environmental microbiology

Environmental Microbiology (2015) **17**(12), 4954–4964 doi:10.1111/1462-2920.13006

Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults

Julien Tap,[1] Jean-Pierre Furet,[1] Martine Bensaada,[1] Catherine Philippe,[1] Hubert Roth,[2,3] Sylvie Rabot,[1] Omar Lakhdari,[1] Vincent Lombard,[4] Bernard Henrissat,[4] Gérard Corthier,[1] Eric Fontaine,[2,3] Joël Doré[1] and Marion Leclerc[1]*

[1]*INRA, Micalis Institute, UMR1319, Jouy-en-Josas 78350, France.*

[2]*CHU Grenoble, Bioénergétique-INSERM U1055, Université J. Fourier, Grenoble 38000, France.*

[3]*CRNH Rhône-Alpes, Pierre Benite 69310, France.*

[4]*Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 7257, 163 Avenue de Luminy, Marseille F-13288, France.*

Summary

Gut microbiota richness and stability are important parameters in host–microbe symbiosis. Diet modification, notably using dietary fibres, might be a way to restore a high richness and stability in the gut microbiota. In this work, during a 6-week nutritional trial, 19 healthy adults consumed a basal diet supplemented with 10 or 40 g dietary fibre per day for 5 days, followed by 15-day washout periods. Fecal samples were analysed by a combination of 16S rRNA gene pyrosequencing, intestinal cell genotoxicity assay, metatranscriptomics sequencing approach and short-chain fatty analysis. This short-term change in the dietary fibre level did not have the same impact for all individuals but remained significant within each individual gut microbiota at genus level. Higher microbiota richness was associated with higher microbiota stability upon increased dietary fibre intake. Increasing fibre modulated the expression of numerous microbiota metabolic pathways such as glycan metabolism, with genes encoding carbohydrate-active enzymes active on fibre or host glycans. High microbial richness was also associated with high proportions of Prevotella and Coprococcus species and high levels of caproate and valerate. This study provides new insights on the role of gut microbial richness in healthy adults upon dietary changes and host microbes' interaction

Received 9 May, 2015; accepted 27 July, 2015. *For correspondence. E-mail marion.leclerc@jouy.inra.fr; Tel. (33) 1 3465 2306; Fax (33) 1 3465 2492.

Introduction

The gut microbiota has been shown to be involved in intestinal homeostasis (Hooper and Gordon, 2001), and it is now widely believed that this complex microbial ecosystem plays a key role in human health. Metabolic syndrome, inflammatory bowel diseases and obesity are among the health impairments for which alterations in microbiota composition and gene content have been observed (Ley *et al.*, 2005; Mazmanian *et al.*, 2005; Qin *et al.*, 2010). Most of the studies examining the human gut microbiota are aimed at demonstrating the health benefits that arise from modulating the microbial communities in the gastrointestinal tract (Gibson and Roberfroid, 1995; Looijer-van Langen and Dieleman, 2009). However, due to the complexity of the GI tract ecosystem, its microbial diversity, spatial organization and relationships between organisms through intricate anaerobic metabolic networks, the primary outcome of targeted modulation of the microbiota remains elusive (Flint *et al.*, 2008). Dietary habits and food ingredients have long been recognized to influence the colonic ecosystem (Finegold *et al.*, 1974; Flint *et al.*, 2007), and 35 years ago, Burkitt and colleagues proposed the dietary fibre hypothesis that 'dietary fibre has a role in the prevention of certain large bowel and other diseases present in Western countries' (Burkitt *et al.*, 1972). Since that time, many publications have dealt with this subject, including revisions to the definition of what constitutes dietary fibres and the recommended daily intake of fibres (Topping and Clifton, 2001). Dietary fibres are widely believed to be beneficial for human health by providing vitamins, short-chain fatty acids (notably butyrate) and other nutrients through microbial fermentation while helping with weight management and gut health because of their bulking effects on transit time and stool frequency (Pryde *et al.*, 2002). The preventive effects of fibres on several cancers have also been reported in the literature (Cummings *et al.*, 1992).

There are a number of human colonic bacteria that have been described as fibrolytic (Salyers *et al.*, 1977a,b; Chassard *et al.*, 2007). However, with the exception of starches, fibrolytic systems have been more extensively

described for the rumen ecosystem (Miron *et al.*, 2001). In humans, despite metagenomic analyses providing more evidence of fibrolytic potential (Gill *et al.*, 2006; Kurokawa *et al.*, 2007; Turnbaugh *et al.*, 2008; Tasse *et al.*, 2010), very few mechanisms are precisely described and the corresponding pathways are not well understood for most of the candidate species (Flint *et al.*, 2008).

To better understand the impact of a short-term increase in dietary fibre intake on the gut microbial community, we undertook a randomized nutritional intervention in which healthy adults received controlled normal diets providing an average dietary fibre level of 10 or 40 g per day based on actual western diet intake and nutritionist recommendations respectively. The microbiota composition and functions were assessed through a panel of complementary assays: quantitative polymerase chain reaction (qPCR) and 16S rRNA gene targeted pyrosequencing for determining microbiota structure; metatranscriptomics for evaluating gene expression changes, liquid chromatography for determining short-chain fatty acid (SCFA) profiles and comet assays to assess genotoxicity towards intestinal cells. We used a statistical framework combining a supervised approach and a multi-table analysis to integrate all data (Chessel *et al.*, 2004). Such a framework allowed us to decipher how factors, dietary fibre level and subject specificity, impacted the gut microbiota dynamics and how the different datasets were linked during the clinical trial.

Results

A short-term change in dietary fibre impacts gut microbiota differently within subjects

Nineteen young healthy adults fully completed the randomized crossover study and complied to the dietary interventions (Fig. S1). A supervised multi-table approach, using diet and host subject as instrumental variables into a between-class analysis (BCA), allowed us to detect a high individual specificity (BCA inertia > 40%, $P < 0.05$) whatever the taxonomic level (Fig. 1 and Text S1 for details about the statistical analysis). In fact, host specificity represented a major driver of the microbiota dynamics during the nutritional intervention. In addition, increasing the level of taxonomic resolution, from phylum to species, increased the variability linked to the individual (Fig. 1). Changing dietary fibre amount did not have the same impact for all individuals but remained significant within each individual (BCA inertia > 10%, $P < 0.05$). In addition, qPCR probes did not allow us to observe any significant change during the clinical trial, except for *Escherichia coli*, which decreased significantly after 40 g fibre per day for the 'r40-10' diet run (Fig. S2). A partial triadic analysis (PTA) performed at the genus level indicated that each time point from the dietary intervention

