NURECON: A Novel Online System for Determining Nutrition Requirements Based on Microbial Composition

Zhao-Qi Hu, Yuan-Mao Hung, Li-Han Chen, Liang-Chuan Lai, Min-Hsiung Pang, Eric Y. Chuang and Mong-Hsun Tsai

***Abstract*—Dietary habits have been proven to have an impact on the microbial composition and health of the human gut. Over the past decade, researchers have discovered that gut microbiota can use nutrients to produce metabolites that have major implications for human physiology. However, there is no comprehensive system that specifically focuses on identifying nutrient deficiencies based on gut microbiota, making it difficult to interpret and compare gut microbiome data in the literature. This study proposes an analytical platform, NURECON, that can predict nutrient deficiency information in individuals by comparing their metagenomic information to a reference baseline. NURECON integrates a next-generation bacterial 16S rRNA analytical pipeline (QIIME2), metabolic pathway prediction tools (PICRUSt2 and KEGG), and a food compound database (FooDB) to enable the identification of missing nutrients and provide personalized dietary suggestions. Metagenomic information from total number of 287 healthy subjects was used to establish baseline microbial composition and metabolic profiles. The uploaded data is analyzed and compared to the baseline for nutrient deficiency assessment. Visualization results include gut microbial composition, related enzymes, pathways, and nutrient abundance. NURECON is a user-friendly online platform that provides nutritional advice to support dietitians' research or menu design. The system is freely available at https://nurecon.cgm.ntu.edu.tw/.**

***Index Terms*—Metagenomics, 16S rRNA, Web-based System, User-friendly Design, Precision Nutrition, Database**

Z. Hu is with the Institute of Biotechnology and Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan.

Y. Hung. is with the Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan.

L. Chen is with the Institute of Fisheries Science and Departmemt of Life Science, National Taiwan University, Taipei, Taiwan.

L. Lai is with the Graduate Institute of Physiology, College of Medicine and Bioinformatics and Biostatistics Core, Center of Genomic and Precision Medicine, National Taiwan University, Taipei, Taiwan.

M. Pang is with the Institute of Food Sciences and Technology, National Taiwan University, Taipei, Taiwan.

E. Chuang is with the Graduate Institute of Biomedical Electronics and Bioinformatics, Taipei, Taiwan and Bioinformatics and Biostatistics Core, Center of Genomic and Precision Medicine, National Taiwan University, Taipei, Taiwan, and also with the Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan.

M. Tsai is with the Institute of Biotechnology, National Taiwan University, Taipei, Taiwan.

* Z. Hu and Y. Hung are authors with equal contribution.

I. Introduction

WITH the rapid advancement of next-generation sequencing (NGS) technologies over the past few decades, NGS has been applied to many metagenomic research samples such as the human intestine, soil, water, air, and microbiota from different locations [1]–[3]. The prokaryotic 16S ribosomal RNA (rRNA) gene fragment (1500 bp) encompasses nine hypervariable regions (V1∼V9), which are rich in taxonomic information. Therefore, prokaryotic 16S sequences are suitable for elucidating microbial profiles [4]. Different hypervariable regions provide different classification accuracy. For NGS short read sequences, the V4 region performs best in phylum-level assignment [5]. However, a single hypervariable region is too short to have genus-level resolution [6]. Therefore, regions V3∼V4 or V3∼V5 have been widely applied to ensure correct assignment at the genus level [7], [8].

Many studies have shown that the human gut microbiota play an important role in human health. Gut microbial composition is largely influenced by our daily diet, medication use, probiotic intake, and exercise [9]–[11]. Gut microbial dysbiosis is also closely associated with immunity and with diseases including inflammatory bowel disease, obesity, type II diabetes, and hypertension [12]–[15]. Previous studies showed that a high-fat diet in diabetic patients increased the proportion of lipopolysaccharide-related bacteria in the human gut, thereby damaging the intestinal mucosal barrier and impairing the immune response [15].

The nutrient intake in our daily diet shapes our gut microbial composition [16], and the metabolites produced by the gut microbiome have a substantial impact on our digestion and endocrinology [17]. For example, complex carbohydrates in the intestine are broken down mainly through the fermentation process of the gut microbiota, involving species such as *Bacteroides thetaiotaomicron* and *Anaerostipes caccae* [18], [19]. *B. thetaiotaomicron* encodes a repertoire of carbohydrate-degrading genes that allow the bacteria to switch between host-derived and dietary carbohydrates [18]. *A. caccae*, a butyrate-producing bacterium, plays a key role in regulating food allergies early in life. Its growth is supported by *B. thetaiotaomicron*, especially in the presence of carbohydrates. The fermentation process produces short-chain fatty acids (SCFAs)

such as acetate, propionate, butyrate, and methane [20], [21]. SCFAs are involved in the differentiation of epithelial cells, mucosal integrity, and the reduction of inflammatory symptoms [22], [23]. In addition, excess intake of animal proteins can increase the genera *Bacteroides* and *Alistipes*, which increase the level of trimethylamine N-oxide metabolites [24], [25]. Trimethylamine N-oxide is one of the factors causing proatherogenesis [26], and stimulates sulfate-reducing microorganisms to produce more sulfated amino acid metabolites, which could potentially induce inflammatory bowel disease and colorectal cancer [27], [28].

Recently, researchers also found that probiotics could stabilize beneficial microorganisms in the human intestine and help lower blood sugar and cholesterol in obese patients [29]. Human gut microorganisms such as *Bifidobacteria* and lactic acid-producing bacteria can utilize indigestible prebiotic ingredients such as β-glucans, pectin, and insulin to promote SCFA formation. [30]–[32]. Previous studies attempted to improve gut microbiota through prebiotics and dietary fiber [31], [33]. In the field of precision nutrition research, blood sugar and dietary control were used to improve nutrition and glycemic control [34]–[36].

Bioinformatic tools are widely used to analyze microbiome data, including QIIME2, a 16S analytical pipeline [37]; EasyMAP, an online graphical user interface website [38]; SPINGO, a high-resolution sequence classifier [39]; PICRUSt2, a microbiome gene function predictor [40]; Kyoto Encyclopedia of Genes and Genomes (KEGG), a metabolic pathway database [41]; and microbial 16S sequence databases such as SILVA, Greengenes, and RDP [42]–[44]. These tools support researchers in performing advanced analysis and provide diverse perspectives for microbiome research. Although many studies have been published in the research field of nutrition and metagenomics [45], there is still a lack of integrated tools for illustrating nutrient deficiencies based on gut microbial metabolism.

In this study, we describe a user-friendly online analytical platform, NURECON (dertermining NUtrition REquirements based on microbial COmpositioN), which was designed to predict nutrient and metabolic pathway deficiencies based on gut microbial composition by comparison with a healthy gut microbial baseline. The FooDB database was integrated to provide food recommendations to supply the deficient nutrients in the human gut. To build a system capable of providing dietary recommendations, the first step was to analyze the metabolites produced by gut microbiota. The gut microbial profile could be used to predict metabolic pathways for each subject. The pathway difference was then calculated by comparing the predicted pathways to the healthy baseline (established in the system). The differences in the metabolic pathway indicated the potentially deficient nutrients, which were mapped to the nutrition database to identify foods that provide the deficient nutrients and make dietary recommendations. NURECON does not require users to register an account. The only required information is a user email address, which is only used for identification purpose. Notice that NURECON does not send any email to users. The online system is freely available at https://nurecon.cgm.ntu.edu.tw/

II. Methods

A. System implementation

NURECON's backend was developed based on Python (version 3.7.12) and Django (version 3.2.16), while JavaScript and jQuery (version 3.6.0) were used to develop the graphical user interface. SQLite (version 3.33.0) was used to store the baseline gut microbial profiles, enzyme abundance information, KEGG data (compound number, enzyme number, and name list) [41], and FooDB data (compound number, compound name, synonym, and food characteristic data) [46], [47]. NURECON identified the deficient nutrients and recommended foods based on users' sequence data. The design framework is shown in (Fig. 1). Two input files are required to initiate the analysis: a biological observation matrix file (.biom) and a DNA sequence alignment file (in FASTA format). The biological observation matrix file contains the operational taxonomic unit or amplicon sequence variant (ASV) information, while the DNA sequence alignment file contains the representative sequence information. These two files can be obtained from EasyMAP [38] or QIIME2 [37]. Users can view and download their analysis results or remove the projects after completing their analysis.

B. Built-in bioinformatic software and parameters

SPINGO (version 1.2) [39] and PICRUSt2 [40] were incorporated into the system for microbial composition analysis, enzyme and pathway prediction. All default parameters are listed below.

1) Taxonomy assignment: SPINGO is a classifier that can accurately assign 16S rRNA sequences to the species level [39], [48]. SPINGO was used for sequence analysis. The setting for kmer size (--kmersize) was 8. The number of bootstrap samples (--bootstrap) was 10. The number of kmers used for each bootstrap subsample (--subsample) was 8. The reference sequence database provided by SPINGO (RDP_11.2) was used for taxonomy assignment tasks.

