Second, few diet questionnaires used the information obtained from a similar pilot population to reformulate the questionnaires.

Third, regarding the study design, our study is one of the few that matched the 24HR with stool collection. Indeed, the collection of stool samples was carried out 1 or 2 days after the food recall, which may reflect the transit time appropriately. (143).

Fourth, the response time of the sFFQ of approximately 20 min, much shorter than the response to any other FFQs, could be easily implemented in future designs of a population study that relate diet to the microbiome.

Finally, we also showed the feasibility of relating microbiome profiles or diversity with dietary and demographic data in the present work. Indeed, compared to previously published studies, the newly developed molecular, bioinformatics, and statistical tools have made these analyses possible.

7. CONCLUSIONS

CONCLUSIONS

The results of the present doctoral thesis, despite its limitations, offer a validated dietary assessment tool and valuable insights into the relationship between dietary data and gut microbiome, leading to the following conclusions:

1. The validated new sFFQ allows to evaluate the habitual diet of healthy adults, as shown by the good and acceptable outcomes of at least three different statistical methods compared with the data estimated by three 24 hour dietary records.

2. The sFFQ, developed as a web-based tool, was shown to be a fast, simple and adequate tool to extract food groups and nutrients from the habitual intake and demographic data with relevance to the gut microbiota modulation in healthy adults.

3. The application of the sFFQ showed that higher fruit consumption was associated with an increase in the gut microbiome diversity. Inversely, the higher intake of processed foods, processed meats, short fatty acids, and sodium was associated with a decreased diversity.

4. Age and vaginal delivery were found positively correlated with microbial diversity, whereas, individuals classified as pre-obese and obese were associated with low diversity.

8. FUTURE PERSPECTIVES

FUTURE PERSPECTIVES

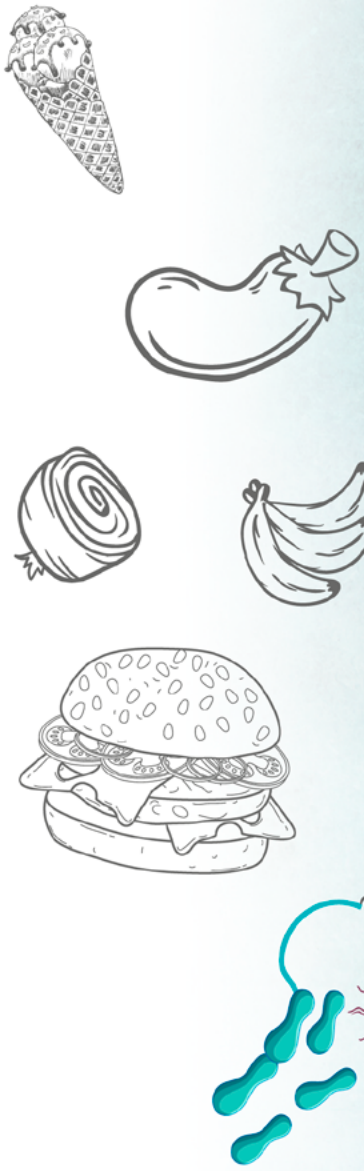
Our study provides a new validated tool to assess the consumption of the habitual diet in studies of the gut microbiome in humans. Nevertheless, it will be necessary to develop specific tools to measure specific food and nutrients intake, such as particular fibre, polyphenols, additives, etc. relevant to microbiome modulation.

Future studies with large cohorts that repeatedly measure diet and the microbiome could provide a more powerful validation study and a deeper insight into the relationships between diet and the microbial ecosystem. Moreover, future interventional and longitudinal studies could isolate the impact of certain changes in dietary intake on gut communities in human hosts.

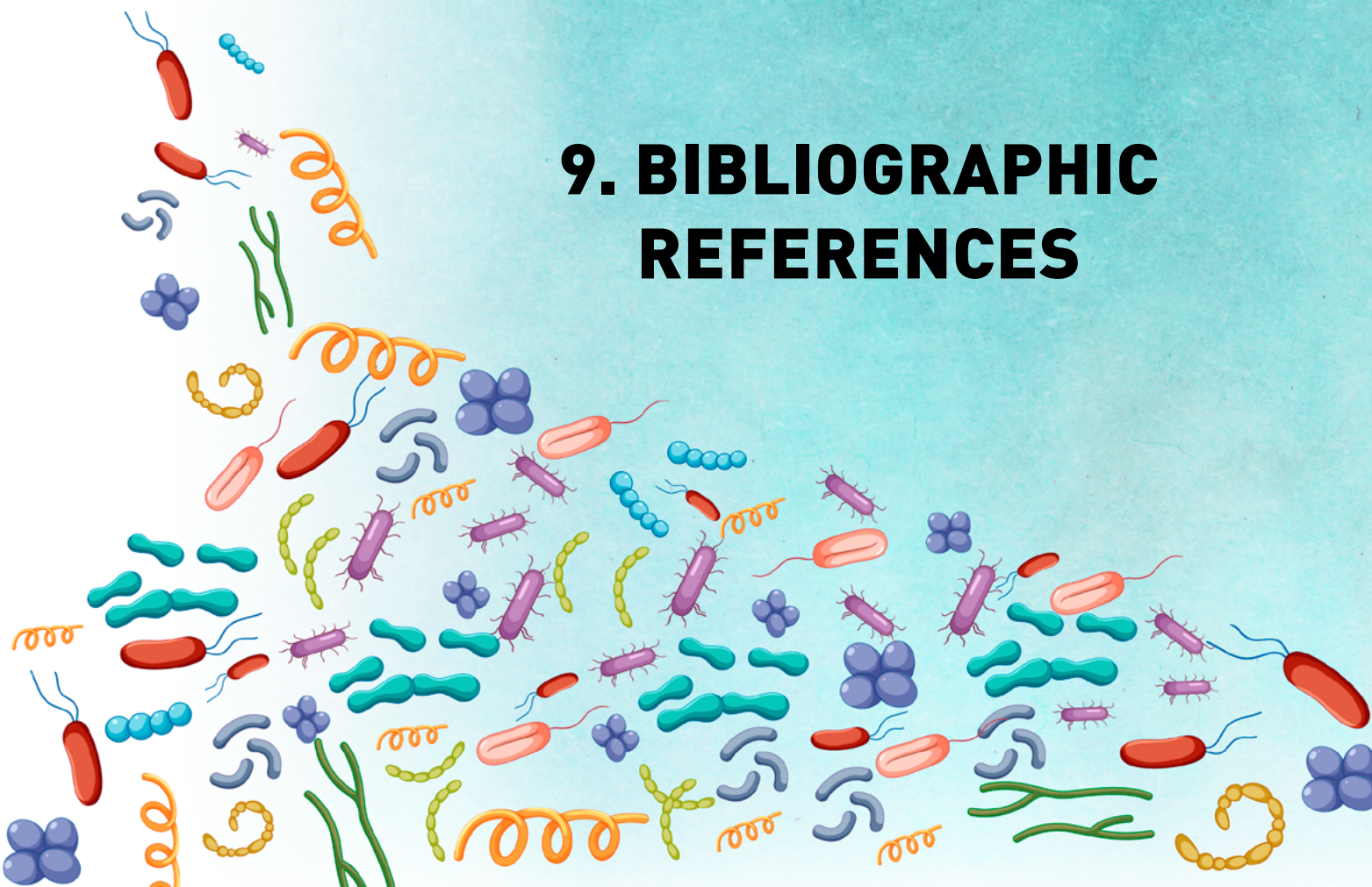
The update of food composition databases with relevant nutritional and non-nutritive components could be key component to better characterise nutrients and new chemical compounds derived from foods. The emergence of new methodologies to determine the composition of foods (e.g. mass spectrometry) would likely reveal significant associations between foods and could be used to improve our understanding of the role of foods in our health from a biochemical perspective.

Future diet-microbiome studies should utilise multiple omic measurements such as metagenomic, metatranscriptomic, metaproteomics, and metabolomics to provide mechanistic insights that identify the distinct diet-driven microbial alterations beneficial for human health. However, there are still some challenges in statistics and bioinformatics tools as well as financial costs to combine the multiple dietary and omic data, in particular for population studies.

Finally, the collaboration of nutritional science, microbiology, statistics, and bioinformatics is critical to elucidate the interactions between the human diet and the gut microbiome.



9. BIBLIOGRAPHIC REFERENCES



BIBLIOGRAPHIC REFERENCES

1. Kuklinski C. Nutrición y Bromatología. Spain: Omega; 2003. 1–432 p.
2. Beardsworth A, Keil T. Sociological Perspectives on Food and Eating. Sociology on the Menu: an invitation to the study of food and society. London; New York: Routledge; 1997. 1–288 p.
3. Peña GM, Jiménez ML, García AR. Hábitos alimentarios y salud. In: Tratado de Nutricion, de Ángel Gil Hernández. 2 ed. Spain; 2010.
4. Willett W. Nutritional Epidemiology. 3 edition. New York. USA: Oxford University Press; 2013. 34–133 p.
5. Kirkpatrick SI, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. J Acad Nutr Diet. 2019;119(11):1801–16.
6. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, *et al.* Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72.
7. Reedy J, Subar AF, George SM, Krebs-Smith SM. Extending methods in dietary patterns research. Nutrients. 2018;10(5):1–8.
8. Greenfield H, Southgate DA. Food Composition Data: Production, Management and Use. Food Nutr Bull. 2004;25(3):307–307.
9. Gibson R. Principles of Nutritional Assessment. 2nd ed. New York: Oxford University Press; 2005. 1–908 p.
10. FAO. Dietary Assessment: A resource guide to method selection and application in low resource settings. Nutrition for Sport, Exercise and Performance. Rome; 2018. 1–113 p.
11. Salas-Salvadó J, Anna Bonada, Trallero R, Saló E, Burgos R. Nutrición y dietética clínica. 4º ed. Barcelona: Elsevier España; 2019. 79–94 p.
12. Slimani N, Freisling H, Illner A-K, Huybrechts I. Methods to Determine Dietary Intake. Nutr Res Methodol. 2015;48–70.
13. Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: Estimating the number of days required to rank dietary intakes with desired precision. Am J Clin Nutr. 1989;50(1):155–67.

14. Emmett PM, Jones LR, Northstone K, Pounis G, Taylor CM. Collection and management of dietary data. Analysis in Nutrition Research: Principles of Statistical Methodology and Interpretation of the Results. Elsevier Inc.; 2018. 43–73 p.

15. Cynthia B, Alanna J. Moshfegh, Baer DJ, Mary J. Kretsch. The USDA Automated Multiple pass method accurately estimates group total energy and nutrient intake. J Nutr. 2006;136:2594–9.

16. Raper N, Perloff B, Ingwersen L, Steinfeldt L, Anand J. An overview of USDA's Dietary Intake Data System. J Food Compos Anal. 2004;17(3–4):545–55.

17. Willett W. Nutritional Epidemiology. 3 edition. Nutritional Epidemiology. New York. USA: Oxford University Press; 2013. 260–286 p.

18. Cade JE, Burley VJ, Warm DL, Thompson RL, Margetts BM. Food-frequency questionnaires: a review of their design, validation and utilisation. Nutr Res Rev. 2004;17(1):5–22.

19. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - A review. Public Heal Nutr. 2002;5(4):567–587.

20. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, *et al.* A 14-item mediterranean diet assessment tool and obesity indexes among high-risk subjects: The PREDIMED trial. PLoS One. 2012;7(8).

21. Papadaki A, Johnson L, Toumpakari Z, England C, Rai M, Toms S, *et al.* Validation of the English version of the 14-item Mediterranean Diet Adherence Screener of the PREDIMED study, in people at high cardiovascular risk in the UK. Nutrients. 2018;10(2).

22. Susan M. Krebs-Smith, TusaRebecca E. Pannucci, P Amy F. Subar Sharon I. Kirkpatrick, P, Jennifer L. Lerman, Janet A. Tooze, Magdalena M. Wilson JR. Update of the Healthy Eating Index: HEI-2015. Physiol Behav. 2017;176(3):139–48.

23. Navarro A. N, Moncada RO. Calidad de la dieta española según el índice de alimentación saludable. Nutr Hosp. 2011;26(2):330–6.

24. Kant AK, Block G, Schatzkin A, Ziegler RG, Nestle M. Dietary diversity in the US population, NHANES II, 1976-1980. J Am Diet Assoc. 1991;91(12):1526–31.

25. Stumbo J Phyllis. New technology in dietary assessment: A review of digital methods in improving food record accuracy. Proc Nutr Soc. 2013;72:70–6.

26. McClung HL, Sigrist LD, Smith TJ, Karl JP, Rood JC, Young AJ, *et al.* Monitoring Energy Intake: A Hand-Held Personal Digital Assistant Provides Accuracy Comparable to Written Records. J Am Diet Assoc. 2009;109(7):1241–5.

27. Fukuo W, Yoshiuchi K, Ohashi K, Togashi H, Sekine R, Kikuchi H, *et al.* Development of a Hand-Held Personal Digital Assistant-Based Food Diary with Food Photographs for Japanese Subjects. J Am Diet Assoc. 2009;109(7):1232–6.

28. Shang J, Pepin E, Johnson E, Hazel D, Teredesai A, Kristal A, *et al.* Dietary intake assessment using integrated sensors and software. Multimed Mob Devices 2012; Multimed Content Access Algorithms Syst VI. 2012;8304:830403.

29. Ruiz E, Ávila JM, Castillo A, Valero T, Del Pozo S, Rodriguez P, *et al.* The ANIBES study on energy balance in Spain: Design, protocol and methodology. Nutrients. 2015;7(2):970–98.

30. Higgins JA, LaSalle AL, Zhaoxing P, Kasten MY, Bing KN, Ridzon SE, *et al.* Validation of photographic food records in children: Are pictures really worth a thousand words? Eur J Clin Nutr. 2009;63(8):1025–33.

31. Dahl Lassen A, Poulsen S, Ernst L, Andersen KK, Biltoft-Jensen A, Tetens I. Evaluation of a digital method to assess evening meal intake in a free-living adult population. Food Nutr Res. 2010;54(7):1–9.

32. Baranowski T, Baranowski JC, Watson KB, Martin S, Beltran A, Islam N, *et al.* Children's accuracy of portion size estimation using digital food images: Effects of interface design and size of image on computer screen. Public Health Nutr. 2011;14(3):418–25.

33. Vereecken CA, Covents M, Matthys C, Maes L. Young adolescents' nutrition assessment on computer (YANA-C). Eur J Clin Nutr. 2005;59(5):658–67.

34. Subar AF, Kirkpatrick SI, Mittl B, Palmer ZT. The Automated Self-Administered 24-Hour Dietary Recall (ASA24): A Resource for Researchers, Clinicians and Educators from the National Cancer Institute. J Acad Nutr Diet. 2012;112(8):1134–7.

35. Makhsous S, Bharadwaj M, Atkinson BE, Novosselov I V., Mamishev A V. Dietsensor: Automatic dietary intake measurement using mobile 3D scanning sensor for diabetic patients. Sensors (Switzerland). 2020;20(12):1–19.

36. Wenyan J, Chen H-C, Yue Y, Li Z, Fernstrom J, Bai Y, *et al.* Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. Public Heal Nutr. 2014;17(8):1671–81.

37. US Department of Agriculture ARS. Nutrient Data Laboratory. USDA National Nutrient Database for Standard Reference, Release 27.

38. ANSES. 2020. Ciqual Table de composition nutritionnelle des aliments [Internet]. [cited 2021 Aug 1]. Available from: https://ciqual.anses.fr/

39. AUSNUT 2011-13. Australian Food and Nutrient Database 1999 [Internet]. [cited 2021 Aug 1]. Available from: https://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/ausnutdatafiles/Pages/foodnutrient.aspx

40. Lupiañez-Barbero Ascencio, González Cintia de LA. Tablas y bases de datos de composición de alimentos españolas: necesidad de un referente para los profesionales de salud. Endocrinol Diabetes Nutr. 2018;65(6):361–73.

41. Farran A, Zamora R, Cervera P, Centre d´Ensenyament Superior de Nutrició i Dietètica. Tabla de composición de alimentos del CESNID. Edicions U. McGraw-Hill/Internamericana de España SAU, editor. Barcelona; 2004. 1–247 p.

42. Mataix Verdú et. al. Tabla de composición nutricional de alimentos. 5º ed. Universidad de Granada; 2009. 1–560 p.

43. Moreiras O, Carbajal A, Cabrera L, Pirámide CC-ME, 2019 U. Tabla de composición de alimentos. 19 ed. Pirámide, editor. Madrid. Spain: Pirámide; 2019. 1–494 p.

44. Peña M. Tabla de composición de alimentos. Madrid: Nutricia; 1997.

45. Ortega RM, López AM, Requejo AM, Carvajales PA. La composición de los alimentos. Herramienta básica para la valoración nutricional [Internet]. 1st ed. Madrid: Ed. Complutense; 2004 [cited 2021 Aug 14]. Available from: https://www.sennutricion.org/es/2013/05/12/la-composicin-de-los-alimentos-ortega-et-al

46. AESAN/BEDCA. Base de datos española de composición de alimentos v1.0 [Internet]. 2007. Available from: http://www.bedca.net

47. Heimler D, Romani A, Ieri F. Plant polyphenol content, soil fertilization and agricultural management: a review. Eur Food Res Technol. 2017;243(7):1107–15.

48. Abdul Khalil HPS, Hossain MS, Rosamah E, Azli NA, Saddon N, Davoudpoura Y, *et al.* The role of soil properties and it's interaction towards quality plant fiber: A review. Renew Sustain Energy Rev. 2015;43:1006–15.

49. Shewfelt RL. Sources of Variation in the Nutrient Content of Agricultural Commodities From the Farm To the Consumer. J Food Qual. 1990;13(1):37–54.

50. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutr J. 2012;11.

51. Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J. 2015;14(1).

52. Natarajan L, Rock CL, Major JM, Thomson CA, Caan BJ, Flatt SW, *et al.* On the importance of using multiple methods of dietary assessment. Epidemiology. 2004;15(6):738–45.

53. Grosso G, Bella F, Godos J, Sciacca S, Del Rio D, Ray S, *et al.* Possible role of diet in cancer: Systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. Nutr Rev. 2017;75(6):405–19.

54. Hills RD, Pontefract BA, Mishcon HR, Black CA, Sutton SC, Theberge CR. Gut microbiome: Profound implications for diet and disease. Nutrients. 2019;11(7):1–40.

55. Giavarina Davide. Understanding Bland Altman analysis. Biochem Medica. 2015;25(2):141–51.

56. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285):59–65.

57. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Host-bacterial mutualism in the human intestine. Vol. 307, Science. 2005. p. 1915–20.

58. Rosner J. Ten Times More Microbial Cells than Body Cells in Humans? Microbiome. 2014;9(2):2014.

59. Sender R, Fuchs S, Milo R. Commentary Are We Really Vastly Outnumbered ? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell. 2016;164(3):337–40.

60. Zhang X, Li L, Butcher J, Stintzi A, Figeys D. Advancing functional and translational microbiome research using meta-omics approaches. Microbiome. 2019;7(1):1–12.

61. Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Res. 2009;19(7):1141–52.

62. Neefs J marc, Van De Peer Y, De Rijk P, Chapelle S, de Wachter R. Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res. 1993;21(13):3025–49.

63. Zhang J, Ding X, Guan R, Zhu C, Xu C, Zhu B, *et al.* Evaluation of different 16S rRNA gene V regions for exploring bacterial diversity in a eutrophic freshwater lake. Sci Total Environ. 2018;618:1254–67.

64. Illumina I. High-Speed, Multiplexed 16S Microbial Sequencing on the MiSeq® System [Internet]. 2000. Available from: http://files/998/High-Speed, Multiplexed 16S Microbial Sequencing o.pdf

65. Adak A, Khan MR. An insight into gut microbiota and its functionalities. Cell Mol Life Sci. 2019;76(3):473–93.

66. Kleine Bardenhorst S, Berger T, Klawonn F, Vital M, Karch A, Rübsamen N. Data Analysis Strategies for Microbiome Studies in Human Populations—a Systematic Review of Current Practice. mSystems. 2021;6(1).

67. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852–7.

68. Navas-Molina JA, Peralta-Sánchez Juan M., González A, McMurdie PJ, Vázquez-Baeza Y, Caporaso JG, Knight R. Advancing our understanding of the human microbiome using. Methods Enzym. 2013;531:371–444.

69. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537–41.

70. Edgar RC. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10(10):996–8.

71. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3.

72. Mc. Donald Daniel, Price MN, Goodrich JK, Nawrocki E, Todd DS, Alexander P, *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 2012;6:610–8.

73. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 2013;41(D1):590–6.

74. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, *et al.* Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol. 2017;67(5):1613–7.

75. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, *et al.* PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res. 2014;42(D1):581–91.

76. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, *et al.* Microbiology: Diversity of the human intestinal microbial flora. Science (80-). 2005;308(5728):1635–8.

77. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature. 2012;490(7418):55–60.

78. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207–14.

79. Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. Vol. 8, Genome Medicine. 2016. p. 1–11.

80. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, *et al.* Strains, functions and dynamics in the expanded Human Microbiome Project. Nature. 2017;550(7674):61–6.

81. Jeremiah J. Faith, Guruge JL, Charbonneau M, Sathish Subramanian HS, Goodman AL, Clemente JC, *et al.* The long-term stability of the human gut microbiota. Science (80-). 2013;341(6141).

82. Donaldson GP, Lee SM, Mazmanian SK. Gut biogeography of the bacterial microbiota. Nat Rev Microbiol. 2017;14(1):20–32.

83. Lasken RS, McLean JS. Recent advances in genomic DNA sequencing of microbial species from single cells. Nat Rev Genet. 2014;15(9):577–84.

84. Sekirov I, Russell SL, Caetano M Antunes L, Finlay BB. Gut microbiota in health and disease. Vol. 90, Physiological Reviews. 2010. p. 859–904.

85. Kho ZY, Lal SK. The human gut microbiome - A potential controller of wellness and disease. Vol. 9, Frontiers in Microbiology. 2018. p. 1–23.

86. Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, *et al.* Gut microbiota functions: metabolism of nutrients and other food components. Eur J Nutr. 2018;57(1):1–24.

87. Valdes *et al.* Role of the gut microbiota in nutrition and health. BMJ. 2018;361(J2179).

88. Den Besten G, Van Eunen K, Groen AK, Venema K, Reijngoud DJ, 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(9):2325–40.

89. Rooks MG, Garrett WS. Gut microbiota, metabolites and host inmunity. Nat Publ Gr. 2016;16(6):341–52.

90. Lin H V, Frassetto A, Jr EJK, Nawrocki AR, Lu MM, Kosinski JR, *et al.* Butyrate and Propionate Protect against Diet-Induced Obesity and Regulate Gut Hormones via Free Fatty Acid Receptor 3-Independent Mechanisms. 2012;7(4):1–9.