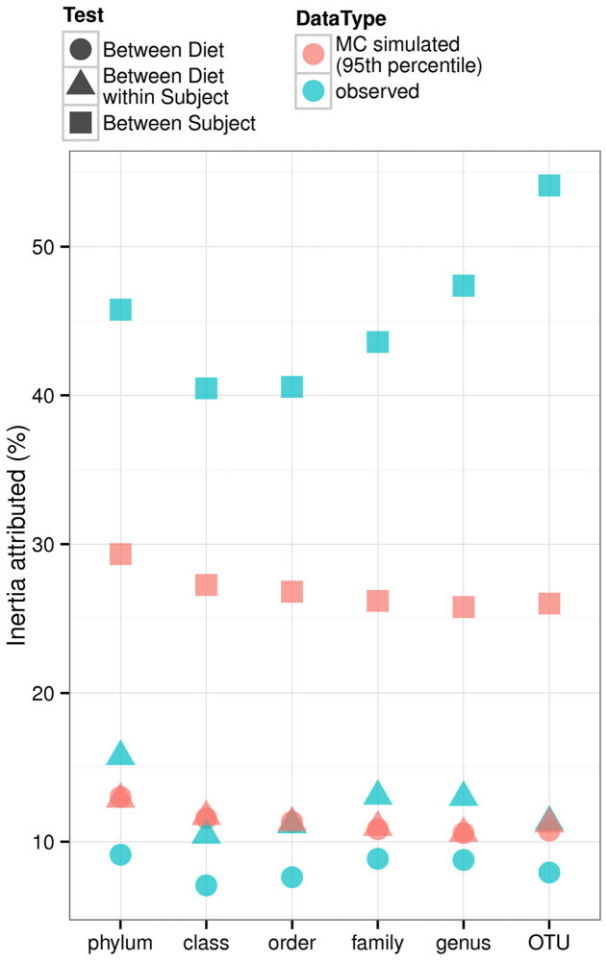

Fig. 1. Gut microbiota variations attributed to diet and subjects at different taxonomic levels. Between-class analysis was used to measure the microbiota composition variations (inertia) attributed to, alternatively, diet and subject as instrumental variable. Monte Carlo (MC) simulated inertia ($n = 1000$ microbiota composition simulations using permutations) was used to assess the significance of inertia attributed to different factors (diet and subject specificity). Significance was obtained when the observed inertia was higher than the 95th percentile of MC simulated inertia (see Text S1 for more information about BCA and MC). Subject specificity explained significantly more than 40% of inertia regardless taxonomic levels. After subject specificity removal (so-called between diet within subject), fibre diet change explained significantly more than 10% of inertia at phylum, family and genus levels.

had the same contribution to the detected variation (see Text S1 for more details about PTA). In other words, PTA showed that the magnitude of gut microbial variation between individuals was conserved whatever the time point, reinforcing individual specificity as the main factor in the observed variability. In addition, gut microbial dynamics showed different gut microbiota stability pattern among each individual (Fig. S3).

In summary, due to the high individual specificity, diet change did not have the same global effect on all indi-

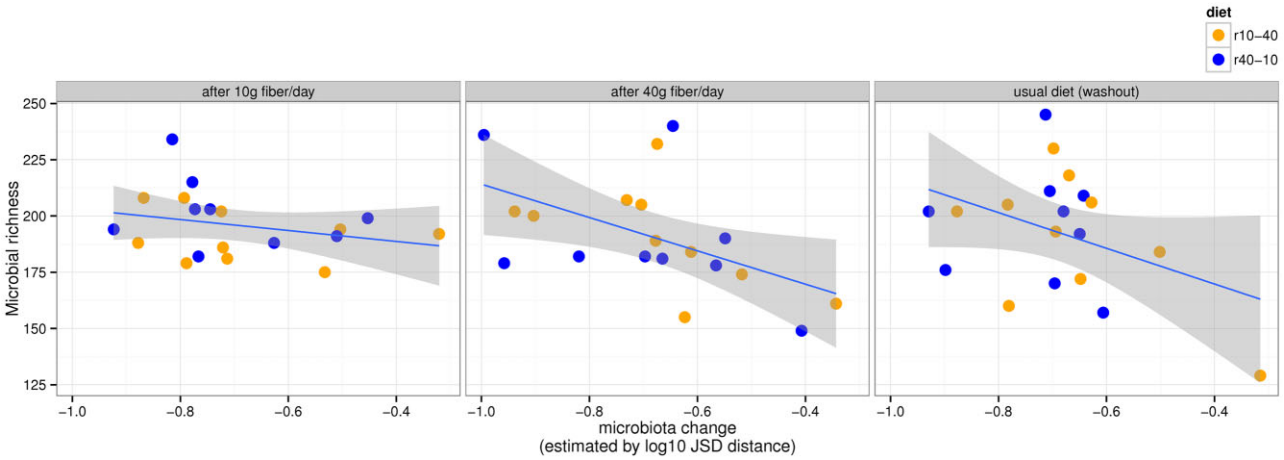

Fig. 2. Gut microbiota change over time associated with OTU richness. Microbial change within subject was assessed using the JSD distance metrics after 10 g fibre per day, 40 g fibre per day and the usual diet (washout period). Microbial richness accounted for the number of OTUs rarefied at the same level of 16S rRNA genes sequences for each sample before the diet. The two diet runs, 10 to 40 g and 40 g to 10 g, were pictured by orange and blue dots respectively. Linear regression is represented by a blue line and standard errors (two standard deviations) by the grey shade. Spearman correlation only showed significant association between microbiota change and richness after 40 g fibre per day (Spearman test, rho = –0.55, P value < 0.05, see Table S1 for more information).

viduals, regardless of taxonomic level, but had a significant impact within each individual at phylum, family and genus levels. In order to keep the most resolutive taxonomic level at which the fibre amount had an impact on individuals, the remaining analysis were done at the genus level to assess diet impact on microbiota stability and richness.

High species richness microbiota are more stable upon dietary fibre changes

To assess subject response to diet changes, species richness was calculated using the number of operational taxonomic units (OTUs) rarefied at the same sequencing depth. Species richness did not differ over time during the trial and 'between-diet' run ('r10-40' versus 'r40-10') (Fig. S2). However, using Jensen Shannon Distances (JSD) metrics (Arumugam *et al.*, 2011), a proxy to measure stability between sampling points at genus level, we were able to correlate the species richness and gut microbiota stability (Fig. 2). Species richness was negatively correlated with microbiota change, after the 40 g fibre per day diet (rho = –0.55, $P < 0.05$; see Table S1). During this short-term high-fibre diet, the low species richness microbiota was stochastically impacted, while the high richness microbiota exhibited more stability (i.e. low JSD distance). This observation was not seen when subjects consumed neither the low fibre diet nor their usual diet. Also, this correlation was independent from the crossover design (see Fig. S1, 'r10-40' versus 'r40-10' diet run). In addition, paired comparisons by subject did not show that microbiota stability globally increased or decreased after 10 g or 40 g fibre per day diet (Wilcoxon paired test, P value > 0.05).