PICRUSt2 (version 2.5.1) [40] was applied to analyze bacterial enzymes and pathways. The workflow includes sequence placement (place_seqs.py), hidden-state prediction (hsp.py), metagenome prediction (metagenome_pipeline.py), inferring pathway abundance (pathway_pipeline.py), and pathway feature descriptions (add_descriptions.py). Default parameters were applied for all steps.

2) Optimize nutrients supplement - Binary Integer Linear Programming: In order to use the fewest number of foods to meet nutrient needs, an optimization method was applied to address the problem. Because food categories and the deficient nutrients were many-to-many relationships, we used binary integer linear programming (ILP)

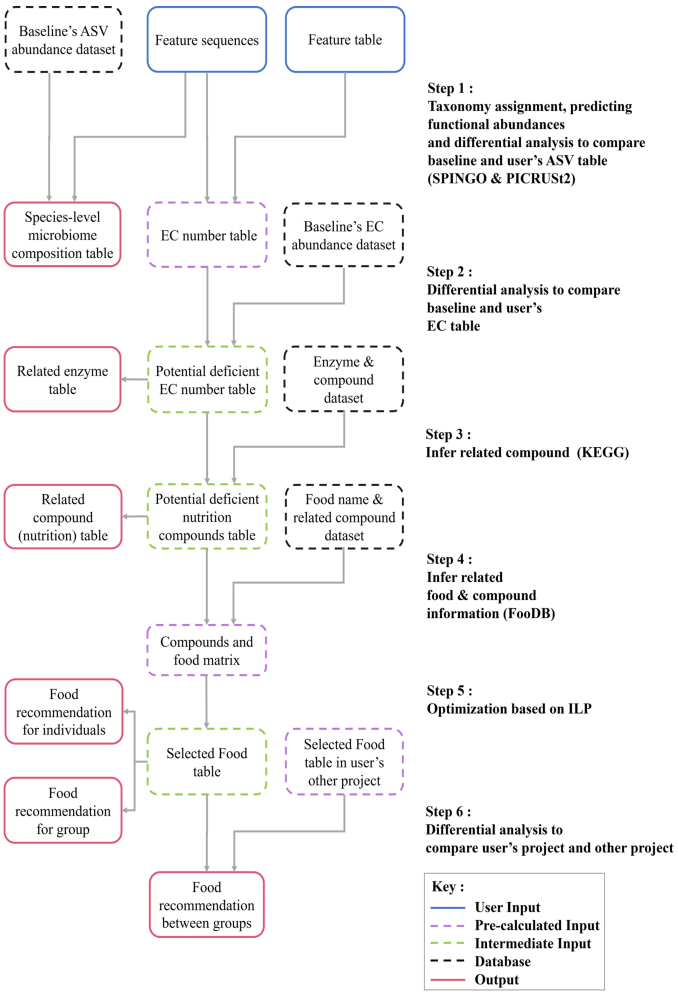


Fig. 1. The design framework of NURECON. The analytical procedure of NURECON consists of taxonomy assignment, metabolic pathway abundance prediction, nutrient coverage optimization, and mutiple differential analysis (compared to the healthy people in the baseline). The inputs are an ASV feature table (.biom) and a representative sequence file in FASTA format (.fna). The output includes a species-level microbial composition table, an enzyme and compound table, and food recommendations.

for optimization. PuLP (version 2.6.0) was used for the binary ILP calculation. The connection between food and deficient nutrients was initiated with a matrix. An example are shown in (1) and (Fig. 2).

$$A = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{array} \begin{array}{c} \begin{array}{ccccc} F_1 & F_2 & F_3 & F_4 & F_5 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \begin{array}{l} F_1 + F_3 \geq 1 \\ F_1 + F_2 \geq 1 \\ F_2 + F_3 + F_4 \geq 1 \\ F_5 \geq 1 \\ F_2 + F_3 \geq 1 \\ F_4 \geq 1 \end{array} \tag{1}$$

Each type of food could supply one or more nutrients, which meant that one food could cover multiple deficient nutrient compounds. The nutrients provided by the corresponding food were marked as "1" in the matrix, and "0" for uncovered compounds. For example, in (1), food F_2 provides the deficient nutrients C_2, C_3, and C_5. Therefore,

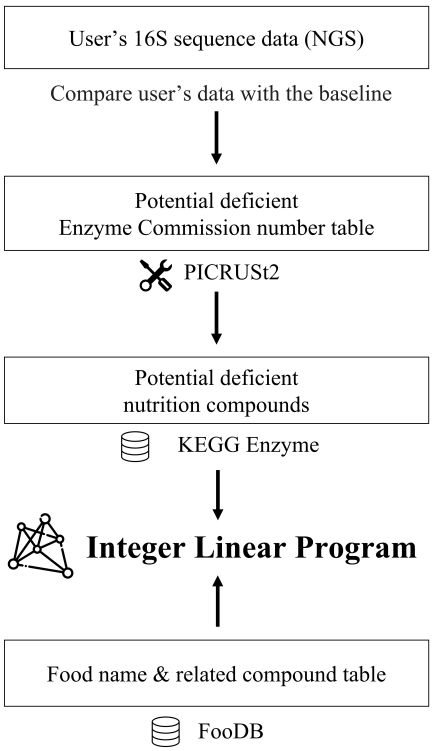


Fig. 2. The integration workflow of the PICRUSt2 enzyme commission table (EC table), the KEGG enzyme database, and the FoodDB database.

in column F_2, the corresponding positions of C_2, C_3, and C_5 are marked as "1", while C_1, C_4, and C_6 are marked "0". The goal was to cover all deficient nutrient compounds with the fewest number of foods. Each deficient compound must be covered by at least one food item, which was also the main constraint of the optimization problem, as shown in (2, 3).

$$min \quad \sum_i F_i, F_i \in \{0, 1\} \tag{2}$$

$$s.t. \quad A \cdot F \geq \quad 1 \tag{3}$$

F_i is a binary vector representing all potential foods that could be used to supply the deficient nutrient compound. i is the index of the food candidate. In the constraint, A is an $M \times N$ matrix that describes the coverage relationship between the candidate foods and the deficient nutrient compounds. M is the number of foods, while N is the number of nutrient compounds. F is a binary vector that indicates whether the food will be selected for nutrient supply. The constraint must be at least 1, meaning that each deficient nutrient is covered with at least one food.

C. Reference database construction for binary ILP optimization

Several databases were integrated into NURECON's analytical pipeline (Fig. 2). We collected 5,236 enzymes and 4,106 compounds from the KEGG enzyme database [41]. To link the analytical results of PICRUSt2 to

KEGG, PICRUSt2's enzyme commission (EC) numbers were mapped to the KEGG database. The EC abundance was compared to the selected baseline (described in the next section). Next, the absent enzymes, which were defined as being deficient, were linked to the corresponding compound ID, allowing NURECON to find the deficient nutrient compounds. The deficient compounds were used to build the binary ILP matrix (the rows in (1). In addition, each food in the FooDB [46] has its own ID and a corresponding nutrient composition. We filtered the foods with repeated names or with definitions that were too broad (e.g., bison, wild boar, buffalo, mule deer, mallard duck, elk, emu, greylag goose, Guinea hen, horse, ostrich, velvet duck, pheasant, European rabbit, and squab) from FooDB because these general food terms related to almost all nutrients, and it is impractical to consume a whole animal for nutrient supply. The remaining foods of FooDB were selected to initialize the ILP matrix (the columns in (1). There were 962 foods and 23,000 compounds in FooDB available for optimization. The concept of database integration is shown in (Fig. 2).

D. Compare user data with the baseline

The baseline data was provided by 287 healthy Taiwanese adults (18-80 years old). Among this population, 119 subjects were from the Taiwanese Microbiome Database (http://twnbiome.cgm.ntu.edu.tw/) [49], which focused on providing samples with healthy weight status. Individuals with a body mass index (BMI) below 18.5 or above 27 were excluded. Further exclusion criteria included people with autoimmune, neurological, or gastrointestinal diseases, cancer, or people who had taken antibiotics, antihypertensives, hypolipidemic agents, steroids, antipsychotics, or gastrointestinal drugs. 118 subjects were from the control group of a published colorectal cancer (CRC) research project [50], and 50 healthy Taiwanese subjects data were from a published research project [51]. All samples were sequenced by an Illumina platform, using 16S V3∼V4 regions. Notice that the users can select different subgroups from the aforementioned control groups based on age interval and gender.

Users of NURECON can obtain ASV, EC, and pathway abundance tables after analyzing their uploaded data. Each User's EC and pathway abundance tables are subtracted from the median value of the baseline's EC and pathway abundance tables, respectively. A negative value for an enzyme or pathway indicates potential deficiency, which means that the deficient enzyme or pathway is included in the subsequent analysis. The built-in baseline included 2,184, 2,064, and 1,978 enzymes; 375, 398, and 416 pathways from Taiwanese Microbiome Database [49], colorectal cancer control group [50], and 50 healthy Taiwanese subjects [51], respectively.