91. Magnúsdóttir S, Ravcheev D, Crécy-lagard V De, Thiele I. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. Front Genet. 2015;6:148(April 20).

92. Sommer F, Anderson JM, Bharti R, Raes J. The resilience of the intestinal microbiota influences health and disease. Nat Rev Microbiol. 2017;

93. Goodrich J, Waters J, Poole A, Sutter J, Koren O, Blekhman R, *et al.* Human genetics shape the gut microbiome . PubMed Commons. Cell. 2014;159(4):789–99.

94. Benson AK, Kelly SA, Legge R, Ma F, Jen S, Kim J, *et al.* Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. 2010;107(44):18933–8.

95. Dabrowska K, Witkiewicz W. Correlations of host genetics and gut microbiome composition. Front Microbiol. 2016;7(AUG):1–7.

96. Al Nabhani Z, Dietrich G, Hugot JP, Barreau F. Nod2: The intestinal gate keeper. PLoS Pathog. 2017;13(3):1–19.

97. Fallani M, Young D, Scott J, Norin E, Amarri S, Aguilera M, *et al.* Intestinal Microbiota of 6-week-old Infants Across. J Pediatr Gastroenterol Nutr. 2010;51(1):77–84.

98. Yatsunenko T, Rey F, Manary M, I T, MG D-B, Contreras M, *et al.* Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222–7.

99. De Filippo C, Cavalieri D, Di M, Ramazzotti M, Baptiste J. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A. 2010;107(33):14691–6.

100. Scotti E, Boué S, Sasso G Lo, Zanetti F, Belcastro V, Poussin C, *et al.* Exploring the microbiome in health and disease. Toxicol Res Appl. 2017;1:239784731774188.

101. Dominguez-Bello MG, Godoy-Vitorino F, Knight R, Blaser MJ. Role of the microbiome in human development. Gut. 2019;68:1108–14.

102. Dominguez-Bello MG, Blaser MJ, Knight R. Development of the human Gastrointestinal Microbiota and Insights From High-Throughput Sequencing. Gastroenterology. 2011;140(6):1713–9.

103. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, *et al.* Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A. 2011;108 Suppl:4578–85.

104. Albenberg LG, Wu GD. Diet and the Intestinal Microbiome: Associations, Functions and Implications for Health and Disease. Gastroenterology. 2014;146(6):1564–72.

105. Guaraldi F, Salvatori G. Effect of breast and formula feeding on gut microbiota shaping in newborns. Front Cell Infect Microbiol. 2012;2(October):1–4.

106. Campbell SC, Wisniewski PJ, Noji M, Mcguinness LR, Häggblom MM, Lighfoot SA, *et al.* The Effect of Diet and Exercise on Intestinal Integrity and Microbial Diversity in Mice. PLoS One. 2016;11(3).

107. Hsu YJ, Chiu CC, Peng Li Y, Ching Huang W, Li Chuang H. Effect of Intestinal Microbiota on Exercise Permance in Mice. J Strength Cond Res. 2015;29(2):552–8.

108. Queipo-Ortuño I, Seoane ML, Murri M, Pardo M, Gomez-Zumaquero JM, Cardona F, *et al.* Gut Microbiota Composition in Male Rat Models under Different Nutritional Status and Physical Activity and Its Association with Serum Leptin and Ghrelin Levels. PLoS One. 2013;8(5).

109. Allen JM, Mailing LJ, Niemiro GM, Moore R, Cook MD, White BA, *et al.* Exercise Alters Gut Microbiota Composition and Function in Lean and Obese Humans. Med Sci Sport Excercise. 2017;50(4):747–57.

110. Clarke SF, Murphy EF, Sullivan OO, Lucey AJ, Humphreys M, Hogan A, *et al.* Exercise and associated dietary extremes impact on gut microbial diversity. Gut. 2014;63:1913–20.

111. Estaki M, Pither J, Baumeister P, Little JP, Gill SK, Ghosh S, *et al.* Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions. Microbiome. 2016;4:42.

112. Codella R, Luzi L, Terruzzi I. Exercise has the guts: How physical activity may positively modulate gut microbiota in chronic and immune-based diseases. Dig Liver Dis. 2018;50(4):331–41.

113. Biedermann L, Zeitz J, Mwinyi J, Sutter-minder E, Rehman A, Ott SJ, *et al.* Smoking Cessation Induces Profound Changes in the Composition of the Intestinal Microbiota in Humans. PLoS One. 2013;8(3).

114. Biedermann L, Brülisauer K, Zeitz J, Frei P, Scharl M, Vavricka SR, *et al.* Smoking Cessation Alters Intestinal Microbiota: Insights from Quantitative Investigations on Human Fecal Samples Using FISH. Inflamm Bowel Dis. 2014;20(9):1496–501.

115. Hildebrand F, Moitinho-Silva L, Blasche S, Jahn MT, Gossmann TI, Huerta-Cepas J, *et al.* Antibiotics-induced monodominance of a novel gut bacterial order. Gut. 2019;68(10):1781–90.

116. Palleja A, Mikkelsen KH, Forslund SK, Kashani A, Allin KH, Nielsen T, *et al.* Recovery of gut microbiota of healthy adults following antibiotic exposure. Nat Microbiol. 2018;3(11):1255–65.

117. Vich Vila A, Collij V, Sanna S, Sinha T, Imhann F, Bourgonje AR, *et al.* Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. Nat Commun. 2020;11(1):1–11.

118. Iizumi T, Battaglia T, Ruiz V, Perez GIP. Gut Microbiome and Antibiotics. Arch Med Res. 2017;48(8):727–34.

119. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity , stability and resilience of the human gut microbiota. Nature. 2012;489:220–30.

120. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. Nature. 2018;555(7698):623–8.

121. Kolodziejczyk AA, Zheng D, Elinav E. Diet–microbiota interactions and personalized nutrition. Nat Rev Microbiol. 2019;17(12):742–53.

122. Zmora N, Suez J, Elinav E. You are what you eat: diet, health and the gut microbiota. Nat Rev Gastroenterol Hepatol. 2019;16(January):35–56.

123. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, *et al.* Evolution of Mammals and Their Gut Microbes. Science (80-). 2008;777(June):1647–52.

124. Zhang C, Zhang M, Wang S, Han R, Cao Y, Hua W, *et al.* Interactions between gut microbiota , host genetics and diet relevant to development of metabolic syndromes in mice. ISME J. 2010;4:232–41.

125. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, *et al.* Environment dominates over host genetics in shaping human gut microbiota. Nature. 2018;555(7695):210–5.

126. David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, *et al.* Host lifestyle affects human microbiota on daily timescales. Genome Biol. 2014;15(7):R89.

127. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, *et al.* Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505(7484):559–63.

128. Klimenko NS, Tyakht A V., Popenko AS, Vasiliev AS, Altukhov IA, Ischenko DS, *et al.* Microbiome responses to an uncontrolled short-term diet intervention in the frame of the citizen science project. Nutrients. 2018;10(5).

129. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, *et al.* Linking long-term dietary patterns with gut microbial enterotypes. Science (80-). 2011;334(6052):105–8.

130. Leeming ER, Johnson AJ, Spector TD, Roy CIL. Effect of diet on the gut microbiota: Rethinking intervention duration. Nutrients. 2019;11(12):1–28.

131. Tebani A, Bekri S. Paving the way to precision nutrition through metabolomics. Front Nutr. 2019;6(April):1–10.

132. Chong PP, Chin VK, Looi CY, Wong WF, Madhavan P, Yong VC. The microbiome and irritable bowel syndrome - A review on the pathophysiology, current research and future therapy. Front Microbiol. 2019;10(JUN):1–23.

133. Halvorson HA, Schlett CD, Riddle MS, Al-Haddad M. Postinfectious irritable bowel syndrome - A meta-analysis. Am J Gastroenterol. 2006;101(8):1894–9.

134. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science (80-). 2013;341(6150).

135. Sanmiguel C, Gupta A, Mayer EA. Gut Microbiome and Obesity: A Plausible Explanation for Obesity. Curr Obes Rep. 2015;4(2):250–61.

136. Bender E.A., Case T.J., Gilpin M.E. Perturbation experiments in community ecology: theory and practice. Ecology. 1984;65(1):1–13.

137. Xu Z, Knight R. Dietary effects on human gut microbiome diversity. Br J Nutr. 2015;113:S1–5.

138. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, *et al.* Population-level analysis of gut microbiome variation. Science (80-). 2016;352(6285):560–4.

139. Vangay P, Johnson AJ, Ward TL, Al-ghalith GA, Robin R, Hillmann BM, *et al.* U.S. inmigration westernizes the human gut microbiome. Cell. 2019;175(4):962–72.

140. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Chehoud C, *et al.* Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. Gut. 2016;65(1):63–72.

141. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, *et al.* High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut. 2015;65(11).

142. Johnson AJ, Vangay P, Al-Ghalith GA, Hillmann BM, Ward TL, Shields-Cutler RR, *et al.* Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. Cell Host Microbe. 2019;25(6):789-802.e5.

143. Johnson AJ, Zheng JJ, Kang JW, Saboe A, Knights D, Zivkovic AM. A Guide to Diet-Microbiome Study Design. Front Nutr. 2020;7(June):1–16.

144. Hu X, Liu G, Li Y, Wei Y, Lin S, Liu S, *et al.* High-throughput analysis reveals seasonal variation of the gut microbiota composition within forest musk deer (Moschus berezovskii). Front Microbiol. 2018;9(JUL):1–12.

145. Orkin JD, Campos FA, Myers MS, Cheves Hernandez SE, Guadamuz A, Melin AD. Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. ISME J. 2019;13(1):183–96.

146. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, *et al.* Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science (80-). 2017;357(6353):802–5.

147. Thaiss CA, Zeevi D, Levy M, Zilberman-Schapira G, Suez J, Tengeler AC, *et al.* Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell. 2014;159(3):514–29.

148. Kaczmarek JL, Musaad SMA, Holscher HD. Time of day and eating behaviors are associated with the composition and function of the human gastrointestinal microbiota. Am J Clin Nutr. 2017;106(5):1220–31.

149. Collado MC, Engen PA, Bandín C, Cabrera-Rubio R, Voigt RM, Green SJ, *et al.* Timing of food intake impacts daily rhythms of human salivary microbiota: A randomized, crossover study. FASEB J. 2018;32(4):2060–72.

150. Gentile CL, Weir TL. The gut microbiota at the intersection of diet and human health. Science (80-). 2018;362(6416):776–80.

151. Rinninella E, Cintoni M, Raoul P, Lopetuso LR, Scaldaferri F, Pulcini G, *et al.* Food components and dietary habits: Keys for a healthy gut microbiota composition. Nutrients. 2019;11(10).

152. Seo YS, Lee H Bin, Kim Y, Park HY. Dietary carbohydrate constituents related to gut dysbiosis and health. Microorganisms. 2020;8(3).

153. Eid N, Enani S, Walton G, Corona G, Costabile A, Gibson G, *et al.* The impact of date palm fruits and their component polyphenols, on gut microbial ecology, bacterial metabolites and colon cancer cell proliferation. J Nutr Sci. 2014;3(22):1–9.

154. Ruiz-Ojeda FJ, Plaza-Díaz J, Sáez-Lara MJ, Gil A. Effects of Sweeteners on the Gut Microbiota: A Review of Experimental Studies and Clinical Trials. Adv Nutr. 2019;10:S31–48.

155. 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(2):1–57.

156. Holscher HD, Gregory Caporaso J, Hooda S, Brulc JM, Fahey GC, Swanson KS. Fiber supplementation influences phylogenetic structure and functional capacity of the human intestinal microbiome: Follow-up of a randomized controlled trial. Am J Clin Nutr. 2015;101(1):55–64.

157. Candela M, Biagi E, Soverini M, Consolandi C, Quercia S, Severgnini M, *et al.* Modulation of gut microbiota dysbioses in type 2 diabetic patients by macrobiotic Ma-Pi 2 diet. Br J Nutr. 2016;116(1):80–93.

158. Tap J, Furet JP, Bensaada M, Philippe C, Roth H, Rabot S, *et al.* Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults. Environ Microbiol. 2015;17(12):4954–64.

159. Cronin P, Joyce SA, O'toole PW, O'connor EM. Dietary fibre modulates the gut microbiota. Nutrients. 2021;13(5):1–22.

160. Martínez I, Kim J, Duffy PR, Schlegel VL, Walter J. Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. PLoS One. 2010;5(11).

161. Singh RK, Chang H-W, 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):73.

162. Sonnenburg ED, Sonnenburg JL. Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell Metab. 2014;20(5):779–86.

163. O'Grady J, O'Connor EM, Shanahan F. Review article: dietary fibre in the era of microbiome science. Aliment Pharmacol Ther. 2019;49(5):506–15.

164. Zhang C, Zhang M, Pang X, Zhao Y, Wang L, Zhao L. Structural resilience of the gut microbiota in adult mice under high-fat dietary perturbations. ISME J. 2012;6(10):1848–57.

165. Hildebrandt MA, Christian H, Scott A. S-M, Sue A. K, Micah H, Ying-Yu C, *et al.* High Fat Diet Determines the Composition of the Murine Gut Microbiome Independently of Obesity. Gastroenterology. 2009;137(5):1–7.

166. Wolters M, Ahrens J, Romaní-Pérez M, Watkins C, Sanz Y, Benítez-Páez A, *et al.* Dietary fat, the gut microbiota, and metabolic health – A systematic review conducted within the MyNewGut project. Clin Nutr. 2019;38(6):2504–20.

167. Wan Y, Wang F, Yuan J, Li J, Jiang D, Zhang J, *et al.* Effects of dietary fat on gut microbiota and faecal metabolites, and their relationship with cardiometabolic risk factors: a 6-month randomised controlled-feeding trial. Gut. 2019;68(8):1417–29.

168. Fava F, Gitau R, Griffin BA, Gibson GR, Tuohy KM, Lovegrove JA. The type and quantity of dietary fat and carbohydrate alter faecal microbiome and short-chain fatty acid excretion in a metabolic syndrome "at-risk" population. Int J Obes. 2013;37(2):216–23.

169. Watson H, Mitra S, Croden FC, Taylor M, Wood HM, Perry SL, *et al.* A randomised trial of the effect of omega-3 polyunsaturated fatty acid supplements on the human intestinal microbiota. Gut. 2018;67(11):1974–83.

170. Lang JM, Pan C, Cantor RM, Wilson Tang WH, Garcia-Garcia JC, Kurtz I, *et al.* Impact of individual traits, saturated fat, and protein source on the gut microbiome. MBio. 2018;9(6):1–14.

171. Cotillard A, Kennedy SP, Kong LC, Prifti E, Pons N, Le Chatelier E, *et al.* Dietary intervention impact on gut microbial gene richness. Nature. 2013;500(7464):585–8.

172. Reddy BS, Weisburger JH, Wynder EL. Effects of high risk and low risk diets for colon carcinogenesis on fecal microflora and steroids in man. J Nutr. 1975;105(7):878–84.

173. Jantchou P, Morois S, Clavel-Chapelon F, Boutron-Ruault MC, Carbonnel F. Animal protein intake and risk of inflammatory bowel disease: The E3N prospective study. Am J Gastroenterol. 2010;105(10):2195–201.

174. Dostal Webster A, Staley C, Hamilton MJ, Huang M, Fryxell K, Erickson R, *et al.* Influence of short-term changes in dietary sulfur on the relative abundances of intestinal sulfate-reducing bacteria. Gut Microbes. 2019;10(4):447–57.

175. Dominika Ś, Arjan N, Karyn RP, Henryk K. The study on the impact of glycated pea proteins on human intestinal bacteria. Int J Food Microbiol. 2011;145(1):267–72.

176. Jun S. Dietary Vitamin D, Vitamin D receptor, and microbiome. Curr Opin Clin Nutr Metab Care. 2018;21(6):471–4.

177. Constante M, Fragoso G, Lupien-Meilleur J, Calvé A, Santos MM. Iron Supplements Modulate Colon Microbiota Composition and Potentiate the Protective Effects of Probiotics in Dextran Sodium Sulfate-induced Colitis. Inflamm Bowel Dis. 2017;23(5):753–66.

178. Lopez CA, Skaar EP. The impact of dietary transition metals on host-bacterial interactions. Cell Host Microbe. 2018;23(6):737–48.

179. Wilck N, Matus MG, Kearney SM, Olesen SW, Forslund K, Bartolomaeus H, *et al.* Salt-responsive gut commensal modulates TH17 axis and disease. Nature. 2017;551(7682):585–9.

180. Singh AK, Cabral C, Kumar R, Ganguly R, Pandey AK. Beneficial Effects of Dietary Polyphenols on Gut Microbiota and Strategies to Improve. Nutrients. 2019;11(2216).

181. Ozdal T, Sela DA, Xiao J, Boyacioglu D, Chen F, Capanoglu E. The reciprocal interactions between polyphenols and gut microbiota and effects on bioaccessibility. Nutrients. 2016;8(2):1–36.

182. Cassidy A, Minihane AM. The role of metabolism (and the microbiome) in defining the clinical efficacy of dietary flavonoids. Am J Clin Nutr. 2017;105(1):10–22.

183. Bialonska D, Ramnani P, Kasimsetty SG, Muntha KR, Gibson GR, Ferreira D. The influence of pomegranate by-product and punicalagins on selected groups of human intestinal microbiota. Int J Food Microbiol. 2010;140(2–3):175–82.

184. Tzounis X, Rodriguez-Mateos A, Vulevic J, Gibson GR, Kwik-Uribe C, Spencer JPE. Prebiotic evaluation of cocoa-derived flavanols in healthy humans by using a randomized, controlled, double-blind, crossover intervention study. Am J Clin Nutr. 2011;93(1):62–72.

185. Hidalgo M, Oruna-Concha MJ, Kolida S, Walton GE, Kallithraka S, Spencer JPE, *et al.* Metabolism of anthocyanins by human gut microflora and their influence on gut bacterial growth. J Agric Food Chem. 2012;60(15):3882–90.

186. Laudisi F, Stolfi C, Monteleone G. Impact of food additives on gut homeostasis. Nutrients. 2019;11(10).

187. Gültekin F. Food Additives And Microbiota. North Clin Istanbul. 2019;7(2):192–200.

188. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, *et al.* Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature. 2014;514(7521):181–6.

189. Chassaing B, Bodt J De, Marzorati M, Wiele T Van de, Gewirtz AT. Dietary emulsifiers directly alter human microbiota composition and gene expression ex vivo potentiating intestinal inflammation. Gut. 2017;66(8):1414–27.

190. Rinninella E, Cintoni M, Raoul P, Mora V, Gasbarrini A, Mele MC. Impact of food additive titanium dioxide on gut microbiota composition, microbiota-associated functions, and gut barrier: A systematic review of in vivo animal studies. Int J Environ Res Public Health. 2021;18(4):1–21.

191. Craig WJ, Mangels AR. Position of the American Dietetic Association: vegetarian diets. J Am Diet Assoc. 2009;109(7):1266–82.

192. Ruengsomwong S, La Ongkham O, Jiang J, Wannissorn B, Nakayama J, Nitisinprasert S. Microbial community of healthy Thai vegetarians and non-vegetarians, their core gut microbiota and pathogens risk. J Microbiol Biotechnol. 2016;26(10):1723–35.

193. Matijašić BB, Obermajer T, Lipoglavšek L, Grabnar I, Avguštin G, Rogelj I. Association of dietary type with fecal microbiota in vegetarians and omnivores in Slovenia. Eur J Nutr. 2014;53(4):1051–64.

194. 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(MAR):1–12.

195. Zhang C, Björkman A, Cai K, Liu G, Wang C, Li Y, *et al.* Impact of a 3-months vegetarian diet on the gut microbiota and immune repertoire. Front Immunol. 2018;9(APR):1–13.

196. Drasar BS, Crowther JS, Goddard P, Hawksworth G, Hill MJ, Peach S, *et al.* The relation between diet and the gut microflora in man. Proc Nutr Soc. 2007;32(2):49–52.

197. Ghosh TS, Rampelli S, Jeffery IB, Santoro A, Neto M, Capri M, *et al.* Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: The NU-AGE 1-year dietary intervention across five European countries. Gut. 2020;69(7):1218–28.

198. Meslier V, Laiola M, Roager HM, De Filippis F, Roume H, Quinquis B, *et al.* Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. Gut. 2020;69(7):1258–68.

199. Muralidharan J, Moreno-Indias I, Bulló M, Lopez JV, Corella D, Castañer O, *et al.* Effect on gut microbiota of a 1-y lifestyle intervention with Mediterranean diet compared with energy-reduced Mediterranean diet and physical activity promotion: PREDIMED-Plus Study. Am J Clin Nutr. 2021;00:1–11.

200. Reddel S, Putignani L, Del Chierico F. The impact of low-FODMAPs, Gluten-free, and ketogenic diets on gut microbiota modulation in pathological conditions. Nutrients. 2019;11(2).

201. Day NE, McKeown N, Wong MY, Welch A, Bingham S. Epidemiological assessment of diet: A comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. Int J Epidemiol. 2001;30(2):309–17.

202. Shanahan ER, McMaster JJ, Staudacher HM. Conducting research on diet–microbiome interactions: A review of current challenges, essential methodological principles, and recommendations for best practice in study design. J Hum Nutr Diet. 2021;00:1–14.

203. Shim J-S, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014;36:e2014009.

204. Jennings A, Cashman KD, Gillings R, Cassidy A, Tang J, Fraser W, *et al.* A Mediterranean-like dietary pattern with Vitamin D3 (10 µg/d) supplements reduced the rate of bone loss in older Europeans with osteoporosis at baseline: Results of a 1-y randomized controlled trial. Am J Clin Nutr. 2018;108(3):633–40.

205. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, *et al.* Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: Relationship with gut microbiome richness and ecology. Gut. 2016;65(3):426–36.

206. Bellikci-Koyu E, Sarer-Yurekli BP, Akyon Y, Aydin-kose F, Karagozlu C, Ozgen AG, *et al.* Effects of Regular Kefir Consumption on Gut Microbiota in Patients with Metabolic Syndrome : Nutrients. 2019;11:1–23.

207. Fujio-Vejar S, Vasquez Y, Morales P, Magne F, Vera-Wolf P, Ugalde JA, *et al.* The gut microbiota of healthy Chilean subjects reveals a high abundance of the phylum Verrucomicrobia. Front Microbiol. 2017;8(JUN):1–11.

208. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, *et al.* Personalized Nutrition by Prediction of Glycemic Responses. Cell. 2015;163(5):1079–94.

209. Bolte LA, Vich Vila A, Imhann F, Collij V, Gacesa R, Peters V, *et al.* Long-term dietary patterns are associated with pro-inflammatory and anti-inflammatory features of the gut microbiome. Gut. 2021;70(7):1287–98.

210. Asnicar F, Berry SE, Valdes AM, Nguyen LH, Piccinno G, Drew DA, *et al.* Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. Nat Med. 2021;27(2):321–32.