Interestingly, we observed a link between microbiota richness and diversity of vegetables the subjects were eating before the trial (See Table S2 for the vegetables-based 3-day food questionnaire and Fig. S4). From the food diversity index, the sum of the different items subjects were eating per day, we could infer the microbial richness at baseline (Fig. S4).

Gut microbiota richness is linked to faecal SCFA but not to genotoxicity

Univariate spearman correlations tests were undertaken between microbiota richness, SCFAs amounts and bacterial genera frequencies. After multiple testing corrections, using false discovery rate, no significant associations were found. Then, co-inertia analyses (COIA) between the dominant gut microbiota at the genus level and SCFA measurements showed that those datasets shared the same co-structure (supported by a significant Monte Carlo test; Text S1 for details on the statistical analysis). The microbial richness seemed a main factor determining how the microbiota responded to the diet, and we observed that richness was significantly correlated with the first principal component (PC1) of the co-inertia analysis, which linked SCFA measurement and bacterial genera (Spearman correlation test, $P < 0.05$). Indeed, higher amounts of caproate and valerate were positively associated with the proportion of *Prevotella*, *Dorea* and *Coprococcus* genera and negatively associated with the proportion of the *Bacteroides* genus in stool

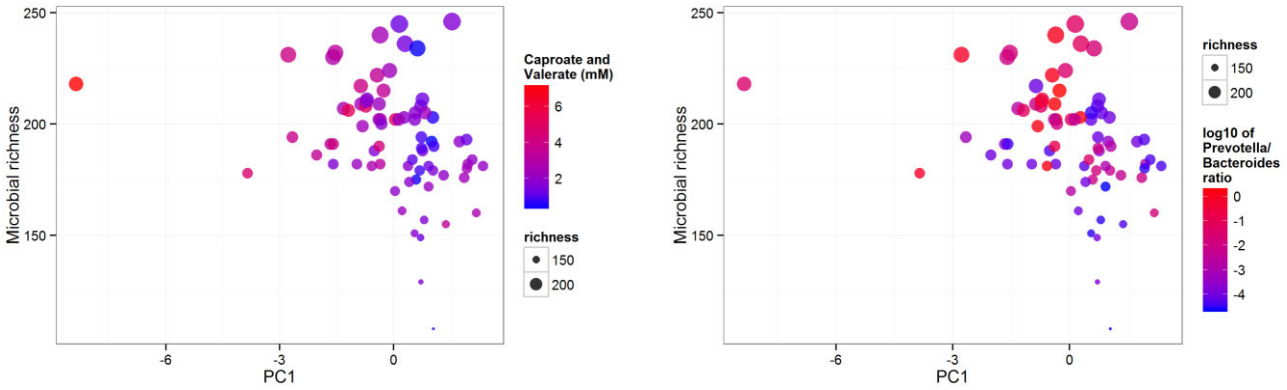

Fig. 3. Association between microbial OTU richness, SCFA and bacterial genera. The first principal component (PC1) from the co-inertia analysis between SCFA and microbial genera was associated significantly with microbial richness. Circle size accounted for OTU richness and colour gradient accounted for caproate and valerate (mm) amount (left panel) and *Prevotella/Bacteroides* ratio (right panel).

samples (all Spearman correlation tests with PC1 of COIA, $P < 0.05$, Fig. 3 and Fig. S5). To investigate how diet and microbiota profiles affected host physiology, the genotoxicity of individual faecal waters from all the subjects was measured by the comet assay, a single-cell gel electrophoresis recognized as a standard procedure to measure DNA damage in eukaryotic cells (here HT-29 cells). The images obtained looked like 'comets' with a distinct head, comprising intact DNA and a tail, which length is a measure of DNA strand breaks or damage (for more information see http://www.cometassay.com/). First, the DNA-damaging effect correlated with neither the amount of fibre in the diet nor the diet time points. In addition, no links between faecal water SCFA concentrations, gut microbial richness and comet tail phenotypes were found.

Increasing dietary fibre modulates the expression of metabolic pathways

A metatranscriptomics approach (see Table S3 and Text S1 for complementary (c)DNA library preparation details) was used to examine how the shift from 10 to 40 g of fibre per day changed the microbiota gene expression in healthy young adults. The eight cDNA libraries were prepared from samples from subjects 1, 10, 11, 17, which represented different microbial responses to the diet (Fig. S3) and different food habits regarding vegetables (Fig. S4). Approximately 23% of the total reads, representing bacterial messenger (m)RNA, remained after cleaning, filtering and removal of the ribosomal RNA sequences (Table S4). Assembly bins resulted in 5006 contigs with 58 858 sequences and 59 443 singletons sequences. There were 64 449 gene clusters generated based on the CAP3 assemblies. Only 15 082 blast hits (e value < 10–5) representing 50% of the reads mass, were obtained against the METAHIT database. Furthermore, there were 1709 clusters of orthologous groups (COG), 1766 KEGG orthology (KO) and 390 non-supervised orthologous groups (NOG) clusters distributed among the samples. In all individuals, the genes related to glycan and lipid metabolism were significantly overrepresented at 40 g fibre per day (Fig. 4). Because of the KEGG database structure, genes related to lipids or carbohydrates metabolism can appear in distinct KEGG categories. In fact, genes involved in carbohydrate metabolism and in glycan pathways can be part of the same pathway (Fig. 4). Based on KEGG and COG categories, multiple glycoside transporters (K10112) were significantly overrepresented at 40 g fibre/day. The expressed carbohydrate-active enzymes represented 73 protein families from the CAZy database: 50 glycoside hydrolases, 12 glycosyltransferases, eight carbohydrate-binding modules and three polysaccharide lyases families (Table S5). Interestingly, at 40 g fibre, three genes involved in mucin degradation were downregulated: CAZy GH101 family, encoding endo-a-N acetylgalactosaminidases involved in mucin degradation, α-L-arabinofuranosidases (GH43) and the recently described GH115 family of α-glucuronidases (Table S5). In contrast, amylase (GH13) and its dockerin domain (GH13-DOC) were significantly upregulated. Consistent with a higher fermentation, archaeal genes involved in methanogenesis, such as methyl-CoM reductase (COG4058) and formylmethanofuran dehydrogenase (K00205) were overexpressed at 40 g of fibre. Among the bacterial machinery, genes for molecular chaperones and heat shock proteins were underexpressed at 40 g/ fibre, while ribosomal protein genes were overexpressed (Table S5). Also overexpressed with high-fibre diet were phage response genes (Table S5).