E. Validation datasets

To validate system performance, colorectal cancer (CRC) and chronic kidney disease (CKD) [50], [52] datasets were used as validation datasets, as there are established associations between these two diseases, gut microbiota, and diet [53], [54]. The validation datasets were gut microbial profiles from Taiwanese CRC patients (accession number: SRP131074) and CKD patients (accession number: SRP279052). For the CRC dataset, the control and the patient groups were assigned 118 and 34 samples, respectively. The age distribution was 60.71 ± 10.44 years for the control group and 61.09±10.27 years for the patient group. In a previous study [50], the sequence data was used to analyze the difference in enterotype-based gut microbiome between CRC patients and healthy people. For the CKD dataset, 100 middle-aged or elderly Chinese CKD patients were used for validation. A total of 137 samples were allocated to the control group. The age distribution was 60.64 ± 16.51 years for the control group and 56.64 ± 17.25 years for the patient group. In a previous study [52], the sequence data was used to examine the relationship between gut dysbiosis and serum-free immunoglobulin light chains in CKD. Both the CRC and the CKD datasets were compared to the 50 healthy Taiwanese subjects dataset (baseline).

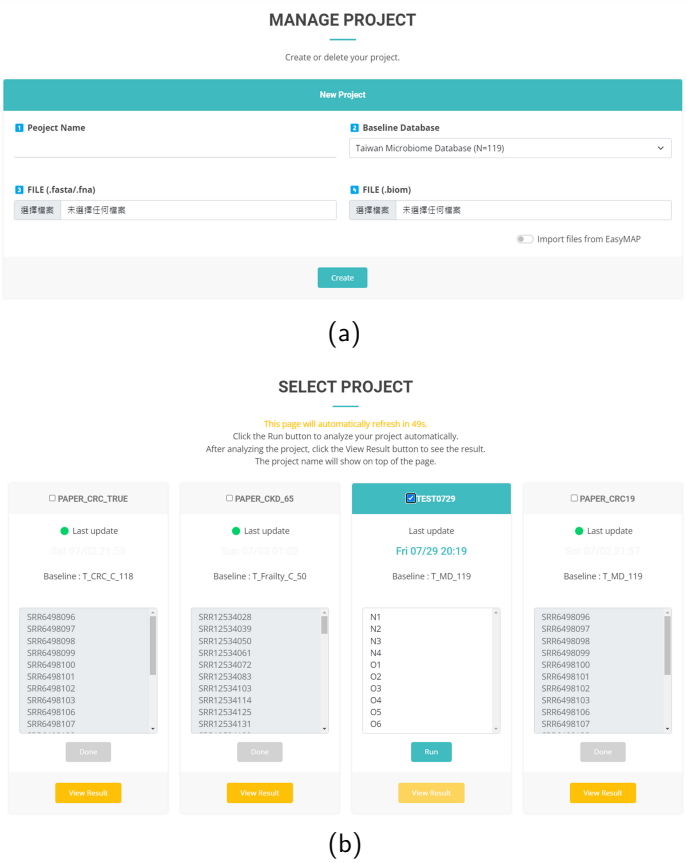


Fig. 3. (a) The data uploading page. (b) The project selection page. This page allows users to select the project for analysis. The uploaded sample information is displayed in the corresponding project panel. Users can start their analysis by simply clicking the "RUN" button.

Fig. 4. Results of the analyzed microbial profile compared to the baseline. (a) Forest plot. The horizontal axis indicates species abundance (%), while the vertical axis lists species names. The column on the right-hand side shows the sample name ("N2" in this case.) with the individual abundance (%) and the abundance distribution from baseline database (median ± 1 standard deviation (SD). The median ± 1 SD column includes two numbers in the parantheses. The left number indicates "median − 1 SD", while the right number means "median + 1 SD". Notice that the left number need to be ≥ 0 since negative values are meaningless to abundance. Therefore, the negative values of the left number are replaced with zero. The black line represents the median ± 1 SD distribution in the baseline database. The black dot represents abundance value of the corresponding sample.) (b) The sorted top 10, 20, or 30 abundant species. Selecting the 10, 20, and 30 tab indexes at the top of the page will sort the table based on microbiome abundance value in the baseline database. In the table, "Low" means the species abundance of user's data is below the median of the baseline ± 1 SD, while "High" means it is above. "Normal" indicates that the species abundance of the user's data is in the range of the baseline control group. "--" means the species is missing from the user's data. (c) Pie chart diagrams provide the food group ratio based on the food recommendation table in each subject. (d) Circle packing visualization. A larger area shows a larger food group fraction (%).

III. Results

A. User interface

NURECON's graphical user interface is shown in (Fig. 3). On the data uploading page, users can assign the project name, select a baseline database for reference, and upload their ASV table (.biom) and representative sequence file (.fasta or .fna). Users can also import their data from EasyMAP [38] (Fig. 3a). After uploading the analysis data (ASV table and sequence data), users can initiate NURECON's pipeline by simply clicking the "RUN" button. (Fig. 3b).

The results include forest plots, pie charts, and circle packing diagrams, all of which can be downloaded by users (Fig. 4). The forest plot displays the personal microbial profile (Fig. 4a) accompanied by a table listing the top 10, 20, and 30 most commonly found species (Fig. 4b). In (Fig. 4a), each black dot in the forest plot indicates the abundance of each bacterial species, while the black line represents the abundance range for the corresponding species from the baseline database. The pie chart shows the ratio of different food categories based on the food recommendation table for each subject (Fig. 4c). In the circle packing diagram (Fig. 4d), greater area means a bigger food fraction. The circle packing diagram illustrates the subjects' suggested dietary profile. The results of personal nutrition recommendations are presented in (Fig. 5). The analysis results described above include gut microbial profiles, deficient enzymes, deficient nutritional compounds, and food recommendation results.

B. Validation results

For the CRC validation dataset, the microbial profiles of the patient group (N=34) and the control group (N=118) showed that several species of the CRC-associated bacterial genus *Fusobacterium* have higher average abundance in the CRC patient group: *Fusobacterium mortiferum* (CRC: 4.08% and control: 3.13%), *F. nucleatum* (CRC: 0.73% and control: 0.10%) and *Morganella morganii* (CRC: 10.88% and control: 0.31%). Among these, the

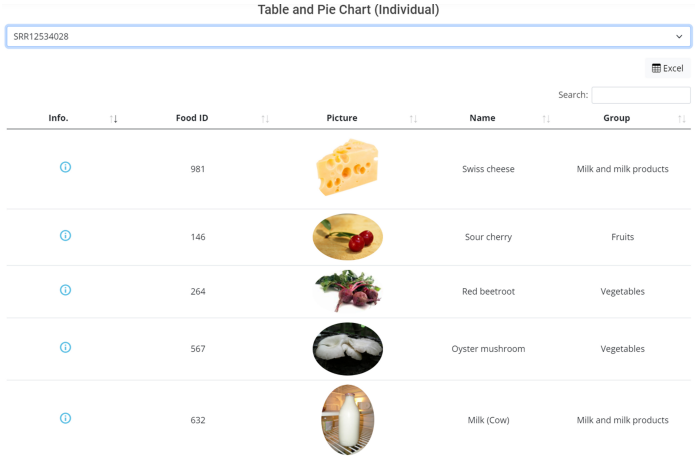


Fig. 5. Results of food recommendation for individuals (each sample). Food recommendation table to supply the deficient nutrients for each individual.

average abundance of *Morganella morganii* in the patient group showed particularly high abundance compared to both the CRC controls and to the 50 healthy Taiwanese subjects (0.637%) (Tables S1-S2). For the CKD validation dataset, the microbial profile in the patient group (N=100) and the control group (N=137) showed that several species in the CKD-associated bacterial genus *Bifidobacterium* have higher average abundance in the CKD group: *Bifidobacterium bifidum* (CKD: 0.99%, control: 0.96%), *B. longum* (CKD: 5.49%, control: 1.98%), *B. dentium* (CKD: 1.43%, control: 0.41%) and *B. breve* (CKD: 3.35%, control: 0.02%) (Tables S3-S4).

To clarify the connection between microbiota, metabolic pathways, and nutrition, NURECON also provides metabolic pathway analysis based on EC results. Compared to baseline, CRC patients have high abundance in tricarboxylic acid (TCA) cycle and glycolysis pathways, such as TCA cycle IV, superpathway of glycolysis, pyruvate dehydrogenase, TCA, and glyoxylate bypass, and the superpathway of glyoxylate bypass and TCA. Some pathways related to the metabolism of lipids and carbohydrates are deficient in CRC patients, such as gondoate biosynthesis, cis-vaccenate biosynthesis, pentose phosphate pathway, starch degradation and pentose phosphate pathway (Fig. 6a). As for CKD metabolic pathway results, the patients exhibit deficiency in fatty acid-related metabolism, including mycolate biosynthesis, oleate biosynthesis IV, stearate biosynthesis II, palmitoleate biosynthesis I and superpathway of fatty acid biosynthesis initiation (Fig. 6b).