211. Gacesa R, Kurilshikov A, Vich Vila A, Sinha T, Klaassen MAY, Bolte LA, *et al.* The Dutch Microbiome Project defines factors that shape the healthy gut microbiome. bioRxiv. 2020;1–33.

212. Taylor BC, Lejzerowicz F, Poirel M, Shaffer JP, Jiang L, Aksenov A, *et al.* Consumption of Fermented Foods Is Associated with Systematic Differences in the Gut Microbiome and Metabolome. mSystems. 2020;5(2).

213. Le Roy CI, Wells PM, Si J, Raes J, Bell JT, Spector TD. Red Wine Consumption Associated With Increased Gut Microbiota α-Diversity in 3 Independent Cohorts. Gastroenterology. 2020;158(1):270-272.e2.

214. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, *et al.* Association of dietary patterns with the gut microbiota in older, community-dwelling men. Am J Clin Nutr. 2019;110(4):1003–14.

215. McDonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G, *et al.* American Gut : an Open Platform for Citizen Science. mSystems. 2018;3(3):1–28.

216. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, *et al.* Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;488(7410):178–84.

217. Ruiz, María Dolores; Artacho R. Guía para estudios dietéticos: álbum fotográfico de alimentos. 1º ed. Universidad de Granada, editor. 2011. 1–128 p.

218. SU-VI-MAX (France). Portions alimentaires : manuel-photos pour l'estimation des quantités. Polytechnica; 2002.

219. Fundación Iberoamericana de Nutrición. EVALFINUT [Internet]. 2015. Available from: https://eval.finut.org/

220. Cuadernos de ejercicios para desarrollar etiquetas nutricionales de alimentos | Santiago de Chile; Chile. Ministerio de Salud; oct. 2006. 82 p. | LILACS [Internet]. [cited 2020 Jun 13]. Available from: https://pesquisa.bvsalud.org/portal/resource/es/lil-669788

221. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol. 1985;122(1):51–65.

222. Marcos SV, Rubio MJ, Sanchidrián FR RD. Spanish National dietary survey in adults, elderly and pregnant women. EFSA Support Publ. 2016;13(6).

223. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, *et al.* Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr. 2010;103(12):1808–16.

224. Rodríguez IT, Fernández Ballart J, Cucó Pastor G, Biarnés Jordà E, Arija Val V. Validación de un cuestionario de frecuencia de consumo alimentario corto: Reproducibilidad y validez. Nutr Hosp. 2008;23(3):242–52.

225. Verger EO, Armstrong P, Nielsen T, Chakaroun R, Aron-Wisnewsky J, Gøbel RJ, *et al.* Dietary Assessment in the MetaCardis Study: Development and Relative Validity of an Online Food Frequency Questionnaire. J Acad Nutr Diet. 2017;117(6):878–88.

226. Haftenberger M, Heuer T, Heidemann C, Kube F, Krems C, Mensink GB. Relative validation of a food frequency questionnaire for national health and nutrition monitoring. Nutr J. 2010;9(1):36.

227. Kroke A, Klipstein-Grobusch K, Voss S, Möseneder J, Thielecke F, Noack R, *et al.* Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water,. Am J Clin Nutr. 1999;70(4):439–47.

228. Rinninella E, Cintoni M, Raoul P, Lopetuso LR, Scaldaferri F, Pulcini G, *et al.* Health Gut Microbiota. Nutrients. 2019;1–23.

229. Battaglioli EJ, Kashyap PC. Diet Effects on Gut Microbiome Composition, Function, and Host Physiology. Sixth Edit. Vols. 1–2, Physiology of the Gastrointestinal Tract: Sixth Edition. Elsevier Inc.; 2018. 755–766 p.

230. Cani PD, Everard A. Talking microbes: When gut bacteria interact with diet and host organs. Mol Nutr Food Res. 2016;60(1):58–66.

231. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: An integrative view. Cell. 2012;148(6):1258–70.

232. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Gordon JI. Metagenomic Analysis in Humanized Gnotobiotic Mice. Sci Transl Med. 2009;1(6):1–19.

233. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A. 2010;107(33):14691–6.

234. Sociedad Española Nutrición Comunitaria. Consejos para una Alimentación saludable [Internet]. 2016. p. 42. Available from: https://www.semfyc.es/wp-content/uploads/2016/07/guia_alimentacion.pdf

235. Basulto J, Moñino M, Farran A, Baladia E, Manera M, Cervera P, *et al.* Establecimiento del tamaño de raciones de consumo de frutas y hortalizas para su uso en guías alimentarias en el entorno español: propuesta del Comité Científico de la Asociación 5 al día. Rev Española Nutr Humana y Dietética [Internet]. 2019;18(2):100–15. Available from: https://dialnet.unirioja.es/servlet/articulo?codigo=7283395

236. Posición del comité Cientiíico 5 al día. Raciones de frutas y hortalizas en España [Internet]. 2010. Available from: https://www.5aldia.org/datos/60/Documento_Raciones_de_Frutas_y_Hortalizas_8944.pdf

237. Nagpal R, Mainali R, Ahmadi S, Wang S, Singh R, Kavanagh K, *et al.* Gut microbiome and aging: Physiological and mechanistic insights. Nutr Heal Aging. 2018;4(4):267–85.

238. Kim YS, Unno T, Kim BY, Park MS. Sex differences in gut microbiota. World J Mens Health. 2020;38(1):48–60.

239. Gao X, Zhang M, Xue J, Huang J, Zhuang R, Zhou X, *et al.* Body mass index differences in the gut microbiota are gender specific. Front Microbiol. 2018;9(JUN):1–10.

240. Arnolds KL, Martin CG, Lozupone CA. Blood type and the microbiome- untangling a complex relationship with lessons from pathogens. Curr Opin Microbiol. 2020;56:59–66.

241. Reyman M, Van Houten MA, Van Baarle D, Bosch AT, Man WH, Chu ML, *et al.* Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. Nat Commun. 2019;10(1):1–12.

242. Zimmer J, Lange B, Frick JS, Sauer H, Zimmermann K, Schwiertz A, *et al.* A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. Eur J Clin Nutr. 2012;66(1):53–60.

243. Karl JP, Armstrong NJ, McClung HL, Player RA, Rood JC, Racicot K, *et al.* A diet of U.S. military food rations alters gut microbiota composition and does not increase intestinal permeability. J Nutr Biochem. 2019;72:108217.

244. Subar AF, Thompson FE, Smith AF, Jpbe JB, Ziegler RG, Potischman N, *et al.* Improving Food Frequency Questionnaires. A Qualitative Approach Using Cognitive Interviewing. Vol. 95, Journal of the American Dietetic Association. 1995. p. 781–8.

245. Bostanci N, Krog MC, Hugerth LW, Bashir Z, Fransson E, Boulund F, *et al.* Dysbiosis of the Human Oral Microbiome During the Menstrual Cycle and Vulnerability to the External Exposures of Smoking and Dietary Sugar. Front Cell Infect Microbiol. 2021;11(March):1–14.

246. Vilela S, Lopes C, Guiomar S, Severo M, Rangelova L, Petrova S, *et al.* Validation of a picture book to be used in a pan-European dietary survey. Public Health Nutr. 2018;21(9):1654–63.

247. Lazarte CE, Encinas ME, Alegre C, Granfeldt Y. Validation of digital photographs, as a tool in 24-h recall, for the improvement of dietary assessment among rural populations in developing countries. Nutr J. 2012;11(1):1–14.

248. Subar A, Crafts J, Palmer ZT, Mcnutt S, Guenther PM, Willis G, *et al.* Assessment of the Accuracy of Portion Size Reports Using Computer-Based Food Photographs Aids in the Development of an Automated Self-Administered 24-Hour Recall. J Am Diet Assoc. 2010;110(1):55–64.

249. Nelson M, Atkinson M, Darbyshire S. Food Photography II : Portion Size And The Nutrient Content of Meals. Br J Nutr. 1996;76:31–49.

250. Nelson M, Atkinson M, Darbyshire S. Food Photography I: The perception of food portion size from photographs. Br J Nutr. 1994;72(5):649–63.

251. Ireland J, Erp-Baart A Van, A. Møller, Smithers G, Trichopoulou A. Selection of a food classification system and a food composition database for future food consumption surveys. Eur J Clin Nutr. 2002;56:33–45.

252. Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, *et al.* Comparative validation of the Block, Willett, and National Cancer Institute Food Frequency Questionnaires: The Eating at America's Table Study. Am J Epidemiol. 2001;154(12):1089–99.

253. Brown CC, Kipnis V, Freedman LS, Hartman AM, Schatzkm A, Wacholder S. Energy adjustment methods for nutritional epidemiology: The effect of categorization. Am J Epidemiol. 1994;139(3):323–38.

254. Bland JM, Altman GD. Statistical methods for assessing agreement between two methods of clinical measurament. Lancet. 1986;327(8476):307–10.

255. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63.

256. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.

257. Masson L, MCNeill G, Tomany J, Simpson J, Peace H, Wei L, *et al.* Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. Public Health Nutr. 2003;6(3):313–21.

258. Nickerson C. A Note On "A Concordance Correlation Coefficient to Evaluate Reproducibility." Biometrics. 1997;53(4):1503.

259. Científico C. Informe del Comité Científico de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) sobre Ingestas Nutricionales de Referencia para la población española. Rev del Com científico [Internet]. 2019; Available from: https://www.aesan.gob.es/AECOSAN/docs/documentos/seguridad_alimentaria/evaluacion_riesgos/informes_comite/INR.pdf

260. Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary. Using Dietary Reference Intakes for Nutrient Assessment of Individuals. 2000 [cited 2021 Feb 17]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK222891/

261. Food E, Authority S. Guidance on the EU Menu methodology 1. EFSA J. 2014;12(12).

262. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake : basal metabolic rate . A practical guide to its calculation , use and limitations. Int J Obes. 2000;24:1119–30.

263. Livingstone MBE, Black AE. Biomarkers of Nutritional Exposure and Nutritional Status Markers of the Validity of Reported Energy Intake 1. J Nutr. 2003;133(Mar):Suppl 3:895S-920S.

264. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr. 1985;39 Suppl 1:5–41.

265. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, *et al.* Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol. 2017;35(11):1069–76.

266. Santiago A, Panda S, Mengels G, Martinez X, Azpiroz F, Dore J, *et al.* Processing faecal samples: A step forward for standards in microbial community analysis. BMC Microbiol. 2014;14(1):1–9.

267. Illumina Inc. Illumina sequencing introduction. Illumina Seq Introd [Internet]. 2017;(October):1–8. Available from: https://www.illumina.com/documents/products/illumina_sequencing_introduction.pdf

268. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Tickle TL, *et al.* Multivariable Association Discovery in Population-scale Meta-omics Studies 3. bioRxiv. 2021;2021.01.20.427420.

269. Kristal AR, Peters U, Potter JD. Is it time to abandon the food frequency questionnaire? Cancer Epidemiol Biomarkers Prev. 2005;14(12):2826–8.

270. Toorang F, Sasanfar B, Esmaillzadeh A, Ebrahimpour-Koujan S, Zendehdel K. Comparison of validity of the food frequency questionnaire and the diet history questionnaire for assessment of energy and nutrients intakes in an Iranian population. East Mediterr Heal J. 2020;26(9):1062–9.

271. Thompson FE, Subar AF. Dietary assessment methodology [Internet]. Fourth Edi. Nutrition in the Prevention and Treatment of Disease. Elsevier Inc.; 2017. 5–48 p. Available from: http://dx.doi.org/10.1016/B978-0-12-802928-2.00001-1

272. Steinemann N, Grize L, Ziesemer K, Kauf P, Probst-Hensch N, Brombach C. Relative validation of a food frequency questionnaire to estimate food intake in an adult population. Food Nutr Res. 2017;61(1305193).

273. Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, *et al.* Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. Am J Epidemiol. 2003 Jul 1;158(1):1–13.

274. Villena-Esponera MP, Moreno-Rojas R, Romero-Saldaña M, Molina-Recio G. Validation of a Food Frequency Questionnaire for the indigenous Épera-Siapidara people in Ecuador. Nutr Hosp. 2017;34(6):1368–75.

275. Harmouche-Karaki M, Mahfouz M, Obeyd J, Salameh P, Mahfouz Y, Helou K. Development and validation of a quantitative food frequency questionnaire to assess dietary intake among Lebanese adults. Nutr J. 2020;19(1):1–19.

276. Leong C, Taylor RW, Haszard JJ, Fleming EA, Tannock GW, Szymlek-Gay EA, *et al.* Relative validity and reproducibility of a food frequency questionnaire to assess nutrients and food groups of relevance to the gut microbiota in young children. Nutrients. 2018;10(11).

277. Maurer J, Taren DL, Teixeira PJ, Thomson CA, Lohman TG, Going SB, *et al.* The psychosocial and behavioral characteristics related to energy misreporting. Nutr Rev. 2006;64(2):53–66.

278. Tooze JA, Subar AF, Thompson F, Troiano RP, Schatzkin A, Kipnis V. Psychosocial predictors of energy underreporting in a large doubly labeled water study. Am J Clin Nutr. 2004;79:795–804.

279. Hebert JR, Hurley TG, E. PK, Ken R, Thompson FE, Yaroch AL, *et al.* Social desirability trait influences on self-report dietary measures among diverse participants in a multicenter multiple risk factor trial. J Nutr. 2008;138:226S-34S.

280. Cook A, Pryer J, Shetty P. The problem of accuracy in dietary surveys. Analysis of the over 65 UK National Diet and Nutrition Survey. J Epidemiol Community Health. 2000;54(8):611–6.

281. Miller TM, Abdel-Maksoud MF, Crane LA, Marcus AC, Byers TE. Effects of social approval bias on self-reported fruit and vegetable consumption: A randomized controlled trial. Nutr J. 2008;7(1):1–7.

282. Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German Part of the EPIC project. Int J Epidemiol. 1997;26(Suppl 1):S59-70.

283. Sierra-Ruelas É, Bernal-Orozco MF, MacEdo-Ojeda G, Márquez-Sandoval YF, Altamirano-Martínez MB, Vizmanos B. Validation of semiquantitative FFQ administered to adults: A systematic review. Public Health Nutr. 2020;(3).

284. Molag ML, De Vries JHM, Ocké MC, Dagnelie PC, Van Den Brandt PA, Jansen MCJF, *et al.* Design characteristics of food frequency questionnaires in relation to their validity. Am J Epidemiol. 2007;166(12):1468–78.

285. Ruiz E, Ávila JM, Valero T, Pozo S Del, Rodriguez P, Aranceta-Bartrina J, *et al.* Energy intake, profile, and dietary sources in the spanish population: Findings of the ANIBES study. Nutrients. 2015;7(6):4739–62.

286. Ruiz E, Ávila JM, Valero T, del Pozo S, Rodriguez P, Aranceta-Bartrina J, *et al.* Macronutrient distribution and dietary sources in the spanish population: Findings from the ANIBES study. Nutrients. 2016;8(3).

287. Partearroyo T, Samaniego-Vaesken M de L, Ruiz E, Aranceta-Bartrina J, Gil Á, González-Gross M, *et al.* Current food consumption amongst the spanish anibes study population. Nutrients. 2019;11(11):1–15.

288. Olza J, Aranceta-Bartrina J, González-Gross M, Ortega RM, Serra-Majem L, Varela-Moreiras G, *et al.* Reported dietary intake and food sources of zinc, selenium, and vitamins A, E and C in the spanish population: Findings from the anibes study. Nutrients. 2017;9(7).

289. Olza J, de Victoria EM, Aranceta-Bartrina J, González-Gross M, Ortega RM, Serra-Majem L, *et al.* Adequacy of critical nutrients affecting the quality of the Spanish diet in the ANIBES study. Nutrients. 2019;11(10):1–17.

290. Olza J, Aranceta-Bartrina J, González-Gross M, Ortega RM, Serra-Majem L, Varela-Moreiras G, *et al.* Reported dietary intake, disparity between the reported consumption and the level needed for adequacy and food sources of calcium, phosphorus, magnesium and vitamin D in the Spanish population: Findings from the ANIBES study. Nutrients. 2017;9(2).

291. Aranow C. Vitamin D and the immune system. J Investig Med. 2011;59(6):881–6.

292. Palacios C, Gonzalez L. Is vitamin D deficiency a major global public health problem? J Steroid Biochem Mol Biol. 2014;October(114PA):138–45.

293. González-Molero I, Morcillo S, Valdés S, Pérez-Valero V, Botas P, Delgado E, *et al.* Vitamin D deficiency in Spain: A population-based cohort study. Eur J Clin Nutr. 2011;65(3):321–8.

294. Thomas RL, Jiang L, Adams JS, Xu ZZ, Shen J, Janssen S, *et al.* Vitamin D metabolites and the gut microbiome in older men. Nat Commun. 2020;11(1).

295. Bowyer RCE, Jackson MA, Pallister T, Skinner J, Spector TD, Welch AA, *et al.* Use of dietary indices to control for diet in human gut microbiota studies. Microbiome. 2018;6(1):77.

296. García-Vega ÁS, Corrales-Agudelo V, Reyes A, Escobar JS. Diet quality, food groups and nutrients associated with the gut microbiota in a nonwestern population. Nutrients. 2020;12(10):1–21.

297. Maldonado-Contreras A, Noel SE, Ward D V., Velez M, Mangano KM. Associations between Diet, the Gut Microbiome, and Short-Chain Fatty Acid Production among Older Caribbean Latino Adults. J Acad Nutr Diet. 2020;120(12):2047-2060.e6.

298. Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, *et al.* Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat Genet. 2021;53(2):156–65.

299. Bokulich NA, Chung J, Battaglia T, Henderson N, Jay M, Li H, *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Transl Med. 2016;8:343ra82.

300. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A. 2010;107(26):11971–5.

301. Jakobsson HE, Abrahamsson TR, Jenmalm MC, Harris K, Quince C, Jernberg C, *et al.* Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by Caesarean section. Gut. 2014;63(4):559–66.

302. Busi SB, de Nies L, Habier J, Wampach L, Fritz J V., Heintz-Buschart A, *et al.* Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. ISME Commun. 2021;1(1):1–12.

303. De la Cuesta-Zuluaga J, Kelley ST, Chen Y, Escobar JS, Mueller NT, Ley RE, *et al.* Age- and Sex-Dependent Patterns of Gut Microbial Diversity in Human Adults. mSystems. 2019;4(4).

304. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, *et al.* Richness of human gut microbiome correlates with metabolic markers. Nature. 2013;500(7464):541–6.

305. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, *et al.* Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature. 2014;514(7521):181–6.

306. Serrano J, Smith K, Crouch A, Sharma V, Yi F, Vargova V, *et al.* High dose saccharin supplementation does not induce gut microbiota dysbiosis or glucose intolerance in healthy humans and mice. Microbiome. 2020;9:11.

307. 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(6):705–15.

308. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. Diet-induced extinctions in the gut microbiota compound over generations. Nature. 2016;529(7585):212–5.

309. Lin D, Peters BA, Friedlander C, Freiman HJ, Goedert JJ, Sinha R, *et al.* Association of dietary fibre intake and gut microbiota in adults. Br J Nutr. 2018;120(9):1014–22.

310. Carmody RN, Bisanz JE, Bowen BP, Maurice CF, Lyalina S, Louie KB, *et al.* Cooking shapes the structure and function of the gut microbiome. Nat Microbiol. 2019;4(12):2052–63.

311. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, *et al.* Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45.

312. Willett WC, Hu FB. Not the time to abandon the food frequency questionnaire: Point. Cancer Epidemiol Biomarkers Prev. 2006;15(10):1757–8.

313. Beaton GH, Burema J, Ritenbaugh C. Errors in the interpretation of dietary assessments. Am J Clin Nutr. 1997;65((suppl)):1100S-7S.

314. Barabási AL, Menichetti G, Loscalzo J. The unmapped chemical complexity of our diet. Nat Food. 2020;1(1):33–7.

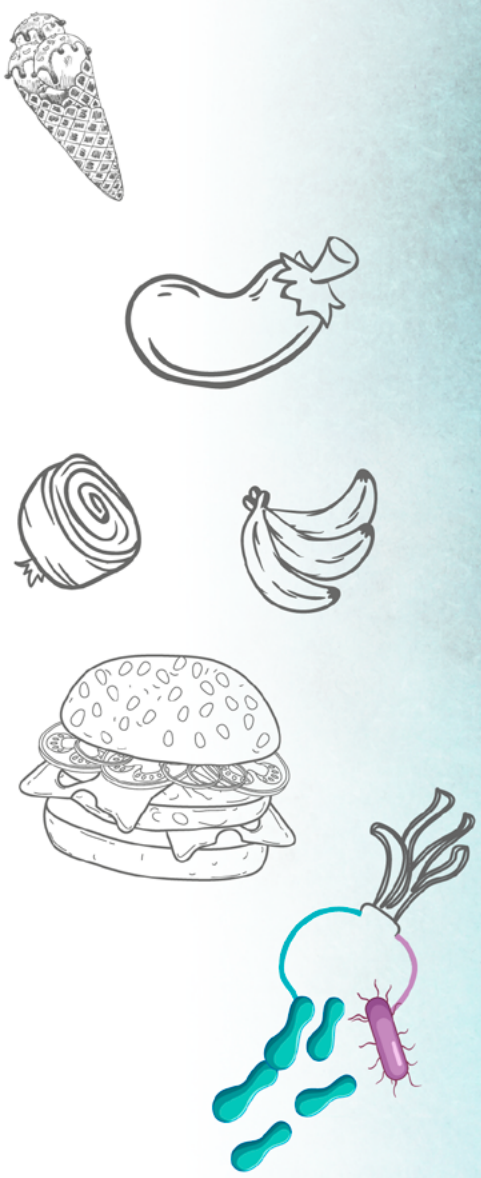
315. Fahey GC, Novotny L, Layton B, Mertens DR. Critical factors in determining fiber content of feeds and foods and their ingredients. J AOAC Int. 2019;102(1):52–62.

316. Mañas E, Bravo L, Saura-Calixto F. Sources of error in dietary fibre analysis. Food Chem. 1994;50(4):331–42.

317. Emmett P, Symes C, Braddon F, Heaton K. Validation of a new questionnaire for assessing habitual intake of starch, non-starch polysaccharides, sugars and alcohol. J Hum Nutr Diet. 1992;5(4):245–53.

318. Healey G, Brough L, Murphy R, Hedderley D, Butts C, Coad J. Validity and reproducibility of a habitual dietary fibre intake short food frequency questionnaire. Nutrients. 2016;8(9):3–9.

319. Dunn S, Datta A, Kallis S, Law E, Myers CE, Whelan K. Validation of a food frequency questionnaire to measure intakes of inulin and oligofructose. Eur J Clin Nutr. 2011;65(3):402–8.