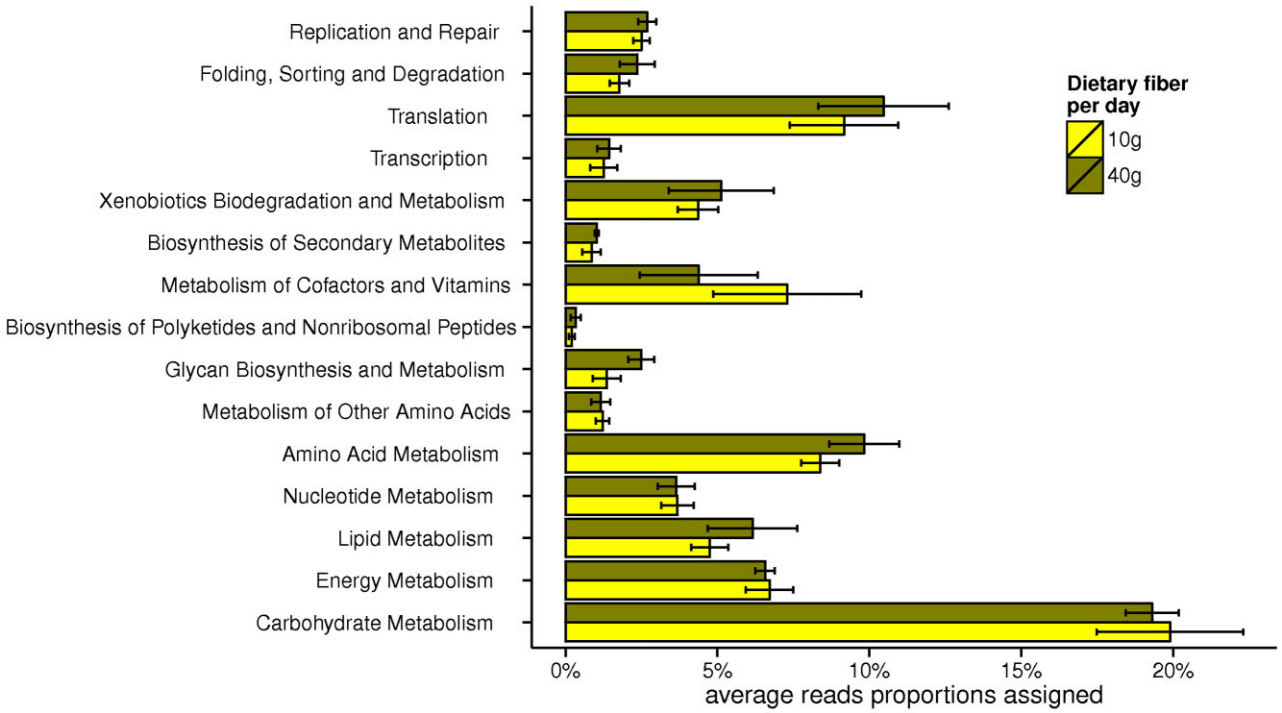

Fig. 4. Impact of diet on microbial gene expression. The graph pictures the relative abundance of gut microbial cDNA reads assigned to KEGG database subcategories according to diet fibre content. Yellow bars account for a diet with 10 g fibre/day (point 3 on Fig. S1). Green bars account for the 40 g fibre/day diet (point 5 on Fig. S1). The standard error mean were indicated using error bars.

The taxonomic assignment of the cDNA reads provided valuable information on the quality of the RNA extraction protocol and on the active fraction of the microbiota. We were able to compare the cDNA taxonomy with the 16S rRNA-based bacterial community structure (Fig. S6). Both datasets showed the abundance and activity of the *Firmicutes*, and members of the *Lachnospiraceae* appeared as active members of the community. Among the *Bacteroidetes*, cDNA affiliated to the *Bacteroides* genus seemed to represent a higher fraction in the transcriptome. Sequences assigned to the *Bifidobacterium* were equally abundant and active. Noticeably, a higher proportion of reads from the *Proteobacteria* were detected in the cDNA. Finally, consistent with the overexpressed genes detected, archaea methanogens were detected in the transcriptome. Several metabolic pathways could be correlated with taxa (Fig. S6 right panel), and this link was stronger at 10 g/fibre, which might indicate that at 40 g, the microbiota was more perturbed.

Discussion

This study investigated how a large increase of dietary fibre (40 g per day) in an otherwise balanced identical diet affected the microbiota structure-function profiles in young healthy normal weight adults. The results presented here suggest that changes in dietary fibre intake had an impact on the gut microbiota populations, but this was dependent upon microbiota richness. The genus appeared to be the most precise taxonomic level to assess the impact of dietary fibre within individuals. Low OTU microbial richness was associated with a higher microbiota change over time due to high dietary fibre change. This was particularly true when subjects switched from their normal diet to the 40 g fibre per day diet. There was no relationship between microbiota change and richness when subjects consumed 10 g fibre per day indicating that, during low fibre diet nutritional intervention, other factors might play a role for the gut microbiota dynamics. Higher microbiota stability was associated with higher richness, which was also associated with higher proportion of caproate and valerate and a higher *Prevotella*:*Bacteroides* ratio.

A 5-day diet specifically modulated the structural and functional dynamics of the human intestinal microbiota. The observed fast reactivity of the human gut microbiota composition has recently been highlighted (Walker *et al*., 2011; David *et al*., 2014), but our current work added that this effect is dependent upon the baseline gut microbial richness. Cotillard and colleagues observed 25% increase in gene diversity by dietary intervention with high fibre diversity in obese and overweight subjects (Cotillard *et al*., 2013). Similarly, Salonen and colleagues demonstrated that in obese men following a fibre-based diet to

promote weight loss, the dietary responsiveness varied substantially and associated inversely with individuals' microbiota diversity (Salonen *et al*., 2014). Another study linked low diversity with most pejorative phenotype in obese individuals (Le Chatelier *et al*., 2013). Interestingly, we found that gut microbial richness in healthy lean adults is positively correlated with higher concentrations of caproate and valerate together with a higher proportion of *Prevotella*, *Coprococcus* and *Dorea* (with unassigned species from *Lachnospiraceae* family) and lower levels of *Bacteroides*. *Prevotella* isolates are known to produce caproate and, to a lesser extent, valerate (Hayashi *et al*., 2007). The association between gut microbiota changes, low species richness and *Bacteroides* in response to the fibre interventions might be a reflection of an expanded metabolic niche for *Prevotella* spp., especially those capable of utilizing xylans for growth. *Prevotella bryantii* B(1)4 (Dodd *et al*., 2011) and *P. ruminicola* (Cantarel *et al*., 2009) are rumen bacteria known for efficiently degrading xylan, but none of the human gut *Prevotella* studied so far have been shown to degrade xylan. However, the enrichment in the human metagenomes of *Prevotella*-affiliated reads has been shown (Qin *et al*., 2010; Tasse *et al*., 2010). In children from Burkina Faso consuming a diet rich in fibre, *Prevotella* spp. represented up to 53% of the microbiota (De Filippo *et al*., 2010). Wu and colleagues also linked long-term low and high-fibre diets with *Bacteroides* and *Prevotella*-driven enterotypes respectively (Wu *et al*., 2011). Importantly, our study also suggests that a fibre-rich diet, as part of a normal diet consumed by healthy young adults, does not increase faecal butyrate concentrations, nor exhibit a 'bifidogenic' effect (Gibson and Roberfroid, 1995; Topping and Clifton, 2001), but significantly affects subjects with lower richness microbiota.