In the CRC food recommendation reports, 13 foods were suggested to half or more of the patients, including milk, herbs and spices, garden tomato, red beetroot, cucumber, breakfast cereal, eggs, mung bean, brassicas, cocoa bean, soybean, beer, and blue cheese (Fig. 7a). After classification, dairy products and vegetables were the most commonly recommended food groups (Fig. 7a). By comparing the CRC control group and the patient group, NURECON recommended foods such as brassicas, soybean, red beetroot, green vegetables, breakfast cereal, garden tomato, herbs and spices, blue cheese, almond milk, mozzarella cheese, common buckwheat, green bean, Chinese cinnamon, cucumber, milk (cow), fig, peach, oyster mushroom, mushrooms, black currant, Swiss cheese, fish, grape wine, beer, and Japanese sea bass in the patient group (Fig. 7b). For the control group, NURECON recommended cheddar cheese, clam, mung bean, guava, yellow wax bean, eggs, garden onion, cocoa and cocoa products, cocoa bean, flaxseed, caraway, carrot, rice, red wine, cottage cheese, Ceylon cinnamon, common beet, peanut, coconut milk, bog bilberry, passion fruit, Pacific halibut, corn, fats and oils, common mushrooms, cabbage, nuts, and tart cherry, as shown in (Fig. 7b). In the CKD food recommendation reports, 5 foods such as milk, mung bean, brassicas, eggs, and cucumber were suggested to half or more of the patients (Fig. 7c). After classification, the dairy and vegetable groups were also the most commonly recommended food groups (Fig. 7c). After comparing the control group and the patient group, NURECON recommended foods such as soybean, Japanese sea bass, garden onion, common mushroom, beer, breakfast cereal, eggs, milk, cucumber, mozzarella cheese, mung bean, fishes, French plantain, common buckwheat, yellow wax bean, tamarind, passion fruit, common beet, apple, sour cherry, oat, cottage cheese, coconut milk, German camomile, blue cheese, cocoa bean, brassicas, mushrooms and sweet orange for CKD patients (Fig. 7d). For the control group, NURECON recommended Greek feta cheese, peanut, nuts, fig, flaxseed, sunflower oil, cauliflower, red wine, chicory, calabash, bitter gourd, orange bell pepper, Chinese cinnamon, common pea, pineapple, almond milk, green bean, black currant, rice, potato, white mustard, capers, caraway, parmesan cheese, cocoa and cocoa products, cabbage, guava, green vegetables, Swiss cheese, grape wine, corn, Pacific halibut, clam, oyster mushroom, white wine, cheddar cheese, red beetroot, fats and oils (Fig. 7d).

IV. Discussion

NURECON is a user-friendly online platform for the illustration of nutritional needs based on microbial 16S rRNA sequencing data. The platform supports microbial composition analysis, prediction of microbial gene function, profiling of deficient nutrients, and corresponding dietary recommendations. All results can be viewed and downloaded from the website.

In food recommendation reports for the CRC validation cases, the vegetable class was strongly recommended to the patient group, including brassicas, red beetroot, green vegetables, garden tomato, and cereals, all of which had frequency values above 10%. The brassicas recommendation was 34% higher than in the control group, while soybean reached almost 23% higher (Fig. 7b). The genus brassica includes cauliflower and broccoli, which are rich in glucosinolates and thus have anti-inflammatory effects [55]. Previous studies discovered that human gut microbiomes can utilize glucosinolates to produce sulforaphane, which can prevent CRC [55], [56]. Cereals and soybeans,

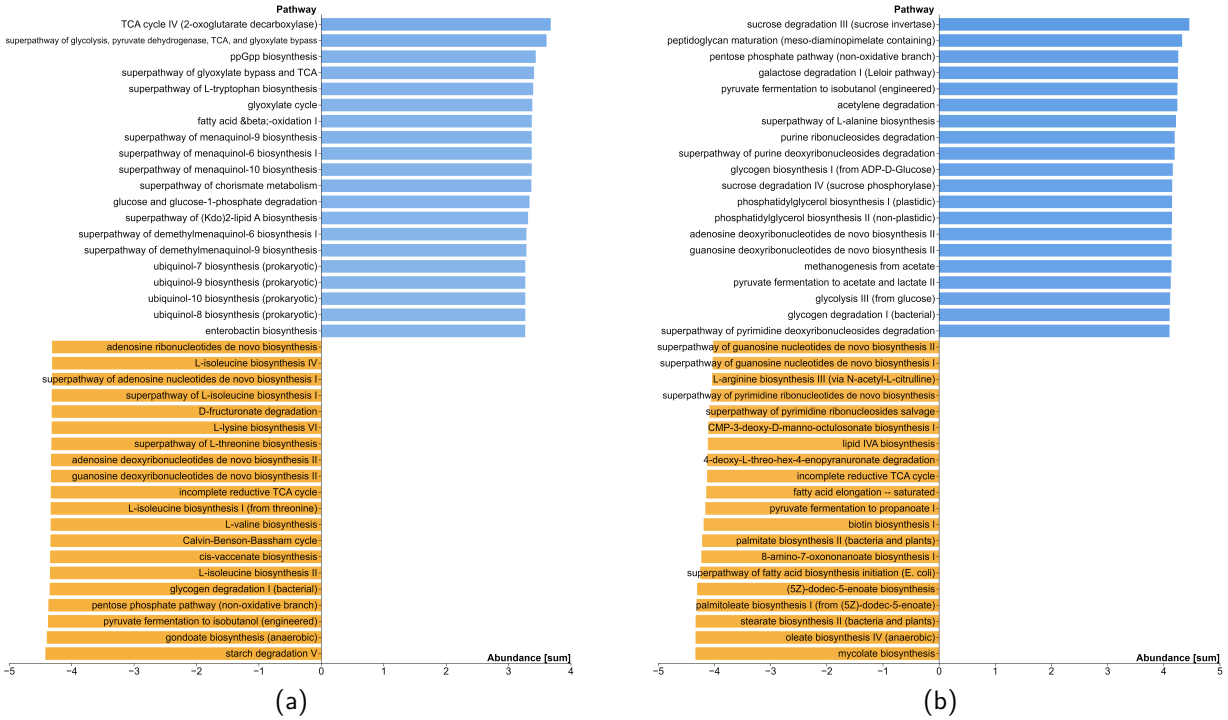


Fig. 6. Results of the top 20 metabolic pathways from (a) CRC and (b) CKD validation datasets. The horizontal axis indicates the pathway abundance difference values (Log_{10}), while the vertical axis lists pathway names. For each pathway, positive values mean that the patient group has higher pathway abundance than the control group (baseline), while negative values indicate that the patient group has lower pathway abundance than the control group (baseline). The 50 healthy Taiwanese subjects baseline was used as the control group in this example.

such as *Avena sativa* L. and *Glycine max* L., are rich in polyunsaturated fatty acids (PUFAs), which includes anti-inflammatory and anticancer agents, and may therefore help prevent bowel disease and colorectal cancer [57]. Breakfast cereal was recommended at 14% in CRC patients, and common buckwheat at almost 8%. In addition, soybean is rich in equol, which can reduce proliferative lesions of the colonic mucosa [58].

For the CKD validation cases, high-frequency proportions of soybean (15%), Japanese sea bass (12%), garden onion (12%), and mushroom (11%) were observed in the patient group (Fig. 7d). Because CKD patients are subject to dietary restrictions, nutrients containing PUFAs, especially eicosatetraenoic acid (ETA) and docosahexaenoic acid (DHA), could slow the progression of kidney damage and are therefore considered beneficial nutrients [59], [60]. Bass fish with high DHA/EPA and less connective tissue is easier to eat and chew for older patients, expanding dietary options for CKD patients. Soybeans are also rich in PUFAs and in isoflavones, which may delay renal tubule expansion and reduce interstitial inflammation [61]. In addition, there are seven types of fruit (cucumber, French plantain, tamarind, passion fruit, apple, sour cherry, sweet orange) that appear in (Fig. 7d). These fruits belong to a group of alkaline-forming foods that could help to balance blood acidity [62]. CKD patients have some limitations in potassium and phosphorus intake due to dysbiosis. All CKD patients in the validation datasets were found to have dysbiosis after controlling for confounders [52]. However, the CKD results included foods or drinks that might be prohibited, such as beer. Although beer is a type of fermented beverage with 279 compounds listed in the FooDB database, this recommendation should be evaluated with care for patients with specific conditions. We also observed that the results of both validation datasets included foods rich in PUFAs, which are commonly found in gut microbiome regulatory processes and in chemopreventive action to improve dysbiosis, regulation of tumor necrosis factor, and COX-2 enzyme function [63].