10. ANNEXES

ANNEX 1. 24HR COLLECTION FORM

Encuesta alimentaria de Recordatorio de 24 horas

N° de Encuesta (1:1° día / 2:2° día / 3:3° día): ☐

Nombre encuestado..... No. Identificación .....

Nombre encuestador.....

Día de semana: Lu ☐ Ma ☐ Mi ☐ Ju ☐ Vi ☐ Sa ☐ Do ☐ Fecha ____/____/____

Hora	Menu (alimento o preparaciones)	Ingredientes	Código	Cantidad medidas caseras	Cantidad gr.total	Observaciones
Desayuno						
: hrs.						
Lugar:						
Media Mañana						
: hrs.						
Lugar						
Comida						
: hrs.						
Lugar						

Hora	Menu (alimento o preparaciones)	Ingredientes	Código	Cantidad medidas caseras	Cantidad gr.total	Observaciones
Merienda						
: hrs.						
Lugar						
Cena						
: hrs.						
Lugar						
Entre Horas						
: hrs.						
Lugar						

PREGUNTAS ADICIONALES

Usted utiliza regularmente suplementos de vitaminas o minerales? **Si ( ) No ( )**

En el caso afirmativo, cuantas veces **Por día?** _____ **Por semana?** _____

¿Cuál es la cantidad consumida/vez? _____

En el caso afirmativo, qué tipo de suplemento utiliza (nombre, marca, etc)?

_____

ANNEX 2. STANDARD OPERATING PROCEDURE TO DATA COLLECTION OF THE 24HRs

Vall d'Hebron Institut de Recerca VHIR	**Procedimiento de registro de recordatorio alimentario de 24 horas**	Code : Date: 20/01/2017 Updated: 21/02/2020 Pages: 3
Date: 20/01/2017 **Updated: 21/02/2020**	**MICROBIOME LAB**	Prepared by: Msc. Francisca Yáñez Reviewed by:

1.- OBJETIVO
Establecer un procedimiento que estandarice la realización y registro de un recordatorio alimentario de 24 hrs a participantes del estudio de dieta y microbiota intestinal.

2.- ALCANCE
Aplicable a todos los participantes que cumplan con los criterios de inclusión y hayan firmado el consentimiento informado de participación en el estudio.

3.- DOCUMENTOS ASOCIADOS
"Guía para estudios dietéticos". Álbum fotográfico de alimentos, Fotografías de alimentos del sFFQ, "Tabla estandarizada de medidas caseras", "Base de datos de composición de alimentos", "Tabla de factores de conversión de peso de alimentos cocidos a crudos"

4.- RESPONSABILIDAD
Responsable ejecución: Personal previamente capacitado en la realización del recordatorio alimentario de 24 horas.
Responsable monitoreo: Dietista/Nutricionista a cargo del estudio.

5.- DEFINICIONES (1,2)
Alimento: toda sustancia, elaborada, semielaborada o bruta, que se destina al consumo humano, incluyendo las bebidas, goma de mascar y cualesquiera otras sustancias que se utilicen en la fabricación, preparación o tratamiento de los alimentos, pero no incluye los cosméticos ni el tabaco ni las sustancias utilizadas solamente como medicamentos.
Ingrediente: cualquier sustancia, incluidos los aditivos alimentarios, que se empleen en la fabricación o preparación de un alimento y esté presente en el producto final, aunque posiblemente en forma modificada.
Preparación culinaria: mezcla o unión de determinados ingredientes que se combinan o se consumen en un mismo plato o recipiente o preparación; ya sea que han sido sometidos a cocción o no.
Minuta o menú: conjunto de alimentos y/o preparaciones consideradas en un tiempo comida.
Medida casera: utensilios comúnmente utilizados por el consumidor para medir alimentos.
Ración: cantidad estándar de alimento que se aconseja consumir en las guías alimentarias tratando de que aporte una cantidad similar de los nutrientes principales
Porción: cantidad de un alimento, expresada en medidas caseras, que generalmente es consumida por una persona en una oportunidad y que varía de una comida a otra y de una persona a otra.
Peso bruto: peso de alimento entero tal y como se compra
Peso neto: peso de la parte comestible del alimento.

6.- INSUMOS
Bolígrafo, Formulario de registro de 24h, "Guía para estudios dietéticos Álbum fotográfico de alimentos", Fotografías de alimentos del FFQ, "Tabla estandarizada de medidas caseras" y "Base de datos de composición de alimentos"

7.- PROCEDIMIENTO (3)

1º Paso: Información básica
- Completar datos del participante que permita su identificación (código numérico específico para cada participante) y código del entrevistador
- Completar datos que permitan identificar "*FECHA DE LA ENTREVISTA*" y "DÍA DE LA SEMANA" al que corresponde el recordatorio alimentario.

2º Paso: Lista rápida de alimentos
- En la columna "*MENÚ*", registrar todos los alimentos o preparaciones consumidas por el participante durante el día anterior (desde que se levantó hasta que se acostó). Preguntar directamente por el consumo de grupo de alimentos como frutas, vegetales, lácteos. etc
- En la columna "*HORA*" registrar los horarios y ocasión en que fueron consumidas dichas preparaciones.
- Dejar que participante indique todo lo que consumió sin intervenir.

3ª Paso: Registro detallado de cada alimento/preparación/bebida reportado
- En la columna *"INGREDIENTES*" registrar los alimentos utilizados en cada preparación detallando si fueron consumidos en crudo, congelado o cocido, con o sin piel, con o sin cascara, con o sin hueso, enlatados con o sin aceite, en conserva, con o sin alcohol.
- Con apoyo de la "Guía para estudios dietéticos Álbum fotográfico de alimentos" y Fotografías de alimentos del sFFQ, completar la columna "CANTIDAD EN MEDIDAS CASERAS" utilizadas en cada alimento/ingrediente
- En "*OBSERVACIONES"* registrar el lugar de consumo, marca comercial u otra característica no mencionada.
- En el caso, que se haya realizado una preparación casera como pastel, bizcocho, etc registrar el número de porciones totales que se obtienen de la receta e ingredientes con sus respectivas medidas caseras o gramajes.
- Para evitar olvidar algún alimento/preparación, ir en orden de consumo en el día.
- Dejar que el participante indique todo lo que consumió sin intervenir, excepto para reforzar si hubo algo más que consumió o extrapolar una cantidad x de alimento a medida casera.

4º Paso: Revisión final
- Repasar la lista completa de alimentos consumidos con sus respectivas "HORAS", "INGREDIENTES", "cantidad en medida caseras" y "observaciones" por posible olvido de algún alimento consumido fuera del horario.
- Reforzar si hubo algo más que consumió, como té, café, onza de chocolate, aceite. etc
- Registrar el consumo de suplementos dietéticos, vitaminas o minerales con su respectivo nombre, dosis y frecuencias de consumo en "PREGUNTAS ADICIONALES".

Una vez finalizada la entrevista, el entrevistador deberá completar las columnas "CANTIDAD GR TOTALES" y "CÓDIGO" de cada alimento/preparación, registrada en el Formulario de registro de 24h.

5º Paso: Transformación de medidas caseras a gramajes

- Con apoyo de la "Guía para estudios dietéticos Álbum fotográfico de alimentos" y "Tabla estandarizada de medidas caseras", transformar las medidas caseras de cada alimento a gramajes y registrar en columna "CANTIDAD GR TOTALES".
- Para la transformación de peso cocido a crudo, utilizar Álbum fotográfico de alimentos" y "Tabla estandarizada de medidas caseras" o "Tabla de factores de conversión de peso de alimentos cocidos a crudos"
- En el caso que no aparezca el gramaje de la porción casera de un alimento determinado, buscar en otras encuestas, etiquetado nutricional (en el caso que se allá registrado la marca comercial del alimento) o comprar el alimento y promediar su peso con la medida casera en cuestión. **Ante cualquier, duda preguntar al responsable del monitoreo.**

6º Paso: Codificación de alimentos

- En "Base de datos de composición de alimentos", buscar el "código" correspondiente para cada alimento/preparación y registrar en columna "CÓDIGO" del formulario.
- En el caso que el alimento no esté en la base de datos, elija otro similar teniendo en cuenta el tipo de alimento y su perfil nutricional. Ante cualquier, duda preguntar al responsable del monitoreo

Referencias:

1. Alimentarius C. Norma general para el etiquetado de los laimentos preenvasados. Codex Aliment. 2018;1(1):1–8.
2. Moreiras O, Carbajal A, Cabrera L, Pirámide CC-ME, 2019 undefined. Tabla de composición de alimentos. 19. a edición.
3. Raper N, Perloff B, Ingwersen L, Steinfeldt L, Anand J. An overview of USDA's Dietary Intake Data System. J Food Compos Anal. 2004;17(3–4):545–55.

ANNEX 3. SHORT FOOD FREQUENCY QUESTIONNAIRE

CUESTIONARIO DE FRECUENCIA DE CONSUMO ALIMENTARIO

Indicaciones

- **El cuestionario consta de dos partes: en la primera deberá responder preguntas de información general y en la segunda deberá completar 59 preguntas a cerca de la frecuencia de consumo de alimentos del mes anterior**
- **La presencia de un * indica que se trata de una pregunta con respuesta obligatoria**
- **El tiempo de respuesta de la encuesta va a ser evaluado**
- **En caso de tener dudas con la ración estándar indicada, puede consultar ver imagen (alimento señalado con una flecha roja) a modo de apoyo**
- **Se deben responder todos los grupos de alimentos**

* Datos personales

Fecha (dd/mm/aaaa)	
Código de identificación	
Sexo (H/M)	
Fecha de nacimiento (dd/mm/aaaa)	
Peso (kg)	
Talla (cm)	

* Número de cuestionario realizado

- ◯ 1º cuestionario
- ◯ 2º cuestionario

Si usted es de sexo femenino. ¿Esta menstruando actualmente?

- ◯ SI
- ◯ NO
- ◯ Menopausia

* ¿Cuál fue su vía de nacimiento?

- ◯ Parto Vaginal
- ◯ Cesárea
- ◯ No sabe

* ¿Usted fuma?

- ◯ SI
- ◯ NO
- ◯ En el pasado

* ¿Cuál es su grupo sanguíneo?

- ☐ O
- ☐ AB
- ☐ A
- ☐ B
- ☐ Rh(-)
- ☐ Rh (+)
- ☐ No se sabe

* ¿Usted sigue algún tipo de alimentación específica?

- ◯ NO
- ◯ SI (indique cuál)
- ◯ ¿Cuál? (especifique)

*

¿Usted consume alimentos listos para el consumo o precocinados comprados en tiendas como tortilla de patatas, pizza, lasañas, hamburguesa?

- ◯ SI
- ◯ NO

*

¿Usted consume alimentos con edulcorantes como sacarina, sucralosa, aspartamo, acesulfamo o stevia ?

- ◯ SI
- ◯ NO

*

¿Qué cantidad total de líquidos ingieres de forma aproximada a lo largo del día? Incluyendo agua, infusiones, café, leche, bebidas vegetales, zumos, refrescos, cerveza u otras bebidas alcohólicas.

2

Indique cuánto consume de cada alimento presentado a continuación. Luego, marque con una **"X"** en el recuadro la frecuencia de consumo de ese alimento durante el **MES ANTERIOR.** (en el caso de responder **erróneamente** la frecuencia de consumo, señale con una **"-"** y vuelva a responder con una **"X"**)

VERDURAS, LEGUMINOSAS Y PATATAS

¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
1/59. Verdura de hoja **cruda** como espinaca, lechuga, endibia, escarola, apio, penca, canónigo, hinojo, brote de soja,... (**ración estándar: 1/2 plato**) **ver imagen**						
2/59. Verdura de hoja **cocida** como espinaca, acelga, penca, hinojo, espárrago, brote de soja,...**(ración estándar: 1/2 plato)** **ver imagen**						
3/59. Tomate **(ración estándar: 1 unidad)**						
4/59. Cebolla, cebolleta o puerro **(ración estándar: 1/2 unidad)**						
5/59. Calabacín, berenjena o pepino **(ración estándar: 1/2 unidad)**						
6/59. Zanahoria, calabaza o remolacha **(ración estándar: 1/2 plato)** **ver imagen**						
7/59. Pimiento, pimiento Padrón **(ración estándar: 1/2 plato)**						
8/59. Crucífera como brócoli, coliflor, nabo, col o rúcula **(ración estándar: 1 plato)**						

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
9/59. Verduras en vinagre o escabeche como pepinillos, alcaparras, zanahorias, cebollínes, alcachofas **(ración estándar: 1/2 taza)**							
10/59. Maíz y leguminosas frescas como habas, guisantes **(ración estándar: 1/2 plato)** ver imagen							
11/59. Setas en general **(ración estándar: 1/2 plato)**							
12/59. Patata, boniato excepto patata frita tipo snack **(ración estándar: 1 unidad),** judía verde **(1/2 plato)**							

De las LEGUMBRES, mire la imagen adjunta y elija la opción correspondiente ver imagen

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
13/59. Lenteja, alubia (pinta, blanca o negra), garbanzo **cocido**							

FRUTAS Y FRUTOS SECOS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
14/59. Fruta fresca **(ración estándar: 1 unidad excepto cuando se indique lo contrario)** como naranja, pomelo, plátano, manzana, pera, nectarina, kiwi, mandarina **(2 unidades)**, fresa **(6 unidades)**, sandía o melón **(1 tajada)**, uva **(1 racimo)**, o jugo de fruta natural **(1 vaso)** **ver imagen**							
15/59. Fruta de alto contenido en grasa como aguacate, olivas, coco **(ración estándar: 1/4 plato)** **ver imagen**							
16/59. Fruta deshidrata como pasas, higos secos, arándanos deshidratados **(ración estándar: 1/3 bowl)** **ver imagen**							
17/59. Frutos secos y semillas como nueces, almendras, cacahuetes, avellanas, pistachos, piñones, pipas, semillas,... **(ración estándar: 1/3 bowl)** **ver imagen**							

5

CEREALES Y DERIVADOS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
18/59. Pan blanco como baguete, pan de payés, pan de molde, pan de leche,... **(ración estándar: 2 rebanadas o 1/3 de barra)** ver imagen							
19/59. Pan integral como pan baguete integral, pan molde integral, pan integral cereales varios,... **(ración estándar: 2 rebanadas o 1/3 de barra)** ver imagen							
20/59. Cereal para el desayuno como corn flakes, avena, muesli u otros **(ración estándar: 1/2 bowl)** ver imagen							
21/59. Galleta normal/integral o bizcocho **(ración estándar: 5 unidades o 1 ración)** ver imagen							

De los CEREALES, mire la imagen adjunta y elija la opción correspondiente ver imagen

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
22/59. Cereales y pastas como fideos, macarrones, espaguetis, arroz blanco, cuscús, bulgur y otros granos **cocidos**							
23/59. Cereales y pastas integrales o altas en fibra como fideos integrales, espaguetis integrales, arroz integral, arroz salvaje, quínoa y otros granos **cocidos**							

LÁCTEOS Y DERIVADOS

Recuerda, se está evaluando el consumo del **MES ANTERIOR**

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
24/59. Leche entera **(ración: 1 vaso o taza)** ver imagen							
25/59. Leche semi-desnatada **(ración: 1 vaso o taza)** ver imagen							
26/59. Leche desnatada **(ración: 1 vaso o taza)** ver imagen							
27/59. Bebida y producto de base vegetal como leche de almendras, de arroz, de avena o de soja **(ración estándar: 1 vaso o taza)** ver imagen							
28/59. Queso alto en grasa como queso curado, parmesano, manchego, roquefort, gruyere, gorgonzola, grana padano **(ración estándar: 1 ración)** ver imagen							
29/59. Queso bajo en grasa como la mozzarella, búfala, camembert, cheddar, queso de cabra, requesón **(ración estándar: 1 ración)** ver imagen							
30/59. Lácteo fermentado como yogur, yogur liquido, kéfir **(ración estándar: 1 envase o pote)**							

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
31/59. Postre lácteo como tiramisú, natillas, flan, helado de crema (2 bolas) **(ración estándar: 1 envase o pote)** **ver imagen**							

HUEVO, PRODUCTOS CÁRNICOS Y PESCADOS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
32/59. Huevo de gallina, pato, ganso, codorniz **(ración estándar: 1 unidad)**							
33/59. Carnes grasas como vacuno, ternera, cerdo, cordero, jabalí, ciervo, caballo **(ración estándar: 1 ración)** **ver imagen**							
34/59. Carnes magras como pollo, pavo, conejo, liebre, cabrito, otras aves **(ración estándar: 1 ración)** **ver imagen**							
35/59. Carnes procesadas como salchichón, chorizo, chistorra, morcilla, mortadela, salchicha, butifarra, sobrasada, tocino, bacón, panceta, jamón curado, jamón dulce,... **(ración estándar: 3 lonchas o 1 ración)** **ver imagen**							
36/59. Pescado azul o alto en grasa como anchoa, anguila, angula, atún, bonito, jurel, salmón, sardina,... **(ración estándar: 1 ración)** **ver imagen**							

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
37/59. Pescado blanco o bajo en grasa como bacalao, merluza, lenguado, rape, pescadilla, rodaballo,... **(ración estándar: 1 ración)** ver imagen							
38/59. Pescado enlatado en aceite como atún, jurel, caballa, bonito,... **(ración estándar: 1 lata)**							

De los MOLUSCOS, mire la imagen adjunta y elija la opción correspondiente ver imagen

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
39/59. Molusco y crustáceo como mejillón, almeja, calamar, pulpo, sepia, gamba, langostino							

ACEITES Y GRASAS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
40/59. Aceite de oliva **(ración estándar: 1 cucharada sopera)** **ver imagen**							
41/59. Aceite de girasol **(ración estándar: 1 cucharada sopera)**							
42/59. Otro aceite como de maíz, colza, pepita de uva,... **(ración estándar: 1 cucharada sopera)**							
43/59. Mantequilla y margarina **(ración estándar: 1 cucharada sopera)** **ver imagen**							

BOLLERÍA Y PASTELERÍA

¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
44/59. Bollería como donuts, croissants, palmeras, churros, buñuelos, tartas, pasteles **(ración estándar: 1 unidad de 50g o 2 unidades pequeñas de 25g cada una)** **ver imagen**						
45/59. Chocolate negro (> 50% de cacao) o cacao en polvo **(ración estándar: 2 cucharaditas o 1 onza)** **ver imagen**						
46/59. Confitería como barra de caramelo, gominola, caramelo, goma de mascar, turrón, mazapán **(ración estándar: 1/3 bowl)** **ver imagen**						

SALSAS

¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
47/59. Salsa de tomate envasada, tomate enlatado **(ración estándar: 1 cucharada sopera)**						
48/59. Otras salsas como mayonesa, kétchup, mostaza, pesto, alioli, balsámico **(ración estándar: 1 cucharada sopera)** **ver imagen**						

BEBIDAS NO ALCOHÓLICAS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
49/59. Té (con y sin cafeína) o infusión **(ración estándar: 1 taza 100mL)** **ver imagen**							
50/59. Café (con y sin cafeína) de todo tipo **(ración estándar: 1 taza 50mL)** **ver imagen**							
51/59. Refrescos como bebidas gaseosas (coca cola o pepsi cola), refrescos light, isotónicas o saborizadas **(ración estándar: 1 vaso)**							
52/59. Zumos o néctar de fruta envasado **(ración estándar: 1 vaso)** **ver imagen**							

BEBIDAS ALCOHÓLICAS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
53/59. Vino o cava (rosado, tinto, mosto, blanco, moscatel) **(ración estándar: 1 copa de vino 100mL)** **ver imagen**							
54/59. Cerveza **(ración estándar: 1 lata o vaso 330mL)** **ver imagen**							
55/59. Whisky, vodka, ginebra, coñac, ron,... **(ración estándar: 1 vaso destilado 50mL)** **ver imagen**							

OTROS

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
56/59. Azúcar agregada, miel, mermelada, membrillo **(ración estándar: 1 cucharadita)** **ver imagen**							
57/59. Snack como patata frita, nacho, rosquilleta, pico, colín, cracker **(ración estándar: 1 bolsa 50gr o plato de postre)** **ver imagen**							

De la COMIDA PROCESADA, mire la imagen adjunta y elija la opción correspondiente ver imagen

	¿Cuánto come de la ración **estándar** indicada?	No consumo	1 a 3 veces al mes	1 o 2 veces por la semana	+ de 3 veces a la semana	1 vez al día	+ de 2 veces al día
58/59. Comida procesada como pizza, lasaña, canelones, nuggets, tortilla de patata							

59/59. Si durante el mes anterior consumiste vitaminas y/o minerales (incluyendo calcio) o productos dietéticos especiales (salvado, leche con ácidos grasos omega-3, flavonoides, etc.) o medicamentos, por favor indica la marca y frecuencia con que los tomaste:

ANNEX 4. STANDARD OPERATING PROCEDURE TO sFFQ NUTRITIONAL COMPOSITION QUANTIFICATION

Vall d'Hebron Institut de Recerca VHIR	**Protocolo de cuantificación del sFFQ**	Code : Date: 21/04/2020 Pages: 2
Date: 21/04/2020	**MICROBIOME LAB**	Prepared by: Msc. Francisca Yáñez Reviewed by: Dra. Chaysavanh Manichanh Date:

1.- OBJETIVO
Establecer un procedimiento que estandarice el ingreso de respuestas del FFQ a formulario de datos externos del estudio de dieta y microbiota.

2.- ALCANCE
Aplicable a todos las sFFQ completados por los participantes del estudio de dieta y microbiota.

3.- DOCUMENTOS ASOCIADOS
Tabla de composición nutricional de FFQ

4.- RESPONSABILIDAD
Responsable ejecución: Personal previamente capacitado en ingreso de respuestas del FFQ al formulario de datos externos.
Responsable monitoreo: Dietista/Nutricionista a cargo del estudio.