We focused the metatranscriptomic analysis on the changes observed between low and high-fibre diets because it reflects the difference between an average western diet and dietary recommendations. Given the short half-life of bacterial mRNA, the transcriptional activity measured in the faeces does not reflect exactly the fermentation occurring in the proximal colon. Because of water absorption between the proximal colon and the rectum, microbial metabolism is probably different in the two compartments. However, as expected, glycan metabolism was upregulated at 40 g fibre/day. The functional-based analysis of the carbohydrate active genes provided a more accurate annotation of these enzymes with highly variable modular structure (Cantarel *et al*., 2009) and allowed us to detect a decreased expression of mucin-related genes during high-fibre diet consumption, while glycosidases and amylase were overexpressed. Because the fraction of soluble and insoluble fibres was kept similar in the diet, the observed changes in α-amylase expression may be reflective of changes in the relative abundances of the predominant genera; or alternatively, that resistant starch and cellulose and hemi-cellulose were differentially metabolized. We could associate the taxonomic diversity of the cDNA with metabolic pathways, and this link was weaker at 40 g fibre per day. This might be the result of a more disturbed ecosystem showing a boarder metabolic versatility. Additionally, methanogenic archaea may not be numerically dominant over bacteria, (Miller and Wolin, 1982; Gill *et al*., 2006; Qin *et al*., 2010) or changing in biomass (Walker *et al*., 2011), but our data showed that hydrogenotrophic methanogenesis was increased at 40 g fibre per day. The cDNA taxonomic assignation confirmed the presence of archaea in the active community, a point that has not been highlighted in the metatranscriptomic studies (Gosalbes *et al*., 2011; David *et al*., 2014). Robert and Bernalier-Donadille reported a link between methanogenic archaea and cellulolytic bacteria and observed that *Firmicutes* were more represented in methane-producing subjects (Robert and Bernalier-Donadille, 2003; Chassard *et al*., 2010). This may be relevant as the competition for hydrogen between methanogenic archaea, sulfate reducers and homoacetogens is critical plays a part in human health (Pochart *et al*., 1993).

The subjects recruited for this study were all normal weight and healthy, originated from the same area, and had similar lifestyles. Yet, the individual variability explained more of the microbiome variability than the dietary fibre interventions. However, this observation depended upon the level of the microbiota analysis, which may explain contradictory reports in the literature. Studies investigating the impact of diet on gut microbiome are difficult to compare. Short-term or long-term intervention studies, very few focus on healthy lean individuals from western countries (David *et al*., 2014). Often, the studied cohort consists of obese subjects with or without metabolic syndrome (Walker *et al*., 2011), enrolled within a weight loss programme. Some very interesting cohorts focus on the diet eaten by inhabitants from western countries and compare them with individuals with very different lifestyles (De Filippo *et al*., 2010; Rampelli *et al*., 2015); diet is one of the components of the whole living environments of these populations (O'Keefe *et al*., 2015). Importantly, the methods to study the gut microbiome dramatically differ between studies: from qPCR targeting known groups and emphasizing on few species with high sensitivity, phylogenetic chip, ribotyping methods [Denaturing Gel Gradient Electrophoresis (DGGE), Temporal Temperature Gradient Electrophoresis (TTGE), Terminal Restriction Fragment Length Polymorphism (TRFLP)] to high throughput sequencing with not only various sequencing depth but also downstream bioinformatics analysis, and composition and diversity description (from

phylum to OTUs). Statistical analysis does not always provide the variability explained by the diet versus the individual variability. The review recently published by Graf and colleagues (2015) perfectly recapitulates the challenges and limits of our work and of others studying the gut microbiota and diet. Collectively, this work shows that despite a substantial interpersonal variation among the subjects, gut microbial richness has to be considered as a key parameter to facilitate meaningful comparisons of the subjects' responses to dietary interventions: a high richness should be further studied in a preventive way, to assess if it is maintained despite microbiota changing factors such as, antibiotics, body mass index (BMI), inflammation, aging. Conversely, individuals with a low richness microbiota, if considered as a risk factor, should be the subjects of long-term dietary intervention providing a natural diet with high vegetables diversity in order to try and improve the microbiome richness.

Experimental procedures

Subjects and study design

The study was registered as a clinical trial (http://clinicaltrials.gov/, Calibrated Diets and Human Intestinal Microflora NCT00639561) and carried out at the Center for Clinical Investigation (CCI) of Grenoble University Hospital (France). The protocol was explained to the volunteers and their written informed consent was obtained. Inclusion criteria were: age from 18 to 30, BMI [body weight (kg)/height (m2)] between 18.5 and 25, no intake of antibiotics, laxatives or other gastrointestinal medications 3 months prior to the beginning of the study.

Nineteen healthy young adults (nine males, 10 females, aged from 19 to 25 years) with no history of gastrointestinal problems completed the study. Subjects were randomly assigned to receive each meal plan during the 5 days with a 2-week washout period: 10 subjects followed a first round of 10 g dietary fibre per day for 5 days, followed by 2-week washout, then a second round of 40 g per day for 5 days, followed by a 7 days washout sequence. Nine other volunteers followed the exact same sequence but with 40 g per day in the first round and 10 g per day in the second round (Fig. S1). The two diets differed by the amount of fibres but the ratio of soluble/non-soluble fibre was kept similar. The meals were cooked with the same ingredients, lyophilized for reproducibility between the two groups, checked for sterility, packaged (Joint Research Unit for Food Process Engineering & Microbiology, INRA-AgroParisTech, Grignon, France) and shipped to the CCI. Fruits and vegetables were provided fresh. All 15 meals (three per day, see all meals in Text S1 and Table S6) were controlled and dispensed to the subjects 1 day before each diet period to be eaten at home (except for butter, sugar, tea, coffee and milk for breakfast and oil for cooking). During the diet periods, subjects were not allowed to eat anything but the food provided, or to drink anything besides water. During the washout, the subjects went back to their normal diet without any constraint. The subjects came to the CCI nine times and were monitored throughout the study for weight, heart rate and blood pressure. A plant-based food questionnaire was filled in by the individuals for the 3 days preceding the nutritional intervention (Table S2).