In the pathway analysis of CRC and CKD validation cases (Fig. 6a and 6b), the results showed that the absence of certain lipid and fatty acid metabolic pathways is consistent with the dietary recommendations for CRC and CKD patients (consume PUFAs). This could support anti-inflammatory responses and support normal lipid metabolism [57].

In gut microbiome results, for the CRC validation cases, the genera *Fusobacterium* and *Morganella* showed high abundance in the gut microbial community in patients with dysbiotic traits. Also, the genotoxic indolamines produced by *Morganella morganii* have been shown to be closely related to CRC [64]. The trend towards high abundance of *Fusobacterium* and *Morganella* was consistent with a previous study [50]. For the CKD validation cases, genus *Bifidobacterium* showed higher abundance in the patient group. High levels of *Bifidobacterium* have been

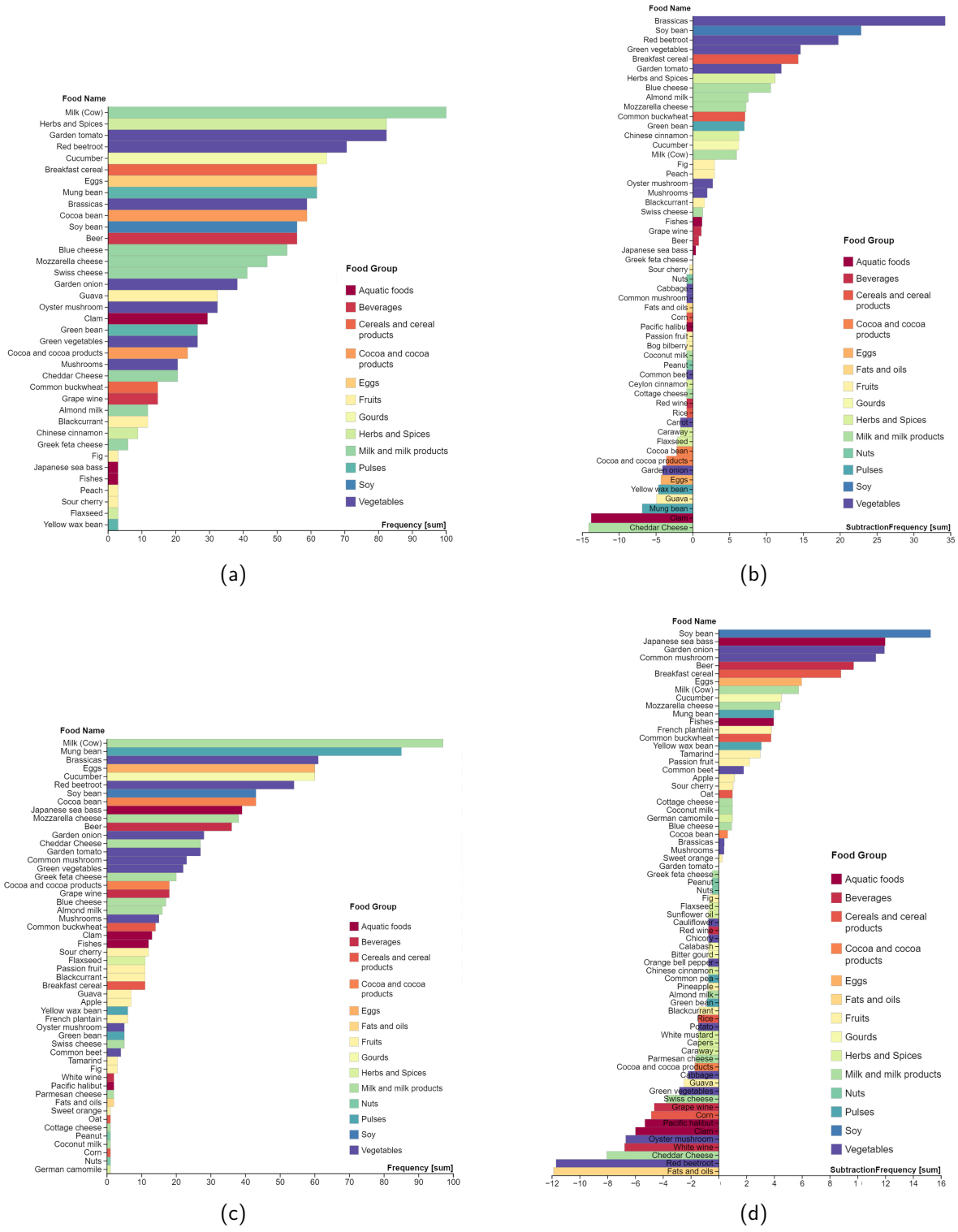


Fig. 7. Validation results of CRC and CKD datasets. Bar chart for the recommended foods groups. (a, c) The horizontal axis is the frequency of the recommended food (%), while the vertical axis lists the food name. Different color annotations indicate different food categories. The bar chart analysis takes into account all user's samples. (b, d) The horizontal axis is the food difference, while the vertical axis lists food names. For each food in the bar chart, positive values mean that the patient group needs more of that food to provide the missing nutrients than the control group. Conversely, negative values indicate that the patient group requires less of the corresponding food than the control group. Zero means that the food in question was suggested to the patient group and control group with equal frequency. The 50 healthy Taiwanese subjects baseline control group was used in this example for comparison to the CRC and CKD datasets.

observed to be a protective response in CKD patients [52].

The comparison of features for NGS analysis and visualization tools is presented in Table I. There are some limitations to the NURECON platform. First, the deficient

TABLE I
COMPARISON OF FEATURES FOR NGS ANALYSIS AND VISUALIZATION TOOLS

Categories	Features	NURECON	EasyMAP	QIIME2	PICRUSt2
Basic information	Interface	Web	Web	CLI, API	CLI
	Backend Language	Python, JavaScript	Python, JavaScript	Python	Python, R
	Tutorial Language	Chinese, English	English	English	English
Preprocessing	Input sequence data format	biom, fasta	fastq	fastq	biom, fasta
	Import sequence data from other tool	✓			
Analysis	One-step button/command	✓			✓
	Built-in baseline database	✓			
	Demultiplex		✓	✓	
	Quality control		✓	✓	
	Taxonomic assignment	✓	✓	✓	
	Alpha/Beta diversity		✓	✓	
	Pathway prediction	✓	✓	✓	✓
	Metabolite and nutrition analysis	✓			
	Food recommendation analysis	✓			
	Differential analysis between projects	✓			
Visualization and results	Taxonomic composition table	✓	✓	✓	
	Taxonomic bar chart		✓	✓	
	Taxonomic forest plot	✓			
	Taxonomic Top10/20/30 table	✓			
	Enzyme Commission number table	✓		✓	✓
	Pathway table	✓	✓	✓	✓
	Pathway bar chart	✓	✓		
	Pathway difference table	✓			
	Pathway difference bar chart	✓			
	Metabolite and nutrition table	✓			
	Food recommendation table	✓			
	Food recommendation pie chart	✓			
	Food recommendation bar plot	✓			
	Food recommendation circle packing	✓			
	Food frequency difference bar chart	✓			

CLI, Command-Line Interface; API, Application Programming Interface.

nutrients illustration is based on qualitative measurement, because the interaction between the gut microbiota and dietary nutrients is complicated and it is difficult to establish an accurate quantitative model [34]. In the future, if a quantitative model is proposed, the system will have the opportunity to be improved and provide more precise results based on better optimization methods. Second, shotgun metagenomics can profile different types of microorganisms. However, shotgun data requires considerable disk space, memory, and CPU threads to complete the analysis in a reasonable time. The current version of NURECON is specifically designed for analyzing 16S metabarcoding data due to limited computing resources. As we expand our computing resources in the future, it is possible to include shotgun data in future updates. Third, the recommended foods are based on comparing metabolites. Patients may avoid eating certain foods due to illness, allergies, and genetic factors. This could be the reason why the patients are deficient in certain nutrients. Therefore, dietitians need to evaluate the recommendation report in light of the patient's overall health when using NURECON to design their menu. In this study, the validation datasets consisted of subjects with CRC and CKD and were selected due to their similarity in age (middle-aged and elderly subjects) and ethnic background (Asians) to those in the reference baseline. It is expected that analytical performance with different age groups could be further improved by including more samples in the future. Furthermore, in recent years, there have been many studies that proposed machine learning or deep learning models to predict CRC based on microbial profiles. The accuracy and AUC of the models are in the range of 0.54 to 0.89 [65], which is not very stable due to the limitation of training sample size from different races or countries. Taking disease prediction models into consideration is a possible application in future updates.