5.- DEFINICIONES
Frecuencia de consumo: cantidad de alimento consumido al día, mes o año
Ración: cantidad estándar de alimento que se aconseja consumir en las guías alimentarias tratando de que aporte una cantidad similar de los nutrientes principales

6.- INSUMOS
Ordenador con Excel, Formulario de datos externos asociado a tabla de composición nutricional FFQ. Respuestas del FFQ

7.- PROCEDIMIENTO
- Abrir pdf de FFQ a cuantificar
- Abrir archivo Excel **"TCA.FFQ_FranY_v10.20**
- Clickear pestaña ***"Comp_Nut_Transp_FFQ"***

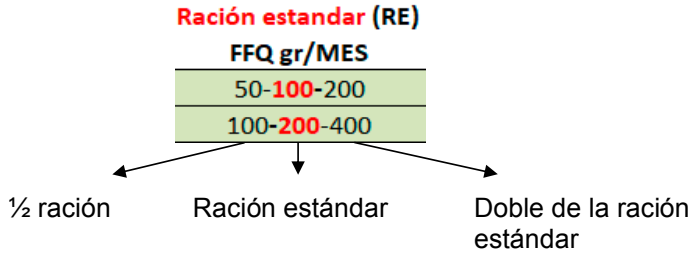
2	Lista de Alimentos	Ración estandar (RE) FFQ gr/MES	¿Cuanto come de la ración estandar?	Código frecuencia/mes
3	1.Verdura de hoja cruda co	50-100-200	100	2
4	2. Verdura de hoja cocida c	100-200-400	200	1
5	3. Tomate	75-150-300	150	3

-

- En columna "***Código frecuencia/mes"*** ingresar el código de respuestas de los 48 ítems del FFQ (desde la casilla B3 a B60), según si participante selecciono una de las siguientes frecuencias de consumo:

Frecuencia de consumo	Código de respuesta
No consumo	0
1 a 3 veces al MES	1
1 a 2 veces a la SEMANA	2
+ de 3 veces a la SEMANA	3
1 vez al DÍA	4
+ de 2 veces al DÍA	5

- En columna ***"Cuanto come de la ración estándar***" ingresar el valor de la cantidad consumida de la ración estándar (RE) para cada ítem (señalada en tabla Excel con el nombre "***Ración estándar (RE) FFQ gr/MES"***), en donde;



- Para los ítems 13, 22, 23, 39 y 58 ingresar el valor seleccionado por el participante, según la siguiente tabla

FFQ	Alternativa A	Alternativa B	Alternativa C
Item 13	75 g	150 g	225 g
Item 22 y 23	40 g	80.25 g	120.5g
Item 39	75 g	150 g	225 g
Item 58	100 g	200 g	300 g

- Una vez ingresados todos los valores del sFFQ, clickear en pestaña "***Cuantificación_FFQ***".

- En la fila 63, vera la composición energética y nutrientes de todo el sFFQ

- Para finalizar, copiar toda la tabla de la pestaña ***"Cuantificación_FFQ"*** a archivo Excel **"FFQv2_cuantificación**", en la pestaña correspondiente al código del participante del sFFQ cuantificado.

ANNEX 5. TABLE OF REPRODUCIBILITY PILOT sFFQ. CORRELATION BETWEEN FOOD GROUPS, ENERGY AND NUTRIENTS INTAKE BETWEEN sFFQ1p AND sFFQ2p

	Wilcoxon signed-rank test	Spearman's rank coefficient	
Food groups (g/d)	p value	Un-adjusted	Energy-adjusted
Alcoholic beverage	0.73	0.93	-
Cereals or grains	0.41	0.69	-
Fish and seafood	0.77	0.87	-
Fruits	0.54	0.66	-
Legumes	0.96	0.66	-
Meat and meat products	0.41	0.84	-
Milk and dairy products	0.55	0.56	-
Non alcoholic beverage	0.93	0.87	-
Nuts	0.23	0.80	-
Oils and fat	0.93	0.67	-
Pastries and cake	0.60	0.85	-
Sauces and condiments	0.95	0.78	-
Sausage	0.41	0.92	-
Sugar and sweet	0.36	0.67	-
Vegetables	0.60	0.61	-
Energy and nutrient			
Energy (kcal)	0.25	0.63	-
Water (g)	0.25	0.62	0.69
Total Protein (g)	0.25	0.63	0.83
Protein vegetal (g)	0.30	0.65	0.79
Protein animal (g)	0.25	0.91	0.95
Total Fat (g)	0.25	0.72	0.72
SFA (g)	0.25	0.76	0.80
MUFA (g)	0.25	0.75	0.80
PUFA (g)	0.43	0.79	0.90
Cholesterol (mg)	0.30	0.87	0.89

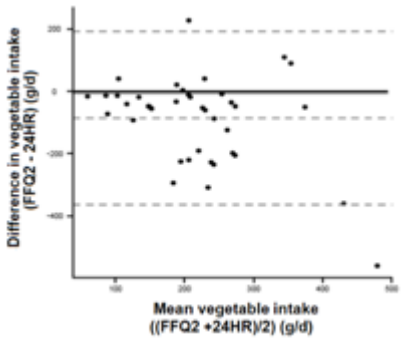
	Intraclass Correlation Coeficient (ICC)		Concordance Correlation Coefficient (CCC)	
	Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
	0.78	-	0.77	-
	0.67	-	0.65	-
	0.89	-	0.89	-
	0.68	-	0.67	-
	0.73	-	0.73	-
	0.98	-	0.98	-
	0.63	-	0.62	-
	0.90	-	0.89	-
	0.77	-	0.75	-
	0.75	-	0.74	-
	0.95	-	0.95	-
	0.53	-	0.53	-
	0.94	-	0.93	-
	0.68	-	0.66	-
	0.64	-	0.63	-
	0.63	-	0.60	-
	0.62	0.67	0.58	0.63
	0.63	0.82	0.60	0.74
	0.65	0.78	0.64	0.75
	0.91	0.95	0.90	0.94
	0.71	0.70	0.70	0.60
	0.75	0.79	0.72	0.75
	0.75	0.80	0.74	0.75
	0.79	0.90	0.79	0.86
	0.86	0.89	0.85	0.88

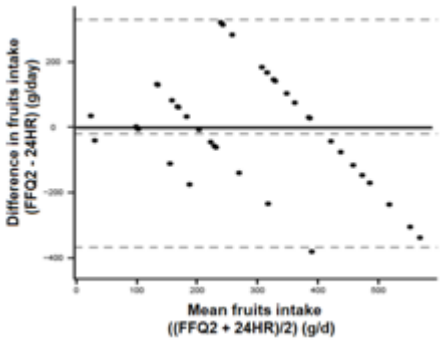
Food groups (g/d)	Wilcoxon signed-rank test p value	Spearman's rank coefficient Un-adjusted	Spearman's rank coefficient Energy-adjusted
Energy and nutrient			
Total Carbohydrates (g)	0.26	0.60	0.63
Sugar digeribles (g)	0.25	0.62	0.59
Starch digeribles (g)	0.38	0.56	0.61
Fibre (g)	0.38	0.75	0.83
Ethanol (g)	0.72	0.79	0.80
Sodium (mg)	0.25	0.70	0.82
Potassium (mg)	0.40	0.74	0.85
Calcium (mg)	0.28	0.72	0.83
Magnesium (mg)	0.38	0.68	0.82
Phosphorus (mg)	0.26	0.57	0.79
Iron (mg)	0.30	0.63	0.77
Zinc (mg)	0.25	0.72	0.83
Vitamin A (mcg.e.r)	0.25	0.73	0.68
Total retinoids (mcg)	0.25	0.77	0.82
Total carotenoids mcg)	0.30	0.68	0.66
Vitamin D (mcg)	0.40	0.81	0.81
Vitamin E (mg.e.t)	0.59	0.77	0.85
Vitamin B1 (mg)	0.38	0.51	0.66
Vitamin B2 (mg)	0.30	0.58	0.82
Niacin (mg)	0.36	0.71	0.77
Vitamin B6 (mg)	0.25	0.65	0.71
Folic (mcg)	0.26	0.68	0.84
Vitamin B12 (mcg)	0.30	0.93	0.95
Vitamin C (mg)	0.25	0.70	0.69

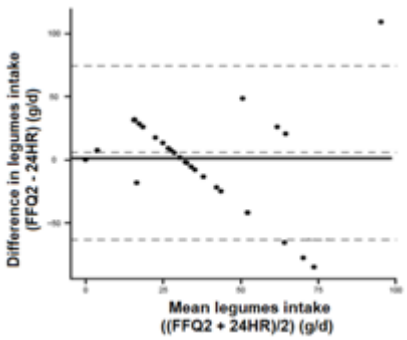
$p<0.05$ between FFQ2 and mean 3-24HR using paired t-test and Wilcoxon test.

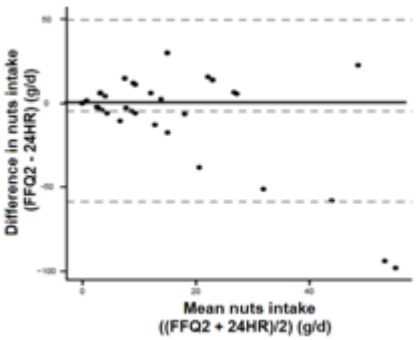
	Intraclass Correlation Coeficient (ICC)		Concordance Correlation Coefficient (CCC)	
	Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
	0.59	0.62	0.56	0.54
	0.61	0.57	0.59	0.55
	0.55	0.60	0.54	0.58
	0.76	0.83	0.75	0.81
	0.80	0.81	0.79	0.80
	0.70	0.82	0.68	0.79
	0.74	0.84	0.74	0.81
	0.71	0.83	0.69	0.82
	0.68	0.82	0.67	0.78
	0.57	0.79	0.55	0.70
	0.63	0.77	0.62	0.70
	0.71	0.83	0.69	0.80
	0.72	0.68	0.70	0.65
	0.77	0.80	0.75	0.79
	0.68	0.66	0.67	0.64
	0.78	0.79	0.78	0.78
	0.76	0.85	0.76	0.80
	0.51	0.66	0.50	0.58
	0.58	0.82	0.56	0.75
	0.71	0.76	0.70	0.74
	0.64	0.70	0.62	0.61
	0.68	0.84	0.66	0.79
	0.93	0.95	0.93	0.95
	0.69	0.67	0.66	0.66

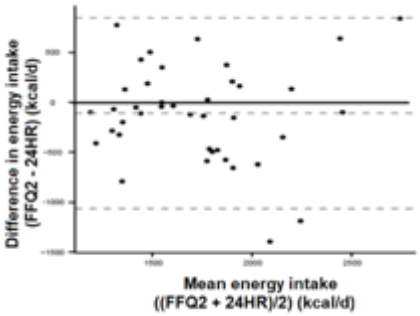
ANNEX 6. DIFFERENCES IN DAILY OF FOOD GROUPS, ENERGY AND NUTRIENTS INTAKE BETWEEN PILOT SFFQ2 AND MEAN 3-24HRS BASED ON BLAND & ALTMAN PLOTS

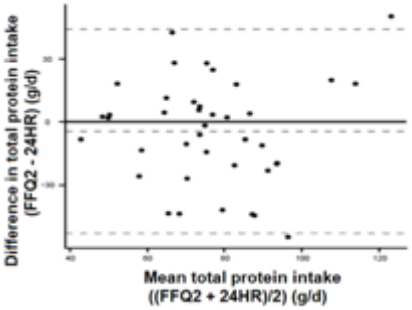


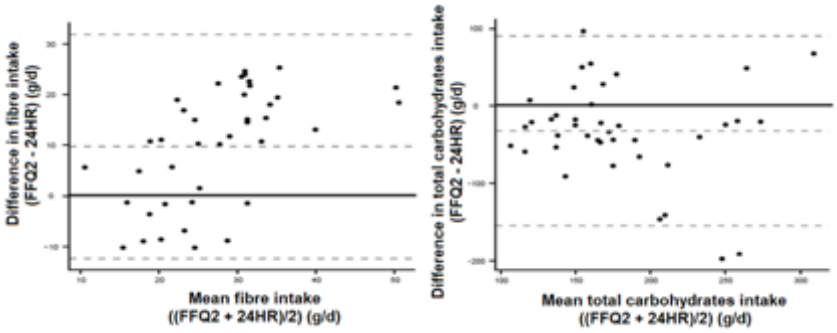






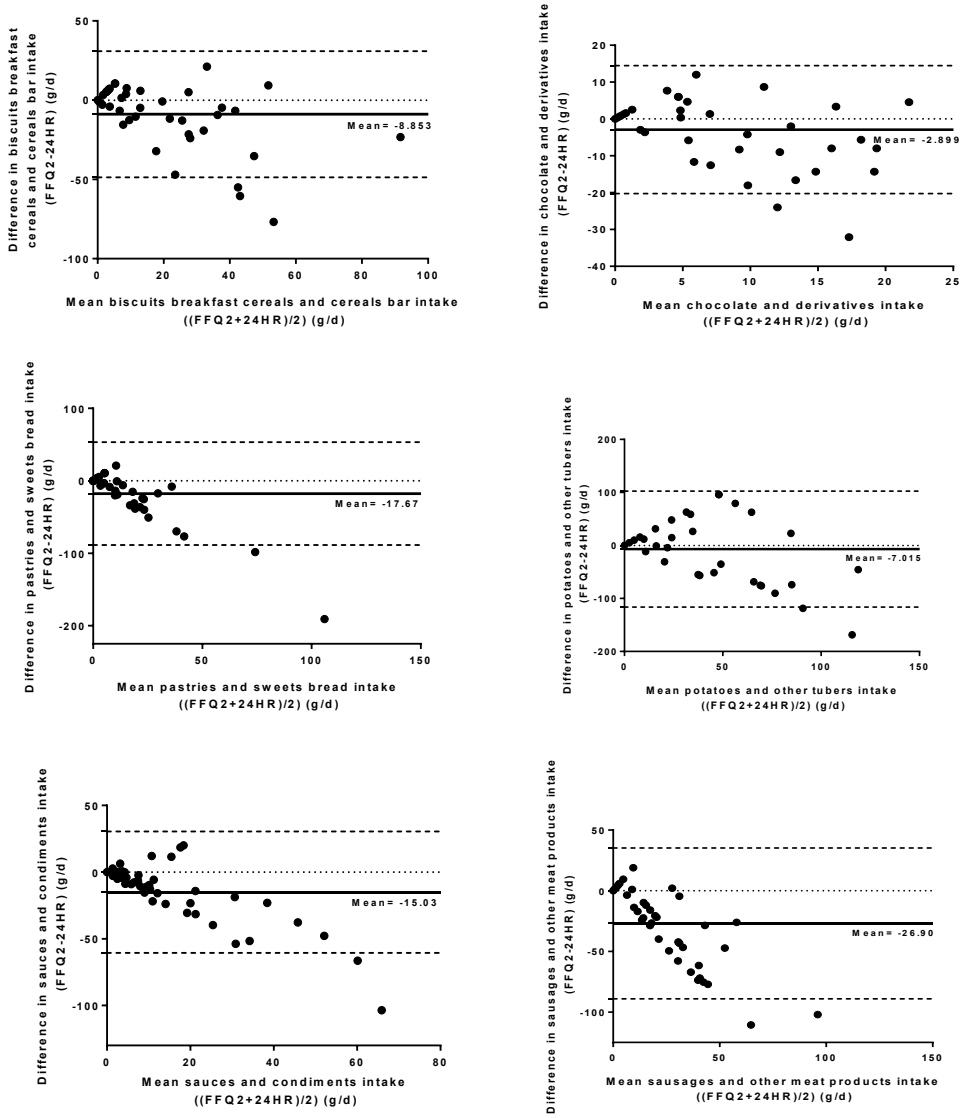






Annex 6, figure 1: Bland & Altman plots show the relationship between the difference in daily intake and the mean daily intake estimated by both methods. Plain lines represent the mean difference, and the dashed line represents lower and upper 95% limits of agreement

ANNEX 7. DIFFERENCES IN DAILY INTAKE BETWEEN FINAL sFFQ2 AND MEAN 3-24HRS BASED ON BLAND & ALTMAN PLOTS



Annex 7, figure 2: Bland & Altman plots show the relations between the difference in daily intake of biscuits, breakfast cereals and cereals bars, chocolates and derivatives, pastries and sweet bread, potatoes and other tuber, sauces and condiments, sausages and other meat products, sugar and other sweets, vegetables and vegetable products, white bread and wholegrain bread estimated by FFQ2 and mean 3-24HR, and the corresponding mean daily intake calculated by the two methods. Plain lines represent the mean difference, and the dashed line represents lower and upper 95% limits of agreement

ANNEX 8. TABLE OF REPRODUCIBILITY OF FINAL sFFQ. CORRELATION BETWEEN FOOD GROUPS, ENERGY AND NUTRIENTS INTAKE BETWEEN sFFQ1 AND sFFQ2

Food group (g/d)	Wilcoxon signed-rank test p value
Alcoholic beverage	0.18
Appetizers	0.18
Biscuits breakfast cereals and cereals bars	0.60
Chocolates and derivatives	0.35
Fats and oils	0.97
Fish and shellfish	0.39
Fruit and fruit products	0.35
Legumes	0.15
Meats and eggs	0.81
Milk and dairy products except fermented milk	0.98
Non Alcoholic beverage	0.71
Nuts and seeds	0.97
Pastries and sweets breads	0.91
Potatoes and other tubers	0.28
Ready to eat meals	0.70
Sauces and condiments	0.05
Sausages and other meat products	0.38
Sugar and other sweets	0.57
Vegetables and vegetable products	0.31
White bread	0.56
White grains and white pastas	0.19
Wholegrain or wholemeal bread	0.65
Wholemeal grains and wholemeal pastas	0.39
Yogurt and fermented milk	0.66
Energy and nutrients	
Energy (kcal/d)	0.48
Total fat (g/d)	0.74

Spearman's rank coefficient		Intraclass Correlation Coeficient (ICC)	
Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
0.79	-	0.83	-
0.33	-	0.37	-
0.69	-	0.62	-
0.61	-	0.45	-
0.71	-	0.65	-
0.73	-	0.45	-
0.62	-	0.26	-
0.58	-	0.57	-
0.82	-	0.68	-
0.85	-	0.72	-
0.89	-	0.88	-
0.72	-	0.68	-
0.61	-	0.64	-
0.46	-	0.21	-
0.78	-	0.71	-
0.71	-	0.56	-
0.45	-	0.45	-
0.76	-	0.90	-
0.85	-	0.85	-
0.84	-	0.55	-
0.66	-	0.67	-
0.64	-	0.56	-
0.90	-	0.39	-
0.89	-	0.89	-
0.78	NA	0.64	NA
0.80	0.74	0.78	0.78

Food group (g/d)	Wilcoxon signed-rank test p value
Energy and nutrients	
Total protein (g/d)	0.45
Total water (g/d)	0.81
Total fiber (g/d)	0.08
Total carbohydrates (g/d)	0.38
Alcohol (g/d)	0.50
MUFA (g/d)	0.57
PUFA (g/d)	0.34
SFA (g/d)	0.54
Cholesterol (mg/d)	0.40
Vitamin A µg retinol (eq/d)	0.27
Vitamin D (µg/d)	0.52
Vitamin E (mg α tocoferol/d)	0.05
Folate total (µg/d)	0.08
Total niacin equivalent (mg/d)	0.84
Riboflavin (mg/d)	0.59
Tiamin (mg/d)	0.17
Vitamin B12 (µg/d)	0.66
Vitamin B6 (mg/d)	0.42
Vitamin C ascorbic acid (mg/d)	0.54
Calcium (mg/d)	0.52
Iron (mg/d)	0.12
Potassium (mg/d)	0.15
Magnesium (mg/d)	0.13
Sodium (mg/d)	0.65
Phosphorus (mg/d)	0.39
Iodine (µg/d)	0.23
Selenium (µg/d)	0.70
Zinc (mg/d)	0.26

ρ<0.05 between FFQ2 and mean 24HR using paired t-test and Wilcoxon test.

Spearman's rank coefficient		Intraclass Correlation Coeficient (ICC)	
Un-adjusted	Energy-adjusted	Un-adjusted	Energy-adjusted
0.80	0.66	0.77	0.73
0.81	0.83	0.73	0.82
0.70	0.62	0.65	0.65
0.67	0.70	0.56	0.81
0.77	0.83	0.76	0.74
0.76	0.70	0.71	0.67
0.78	0.86	0.68	0.73
0.82	0.68	0.82	0.81
0.85	0.74	0.77	0.77
0.81	0.72	0.80	0.82
0.65	0.57	0.71	0.80
0.79	0.78	0.80	0.90
0.68	0.54	0.72	0.74
0.77	0.81	0.68	0.71
0.86	0.79	0.81	0.79
0.74	0.55	0.68	0.69
0.83	0.64	1.00	0.96
0.78	0.70	0.59	0.73
0.69	0.72	0.55	0.67
0.78	0.69	0.76	0.75
0.77	0.59	0.74	0.72
0.70	0.67	0.57	0.73
0.87	0.80	0.83	0.86
0.77	0.63	0.78	0.67
0.78	0.63	0.77	0.72
0.79	0.75	0.62	0.73
0.81	0.68	0.47	0.61
0.78	0.66	0.72	0.58

ANNEX 9. PUBLICATION RELATED TO THE THESIS:

YÁÑEZ F, SOLER Z, OLIERO M, XIE Z, OYARZUN I, SERRANO-GÓMEZ G, *ET AL.* INTEGRATING DIETARY DATA INTO MICROBIOME STUDIES: A STEP FORWARD FOR NUTRI-METAOMICS. NUTRIENTS. 2021;13(9).

nutrients

MDPI

Article

Integrating Dietary Data into Microbiome Studies: A Step Forward for Nutri-Metaomics

Francisca Yáñez [1,2], Zaida Soler [1,2], Manon Oliero [1,†], Zixuan Xie [1,2], Iñigo Oyarzun [1], Gerard Serrano-Gómez [1] and Chaysavanh Manichanh [1,2,3,*]

[1] Gut Microbiome Group, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain; panchayanez@gmail.com (F.Y.); zaida.soler@vhir.org (Z.S.); manon.oliero@umontreal.ca (M.O.); xie.zixuan@vhir.org (Z.X.); inigooyarzunlafuente@gmail.com (I.O.); gerard.s.g@hotmail.com (G.S.-G.)
[2] Departament de Medicina, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain
[3] Gut Microbiome Group, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), 28029 Madrid, Spain
* Correspondence: cmanicha@gmail.com; Tel.: +34-934894036
† Current address: Department of Microbiology and Immunology, Faculty of Medicine, University of Montréal, Montréal, QC H3T 1J4, Canada.

Abstract: Diet is recognised as the main driver of changes in gut microbiota. However, linking habitual dietary intake to microbiome composition and activity remains a challenge, leaving most microbiome studies with little or no dietary information. To fill this knowledge gap, we conducted two consecutive studies (n = 84: a first pilot study (n = 40) to build a web-based, semi-quantitative simplified FFQ (sFFQ) based on three 24-h dietary recalls (24HRs); a second study (n = 44) served to validate the newly developed sFFQ using three 24HRs as reference method and to relate gut microbiome profiling (16S rRNA gene) with the extracted dietary and lifestyle data. Relative validation analysis provided acceptable classification and agreement for 13 out of 24 (54%) food groups and 20 out of 29 nutrients (69%) based on intraclass correlation coefficient, cross-classification, Spearman's correlation, Wilcoxon test, and Bland–Altman. Microbiome analysis showed that higher diversity was positively associated with age, vaginal birth, and intake of fruit. In contrast, microbial diversity was negatively associated with BMI, processed meats, ready-to-eat meals, sodium, and saturated fat. Our analysis also revealed a correlation between food groups or nutrients and microbial composition. Overall, we provide the first dietary assessment tool to be validated and correlated with microbiome data for population studies.