Fecal sample collections, storage and DNA extraction procedures

Fecal samples were collected 7 days before and 7 days after the nutritional intervention period. During the diet periods, faecal samples were collected the day before and days 4 and 5 of the intervention period. (Fig. S1). Fecal samples were kept in anaerobic containers (Anaerocult-A Merck) at 4°C and delivered to the CCI within 24 h for storage at −80°C. After homogenization, faecal aliquots were distributed, anonymously labeled and frozen at −80°C for analyses. Random duplicates were included. To establish SCFA profiles and genotoxicity, faecal waters were prepared by centrifuging 15 g of fresh faecal homogenate for 2 h at 171 000 g at 4°C. The supernatant was collected and stored at −80°C until analysis. All sample aliquots were shipped on dry ice, kept frozen throughout thereafter, and thawed only once for further work or analysis. Metagenomic DNA was extracted from about 200 mg aliquots of faeces as described previously (Furet *et al*., 2009), re-suspended in 150 µL of Tris-EDTA (TE) buffer and stored at −20°C until further analysis.

qPCR assays

Quantitative PCR procedures were applied as previously described [(Furet *et al*., 2009) and Supplementary Notes section 1.1 and 1.2]. The qPCR assays targeted total bacteria, *Clostridium coccoides* group (cluster XIVa), *C. leptum* group (cluster IV), *Bacteroides/Prevotella* group, *Bifidobacterium* genus and *E. coli* species. These data were normalized to account for the differences in water content between faecal samples, as described previously (Furet *et al*., 2009). The all-bacteria results were presented as the mean of the log10 value ± SD/g of stool (i.e. bacteria equivalents per gram) after corrections for the concentrations of each targeted bacterial species or group mentioned above.

16S rRNA genes molecular inventories

Prior to more extensive sample analyses, eight identical DNA samples were submitted to two companies for pyrosequencing, and were found to produce a highly significant correlation in terms of genus distributions ($P < 0.05$, Spearman methods). Second, to compare 454 GS FLX Ti and Sanger sequencing technologies, one sample was analysed by both methods and the datasets were found to share the same trends (Fig. S7). Based on these findings, the DNA samples from all individuals, before and after each diet (time point 2, 3, 4, 5 shown in Fig. S1) were analysed by targeted amplification and pyrosequencing of the V3-V4 region of the 16S rRNA genes using V3F (TACGGRAGGCAGCAG 343–357 *E. coli* position) and V4R (GGACTACCAGGGTATCTAAT 787 – 806 *E. coli* position) primers (For more details, see Text S1 section 1.3). Sequences were clustered into OTU at 97% similarity with LOTUS software (Hildebrand *et al*., 2014) using 150 nucleotides as minimal sequence length (see Fig. S8

for individual rarefaction curves at OTUs levels). The 76 16S rRNA genes sequence files from this study were deposited to European Nucleotide Archive under the reference ERA008162 (http://www.ebi.ac.uk/ena/data/view/ERA008162).

RNA extraction and metatranscriptomics

Total RNA was extracted from about 200 mg aliquots of faeces, using Diethylpyrocarbonate (DEPC) water and RNAse-free reagents. A chemical lysis step with phenol chloroform was combined with a Fast Prep mechanical lysis, before using the High Pure RNA Isolation kit (Roche). The protocol is a slight modification of the method described by van Hijum and colleagues (2005), and is described in Text S1 (section 2). The RNA extracts were suspended in 150 µL of TE buffer and stored at −80°C. Prior to cDNA production, 5S rRNA were removed using RNeasy mini kit (Qiagen), and 16S and 23S rRNA was removed using the MICROBExpress Bacterial mRNA Enrichment kit (Ambion). The cDNA libraries were then prepared using the TransPlex Whole Transcriptome Amplification kit (Sigma-Aldrich), with each step quality assessed by Bioanalyser profile (Agilent), and cDNAs were finally checked with Picogreen assays (Table S3).

Sequencing was performed using 454 GS FLX titanium technology (Beckman genomics, Grenoble, France) with quality control performed as previously described (Huse *et al*., 2007). The cDNA datasets were cleaned of any rRNA sequence reads by using a Hidden Markov Model-based algorithm (Huang *et al*., 2009). From 603 463 reads, 23% (118,301 reads) passed quality control and rRNA removal (Table S4). Assembly bins were made with CAP3 software (Huang and Madan, 1999) using low stringency parameter (option – p 66), (Text S1 section 3). The BLASTN program (Altschul *et al*., 1997) was used to compare the sequences to the MetaHIT project database (http://gutmeta.genomics.org.cn/). The KEGG Orthology, NOG and COG distributions were then retrieved for each sample. To profile the glycan digestion enzymes in the various individuals, cDNA contigs and singletons were assigned to CAZy families (Cantarel *et al*., 2009) using fasty (e-value < 0.001) on sequence libraries built with the isolated catalytic modules of glycoside hydrolases, polysaccharide lyases, carbohydrate esterases and glycosyltransferases, supplemented with the isolated carbohydrate-binding modules borne by these enzymes. The cDNA libraries sequences were registered to EBI under project number ERA008162.

SCFA and genotoxicity assays of faecal samples

The SCFA in both faecal waters and homogenates were analysed by gas–liquid chromatography as described previously (Lan *et al*., 2007) Prior to analysis, the faecal homogenates were water extracted and proteins were precipitated with phosphotungstic acid. Faecal waters were analysed without prior treatment. The single cell gel electrophoresis (SCGE) assay was used to assess the genotoxicity of faecal waters, using the human colon adenocarcinoma cell line HT-29 as a target (see Text S1 section 4).

Statistical analysis

All statistical analyses were computed with the GNU R software (www.r-project.org). Variations analyses were made using multi-table functions from ADE-4 library (Chessel *et al*., 2004). Notably, between (BCA) and within (WCA) class analyses were computed to assess the impact of the diet and the host, used as instrumental variables, on the whole inertia for different datasets. Impacts were then measured on how gut microbiota inertia was attributed to different factors. Co-inertia using RV coefficient was analysed to detect significant co-structure between datasets (Robert and Escoufier, 1976), meaning that different sets of variables (e.g. microbial genera abundance and SCFA profiles) were not independent and shared a fraction of inertia. Based on genera distribution, the JSD between samples were computed, as previously described (Arumugam *et al*., 2011) in order to measures microbiota stability between sampling point. Non-parametric paired Wilcoxon, Spearman correlation and Monte-Carlo tests were used to confirm observed relations between different datasets, assuming a P value < 0.05. Differential expression of metatranscriptomics data was analysed with the SHOTGUNFUNCTIONALIZER library using paired comparisons in a Poisson model (Kristiansson *et al*., 2009). In case of multiple comparisons, P values were adjusted by Benjamini Hochberg False Discovery Rate. GNU R source code and dataset used for this study are available at http://github.com/tapj/AlimIntest. Study design metadata, SCFA profile, qPCR and comet assays, OTUs count table and taxonomic assignation are available in the supporting information as Table S7.