V. CONCLUSION

NURECON is a novel online system that could help researchers and dietitians identify potentially deficient nutrient compounds in the human gut based on microbial 16S metabarcoding data. The validation results showed that ingesting foods rich in PUFAs could help improve the health of CRC and CKD patients, underscoring the value of this system. This platform aims to facilitate the application of precision nutrition.

ACKNOWLEDGMENTS

This work was supported in part by the Center of Genomic and Precision Medicine, National Taiwan University, the Center of Biotechnology, National Taiwan University, Taiwan (grant number: GTZ300), and National

Taiwan University (grant number NTU-CC-112L891406). We thank Melissa Stauffer of Scientific Editing Solutions for editorial assistance.

References

[1] I. Allali, R. E. Abotsi, L. A. Tow, L. Thabane, H. J. Zar, N. M. Mulder, and M. P. Nicol, "Human microbiota research in Africa: a systematic review reveals gaps and priorities for future research," *Microbiome*, vol. 9, no. 1, p. 241, 2021, doi: https://doi.org/10.1186/s40168-021-01195-7.

[2] S. L. McLellan, J. C. Fisher, and R. J. Newton, "The microbiome of urban waters," *International microbiology: the official journal of the Spanish society for microbiology*, vol. 18, no. 3, p. 141, 2015, doi: https://doi.org/10.2436%2F20.1501.01.244.

[3] X. Ding, W. Lan, and J.-D. Gu, "A review on sampling techniques and analytical methods for microbiota of cultural properties and historical architecture," *Applied Sciences*, vol. 10, no. 22, p. 8099, 2020, doi: https://doi.org/10.3390/app10228099.

[4] Y. S. Bukin, Y. P. Galachyants, I. Morozov, S. Bukin, A. Zakharenko, and T. Zemskaya, "The effect of 16S rRNA region choice on bacterial community metabarcoding results," *Scientific Data*, vol. 6, no. 1, pp. 1–14, 2019, doi: https://doi.org/10.1038/sdata.2019.7.

[5] B. Yang, Y. Wang, and P.-Y. Qian, "Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis," *BMC bioinformatics*, vol. 17, no. 1, pp. 1–8, 2016, doi: https://doi.org/10.1186/s12859-016-0992-y.

[6] J. S. Johnson, D. J. Spakowicz, B.-Y. Hong, L. M. Petersen, P. Demkowicz, L. Chen, S. R. Leopold, B. M. Hanson, H. O. Agresta, M. Gerstein *et al.*, "Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis," *Nature communications*, vol. 10, no. 1, p. 5029, 2019, doi: https://doi.org/10.1038/s41467-019-13036-1.

[7] M.-A. Osman, H.-m. Neoh, N.-S. Ab Mutalib, S.-F. Chin, and R. Jamal, "16S rRNA gene sequencing for deciphering the colorectal cancer gut microbiome: current protocols and workflows," *Frontiers in microbiology*, vol. 9, p. 767, 2018, doi: https://doi.org/10.3389/fmicb.2018.00767.

[8] F. de Melo, F. C. Milanesi, P. D. M. Angst, and R. V. Oppermann, "A systematic review of the microbiota composition in various peri-implant conditions: data from 16S rRNA gene sequencing," *Archives of Oral Biology*, vol. 117, p. 104776, 2020, doi: https://doi.org/10.1016/j.archoralbio.2020.104776.

[9] A. E. Mohr, R. Jäger, K. C. Carpenter, C. M. Kerksick, M. Purpura, J. R. Townsend, N. P. West, K. Black, M. Gleeson, D. B. Pyne *et al.*, "The athletic gut microbiota," *Journal of the International Society of Sports Nutrition*, vol. 17, no. 1, p. 24, 2020, doi: https://doi.org/10.1186/s12970-020-00353-w.

[10] S. R. Konstantinov, "Diet, microbiome, and colorectal cancer," *Best Practice & Research Clinical Gastroenterology*, vol. 31, no. 6, pp. 675–681, 2017, doi: https://doi.org/10.1016/j.bpg.2017.09.007.

[11] A. E. Lindell, M. Zimmermann-Kogadeeva, and K. R. Patil, "Multimodal interactions of drugs, natural compounds and pollutants with the gut microbiota," *Nature Reviews Microbiology*, vol. 20, no. 7, pp. 431–443, 2022, doi: https://doi.org/10.1038/s41579-022-00681-5.

[12] J. Li, F. Zhao, Y. Wang, J. Chen, J. Tao, G. Tian, S. Wu, W. Liu, Q. Cui, B. Geng *et al.*, "Gut microbiota dysbiosis contributes to the development of hypertension," *Microbiome*, vol. 5, pp. 1–19, 2017, doi: https://doi.org/10.1186/s40168-016-0222-x.

[13] D. C. Baumgart and S. R. Carding, "Inflammatory bowel disease: cause and immunobiology," *The Lancet*, vol. 369, no. 9573, pp. 1627–1640, 2007, doi: https://doi.org/10.1016/S0140-6736(07)60750-8.

[14] E. Amabebe, F. O. Robert, T. Agbalalah, and E. S. Orubu, "Microbial dysbiosis-induced obesity: role of gut microbiota in homoeostasis of energy metabolism," *British Journal of Nutrition*, vol. 123, no. 10, pp. 1127–1137, 2020, doi: https://doi.org/10.1017/S0007114520000380.

[15] X. Li, K. Watanabe, and I. Kimura, "Gut microbiota dysbiosis drives and implies novel therapeutic strategies for diabetes mellitus and related metabolic diseases," *Frontiers in immunology*, vol. 8, p. 1882, 2017, doi: https://doi.org/10.3389/fimmu.2017.01882.

[16] L. A. Frame, E. Costa, and S. A. Jackson, "Current explorations of nutrition and the gut microbiome: a comprehensive evaluation of the review literature," *Nutrition Reviews*, vol. 78, no. 10, pp. 798–812, 2020, doi: https://doi.org/10.1093/nutrit/nuz106.

[17] J. K. Nicholson, E. Holmes, J. Kinross, R. Burcelin, G. Gibson, W. Jia, and S. Pettersson, "Host-gut microbiota metabolic interactions," *Science*, vol. 336, no. 6086, pp. 1262–1267, 2012, doi: https://doi.org/10.1126/science.1223813.

[18] H. J. Flint, K. P. Scott, S. H. Duncan, P. Louis, and E. Forano, "Microbial degradation of complex carbohydrates in the gut," *Gut microbes*, vol. 3, no. 4, pp. 289–306, 2012, doi: https://doi.org/10.4161/gmic.19897.

[19] L. W. Chia, M. Mank, B. Blijenberg, S. Aalvink, R. S. Bongers, B. Stahl, J. Knol, and C. Belzer, "Bacteroides thetaiotaomicron fosters the growth of butyrate-producing anaerostipes caccae in the presence of lactose and total human milk carbohydrates," *Microorganisms*, vol. 8, no. 10, p. 1513, 2020, doi: https://doi.org/10.3390/microorganisms8101513.

[20] N. T. Baxter, A. W. Schmidt, A. Venkataraman, K. S. Kim, C. Waldron, and T. M. Schmidt, "Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers," *MBio*, vol. 10, no. 1, pp. e02 566–18, 2019, doi: https://doi.org/10.1128/mBio.02566-18.

[21] S. Deleu, K. Machiels, J. Raes, K. Verbeke, and S. Vermeire, "Short chain fatty acids and its producing organisms: An overlooked therapy for IBD?" *EBioMedicine*, vol. 66, p. 103293, 2021, doi: https://doi.org/10.1016/j.ebiom.2021.103293.

[22] T. Tsukahara, Y. Iwasaki, K. Nakayama, and K. Ushida, "Stimulation of butyrate production in the large intestine of weaning piglets by dietary fructooligosaccharides and its influence on the histological variables of the large intestinal mucosa," *Journal of nutritional science and vitaminology*, vol. 49, no. 6, pp. 414–421, 2003, doi: https://doi.org/10.3177/jnsv.49.414.

[23] L. Klampfer, J. Huang, T. Sasazuki, S. Shirasawa, and L. Augenlicht, "Inhibition of interferon γ signaling by the short chain fatty acid butyrate," *Molecular Cancer Research*, vol. 1, no. 11, pp. 855–862, 2003.

[24] L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach *et al.*, "Diet rapidly and reproducibly alters the human gut microbiome," *Nature*, vol. 505, no. 7484, pp. 559–563, 2014, doi: https://doi.org/10.1038/nature12820.

[25] A. M. Sheflin, C. L. Melby, F. Carbonero, and T. L. Weir, "Linking dietary patterns with gut microbial composition and function," *Gut microbes*, vol. 8, no. 2, pp. 113–129, 2017, doi: https://doi.org/10.1080/19490976.2016.1270809.

[26] L. Barrea, G. Annunziata, G. Muscogiuri, D. Laudisio, C. Di Somma, M. Maisto, G. C. Tenore, A. Colao, and S. Savastano, "Trimethylamine N-oxide, Mediterranean diet, and nutrition in healthy, normal-weight adults: also a matter of sex?" *Nutrition*, vol. 62, pp. 7–17, 2019, doi: https://doi.org/10.1016/j.nut.2018.11.015.