Keywords: sFFQ development; relative validation; diet–microbiome relationship

check for updates

Citation: Yáñez, F.; Soler, Z.; Oliero, M.; Xie, Z.; Oyarzun, I.; Serrano-Gómez, G.; Manichanh, C. Integrating Dietary Data into Microbiome Studies: A Step Forward for Nutri-Metaomics. *Nutrients* **2021**, *13*, 2978. https://doi.org/10.3390/nu13092978

Academic Editors: Elad Tako and Alessio Fasano

Received: 12 July 2021
Accepted: 25 August 2021
Published: 27 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

For about 200,000 years, humans followed a hunter-gatherer lifestyle, using fire for cooking and eating wild game, fruits, vegetables, and nuts, with lipid, protein, and carbohydrate content each accounting for 33% of dietary intake [1,2]. Diet is recognized as the key driver of changes in the adult gut microbiota. Indeed, humans have co-evolved with their microbiota along with the development of agriculture [3,4]. The dietary profile of the modern human has changed rapidly over the last 100 years—much faster than normal evolutionary adaptation—such that it has undoubtedly had an impact on shaping our gut microbiota and, consequently, our health. Comparison of rural versus urban populations provides an interesting approach to understand the changes of the microbiome in the context of modern life. The faecal microbiota of individuals from two very different geographical locations with contrasting dietary habits, namely inhabitants of rural village in Burkina Faso and European children (EU), differ significantly with regard to the relative abundance of bacteria known to be involved in cellulose and xylan hydrolysis [5]. The urbanization of regions in China followed by geography, dietary habit, and ethnicity

was shown to have an impact on the variation of the gut fungal microbiome, increasing *Saccharomyces cerevisiae* and depleting *Candida dubliniensis* [6].

Compelling evidence supports an association between changes in the microbiota community and human metabolic disorders, including obesity [7] and type 2 diabetes [8]. Moreover, many intervention studies have shown that diet alters gut microbiota. A shift from plant-based to animal-based diets has been linked to an increase in the abundance of bile-tolerant microorganisms and a decrease in Firmicutes, which metabolize dietary plant polysaccharides [9]. A relatively long-term diet intervention (12 months) showed that the Mediterranean diet improved cognitive function, frailty, and inflammation status in elderly individuals by modifying gut microbiota [10].

To the best of our knowledge, very few population studies have demonstrated an association between habitual diet, gut microbiota, and health status. Most studies related to microbiome profiling in the context of a specific disease or related to dietary intervention did not collect dietary data, while those assessing diet intake did not perform microbiome analysis [11,12]. We believe that the scarcity of studies relating diet, microbiome, and disease is not due to the lack of molecular or bioinformatics tools or even cost but to the unavailability of an appropriate dietary assessment tool that has been tested against microbiome data.

Along with 24-h dietary recalls (24HRs) and food records or diaries, food-frequency questionnaires (FFQ) are one of the most widely used dietary assessment tools. FFQs are usually designed in function of the purpose of the study and can contain from a few questions intended to seek the effect of a specific food or group of food [13,14] to a comprehensive list of between 5 and 350 items to capture habitual diet [15]. FFQs are considered one of the most suitable instruments for epidemiological studies compared to 24HRs or food records, as they are self-administered and therefore do not require the presence of a trained interviewer or considerable time dedication on the part of respondents. Moreover, they are cost-effective. However, FFQs, being usually very long, may lead to the misreporting of habitual dietary intake, and they have sometimes been reported to be unreliable. Additionally, current FFQs have been under-evaluated against microbiome data in epidemiological studies, as most microbiome studies contain few or no dietary data.

Here, we sought to (1) design a new semi-quantitative and simplified FFQ (sFFQ); (2) undertake a relative validation analysis; (3) perform a reproducibility analysis; and (4) correlate dietary intake with microbiome data.

2. Materials and Methods

2.1. Study Population

A total of 84 healthy volunteers (40 participants in the pilot study and 44 in the validation study) were recruited between May 2017 and August 2020 by disseminating an announcement. The study was conducted in accordance with Declaration of Helsinki, and the protocol was approved by the local Ethics Committee of the Vall d'Hebron University Hospital, Barcelona (Project identification code: PR(AG)156/2017). All participants signed a consent form.

Power calculation showed that a minimum of 40 subjects would be needed to give 85% power to detect correlation between sFFQ and mean three 24HRs of 0.45 as significant at the 5% level in order to take into account the small sample size. Exclusion criteria included age under 18 and over 65 years, antibiotic use during the three months prior to entering the study, use of proton pump inhibitor medication, and any disorders that may be associated with altered gut microbiota, such as diabetes, chronic digestive pathology, inflammatory bowel disease, and autoimmune disease.

2.2. 24-h Dietary Recall: The Pilot Study

We conducted a pilot study using three 24HRs on 40 healthy subjects to evaluate their dietary habits and select the food items to be added to new sFFQ. A dietitian and trained staff performed the three interviews, two of them during weekdays and one during the

weekend. These interviews were used to collect data on food consumption on the previous day, from the first intake in the morning to the last meal and beverage consumed during the night. To avoid biases in the 24HR response, we conducted the interviews randomly in time, taking into account the participants' availability. We assigned an alphanumeric code to each participant to maintain anonymity and registered the food in a 24HR collection form.

To objectively evaluate the serving size of each food and beverage, we used photographic albums: "Guide for dietary studies" from the Granada University [16] and the SU.VI.MAX. Portions Alimentaires [17].

To estimate the serving weight provided by the participant, we created a "Standardized household measures table" based on the "Food composition table" published by Moreiras et al. from the Complutense University of Madrid [18] and based on homemade food used for measuring purposes. For the latter, lab staff weighed their food, for which no information from any existing table or from the food industry was available, to recover an average weight for each one. To reduce possible bias introduced by the interviewers, we applied a standard operating procedure for data collection based on the five-step interview proposed by the USDA [19]. The mixed dishes were broken down into simple ingredients; for this, the participants were asked to describe the recipes and cooking procedures in detail.

To facilitate the search and reduce the encoding error for energy and nutrient quantification of food extracted from the 24HR, we created an inhouse food-composition database. For this purpose, we combined several food-composition databases, including the Spanish database (AESAN/BEDCA v 1.0, 2010), which includes a list of 950 foods and 31 nutrients (http://www.bedca.net; accessed on 10 March 2021); Moreiras's table, which includes a list of 900 food items and nutrients [18]; and the USDA National Nutrient Database for Standard Reference [20], which has a list of 8618 foods and 150 nutrients. Our food-composition database currently contains 1104 foods and preparations grouped into 13 food categories and information for 29 nutrients plus energy per 100 g of food.

2.3. Design and Development of the sFFQ

To develop the sFFQ, we first applied a general concept following the recommendations from previous studies [15,21–24]. To select the food items to be included in the sFFQ, we then combined and cross-checked the data collected from the three 24HRs of our pilot study with the food-consumption data reported in the "National Food Survey on adults, the elderly, and pregnant women (ENALIA2)" [25] related to the foods most consumed by the Spanish population. In total, 310 foods were selected on the basis of their higher intake within the population and higher intra- and inter-individual variability of consumption. We added questions that could pinpoint relevant factors with a potential effect on microbiome composition changes, such as blood type [26], mode of delivery at birth [27], consumption of ready-to-eat meal [28], and whether or not the participant was following a specific diet [29] or was excluding a specific food or type of food. We also included other factors potentially associated with changes in the gut microbiota, such as age, BMI, smoking, use of sweeteners, and number of fruits and vegetables consumed [5,9,30–37].

Our resulting sFFQ contained 58 food items (Supplementary Table S1: List of 58 food items as specified in the sFFQ) in which the consumption frequency of the previous month was categorized into six possible responses for each item: "Never", "1 or 3 times per month", "1 or 2 times per week", "3 or more times per week", "once per day", and "2 or more times per day". We estimated the food serving size of each item of the sFFQ based on the results of various surveys and guidelines, including the ENALIA2 Survey [25], the guidelines of the Spanish Society of Community Nutrition [38], and the guidelines of the scientific committee "5 a day" [39]. We also used the serving size assigned by the food industry and the serving obtained from our own pilot study. To further improve estimation of the amounts of food consumed in the sFFQ, we created a support document based on food photographs and added three consumption alternatives for the standard serving size: "1/2 of the standard portion size", "standard portion size" and "double the standard

portion size" [40–42]. To estimate the energy and nutrient intake from the sFFQ, we used our food-composition database.

To recover the nutritional composition of the food items included in the sFFQ, we calculated the weighted mean of nutrients and energy based on the data obtained from the 24HR pilot study. First, the foods collected from this study were classified into 58 items, as presented in the sFFQ. We then recovered the proportions contributed by each food to each of the 58 items from the 24HR pilot study and used them as weighted factors to calculate the energy and nutrient intake for each item in the sFFQ.

2.4. sFFQ Administration

The sFFQ was used as a web-based survey using the SurveyMonkey Inc. (San Mateo, CA, USA) platform two times one month apart (sFFQ1 and sFFQ2). On the day of the first 24HR interview, we provided the participants with the web link or QR code to complete the sFFQ. Once we had obtained the responses, we verified the missing data. We contacted participants if there was a lack of response to any of the items in the sFFQ. A new version of the sFFQ using an independent online survey from our own server is currently being prepared.

2.5. Analysis of the sFFQ Responses

To compare the results of the sFFQ with the reference 24HR method, we transformed the monthly and weekly consumption data into daily consumption frequencies. To this end, we calculated the g/day as follows: a consumption response of 1 to 2 times per week was understood as an average consumption of 1.5 times per week, which, divided by the seven days of the week, gives an average daily consumption of 0.21. This consumption was then multiplied by the weight associated with the selected serving size (for example, for the legumes item with a serving size of 150 g and consumption frequency mentioned above, the final value of grams per day would be 0.21 × 150 g = 31.5 g/d). Using this g/day information, we then calculated the energy and nutritional value of each item in the sFFQ. Foods and beverages from the sFFQs and the 24HRs were then classified into 24 food groups, total energy, and 29 nutrients (Supplementary Table S2: Classification into food groups, energy and nutrients of the sFFQs and the mean 3–24HRs).

2.6. Identification of Unreliable sFFQs and 24HRs

We verified each value obtained from the quantification of the sFFQs and the 24HRs. When a possible outlier was detected, we examined each data entry involved in reaching this value. For instance, we excluded participants with calorie intake values in the sFFQ and means of 24HRs lower than 800 kcal/day or higher than 4200 kcal/day for men and less than 600 kcal/day or more than 3500 kcal/day for women [43,44].

2.7. Statistical Analysis to Evaluate the Validity and Reproducibility of the sFFQ

The median and 25–75 percentile of food, energy, and nutrient consumption were calculated from the mean of the three 24HRs and both sFFQs. Nutrients were adjusted by energy using the density method [45] to control the confounding effect of calories. The validity (sFFQ2 versus mean of the three 24HRs) and reproducibility (sFFQ1 versus sFFQ2) of the newly developed sFFQ were evaluated using a series of statistical tests.

To control for inter-and intra-individual variation, we calculated the intraclass correlation coefficient (ICC) [46,47]. We used cross-classification (CC) to categorize individuals into equal third or opposite third for food group and energy-adjusted nutrient intake extracted from both methods [48]. We used Spearman's correlation coefficient to estimate the strength and direction of the association [22].

We applied the Wilcoxon signed-rank test to assess the differences in food, energy, and nutrient consumption and used the Bland–Altman analysis to check the degree of agreement between the two sFFQs and the three 24HRs. The differences between the two methods (FFQ2-24HR) were plotted against the mean intake of the measures ((FFQ2 + 24HR)/2)

and the limits of the agreement, defined as the mean ± 1.96 SD of the mean between the two methods, were evaluated [49]. In addition, to illustrate the magnitude of the possible systematic difference, we calculated the 95% CI of the mean differences. To reflect the presence of proportional bias, Spearman's correlation was calculated between the mean and the mean difference of the two methods [50]. Statistical analyses were performed in GraphPad Prism (v8) and the RStudio (Version 1.4.1106) package.

2.8. Microbiome and Statistical Analyses

Each of the 84 participants provided a faecal sample at baseline (M0), i.e., 24 to 48 h after the first 24HR, and one month after (M1), i.e., first stool after the third 24HR. Genomic DNA was extracted from 166 samples (two subjects did not provide a second sample), as previously described [51] and following the recommendations of the International Human Microbiome Standards (IHMS, http://www.human-microbiome.org/; accessed on 23 April 2021). The V4 hypervariable region of the 16S rRNA gene was PCR-amplified and sequenced using the MiSeq Illumina platform [52]. Sequence data were analysed using the QIIME 2^{TM}, which is a bioinformatics platform that stands for Quantitative Insights into Microbial Ecology. The sequences were demultiplexed to attribute sequence reads to the appropriate samples and were then denoised and dereplicated into amplicon sequence variants (ASVs) using the dada2 tool, which also filtered out chimeras. Each sequence read was trimmed to a length of 298 bp. A total of 3.1 million sequences of the 16S rRNA gene were generated from the 166 samples, with a mean of 19,000 sequences per sample. A feature table was generated for all samples with a minimum of 9159 sequences per sample. One sample with a very low number of reads was removed for further analysis. The feature table of the 165 remaining samples was then used to perform taxonomic classification, alpha- and beta-diversity analyses, and differential abundance measurements in different experimental groups. Taxonomy was assigned to each ASV using a database that combined the Greengenes (version 13.8) and PATRIC (version 2016) databases. To study the association between the microbiome data and clinical or dietary variables, we then used linear mixed models as implemented in the Microbiome Multivariable Association with Linear Models (MaAsLin2) package [53]. MaAsLin2 was set up with the following parameters: normalization = "TMM", transform = "LOG", correction = "BH", analysis_method = "LM", max_significance = 0.25 (default significance threshold), min_abundance = 0.0001, min_prevalence = 0.1. Age, gender, and other characteristics of the participants as well as dietary data were added as fixed effects. All models were adjusted for gender, and as participant samples from two timepoints were included, the participant identification number was added as a random effect. Results with a false-discovery rate (FDR) lower than 0.25 were considered significant.

2.9. Deposition of Sequences Data

Sequence data have been deposited in the NCBI database with the following access number: PRJNA745527.

3. Results

3.1. Study Design

We built a new semi-quantitative and simplified food-frequency questionnaire (sFFQ) to assess usual dietary intake. This sFFQ would be useful for epidemiological studies seeking to correlate dietary information with microbiome data. To address this knowledge gap, we designed a pilot study (n = 40) to build a sFFQ based on the dietary habits of our population extracted from three 24HRs administered over a period of one month. From this pilot study, we generated a sFFQ, which included several lifestyle-related questions and 58 food items. The food items were classified into 24 food groups and 29 nutrients (Supplementary Table S2: Classification into food groups, energy and nutrients of the sFFQs and the mean 3–24HRs). Participants completed the sFFQ online.

To collect information on dietary intake, participants were interviewed three times by a trained technical staff over a period of one month in a form of three dietary recalls (24HRs) (baseline, day 15, day 30) and were asked to complete the online sFFQ at baseline (sFFQ1) and on day 30 (sFFQ2), as described in Figure 1 and in the method section. Participants collected stool samples at baseline and one month after (first stool after completing each sFFQ) and kept them in their home freezer until they could bring them to the lab, where they were kept at −80 °C.

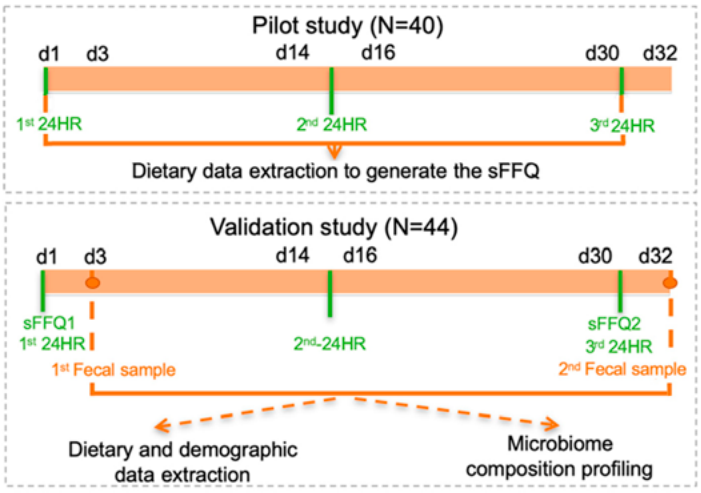


Figure 1. Study design. We first designed a pilot study to extract and quantify the dietary habits of our population. These data were then used to build a simplified food-frequency questionnaire (sFFQ). In the pilot study, participants were interviewed three times by trained staff members over a one-month period in a form of three dietary recalls (three 24HRs) (baseline, day 15, day 30), and in the validation study, they also underwent three 24HRs and were asked to complete two web-based sFFQs, one at baseline and the other a month later (first stool after completing each sFFQ). Participants provided a frozen stool sample on day 3 and day 32 for microbiome analysis.

3.2. Participants' Characteristics

The participants (n = 84) in this study reside mainly in Spain, and 81% hold Spanish nationality. The cohort was recruited among staff from the Vall d'Hebron hospital, as well as their relatives and close friends, via flyers and word of mouth. Females accounted for 55.9% of the participants. The average age was 34.2 years old (from 20 to 64 years old), with 52.4% in the range of 18–29 years old. Among other relevant characteristics, 80% of the cohort presented a normal BMI (18.5–24.9 kg/m^2), 85.7% were born vaginally, and 72.6% were non-smokers. Among the different types of diet, about 80% of the participants followed a conventional diet, but 60% reported the consumption of ready-to-eat meals and 25% the use of artificial sweeteners. Comparison analysis of several parameters, such as age, gender, BMI, nationality, type of birth, and dietary habits, between the pilot and the validation study did not reveal significant differences. More detailed information on the characteristics of the cohort is provided in Table 1.

3.3. Validation of the sFFQ

Dietary data extracted from the 24HRs were converted into 58 food items (Supplementary Table S1: List of 58 food items as specified in the sFFQ) as they were listed in the sFFQ, and all data were also converted into a list of 24 food groups, total energy, and 29 nutrients (Supplementary Table S2: Classification into food groups, energy and nutrients of the sFFQs and the mean 3–24HRs, Figure 2). The participants spent an average of 22 min answering questions on the 58 items (SD = 16.2 min, max = 86 min, min = 4 min).

Table 1. Participants' characteristics.

	Total	Pilot Study	Validation Study
n	**84**	**40**	**44**
Age (years)	34.2 ± 12.7	32.6 ± 11.1	35.7 ± 14.0
18–29 years, *n* (%)	44 (52.4)	21 (52.5)	23 (52.3)
30–39 years, *n* (%)	19 (22.6)	13 (32.5)	6 (13.6)
40–49 years, *n* (%)	7 (8.3)	2 (5.0)	5 (11.4)
50–59 years, *n* (%)	9 (10.7)	3 (7.5)	6 (13.6)
>60 years, *n* (%)	5 (4.2)	1 (2.5)	4 (9.1)
Female gender, *n* (%)	47 (55.9)	26 (65.0)	21 (47.7)
BMI (kg/m^2)	22.5 ± 3.0	22.0 ± 2.6	23.1 ± 3.3
Weight status, *n* (%)			
Underweight (<18.5 kg/m^2)	5 (4.2)	3 (7.5)	2 (4.5)
Normal (18.5–24.9 kg/m^2)	67 (79.8)	35 (87.5)	32 (72.2)
Overweight (25–29.9 kg/m^2)	10 (11.9)	2 (5.0)	8 (18.1)
Obese (>30 kg/m^2)	2 (2.4)	0	2 (4.5)
Nationality, *n* (%)			
Spain	68 (81.0)	29 (72.5)	39 (88.6)
European—non-Spanish	8 (9.5)	6 (15)	2 (4.5)
Others	8 (9.5)	5 (12.5)	3 (6.8)
Birth type, *n* (%)			
Vaginal birth	72 (85.7)	35 (87.5)	37 (84.1)
C-section	12 (14.3)	5 (12.5)	7 (16.0)
Blood type, *n* (%)			
A	26 (30.9)	14 (16.7)	12 (14.3)
B	4 (4.7)	2 (2.4)	2 (2.4)
AB	1 (1.2)	0	1 (1.2)
O	33 (39.3)	15 (17.9)	18 (21.4)
Unknown	20 (23.8)	9 (10.7)	11 (13.1)
Smoking status, *n* (%)			
Non-smoker	61 (72.6)	32 (80)	29 (65.9)
Smoker	9 (10.7)	2 (5)	7 (15.9)
Former smoker	8 (9.5)	0	8 (18.2)
Unknown	6 (7.1)	6 (15)	0
Diet type, *n* (%)			
Conventional	67 (79.8)	31 (77.5)	37 (84.1)
Vegetarian diet	6 (7.1)	4 (10.0)	2 (4.5)
Vegan diet	2 (2.4)	1 (2.5)	1 (2.3)
Organic diet	2 (2.4)	2 (5.0)	0
Others diet	7 (8.3)	2 (5.0)	4 (9.1)
Intake of ready-to-eat meals, *n* (%)			
Yes	51 (60.7)	23 (57.5)	28 (63.6)
No	33 (39.3)	17 (42.5)	16 (36.4)
Intake of sweeteners, *n* (%)			
Yes	-	–	11 (25.0)
No	-	–	33 (75.0)
Intake of supplements or drugs, *n* (%)			
Dietary supplements	20 (23.8)	8 (20.0)	12 (27.3)
Probiotics	1 (1.2)	1 (2.5)	0
Oral contraceptive	6 (7.1)	4 (10.0)	2 (4.5)
ACE inhibidors	3 (3.6)	2 (5.0)	1 (2.3)
Fibrate	1 (1.2)	1 (2.5)	0
Statin	1 (1.2)	1 (2.5)	0
Levothyroxine	2 (2.4)	1 (2.5)	1 (2.3)
Other drugs	8 (9.5)	4 (10.0)	4 (9.1)

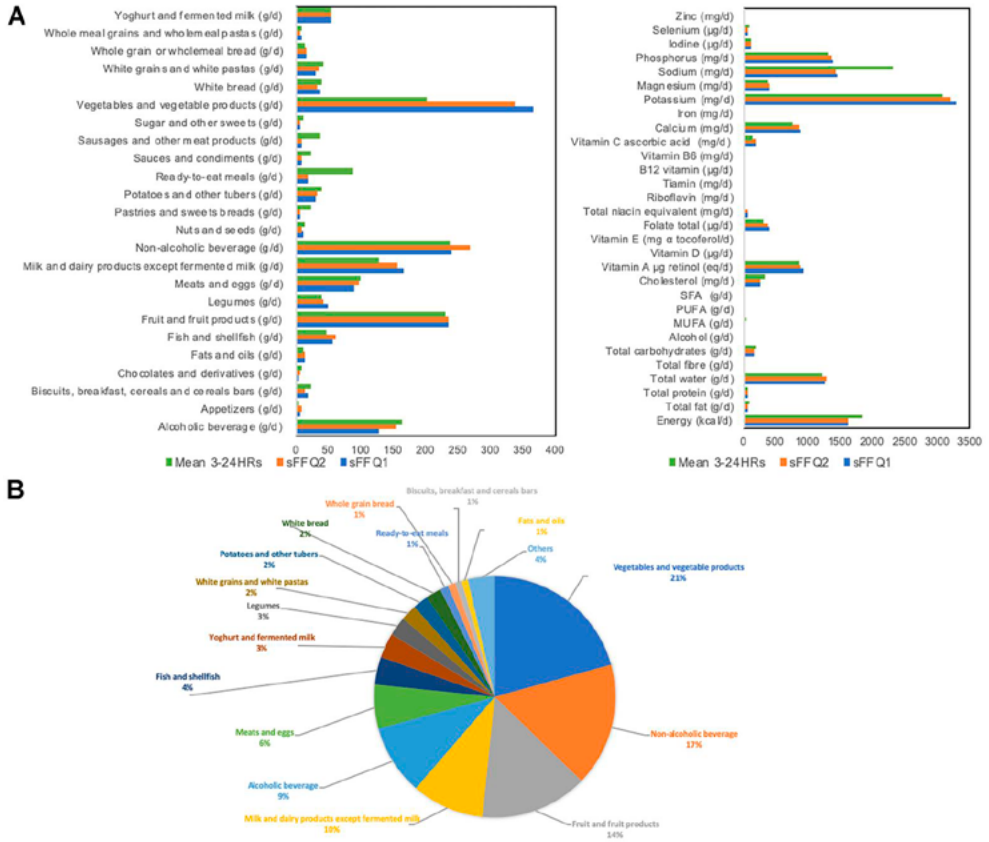


Figure 2. Participants' dietary profile. (**A**) Food groups, energy, and nutrients extracted from the sFFQs and 24HRs. (**B**) Proportions of food groups as extracted from the sFFQ2 (validation study). Food groups with proportions lower than 1% were grouped into "Others".