Description of additional data files

Supplementary information, available at the journal's website, contains supplementary notes (detailed protocol and diet), supplementary figures (study design, technical replicates, multivariate analysis) and supplementary tables (protocol read out, metatranscriptomics approach summaries, plant-based food questionnaire).

Acknowledgements

Julien Tap and this project were supported by the French National Agency for Research, ANR/DEDD/PNRA/PROJ/200206-01-01, within the AlimIntest program. The clinical study was sponsored by University Hospital of Grenoble and carried out in its Clinical Research Center CIC1406. We are very grateful to Marie-France Nguyen Vaillant for designing all the menus. We thank Hervé Blottière for providing us with HT29 cells, Aurelia Bruneau for contributing to faecal sample collection and Deborah Jardet for her technical support. We are grateful to Timothy Swartz for reading the manuscript.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., *et al*. (2011) Enterotypes of the human gut microbiome. *Nature* **473:** 174–180.

Burkitt, D.P., Walker, A.R., and Painter, N.S. (1972) Effect of dietary fibre on stools and the transit-times, and its role in the causation of disease. *Lancet* **2:** 1408–1412.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res* **37:** D233–D238.

Chassard, C., Goumy, V., Leclerc, M., Del'homme, C., and Bernalier-Donadille, A. (2007) Characterization of the xylan-degrading microbial community from human faeces. *FEMS Microbiol Ecol* **61:** 121–131.

Chassard, C., Delmas, E., Robert, C., and Bernalier-Donadille, A. (2010) The cellulose-degrading microbial community of the human gut varies according to the presence or absence of methanogens. *FEMS Microbiol Ecol* **74:** 205–213.

Chessel, D., Dufour, A.B., and Thioulouse, J. (2004) The ade4 package-I-One-table methods. *R News* **4:** 5–10.

Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., *et al*. (2013) Dietary intervention impact on gut microbial gene richness. *Nature* **500:** 585–588.

Cummings, J.H., Bingham, S.A., Heaton, K.W., and Eastwood, M.A. (1992) Fecal weight, colon cancer risk, and dietary intake of nonstarch polysaccharides (dietary fiber). *Gastroenterology* **103:** 1783–1789.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., *et al*. (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505:** 559–563.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., *et al*. (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci USA* **107:** 14691–14696.

Dodd, D., Mackie, R.I., and Cann, I.K. (2011) Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. *Mol Microbiol* **79:** 292–304.

Finegold, S.M., Attebery, H.R., and Sutter, V.L. (1974) Effect of diet on human fecal flora: Comparison of Japanese and American diets. *Am J Clin Nutr* **27:** 1456–1469.

Flint, H.J., Duncan, S.H., Scott, K.P., and Louis, P. (2007) Interactions and competition within the microbial community of the human colon: Links between diet and health. *Environ Microbiol* **9:** 1101–1111.

Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R., and White, B.A. (2008) Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. *Nat Rev Microbiol* **6:** 121–131.

Furet, J.P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., *et al*. (2009) Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS Microbiol Ecol* **68:** 351–362.

Gibson, G.R., and Roberfroid, M.B. (1995) Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *J Nutr* **125:** 1401–1412.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., *et al*. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* **312:** 1355–1359.

Gosalbes, M.J., Durban, A., Pignatelli, M., Abellan, J.J., Jimenez-Hernandez, N., Perez-Cobas, A.E., *et al*. (2011) Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS ONE* **6:** e17447.

Graf, D., Di Cagno, R., Fak, F., Flint, H.J., Nyman, M., Saarela, M., and Watzl, B. (2015) Contribution of diet to the composition of the human gut microbiota. *Microb Ecol Health Dis* **26:** 26164–26175.

Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S., and Benno, Y. (2007) *Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* **57:** 941–946.

van Hijum, S.A., de Jong, A., Baerends, R.J., Karsens, H.A., Kramer, N.E., Larsen, R., *et al*. (2005) A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. *BMC Genomics* **6:** 77–87.

Hildebrand, F., Tadeo, R., Voigt, A.Y., Bork, P., and Raes, J. (2014) LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2:** 30–37.

Hooper, L.V., and Gordon, J.I. (2001) Commensal host-bacterial relationships in the gut. *Science* **292:** 1115–1118.

Huang, X., and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res* **9:** 868–877.

Huang, Y., Gilna, P., and Li, W. (2009) Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25:** 1338–1340.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8:** R143.

Kristiansson, E., Hugenholtz, P., and Dalevi, D. (2009) ShotgunFunctionalizeR: an R-package for functional comparison of metagenomes. *Bioinformatics* **25:** 2737–2738.

Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Toh, H., Toyoda, A., *et al*. (2007) Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* **14:** 169–181.

Lan, A., Bruneau, A., Philippe, C., Rochet, V., Rouault, A., Herve, C., *et al*. (2007) Survival and metabolic activity of selected strains of *Propionibacterium freudenreichii* in the gastrointestinal tract of human microbiota-associated rats. *Br J Nutr* **97:** 714–724.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., *et al*. (2013) Richness of human gut microbiome correlates with metabolic markers. *Nature* **500:** 541–546.

Ley, R.E., Backhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci USA* **102:** 11070–11075.

Looijer-van Langen, M.A., and Dieleman, L.A. (2009) Prebiotics in chronic intestinal inflammation. *Inflamm Bowel Dis* **15:** 454–462.

Mazmanian, S.K., Liu, C.H., Tzianabos, A.O., and Kasper, D.L. (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122:** 107–118.

Miller, T.L., and Wolin, M.J. (1982) Enumeration of *Methanobrevibacter smithii* in human feces. *Arch Microbiol* **131:** 14–18.