[27] P. Jantchou, S. Morois, F. Clavel-Chapelon, M.-C. Boutron-Ruault, and F. Carbonnel, "Animal protein intake and risk of inflammatory bowel disease: The E3N prospective study," *Official journal of the American College of Gastroenterology| ACG*, vol. 105, no. 10, pp. 2195–2201, 2010, doi: https://doi.org/10.1038/ajg.2010.192.

[28] I. Kushkevych, J. Cejnar, J. Treml, D. Dordević, P. Kollar, and M. Vítězová, "Recent advances in metabolic pathways of sulfate reduction in intestinal bacteria," *Cells*, vol. 9, no. 3, p. 698, 2020, doi: https://doi.org/10.3390/cells9030698.

[29] B. T. Beserra, R. Fernandes, V. A. do Rosario, M. C. Mocellin, M. G. Kuntz, and E. B. Trindade, "A systematic review and meta-analysis of the prebiotics and synbiotics effects on glycaemia, insulin concentrations and lipid parameters in adult patients with overweight or obesity," *Clinical Nutrition*, vol. 34, no. 5, pp. 845–858, 2015, doi: https://doi.org/10.1016/j.clnu.2014.10.004.

[30] L. B. Bindels, N. M. Delzenne, P. D. Cani, and J. Walter, "Towards a more comprehensive concept for prebiotics," *Nature reviews Gastroenterology & hepatology*, vol. 12, no. 5, pp. 303–310, 2015, doi: https://doi.org/10.1038/nrgastro.2015.47.

[31] D. Mohanty, S. Misra, S. Mohapatra, and P. S. Sahu, "Prebiotics and synbiotics: Recent concepts in nutrition," *Food bioscience*, vol. 26, pp. 152–160, 2018, doi: https://doi.org/10.1016/j.fbio.2018.10.008.

[32] G. R. Gibson, R. Hutkins, M. E. Sanders, S. L. Prescott, R. A. Reimer, S. J. Salminen, K. Scott, C. Stanton, K. S. Swanson, P. D. Cani *et al.*, "Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics," *Nature reviews Gastroenterology & hepatology*, vol. 14, no. 8, pp. 491–502, 2017, doi: https://doi.org/10.1038/nrgastro.2017.75.

[33] H. D. Holscher, "Dietary fiber and prebiotics and the gastrointestinal microbiota," *Gut microbes*, vol. 8, no. 2, pp. 172–184, 2017, doi: https://doi.org/10.1080/19490976.2017.1290756.

[34] D. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, O. Ben-Yacov, D. Lador, T. Avnit-Sagi, M. Lotan-Pompan *et al.*, "Personalized nutrition by prediction of glycemic responses," *Cell*, vol. 163, no. 5, pp. 1079–1094, 2015, doi: https://doi.org/10.1016/j.cell.2015.11.001.

[35] J. Suez, T. Korem, D. Zeevi, G. Zilberman-Schapira, C. A. Thaiss, O. Maza, D. Israeli, N. Zmora, S. Gilad, A. Weinberger *et al.*, "Artificial sweeteners induce glucose intolerance by altering the gut microbiota," *Nature*, vol. 514, no. 7521, pp. 181–186, 2014, doi: https://doi.org/10.1038/nature13793.

[36] J. Qin, Y. Li, Z. Cai, S. Li, J. Zhu, F. Zhang, S. Liang, W. Zhang, Y. Guan, D. Shen *et al.*, "A metagenome-wide association study of gut microbiota in type 2 diabetes," *Nature*, vol. 490, no. 7418, pp. 55–60, 2012, doi: https://doi.org/10.1038/nature11450.

[37] M. Hall and R. G. Beiko, "16S rRNA gene analysis with QIIME2," *Microbiome analysis: methods and protocols*, pp. 113–129, 2018, doi: https://doi.org/10.1007/978-1-4939-8728-3_8.

[38] Y.-M. Hung, T.-P. Lu, M.-H. Tsai, L.-C. Lai, and E. Y. Chuang, "EasyMAP: A user-friendly online platform for analyzing 16S ribosomal DNA sequencing data," *New Biotechnology*, vol. 63, pp. 37–44, 2021, doi: https://doi.org/10.1016/j.nbt.2021.03.001.

[39] G. Allard, F. J. Ryan, I. B. Jeffery, and M. J. Claesson, "SPINGO: a rapid species-classifier for microbial amplicon sequences," *BMC bioinformatics*, vol. 16, no. 1, pp. 1–8, 2015, doi: https://doi.org/10.1186/s12859-015-0747-1.

[40] G. M. Douglas, V. J. Maffei, J. R. Zaneveld, S. N. Yurgel, J. R. Brown, C. M. Taylor, C. Huttenhower, and M. G. Langille, "PICRUSt2 for prediction of metagenome functions," *Nature biotechnology*, vol. 38, no. 6, pp. 685–688, 2020, doi: https://doi.org/10.1038/s41587-020-0548-6.

[41] M. Kanehisa, "Enzyme annotation and metabolic reconstruction using KEGG," *Protein function prediction: methods and protocols*, pp. 135–145, 2017, doi: https://doi.org/10.1007/978-1-4939-7015-5_11.

[42] C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner, "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools," *Nucleic acids research*, vol. 41, no. D1, pp. D590–D596, 2012, doi: https://doi.org/10.1093/nar/gks1219.

[43] T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen, "Greengenes, a chimera-checked 16s rRNA gene database and workbench compatible with ARB," *Applied and environmental microbiology*, vol. 72, no. 7, pp. 5069–5072, 2006, doi: https://doi.org/10.1128/AEM.03006-05.

[44] B. L. Maidak, G. J. Olsen, N. Larsen, R. Overbeek, M. J. McCaughey, and C. R. Woese, "The RDP (ribosomal database project)," *Nucleic acids research*, vol. 25, no. 1, pp. 109–110, 1997, doi: https://doi.org/10.1093/nar/25.1.109.

[45] R. Krajmalnik-Brown, Z.-E. Ilhan, D.-W. Kang, and J. K. DiBaise, "Effects of gut microbes on nutrient absorption and energy regulation," *Nutrition in clinical practice*, vol. 27, no. 2, pp. 201–214, 2012, doi: https://doi.org/10.1177/0884533611436116.

[46] D. Wishart, "Foodb: the food database," *FooDB version*, vol. 1, 2014.

[47] J. J. Naveja, M. P. Rico-Hidalgo, and J. L. Medina-Franco, "Analysis of a large food chemical database: chemical space, diversity, and complexity," *F1000Research*, vol. 7, 2018, doi: https://doi.org/10.12688/f1000research.15440.1.

[48] Y.-M. Hung, W.-N. Lyu, M.-L. Tsai, C.-L. Liu, L.-C. Lai, M.-H. Tsai, and E. Y. Chuang, "To compare the performance of prokaryotic taxonomy classifiers using curated 16s full-length rRNA sequences," *Computers in Biology and Medicine*, vol. 145, p. 105416, 2022, doi: https://doi.org/10.1016/j.compbiomed.2022.105416.

[49] S. Shivani, C.-Y. Kao, A. Chattopadhyay, J.-W. Chen, L.-C. Lai, W.-H. Lin, T.-P. Lu, I. Huang, M.-H. Tsai, C.-H. Teng *et al.*, "Uremic toxin-producing bacteroides species prevail in the gut microbiota of taiwanese ckd patients: an analysis using the new taiwan microbiome baseline," *Frontiers in Cellular and Infection Microbiology*, vol. 12, p. 414, 2022, doi: https://doi.org/10.3389/fcimb.2022.726256.

[50] T.-W. Yang, W.-H. Lee, S.-J. Tu, W.-C. Huang, H.-M. Chen, T.-H. Sun, M.-C. Tsai, C.-C. Wang, H.-Y. Chen, C.-C. Huang *et al.*, "Enterotype-based analysis of gut microbiota along the conventional adenoma-carcinoma colorectal cancer pathway," *Scientific reports*, vol. 9, no. 1, pp. 1–13, 2019, doi: https://doi.org/10.1038/s41598-019-45588-z.

[51] C.-J. Wen, Y.-C. Koh, Y.-C. Tung, P.-Y. Ho, S.-C. Hsieh, Y.-C. Lo, J.-S. Tsai, and M.-H. Pan, "Cross-sectional study: New approach for diagnostic identification of non-robust older adult," *Molecular Nutrition & Food Research*, vol. 67, no. 13, p. 2300056, 2023, doi: https://doi.org/10.1002/mnfr.202300056.