To validate the newly developed sFFQ, we conducted a second study (n = 44)—a validation study—in which we compared the food items, food groups, and nutrients obtained from the sFFQ with the mean of the three 24HRs. For this comparison, we used data obtained from sFFQ2, which recalled the dietary intake of the previous month and therefore may better correspond to the mean of the three 24HRs of the same month.

We applied several statistical tests to measure the strength and direction of the association between the two different measurements at individual level (ICC, CC, and Spearman's rank correlation tests) and to quantify agreement between the two measures at group level (Wilcoxon test and Bland–Altman plots), as recommended by Lombard et al. [22].

The median ICC coefficient of food groups between the sFFQ2 and the mean of the three 24HRs was 0.35 (range: 0.05–0.83), and the median ICC of energy-adjusted nutrients was 0.55 (range 0.08–0.98) (Supplementary Table S3: Relative validation of the sFFQ (sFFQ2 vs. mean 3–24HRs)). The cross-classification values satisfactorily classified participants on the basis of their intake based on the two methods since more than 50% of the participants were classified in the same tertile for 14 food groups and 13 nutrients, while less than 10% were classified for the opposite tertile in 12 food groups and 15 nutrients. The median Spearman's correlation coefficient for food groups between sFFQ2 and mean of three 24HRs was 0.46 (range 0.18–0.78) and was also 0.46 (range 0.10–0.71) for energy-adjusted nutrients (Supplementary Table S3: Relative validation of the sFFQ (sFFQ2 vs. mean 3–24HRs)).

Wilcoxon tests showed that 15 out of 24 (62.5%) food groups and 19 out of 29 nutrients (63%) were not significantly different between the two diet assessment methods.

Based on Bland–Altman analysis, we observed that the sFFQ2 tended to report a lower intake of biscuits breakfast cereals, chocolate and derivatives, pastries and sweet breads, ready-to-eat meals, sauces-condiments, and sausages and a higher consumption of the food groups of appetizers and vegetables than the 24HR based on the method used in Giavarina [50] (Supplementary Table S3: Relative validation of the sFFQ (sFFQ2 vs. mean 3–24HRs)). Regarding energy and nutrients, the sFFQ2 underestimated energy intake, total fat, polyunsaturated fatty acids (PUFA), saturated fatty acids (SFA), cholesterol, sodium, and selenium and overestimated vitamin D, folate, niacin, and vitamin C compared to the 24HR. Altogether, we observed that 13 out 24 (54%) food groups and 20 out of 29 nutrients (69%) were classified as good or acceptable according to the criteria reported in Lombard et al. [22].

3.4. Reproducibility of the sFFQ

We then evaluated the reproducibility of the sFFQ by comparing the dietary data extracted from the sFFQs administered on two occasions one month apart in the validation study. The median ICC coefficient of food groups was 0.63 (range: 0.21–0.90), and the ICC of energy-adjusted nutrients was 0.73 (range 0.58–0.96) (Supplementary Table S4: Reproducibility analysis of the sFFQ (sFFQ1 vs. sFFQ2)). The median Spearman's correlation for food groups and energy-adjusted nutrients was 0.72 (range 0.32–0.90) and 0.70 (range 0.54–0.86) (Supplementary Table S4: Reproducibility analysis of the sFFQ (sFFQ1 vs. sFFQ2)), respectively. The Wilcoxon test revealed that no food groups or nutrients were significantly different between the two sFFQs.

3.5. Participants' Dietary Profile

The food groups, energy, and nutrients extracted from the sFFQs and 24HRs were quantified (Figure 2 and Supplementary Table S2: Classification into food groups, energy and nutrients of the sFFQs and the mean 3–24HRs). The most consumed food groups in our population were vegetables (21%), non-alcoholic beverages (17%), fruits (14%), milk and dairy products except fermented milk (10%), alcoholic beverages (9%), meat and eggs (6%), fish and shellfish (4%), yoghurt and fermented milk (3%), and legumes (3%).

3.6. Correlation between Participants' Characteristics and Microbial Diversity and Taxa

We collected two faecal samples from participants in the pilot (n = 40) and validation (n = 44) studies at baseline and one month later (first stool after each sFFQ). Microbiome composition was analysed based on the amplification and sequencing of the V4 region of 16S rRNA gene. We evaluated the association between microbial alpha-diversity (richness and evenness) or taxonomic profile (relative abundance of microbial genera) and several characteristics of the participants, including age, BMI, gender, smoking habit, blood type, and type of birth. To this end, we used linear mixed models implemented in the MaAsLin2 tool and took into account the longitudinal setting of the study, as the participant identification number (Subject ID) was added as a random effect.

Vaginal birth was found to be associated with higher microbial diversity (FDR < 0.02 for Chao1 (richness) and Shannon (evenness) indices) and resulted in enrichment in several bacterial genera, including an unclassified genus from the Ruminococcaceae family (FDR = 6.37×10^{-7}), from the Clostridiales order (FDR = 5.80×10^{-5}) and from RF39 (FDR = 0.0006) as compared with C-section births (Figure 3). Age was found positively correlated with diversity (FDR = 0.06 for Shannon index) and negatively correlated with *Bilophila* (FDR = 0.005). The pre-obese and obese group (BMI above 25) was associated with a lower microbial diversity (FDR = 0.018 for Chao1 and FDR = 0.12 for Shannon) and depleted in members of the Clostridiales order, such as *Facecalibacterium* (FDR = 0.10). BMI was classified following the World Health Organization's recommendation as follows: underweight (BMI below 18.5), normal weight (BMI = 18.5–24.9), pre-obesity

(BMI = 25.0–29.9), and obesity (BMI above 30.0). Smoking habit (smoker, non-smoker, or ex-smoker) was not associated with diversity or any microbial taxon. Gender was not found associated with diversity but was associated with a depletion of *Bilophila* (FDR = 0.04) in male participants. Use of sweeteners was negatively associated with *Desulfovibrio* (FDR = 0.07).

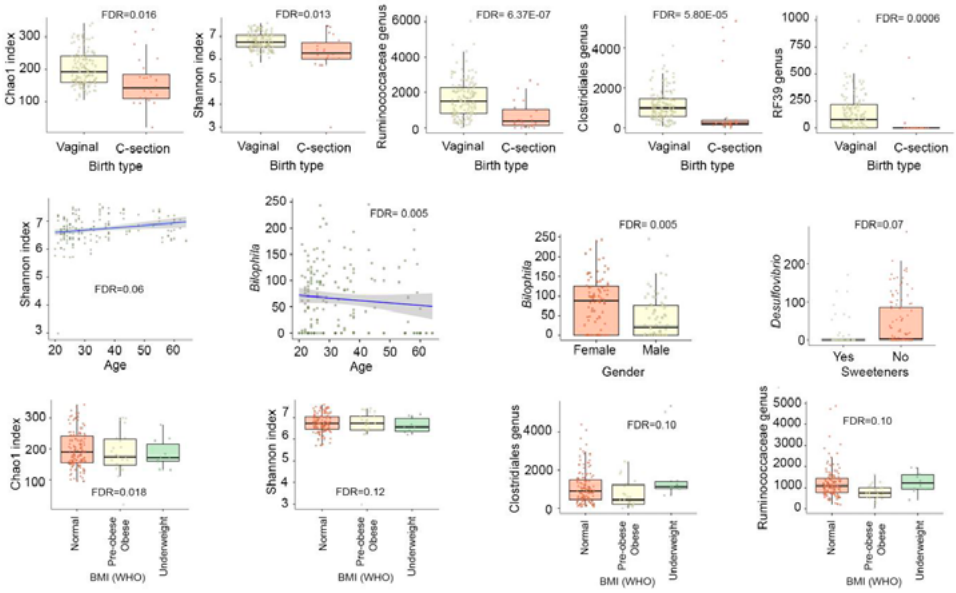


Figure 3. Association between participants' characteristics and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 were used to analyse the microbiome and data on the characteristics of participants in the pilot and validation studies (n = 84, 165 faecal samples). BMI, body mass index; WHO, World Health Organization.

3.7. Correlation between Dietary Intake and Microbial Diversity and Taxa

To correlate microbiome data with dietary intake, we first used the 58 food items from the sFFQs (n = 44, 85 faecal samples; one sample did not provide sufficient high-quality sequences, and two subjects provided only one sample at baseline). The analysis was performed using MaAsLin2. The results are shown in Figure 4. Item 14, which consisted of fresh fruit, was positively associated with richness and evenness (FDR = 0.009 for Chao1 and Shannon) and with the relative abundance of a member of the Ruminococcaceae family (FDR = 0.1). Item 35, which consists of processed meats, was negatively associated with richness (Chao1, FDR = 0.034), evenness (Shannon, FDR = 0.03), and an unclassified genus from the Clostridiales order (FDR = 0.1). Item 58, which comprised process foods, was negatively associated with richness and evenness (FDR = 0.009 for Chao1 and Shannon) and with an unclassified genus from the Clostridiales order (FDR = 0.19).

We then correlated microbiome data with the 24 food groups extracted from the sFFQ. Fruits and fruit products, which encompassed food items 14, 15, and 16, were also found to be positively correlated with microbial diversity (FDR = 0.005 for Chao1 and Shannon indices). "Sausages and other processed meats" and "ready-to-eat meals", which corresponded to food items 35 and 58, respectively, were found to be negatively correlated with microbial diversity and taxa, as mentioned above. No association was observed between microbiome data and the other items or food groups.

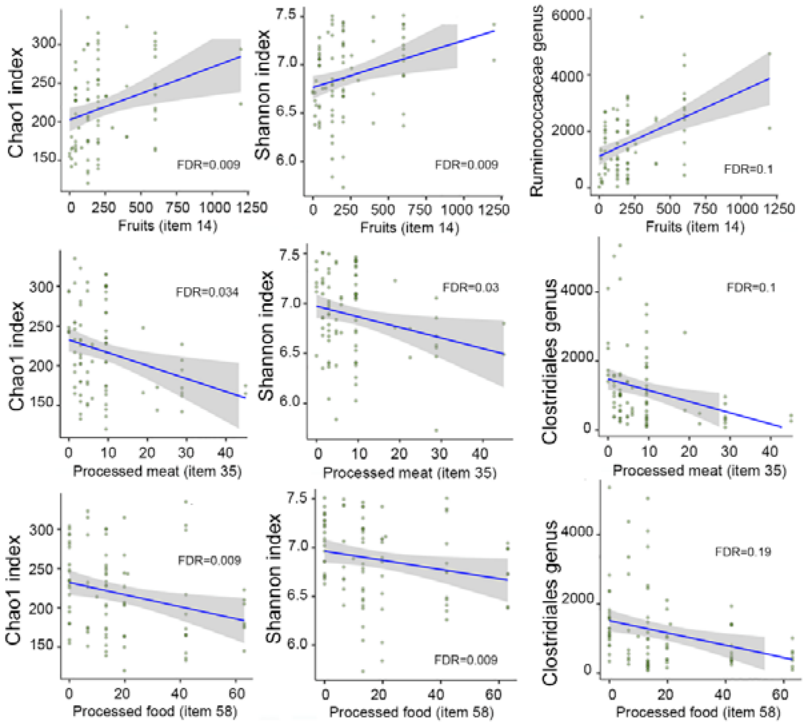
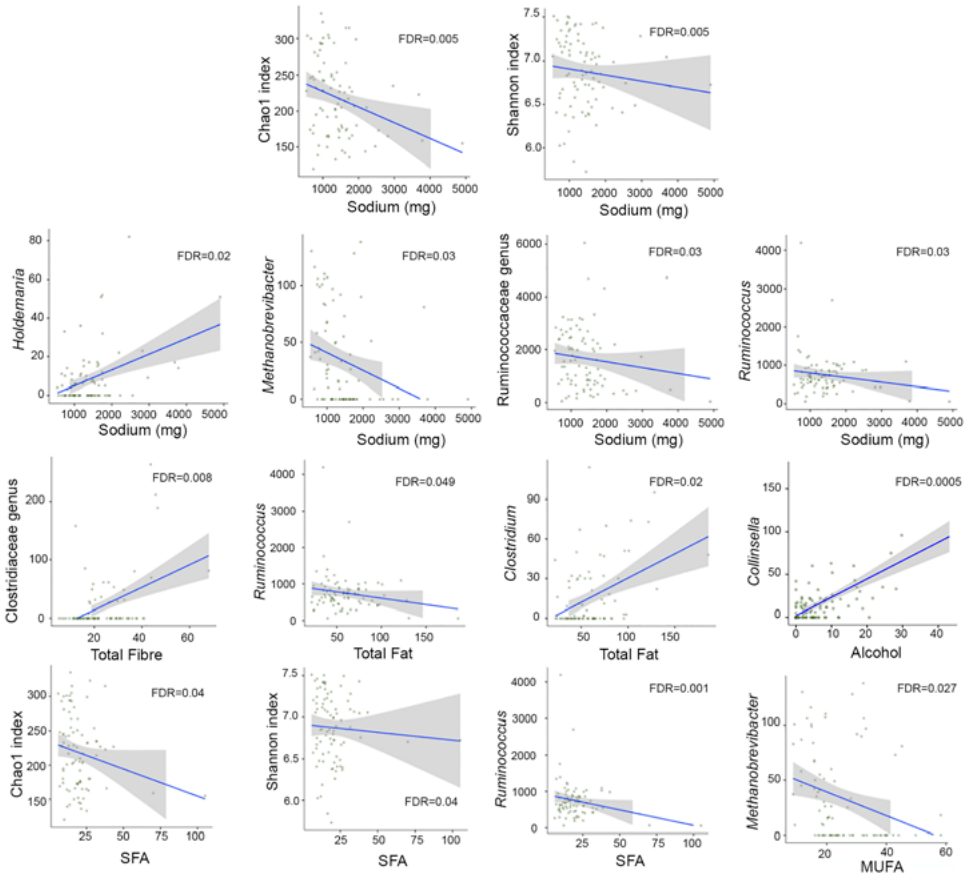


Figure 4. Correlation between food groups and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 were used to analyse the microbiome and dietary data extracted from the validation study (n = 44, 85 faecal samples).

Finally, we did not uncover any association between the 29 nutrients and total energy extracted from the sFFQs and microbial diversity except for sodium (FDR = 0.005) (Figure 5) and saturated fatty acid (SFA) (FDR = 0.04) (Figure 5), whose levels were negatively correlated with both richness and evenness. Sodium was positively correlated with *Holdemania* and negatively correlated with *Ruminococcus* and *Methanobrevicter*, a member of the Ruminococcaceae family (FDR < 0.05). Other nutrients, such as alcohol, total fat, and total fibre, were also associated with several microbial genera (Figure 5). Alcohol was positively correlated with two genera from the Coriobacteriaceae family, one of them being *Collinsella* (FDR = 0.0004), and negatively correlated with a member of the Peptostreptococcaceae family (FDR = 0.03). Total fat and SFA were negatively correlated with *Ruminococcus;* total fat was also positively correlated with *Clostridium*. Monounsaturated fatty acids (MUFA) were negatively correlated with *Methanobrevibacter* (FDR = 0.03). Total fibre was positively correlated with a member of the Clostridiaceae family (FDR = 0.009).



ure 5. Correlation between nutrients and microbial diversity and taxa. Linear mixed models implemented in MaAsLin2 ·e used to analyse the microbiome diversity and taxonomic profile and nutrients data extracted from the validation study : 44, 85 faecal samples). SFA, saturated fatty acids; MUFA, monounsaturated fatty acids. All nutrients except for sodium mg) are displayed in grams.

4. Discussion

This study describes the development and validation of a semi-quantitative and simplified FFQ and the integration of demographic and dietary data into microbiome data. The newly developed online sFFQ, which contains 58 food items converted into 24 food groups and 29 nutrients, reports the dietary intake of an adult population in the last month. The reduced number of items was chosen to lessen the burden on respondents and to maximize their full attention. The time frame of one month is an attempt to match usual dietary consumption with changes in the microbiome community [54]. Furthermore, respondents completed the sFFQ in an average of 22 min, which is much less time than that needed for the most commonly used FFQs (from 30 to 60 min) [14]. We consider that a short FFQ will attract more volunteers who will be more willing to repeat the experiment several times in a year to cover, for instance, every season.

The Bland–Altman analysis reflected good levels of agreement between the sFFQ and DR, and the graphs showed that most of the data fell within the limits of agreement. The sFFQ reported 20 out of 29 (69%) nutrients and 13 out of 24 (54%) food groups with good or acceptable outcome compared to the 24HR reference method, as evaluated by

at least three distinct statistical methods (Supplementary Table S3: Relative validation of the sFFQ (sFFQ2 vs. mean 3–24HRs). The sFFQ showed only 24% underestimation and 11% overestimation of the food, energy, and nutrient group. The results obtained for several food items (biscuits, breakfast, cereals, chocolate and derivatives, pastries and sweet breads, ready-to-eat meal, sauces-condiments, and sausages) for energy and for several nutrients, such as total fat, PUFA, SFA, cholesterol, sodium, and selenium, should be interpreted with caution since they were underreported, whereas vitamin D, folate, niacin, and vitamin C were overreported in the sFFQ compared to the three 24HRs. The under- or over-estimation trends of certain foods or nutrients could be explained by a series of characteristics of the participants and by social approval of certain foods [55,56]. Indeed, the consumption of foods considered "beneficial for health", such as fruits and vegetables, are usually reported more frequently while that of "bad" foods, such as foods high in fat or sugar, are usually less frequently reported [57–59]. In addition, since the main objective of this dietary assessment was not to measure energy intake, the underestimation of energy should not impact the overall design of our sFFQ. At the group level, the correlation values for food groups (Spearman = 0.177–0.78 and ICC = 0.049–0.83) were within the ranges observed in previous validation studies in adults [60,61]. The energy-adjusted correlation values were similar to those reported in several validation studies [62,63]. The greatest discrepancies in cross-classification were observed especially for foods eaten sporadically (fish and shellfish, legumes, pastries, ready-to-eat meals, sausages), possibly due to the low probability of encountering these foods in the three 24HRs.

The very high correlation of our extracted data, with two published epidemiological studies (ANIBES and ENIDE) investigating the usual diet of the Spanish population, suggests that the newly built sFFQ could be applied at the population level (data not shown). Moreover, the very high repeatability of the questionnaire indicates that only one sFFQ would be needed to cover food intake over one month. However, to adapt the questionnaire to another population, validation using a reference such as a 24HR on a subpopulation would be needed to make additional and necessary changes to the sFFQ.

Our study, using three 24HRs on 84 healthy individuals, captured certain effects of BMI as well as lifestyle on the diversity and composition of the gut microbial community. The association between microbiome and demographic data was achieved with participants from both the pilot and validation study (n = 84, 165 faecal samples). The impact of type of birth on the gut microbial ecosystem has been widely studied during early life [64,65]. A persistent effect of mode of delivery on the microbiome composition, the host immune system [66], and the biosynthesis of natural antibiotics [67] has been reported during the first years of life. However, to the best of our knowledge, no relevant study has reported this effect in adult subjects. Our findings regarding relationship between age and diversity, depletion of Clostridiales, and overweight-obesity corroborate previous findings [68,69].

Contradictory findings have been published about the effect of sweeteners, such as saccharin, on glucose tolerance and dysbiosis in healthy individuals [70,71]. We observed that sweeteners, which consisted mainly of aspartame and saccharin in our study, decreased *Desulfovibrio*, which is a sulphate-reducing bacterium.

On the one hand, through the sFFQ, we were able to associate high microbial diversity, which is considered a health-promoting factor [72], with the intake of fruits and low diversity with processed meat, ready-to-eat meals, total fat, saturated fatty acids, and sodium intake. On the other hand, the sFFQ allowed us to correlate food items or nutrients with specific groups of microorganisms. Some members of the Clostridiales order were positively associated with fruits and total fibre, whereas others, including *Ruminococcus*, were negatively associated with total fat, saturated fatty acids, and sodium intake. The association with the latter should be interpreted with caution given that the comparison between sFFQ2 and the mean of the three 24HRs based on Bland–Altman showed that sodium was underreported by participants.

A low intake of dietary fibre has been related to loss of diversity and loss, in particular of members of Clostridiales [73] and the class Clostridia [74]. The non-association observed

between vegetables and microbial diversity could be explained by the cooking method used. Indeed, fruits and vegetables may contain similar nutrients that could be lost during cooking procedures involving boiling, steaming, or stir-frying, methods commonly use to prepare vegetables. Cooking, which transforms fibre and starch, increasing their absorption in the small intestine and thus reducing their fraction in the colon, has been shown to reshape the structure and function of gut microbiota [75].