Miron, J., Ben-Ghedalia, D., and Morrison, M. (2001) Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J Dairy Sci* **84:** 1294–1309.
O'Keefe, S.J., Li, J.V., Lahti, L., Ou, J., Carbonero, F., Mohammed, K., *et al*. (2015) Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun* **6:** 6342.
Pochart, P., Lemann, F., Flourie, B., Pellier, P., Goderel, I., and Rambaud, J.C. (1993) Pyxigraphic sampling to enumerate methanogens and anaerobes in the right colon of healthy humans. *Gastroenterology* **105:** 1281–1285.
Pryde, S.E., Duncan, S.H., Hold, G.L., Stewart, C.S., and Flint, H.J. (2002) The microbiology of butyrate formation in the human colon. *FEMS Microbiol Lett* **217:** 133–139.
Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., *et al*. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464:** 59–65.
Rampelli, S., Schnorr, S.L., Consolandi, C., Turroni, S., Severgnini, M., Peano, C., *et al*. (2015) Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Curr Biol* **25:** 1682–1693.
Robert, C., and Bernalier-Donadille, A. (2003) The cellulolytic microflora of the human colon: evidence of microcrystalline cellulose-degrading bacteria in methane-excreting subjects. *FEMS Microbiol Ecol* **46:** 81–89.
Robert, P., and Escoufier, Y. (1976) A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Appl Stat* **25:** 257–265.
Salonen, A., Lahti, L., Salojarvi, J., Holtrop, G., Korpela, K., Duncan, S.H., *et al*. (2014) Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. *ISME J* **8:** 2218–2230.
Salyers, A.A., Vercellotti, J.R., West, S.E., and Wilkins, T.D. (1977a) Fermentation of mucin and plant polysaccharides by strains of *Bacteroides* from the human colon. *Appl Environ Microbiol* **33:** 319–322.
Salyers, A.A., West, S.E., Vercellotti, J.R., and Wilkins, T.D. (1977b) Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* **34:** 529–533.
Tasse, L., Bercovici, J., Pizzut-Serin, S., Robe, P., Tap, J., Klopp, C., *et al*. (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res* **20:** 1605–1612.
Topping, D.L., and Clifton, P.M. (2001) Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* **81:** 1031–1064.
Turnbaugh, P.J., Backhed, F., Fulton, L., and Gordon, J.I. (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **3:** 213–223.
Walker, A.W., Ince, J., Duncan, S.H., Webster, L.M., Holtrop, G., Ze, X., *et al*. (2011) Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J* **5:** 220–230.
Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., *et al*. (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334:** 105–108.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Text S1. Supplementary notes on microbiota composition analysis (quantitative PCR, pyrosequencing and statistical methods), RNA extraction procedure from faecal microbiota sample, cDNA libraries preparation and metatranscriptomic *in silico* analysis, SCGE assay detailed protocol and detailed menus of each diet plan.
Fig. S1. Crossover nutritional intervention of this study. Each disc represents a collection point. The blue group was composed of nine individuals, who first had 40 g fibre/day; the orange group of 10 individuals who started with 10 g fibre/day. During the 15-day washout period, the individuals were on their usual personal diet without any constraint.
Fig. S2. Gut microbiota richness and composition as a function of diet over time. 454 pyrosequencing was used to assess the microbial richness (top panel) and qPCR was used to quantify major microbial groups overtime including the basal time (bottom panel). The all-bacteria qPCR results were presented as the mean of the log10 value ± SD/g of stool (i.e. bacteria equivalents per gram). Each other targeted-bacterial group was then normalized by the all bacteria. The numbers above the graphs correspond to the intervention time points (see Fig. S1).
Fig. S3. Dynamics paired impact of host and diet on microbiota structure. This PTA deciphered the structure common to four contingency tables corresponding to four collections points with the same subjects and same bacterial genera. Partial triadic analysis allowed a mapping of both microbiota and individual paired as function of time. Each point represented a collect point for each individual linked, in time, to the next collected point. Blue dots accounted for the 40–10 g diet sequence, orange for the 10–40 g sequence (see Fig. S1). Microbiota variables, the 14 most abundant bacterial genera obtained by 16S rRNA gene pyrosequencing, were pictured as a compromise of the variation in time. The small boxes represented the nutritional intervention periods (see Text S1 for more information about PTA).
Fig. S4. Associations between vegetal-based food diversity and microbiota richness. All the subjects filled a plant-based food questionnaire for each of the 3 days preceding their inclusion to the nutritional intervention. (left panel) Between subject comparisons based on multiple correspondence analysis (MCA) using all the vegetal based food questionnaires data. Multiple correspondence analysis' second (PC2) and third (PC3) components were plotted. Each number corresponds to a subject. The three questionnaires were linked by their centroid, where subject ID number is indicated. Between class analysis showed that food questionnaires were subject dependent (41% inertia contribution, $P < 0.001$). Microbial richness (top right panel) and vegetal-based food diversity (bottom right panel, sum of PC2 and PC3 as Y axis) were significantly with linked with PC3 (both Spearman correlation test, rho > 0.4, P value < 0.05).
Fig. S5. Association between microbial OTU richness, SCFA and bacterial genera using COIA. The two first principal components loadings from COIA between SCFA profile and

bacterial genera are plotted. Colour gradient accounts for Spearman rho correlation between SCFA profile and bacterial genera relative abundance, with microbial OTU richness (see Text S1 for more information about COIA).

Fig. S6. Impact of fibre intake on taxonomic and functional activities measured by metatranscriptomics approach. (left panel) Taxonomic and functional comparison between 16S rRNA gene sequencing and metatranscriptomics dataset. (right panel) Co-inertia compromise between taxonomic and functional activities in the gut microbiota with 10 g or 40 fibre diet: Combined with a PTA, the contribution of different fibre amount to the link between taxonomic and function is measured (inset graphic). After 10 g of fibre, the link between taxonomic and functional activities is higher than 40 g of fibre.

Fig. S7. Technical replicates comparisons based on dominant genera relative log10 proportion. Left panel: eight replicates pyrosequenced by two different companies were compared and show high significant correlation based on genera distribution (rho > 0.7, $P < 0.05$, spearman method). Right panel: a sample sequenced by both Sanger and pyrosequencing methods (right panel) with different lane and barcoding configurations showed the same trend (rho > 0.7, $P < 0.05$, Spearman correlation test)

Fig. S8. Rarefaction analysis of the pyrosequencing data by diet time point. Number of unique OTUs identified in the 76 samples relative to the sequencing depth. For time points and nutritional study scheme see Fig. S1.

Table S1. Associations between microbial change and richness before the diet and comet assays after the diet. Associations were assessed by Spearman rho correlation, and *P* values were indicated in parenthesis.

Table S2. Daily plant-based food questionnaire summaries. Questionnaires were filled by all the subjects for the 3 days preceding the nutritional intervention. For each day, a binary vector was obtained with subject as row and plant-based food item as column. '1' meant that the subject took this food item during the day.

Table S3. Nucleic acid quantification during cDNA library preparation from 200 mg of faecal sample.

Table S4. Sequences numbers after quality check and *in silico* ribosomal RNA removal

Table S5. Microbial COGs, NOGs, KOs and CAZy families significantly different between low-fibre diet and high-fibre diet. A highly positive coefficient (computed by SHOTGUNFUNCTIONALIZER R package) means that the gene was overexpressed when individuals shifted from the 10 g to the 40 g fibre/day diet. *P* values were adjusted by Benjamini Hochberg False Discovery Rate.

Table S6. Weekly diet composition with 10 and 40 g of fibre per day.

Table S7. Study design metadata, SCFA profile, qPCR and comet assays, OTUs count table and taxonomic assignation.