[52] F. Liu, X. Xu, L. Chao, K. Chen, A. Shao, D. Sun, Y. Hong, R. Hu, P. Jiang, N. Zhang *et al.*, "Alteration of the gut microbiome in chronic kidney disease patients and its association with serum free immunoglobulin light chains," *Frontiers in Immunology*, vol. 12, p. 609700, 2021, doi: https://doi.org/10.3389/fimmu.2021.609700.

[53] C. V. De Almeida, M. R. de Camargo, E. Russo, and A. Amedei, "Role of diet and gut microbiota on colorectal cancer immunomodulation," *World journal of gastroenterology*, vol. 25, no. 2, p. 151, 2019, doi: https://doi.org/10.3748/wjg.v25.i2.151.

[54] P. Mertowska, S. Mertowski, J. Wojnicka, I. Korona-Głowniak, E. Grywalska, A. Błażewicz, and W. Załuska, "A link between chronic kidney disease and gut microbiota in immunological and nutritional aspects," *Nutrients*, vol. 13, no. 10, p. 3637, 2021, doi: https://doi.org/10.3390/nu13103637.

[55] K. Sikorska-Zimny and L. Beneduce, "The metabolism of glucosinolates by gut microbiota," *Nutrients*, vol. 13, no. 8, p. 2750, 2021, doi: https://doi.org/10.3390/nu13082750.

[56] J. M. Macharia, R. W. Mwangi, N. Rozmann, K. Zsolt, T. Varjas, P. O. Uchechukwu, I. N. Wagara, and B. L. Raposa, "Medicinal plants with anti-colorectal cancer bioactive compounds: Potential game-changers in colorectal cancer management," *Biomedicine & Pharmacotherapy*, vol. 153, p. 113383, 2022, doi: https://doi.org/10.1016/j.biopha.2022.113383.

[57] S. Miccadei, R. Masella, A. M. Mileo, and S. Gessani, "ω3 Polyunsaturated fatty acids as immunomodulators in colorectal cancer: new potential role in adjuvant therapies," *Frontiers in immunology*, vol. 7, p. 486, 2016, doi: https://doi.org/10.3389/fimmu.2016.00486.

[58] L. Polimeno, M. Barone, A. Mosca, M. T. Viggiani, F. Joukar, F. Mansour-Ghanaei, S. Mavaddati, A. Daniele, L. Debellis, M. Bilancia *et al.*, "Soy metabolism by gut microbiota from patients with precancerous intestinal lesions," *Microorganisms*, vol. 8, no. 4, p. 469, 2020, doi: https://doi.org/10.3390/microorganisms8040469.

[59] G. Lore, "The NKF releases its K/DOQI nutrition clinical practice guidelines," *Contemporary dialysis and nephrology*, vol. 21, no. 8, pp. 24–27, 2000.

[60] I. Castro-González, A. Maafs-Rodriguez, J. Silencio-Barrita, C. Galindo-Gómez, and F. Pérez-Gil, "Evaluation of the possible inclusion of certain fish species in chronic kidney disease diets based on their adverse and beneficial nutrient ratios," *International Journal of Food Sciences and Nutrition*, vol. 64, no. 1, pp. 82–88, 2013, doi: https://doi.org/10.3109/09637486.2012.700921.

[61] I.-S. Kim, W.-S. Yang, and C.-H. Kim, "Beneficial effects of soybean-derived bioactive peptides," *International Journal of Molecular Sciences*, vol. 22, no. 16, p. 8570, 2021, doi: https://doi.org/10.3390/ijms22168570.

[62] C.-Y. Yang and D.-C. Tarng, "Diet, gut microbiome and indoxyl sulphate in chronic kidney disease patients," *Nephrology*, vol. 23, pp. 16–20, 2018, doi: https://doi.org/10.1111/nep.13452.

[63] L. A. Mendonça, R. dos Santos Ferreira, R. de Cássia Avellaneda Guimarães, A. P. De Castro, O. L. Franco, R. Matias, and C. M. Carvalho, "The complex puzzle of interactions among functional food, gut microbiota, and colorectal

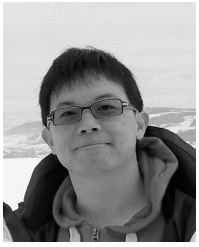
cancer," *Frontiers in Oncology*, vol. 8, p. 325, 2018, doi: https://doi.org/10.3389/fonc.2018.00325.

[64] Y. Cao, J. Oh, M. Xue, W. J. Huh, J. Wang, J. A. Gonzalez-Hernandez, T. A. Rice, A. L. Martin, D. Song, J. M. Crawford *et al.*, "Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites," *Science*, vol. 378, no. 6618, p. eabm3233, 2022, doi: https://doi.org/10.1126/science.abm3233.

[65] F. H. Zwezerijnen-Jiwa, H. Sivov, P. Paizs, K. Zafeiropoulou, and J. Kinross, "A systematic review of microbiome-derived biomarkers for early colorectal cancer detection," *Neoplasia*, vol. 36, p. 100868, 2023, doi: https://doi.org/10.1016/j.neo.2022.100868.

ZHAO-QI HU received the B.S. degree in Biology from National Changhua University of Education (NCUE), Changhua, Taiwan, in 2020. He obtained a double M.S. degree in BioAgri and EECS from the National Taiwan University (NTU) of Biotechnology and Biomedical Electronics and Bioinformatics, Taipei, Taiwan, in 2023. He is currently working as the CEO of PetSci Co., Ltd., in Taiwan. His research interests mainly focus on bioinformatics, metagenomics, and full-stack web development.

YUAN-MAO HUNG received the B.S. degree from National Central University of Electrical Engineering, Zhongli, Taiwan, in 2012, the M.S. degree from National Chiao Tung University of Communication Engineering, Hsinchu, Taiwan, in 2015, and the Ph.D. degree from National Taiwan University of Biomedical Electronics and Bioinformatics, Taipei, Taiwan. He is currently working as a post-doc at Broad institute of MIT and Harvard. His research interests include metagenomics, immunology, deep learning, and computer aided engineering.



LI-HAN CHEN received the Ph.D. degree from Norwegian University of Life Science, in 2015. He is an assistant professor in Institute of Fisheries Science, National Taiwan University. His main research interests include microbiology, immunology, and metagenomics.

LIANG-CHUAN LAI received the M.S. and Ph.D. degree in Molecular and Integrative Physiology from University of Illinois at Urbana-Champaign, Urbana, IL, in 2001 and 2005. He is currently working as a Professor at the Department of Physiology, National Taiwan University, Taipei, Taiwan. His research interests are using genomic approaches, i.e. microarrays & next generation sequencing, to explore the molecular mechanism of carcinogenesis.

MIN-HSIUNG PAN received the Ph.D. degree from Biochemistry and Molecular Biology, College of Medicine, National Taiwan University, in 2000. He is currently a Distinguished Professor at the Institute of Food Science and Technology, National Taiwan University. His main interests include the natural dietary compounds, plant derived exosome-like nanoparticles, cancer chemoprevention, metabolism disorder, gut microbiota, circadian rhythm.

ERIC Y. CHUANG received his doctorate in cancer biology with toxicology and molecular genetics as two sub-specialties from Harvard University in 1997, and his doctoral thesis was to study radiation-induced mutagenesis in human cells. After graduation, he stayed at Harvard as a postdoctoral fellow for one year. He then joined the Radiation Biology Branch of the National Cancer Institute (NCI), National Institutes of Health (NIH), as an IRTA fellow to study radiogenomics in Bethesda, MD, USA. Next, he became the Head of Microarray Laboratory for Radiation Oncology Sciences Program at NCI.

After working at the NIH for several years, he took a faculty position at National Taiwan University (NTU). In 2009, he joined the Radiation Research Program of Division of Cancer Treatment and Diagnosis at NCI as a Program Director to oversee a portfolio of NIH grants that included radiation-induced signaling pathways, molecular mechanisms, and normal tissue injuries, as well as radiation-related genomic studies. In 2011, he returned to NTU and served as the Director of the Graduate Institute of Biomedical Electronics and Bioinformatics (BEBI) from 2012 to 2018.

Currently, he is a Professor of BEBI at NTU and serving as the General Director, Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan. Being an expert in genomic technologies, bioinformatics, cancer, radiation biology & oncology, biomedical engineering, and precision medicine, he has published more than 160 peer-reviewed papers in related fields. Moreover, Dr. Chuang has been serving as an editorial board member of Scientific Reports and the Editor-in-Chief of Translation Cancer Research. Finally, Dr. Chuang received his Executive MBA degree in International Business Management from NTU in June 2017.

MONG-HSUN TSAI received the B.S. degree in Zoology from National Taiwan University, in 1993, the M.S. degree in Radiation Biology from National Tsing Hua University, Hsinchu, Taiwan, in 1995, and the Ph.D. degree in Public Health from National Yang-Ming University, Taipei, Taiwan, in 2001. He is currently working as a Professor at the Institute of Biotechnology, National Taiwan University, Taipei, Taiwan. His research interests include Bioinformatics, Microarray, Cancer biology, and Radiation biology.