Our study showed that the newly developed sFFQ has the potential to capture the usual diet of adult healthy individuals as reflected by its validation with a reference method and a comparison with other epidemiological studies. Nevertheless, we need to stress several limitations. First, this study, which was observational in nature, requires, for instance, human interventional studies to further validate the associations found between diet and the microbial community. This study relies only on 16S rRNA analysis, which reveals only microbiome composition, and it could be complemented by functional analysis through DNA and RNA shotgun sequencing or metabolomics analysis. Moreover, the results obtained on several food items, energy, and several nutrients should be interpreted with caution since they have been shown to be misreported in the sFFQ compared to the three 24HRs. However, the 24-HR used here as a reference method, is not the gold standard. This reporting method also relies on memory and may be biased due to underestimation or overestimation. This limitation could be addressed by using metabolomic biomarkers, although only a few comprehensively validated biomarkers of food intake are available.

5. Conclusions

To the best of our knowledge, this newly developed sFFQ is the first to be validated and tested against microbiome data. This new sFFQ could be adapted and used in future population studies to assess diet in a population from another region of the world and/or to study the effect of diet and metabolic disorders. We expect this new tool to open up new avenues in both nutritional and microbiome fields leading to nutri-metaomics.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/nu13092978/s1, Table S1: List of 58 food items as specified in the sFFQ, Table S2: Classification into food groups, energy and nutrients of the sFFQs and the mean 3–24HRs, Table S3: Validation of the sFFQ (sFFQ2 vs. mean 3–24HRs), Table S4: Reproducibility analysis of the sFFQ.

Author Contributions: Conceptualization, C.M.; formal analysis, F.Y. and C.M.; funding acquisition, C.M.; methodology, F.Y., Z.S., M.O., Z.X. and G.S.-G.; software, Z.X. and I.O.; validation, F.Y.; writing—original draft, F.Y. and C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by grants from by the Instituto de Salud Carlos III/FEDER (PI17/00614; PI20/00130), a government agency. Francisca Yáñez was supported by a fellowship from ANID, BECAS Chile, No. 72190278. Zixuan Xie received a fellowship from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Action, Innovative Training Network: FunHoMic; grant number 812969.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the Vall d'Hebron University Hospital of Barcelona (protocol code: PR(AG)156/2017 and date of approval: 8 May 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Sequence data have been deposited in the NCBI database with the following access number: PRJNA745527.

Acknowledgments: To Ferran Pinsach and Carlos Blazquez for their help with the statistical analyses and Blanca Ruiz de la Torre for assistance in collecting dietary data.

Conflicts of Interest: No conflict of interest.

References

1. Gowlett, J.A. The discovery of fire by humans: A long and convoluted process. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20150164. [CrossRef]
2. Ragir, S. Diet and Food Preparation: Rethinking Early Hominid Behavior. *Evol. Anthropol.* **2000**, *9*, 153–155. [CrossRef]
3. Davenport, E.R.; Sanders, J.G.; Song, S.J.; Amato, K.R.; Clark, A.G.; Knight, R. The human microbiome in evolution. *BMC Biol.* **2017**, *15*, 127. [CrossRef]
4. Pittelkow, C.M.; Liang, X.; Linquist, B.A.; van Groenigen, K.J.; Lee, J.; Lundy, M.E.; van Gestel, N.; Six, J.; Venterea, R.T.; van Kessel, C. Productivity limits and potentials of the principles of conservation agriculture. *Nature* **2015**, *517*, 365–368. [CrossRef]
5. De Filippo, C.; Cavalieri, D.; Di Paola, M.; Ramazzotti, M.; Poullet, J.B.; Massart, S.; Collini, S.; Pieraccini, G.; Lionetti, P. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14691–14696. [CrossRef]
6. Sun, Y.; Zuo, T.; Cheung, C.P.; Gu, W.; Wan, Y.; Zhang, F.; Chen, N.; Zhan, H.; Yeoh, Y.K.; Niu, J.; et al. Population-Level Configurations of Gut Mycobiome Across 6 Ethnicities in Urban and Rural China. *Gastroenterology* **2021**, *160*, 272–286.e11. [CrossRef]
7. Turnbaugh, P.J.; Ley, R.E.; Mahowald, M.A.; Magrini, V.; Mardis, E.R.; Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444*, 1027–1031. [CrossRef] [PubMed]
8. Qin, J.; Li, Y.; Cai, Z.; Li, S.; Zhu, J.; Zhang, F.; Liang, S.; Zhang, W.; Guan, Y.; Shen, D.; et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **2012**, *490*, 55–60. [CrossRef]
9. David, L.A.; Maurice, C.F.; Carmody, R.N.; Gootenberg, D.B.; Button, J.E.; Wolfe, B.E.; Ling, A.V.; Devlin, A.S.; Varma, Y.; Fischbach, M.A.; et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **2014**, *505*, 559–563. [CrossRef]
10. Ghosh, T.S.; Rampelli, S.; Jeffery, I.B.; Santoro, A.; Neto, M.; Capri, M.; Giampieri, E.; Jennings, A.; Candela, M.; Turroni, S.; et al. Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: The NU-AGE 1-year dietary intervention across five European countries. *Gut* **2020**, *69*, 1218–1228. [CrossRef] [PubMed]
11. Johnson, A.J.; Zheng, J.J.; Kang, J.W.; Saboe, A.; Knights, D.; Zivkovic, A.M. A Guide to Diet-Microbiome Study Design. *Front. Nutr.* **2020**, *7*, 79. [CrossRef]
12. So, D.; Whelan, K.; Rossi, M.; Morrison, M.; Holtmann, G.; Kelly, J.T.; Shanahan, E.R.; Staudacher, H.M.; Campbell, K.L. Dietary fiber intervention on gut microbiota composition in healthy adults: A systematic review and meta-analysis. *Am. J. Clin. Nutr.* **2018**, *107*, 965–983. [CrossRef]
13. Gibson, R.S. *Principles of Nutritional Assessment*, 2nd ed.; Oxford University Press: New York, NY, USA, 2005.
14. Thompson, F.E.; Subar, A.F. *Dietary Assessment Methodology*; Chapter 1; Academic Press: San Diego, CA, USA, 2017.
15. Cade, J.E.; Burley, V.J.; Warm, D.L.; Thompson, R.L.; Margetts, B.M. Food-frequency questionnaires: A review of their design, validation and utilisation. *Nutr. Res. Rev.* **2004**, *17*, 5–22. [CrossRef]
16. Ruiz, M.D.; Martín-Lagos, R.A. *Guía Para Estudios Dietéticos. Album Fotográfico de Alimentos*, 1st ed.; Universidad de Granada: Granada, Spain, 2010.
17. Hercberg, S.; Deheeger, M.; Preziosi, P. *Portions Alimentaries: Manuel Photos Pour L'Estimation des Quantités*, 1st ed.; Economica: Paris, France, 2002.
18. Moreiras, O.; Carbajal, A.; Cabrera, L. *Tabla de Composición de Alimentos*, 19th ed.; Pirámide: Madrid, Spain, 2019.
19. Raper, N.; Perloff, B.; Ingwersen, L.; Steinfeldt, L.; Anand, J. An overview of USDA's Dietary Intake Data System. *J. Food Compos. Anal.* **2004**, *17*, 545–555. [CrossRef]
20. The USDA National Nutrient Database for Standard Reference. Release 27 ed.; 2015. Available online: http://www.ars.usda.gov/nea/bhnrc/mafcl (accessed on 25 March 2021).
21. Cade, J.; Thompson, R.; Burley, V.; Warm, D. Development, validation and utilisation of food-frequency questionnaires—A review. *Public Health Nutr.* **2002**, *5*, 567–587. [CrossRef]
22. Lombard, M.J.; Steyn, N.P.; Charlton, K.E.; Senekal, M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. *Nutr. J.* **2015**, *14*, 40. [CrossRef] [PubMed]
23. Walter, W. *Nutritional Epidemiology*, 3rd ed.; Oxford University Press: New York, NY, USA, 2012.
24. Willett, W.C.; Sampson, L.; Stampfer, M.J.; Rosner, B.; Bain, C.; Witschi, J.; Hennekens, C.H.; Speizer, F.E. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol.* **1985**, *122*, 51–65. [CrossRef] [PubMed]
25. Suarez, M.V.; Mañas, R.J.; Fernández, S.R.; de Dios, R. National Food Survey on Adults, the Elderly and Pregnant Women (ENALIA2). 2016. Available online: http://www.aesan.gob.es/AECOSAN/docs/documentos/seguridad_alimentaria/evaluacion_riesgos/Spanish_National_dietary_survey_adults_elderly_pregnant.pdf (accessed on 17 February 2021).
26. Arnolds, K.L.; Martin, C.G.; Lozupone, C.A. Blood type and the microbiome- untangling a complex relationship with lessons from pathogens. *Curr. Opin. Microbiol.* **2020**, *56*, 59–66. [CrossRef]
27. Reyman, M.; van Houten, M.A.; van Baarle, D.; Bosch, A.; Man, W.H.; Chu, M.; Arp, K.; Watson, R.L.; Sanders, E.A.M.; Fuentes, S.; et al. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat. Commun.* **2019**, *10*, 4997. [CrossRef] [PubMed]
28. Karl, J.P.; Armstrong, N.J.; McClung, H.L.; Player, R.A.; Rood, J.C.; Racicot, K.; Soares, J.W.; Montain, S.J. A diet of U.S. military food rations alters gut microbiota composition and does not increase intestinal permeability. *J. Nutr. Biochem.* **2019**, *72*, 108217. [CrossRef]

29. Zimmer, J.; Lange, B.; Frick, J.S.; Sauer, H.; Zimmermann, K.; Schwiertz, A.; Rusch, K.; Klosterhalfen, S.; Enck, P. A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. *Eur. J. Clin. Nutr.* **2012**, *66*, 53–60. [CrossRef]
30. Cani, P.D.; Everard, A. Talking microbes: When gut bacteria interact with diet and host organs. *Mol. Nutr. Food Res.* **2016**, *60*, 58–66. [CrossRef] [PubMed]
31. Claesson, M.J.; Jeffery, I.B.; Conde, S.; Power, S.E.; O'Connor, E.M.; Cusack, S.; Harris, H.M.; Coakley, M.; Lakshminarayanan, B.; O'Sullivan, O.; et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature* **2012**, *488*, 178–184. [CrossRef] [PubMed]
32. Cotillard, A.; Kennedy, S.P.; Kong, L.C.; Prifti, E.; Pons, N.; Le Chatelier, E.; Almeida, M.; Quinquis, B.; Levenez, F.; Galleron, N.; et al. Dietary intervention impact on gut microbial gene richness. *Nature* **2013**, *500*, 585–588. [CrossRef]
33. McDonald, D.; Hyde, E.; Debelius, J.W.; Morton, J.T.; Gonzalez, A.; Ackermann, G.; Aksenov, A.A.; Behsaz, B.; Brennan, C.; Chen, Y.; et al. American Gut: An Open Platform for Citizen Science Microbiome Research. *MSystems* **2018**, *3*, e00031-18. [CrossRef] [PubMed]
34. Rinninella, E.; Cintoni, M.; Raoul, P.; Lopetuso, L.R.; Scaldaferri, F.; Pulcini, G.; Miggiano, G.A.D.; Gasbarrini, A.; Mele, M.C. Food Components and Dietary Habits: Keys for a Healthy Gut Microbiota Composition. *Nutrients* **2019**, *11*, 2393. [CrossRef]
35. Turnbaugh, P.J.; Ridaura, V.K.; Faith, J.J.; Rey, F.E.; Knight, R.; Gordon, J.I. The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **2009**, *1*, 6ra14. [CrossRef]
36. Wu, G.D.; Chen, J.; Hoffmann, C.; Bittinger, K.; Chen, Y.Y.; Keilbaugh, S.A.; Bewtra, M.; Knights, D.; Walters, W.A.; Knight, R.; et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **2011**, *334*, 105–108. [CrossRef]
37. Battaglioli, E.J.; Kashyap, P.C. *Diet Effects on Gut Microbiome Composition, Function, and Host Physiology*; Elsevier Inc.: London, UK, 2018; pp. 755–766.
38. SENC (Sociedad Española de Nutrición Comunitaria). Dietary guidelines for the Spanish population. *Nutr. Hosp.* **2016**, *33*, 1–48.
39. Russolillo, G.; Baladia, E.; Moñino, M.; Marques-Lopes, I.; Farran, A.; Bonany, J.; Gilabert, V.; Astiasarán, I.; Raigón, M.D.; Martínez, A.; et al. Establecimiento del tamaño de raciones de consumo de frutas y hortalizas para su uso en guías alimentarias en el entorno español: Propuesta del Comité Científico de la Asociación 5 al día. *Rev. Esp. Nutr. Hum. Diet.* **2019**, *23*, 205–221. [CrossRef]
40. Nelson, M.; Atkinson, M.; Darbyshire, S. Food photography. I: The perception of food portion size from photographs. *Br. J. Nutr.* **1994**, *72*, 649–663. [CrossRef] [PubMed]
41. Nelson, M.; Atkinson, M.; Darbyshire, S. Food photography II: Use of food photographs for estimating portion size and the nutrient content of meals. *Br. J. Nutr.* **1996**, *76*, 31–49. [CrossRef] [PubMed]
42. Subar, A.F.; Thompson, F.E.; Smith, A.F.; Jobe, J.B.; Ziegler, R.G.; Potischman, N.; Schatzkin, A.; Hartman, A.; Swanson, C.; Kruse, L.; et al. Improving food frequency questionnaires: A qualitative approach using cognitive interviewing. *J. Am. Diet. Assoc.* **1995**, *95*, 781–788, quiz 789–790. [CrossRef]
43. Subar, A.F.; Thompson, F.E.; Kipnis, V.; Midthune, D.; Hurwitz, P.; McNutt, S.; McIntosh, A.; Rosenfeld, S. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: The Eating at America's Table Study. *Am. J. Epidemiol.* **2001**, *154*, 1089–1099. [CrossRef] [PubMed]
44. Verger, E.O.; Armstrong, P.; Nielsen, T.; Chakaroun, R.; Aron-Wisnewsky, J.; Gobel, R.J.; Schutz, T.; Delaere, F.; Gausseres, N.; Clement, K.; et al. Dietary Assessment in the MetaCardis Study: Development and Relative Validity of an Online Food Frequency Questionnaire. *J. Acad. Nutr. Diet.* **2017**, *117*, 878–888. [CrossRef]
45. Brown, C.C.; Kipnis, V.; Freedman, L.S.; Hartman, A.M.; Schatzkin, A.; Wacholder, S. Energy adjustment methods for nutritional epidemiology: The effect of categorization. *Am. J. Epidemiol.* **1994**, *139*, 323–338. [CrossRef]
46. Shrout, P.E.; Fleiss, J.L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* **1979**, *86*, 420–428. [CrossRef]
47. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
48. Masson, L.F.; McNeill, G.; Tomany, J.O.; Simpson, J.A.; Peace, H.S.; Wei, L.; Grubb, D.A.; Bolton-Smith, C. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: Use of correlation coefficients and the kappa statistic. *Public Health Nutr.* **2003**, *6*, 313–321. [CrossRef] [PubMed]
49. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *1*, 307–310. [CrossRef]
50. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **2015**, *25*, 141–151. [CrossRef]
51. Sarrabayrouse, G.; Landolfi, S.; Pozuelo, M.; Willamil, J.; Varela, E.; Clark, A.; Campos, D.; Herrera, C.; Santiago, A.; Machiels, K.; et al. Mucosal microbial load in Crohn's disease: A potential predictor of response to faecal microbiota transplantation. *EBioMedicine* **2020**, *51*, 102611. [CrossRef]
52. Santiago, A.; Sanchez, E.; Clark, A.; Pozuelo, M.; Calvo, M.; Yáñez, F.; Sarrabayrouse, G.; Perea, L.; Vidal, S.; Gallardo, A.; et al. Sequential Changes in the Mesenteric Lymph Node Microbiome and Immune Response during Cirrhosis Induction in Rats. *MSystems* **2019**, *4*, e00278-18. [CrossRef]
53. Mallick, H.; Rahnavard, A.; McIver, L.J.; Ma, S.; Zhang, Y.; Nguyen, L.H.; Tickle, T.L.; Weingart, G.; Ren, B.; Schwager, E.H.; et al. Multivariable Association Discovery in Population-scale Meta-omics Studies. *BioRxiv* **2021**. [CrossRef]
54. Leeming, E.R.; Johnson, A.J.; Spector, T.D.; Le Roy, C.I. Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. *Nutrients* **2019**, *11*, 2862. [CrossRef] [PubMed]

55. Maurer, J.; Taren, D.L.; Teixeira, P.J.; Thomson, C.A.; Lohman, T.G.; Going, S.B.; Houtkooper, L.B. The psychosocial and behavioral characteristics related to energy misreporting. *Nutr. Rev.* **2006**, *64*, 53–66. [CrossRef] [PubMed]
56. Tooze, J.A.; Subar, A.F.; Thompson, F.E.; Troiano, R.; Schatzkin, A.; Kipnis, V. Psychosocial predictors of energy underreporting in a large doubly labeled water study. *Am. J. Clin. Nutr.* **2004**, *79*, 795–804. [CrossRef] [PubMed]
57. Cook, A.; Pryer, J.; Shetty, P. The problem of accuracy in dietary surveys. Analysis of the over 65 UK National Diet and Nutrition Survey. *J. Epidemiol. Community Health* **2000**, *54*, 611–616. [CrossRef]
58. Hebert, J.R.; Hurley, T.G.; Peterson, K.E.; Resnicow, K.; Thompson, F.E.; Yaroch, A.L.; Ehlers, M.; Midthune, D.; Williams, G.C.; Greene, G.W.; et al. Social desirability trait influences on self-reported dietary measures among diverse participants in a multicenter multiple risk factor trial. *J. Nutr.* **2008**, *138*, 226S–234S. [CrossRef]
59. Miller, T.M.; Abdel-Maksoud, M.F.; Crane, L.A.; Marcus, A.C.; Byers, T.E. Effects of social approval bias on self-reported fruit and vegetable consumption: A randomized controlled trial. *Nutr. J.* **2008**, *7*, 18. [CrossRef]
60. Bohlscheid-Thomas, S.; Hoting, I.; Boeing, H.; Wahrendorf, J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int. J. Epidemiol.* **1997**, *26* (Suppl. S1), S59–S70. [CrossRef]
61. Haftenberger, M.; Heuer, T.; Heidemann, C.; Kube, F.; Krems, C.; Mensink, G.B. Relative validation of a food frequency questionnaire for national health and nutrition monitoring. *Nutr. J.* **2010**, *9*, 36. [CrossRef]
62. Harmouche-Karaki, M.; Mahfouz, M.; Obeyd, J.; Salameh, P.; Mahfouz, Y.; Helou, K. Development and validation of a quantitative food frequency questionnaire to assess dietary intake among Lebanese adults. *Nutr. J.* **2020**, *19*, 65. [CrossRef]
63. Sierra-Ruelas, E.; Bernal-Orozco, M.F.; Macedo-Ojeda, G.; Marquez-Sandoval, Y.F.; Altamirano-Martinez, M.B.; Vizmanos, B. Validation of semiquantitative FFQ administered to adults: A systematic review. *Public Health Nutr.* **2020**, *24*, 3399–3418. [CrossRef] [PubMed]
64. Bokulich, N.A.; Chung, J.; Battaglia, T.; Henderson, N.; Jay, M.; Li, H. Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **2016**, *8*, 343ra382. [CrossRef]
65. Dominguez-Bello, M.G.; Costello, E.K.; Contreras, M.; Magris, M.; Hidalgo, G.; Fierer, N.; Knight, R. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11971–11975. [CrossRef] [PubMed]
66. Jakobsson, H.E.; Abrahamsson, T.R.; Jenmalm, M.C.; Harris, K.; Quince, C.; Jernberg, C.; Bjorksten, B.; Engstrand, L.; Andersson, A.F. Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* **2014**, *63*, 559–566. [CrossRef] [PubMed]
67. Busi, S.B.; De Nies, L.; Habier, J.; Wampach, L.; Fritz, J.V.; Heintz-Buschart, A.; May, P.; Halder, R.; de Beaufort, C.; Wilmes, P. Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. *ISME Commun.* **2021**, *1*, 8. [CrossRef]
68. De la Cuesta-Zuluaga, J.; Kelley, S.T.; Chen, Y.; Escobar, J.S.; Mueller, N.T.; Ley, R.E.; McDonald, D.; Huang, S.; Swafford, A.D.; Knight, R.; et al. Age- and Sex-Dependent Patterns of Gut Microbial Diversity in Human Adults. *MSystems* **2019**, *4*, e00261-19. [CrossRef] [PubMed]
69. Le Chatelier, E.; Nielsen, T.; Qin, J.; Prifti, E.; Hildebrand, F.; Falony, G.; Almeida, M.; Arumugam, M.; Batto, J.M.; Kennedy, S.; et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* **2013**, *500*, 541–546. [CrossRef]
70. Serrano, J.; Smith, K.R.; Crouch, A.L.; Sharma, V.; Yi, F.; Vargova, V.; LaMoia, T.E.; Dupont, L.M.; Serna, V.; Tang, F.; et al. High-dose saccharin supplementation does not induce gut microbiota changes or glucose intolerance in healthy humans and mice. *Microbiome* **2021**, *9*, 11. [CrossRef] [PubMed]
71. Suez, J.; Korem, T.; Zeevi, D.; Zilberman-Schapira, G.; Thaiss, C.A.; Maza, O.; Israeli, D.; Zmora, N.; Gilad, S.; Weinberger, A.; et al. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* **2014**, *514*, 181–186. [CrossRef] [PubMed]
72. Makki, K.; Deehan, E.C.; Walter, J.; Backhed, F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell Host Microbe* **2018**, *23*, 705–715. [CrossRef]
73. Sonnenburg, E.D.; Smits, S.A.; Tikhonov, M.; Higginbottom, S.K.; Wingreen, N.S.; Sonnenburg, J.L. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **2016**, *529*, 212–215. [CrossRef] [PubMed]
74. Lin, D.; Peters, B.A.; Friedlander, C.; Freiman, H.J.; Goedert, J.J.; Sinha, R.; Miller, G.; Bernstein, M.A.; Hayes, R.B.; Ahn, J. Association of dietary fibre intake and gut microbiota in adults. *Br. J. Nutr.* **2018**, *120*, 1014–1022. [CrossRef]
75. Carmody, R.N.; Bisanz, J.E.; Bowen, B.P.; Maurice, C.F.; Lyalina, S.; Louie, K.B.; Treen, D.; Chadaideh, K.S.; Maini Rekdal, V.; Bess, E.N.; et al. Cooking shapes the structure and function of the gut microbiome. *Nat. Microbiol.* **2019**, *4*, 2052–2063. [CrossRef]

ANNEX 10. GRANTS

This work was funded by the National Agency for Research and Development (ANID) / Scholarship Program / DOCTORADO BECAS CHILE/2018 - 72190278.

UAB
Universitat Autònoma
de Barcelona

Vall d'Hebron
Institut de Recerca

Agencia
Nacional de
Investigación
y Desarrollo