A randomized controlled trial for response of microbiome network to exercise and diet intervention in patients with nonalcoholic fatty liver disease

Exercise and diet are treatments for nonalcoholic fatty liver disease (NAFLD) and prediabetes, however, how exercise and diet interventions impact gut microbiota in patients is incompletely understood. We previously reported a 8.6-month, four-arm (Aerobic exercise, n = 29; Diet, n = 28; Aerobic exercise + Diet, n = 29; No intervention, n = 29) randomized, singe blinded (for researchers), and controlled intervention in patients with NAFLD and prediabetes to assess the effect of interventions on the primary outcomes of liver fat content and glucose metabolism. Here we report the third primary outcome of the trial—gut microbiota composition—in participants who completed the trial (22 in Aerobic exercise, 22 in Diet, 23 in Aerobic exercise + Diet, 18 in No Intervention). We show that combined aerobic exercise and diet intervention are associated with diversified and stabilized keystone taxa, while exercise and diet interventions alone increase network connectivity and robustness between taxa. No adverse effects were observed with the interventions. In addition, in exploratory ad-hoc analyses we find that not all subjects responded to the intervention in a similar manner, when using differentially altered gut microbe amplicon sequence variants abundance to classify the responders and low/non-responders. A personalized gut microbial network at baseline could predict the individual responses in liver fat to exercise intervention. Our findings suggest an avenue for developing personalized intervention strategies for treatment of NAFLD based on host-gut microbiome ecosystem interactions, however, future studies with large sample size are needed to validate these discoveries. The Trial Registration Number is ISRCTN 42622771.
Nonalcoholic fatty liver disease (NAFLD) is the most common chronic liver disease with prevalence estimates ranging from 25% to 45% in worldwide. NAFLD is closely associated with type 2 diabetes (T2D), most likely via a common pathophysiological mechanism - insulin resistance. Therefore, interventions targeting the synergistically pathogenic mediator of NAFLD and diabetes are likely to be a rational approach in the prevention and treatment of the comorbidity condition.

The gut microbiota, which contributes to the metabolic health of the human host, may work as one of the targets for the treatment of NAFLD. Emerging data demonstrate that dysbiosis of the gut microbiota is linked with NAFLD and T2D. Although the exact mechanism(s) require further elucidation, inflammation, damage to the intestinal membrane, and translocation of bacteria have all been suggested. Drugs with antifibrotic or anti-inflammatory treatments for various stages of NAFLD in human trials is pending. Currently, increased physical activity and dietary modifications are the only effective therapeutic options for NAFLD management, and the mechanisms of these interventions have been associated with modulation of gut microbiota and its metabolites. For example, a low-carbohydrate diet (LCD) has been reported to improve fatty liver metabolism and to promote rapid shifts of the gut microbiota composition in NAFLD patients. Regular aerobic exercise has been shown to reduce hepatic fat in obesity, and some studies have shown that exercise increases gut microbial diversity and alters the composition and functional capacity of gut microbiota. However, the underlying mechanism of the beneficial effects of exercise and diet on NAFLD, as well as the related metabolic disorders through regulation of the gut microbiota, still remain to be elucidated.

One of the challenges in constructing the relationship between the gut microbiota and improvement of NAFLD during exercise and/or dietary interventions is that not all patients respond to the intervention in a similar manner. Earlier studies have shown that there was an individual variation in response to exercise, with some subjects experiencing greater improvement than others. The low/non-responders could be up to 50% in terms of change of gut microbiota after exercise intervention. Some researchers have suggested that the variation in responsiveness might be due to genotypic and phenotypic factors. Thus, large inter-individual variance may mask the changes of the microbiome in studies using traditional statistical analytical methods. Currently, it is widely accepted that, in the microbial community, the keystone taxa are drivers of microbiome structure and function, and in particular, their interaction network, which plays an important role in microbial functions and disease progression. Therefore, to establish an effective intervention strategy, it may be worthwhile to analyze the microbiota correlation network at the microbial community level, while assessing the physiological mechanisms at the individual level and further exploring individual microbial networks that underlie the differences between responders and low/non-responders of various interventions.

In the present study, we analyzed the composition and metabolic pathways of the gut microbiota, constructed a co-occurrence network at the population level, and developed personalized gut microbial networks. This individual network enabled us to identify the microbial signature and interaction of taxa at a personal resolution, and to further differentiate responders from low/non-responders after intervention. Taken together, our study with relatively long-term intervention provides both statistical and personal resolution, and to further differentiate responders from low/non-responders after intervention. Therefore, we stratified the participants into responders (HFC decreased more than 5%) and low/non-responders (HFC decrease less than 5% or increased) (Fig. 1b). For the primary outcome of gut microbiota composition, seventy-six subjects provided paired stool samples at baseline and after the intervention for 16 s rRNA gene sequencing. The demographics and clinical variables of those who have had gut microbiota results are presented to Table S1. In addition, to better understand the intervention impact on the function of microbiome, we selected a subset of the cohort (n = 42) with the best or worst response, in terms of HFC reduction, to the intervention, and analyzed gut microbiota by the metagenomics data (Fig. 1c).

The microbiota data and clinical parameters were used to (1) characterize the change of gut microbiota composition in response to interventions; (2) determine associations of the microbiome with the level of physical fitness, fat mass, serum biomarkers and short-chain fatty acids (SCFAs); (3) analyze the metabolic function shift of the microbiome in response to interventions; (4) discover the co-occurrence network features of the microbiome ecosystem before and after intervention; and (5) predict personalized response to intervention based on the baseline gut microbiota composition and network.

Changes of structure and function of gut microbiota induced by exercise and/or dietary intervention of participants with comorbidity of NAFLD and prediabetes. We first performed 16 S rRNA gene sequencing of fecal samples collected before and after interventions and obtained a total of 5421 amplicon sequence variants (ASVs). We found that the alpha diversity (Shannon index) of gut microbiota was significantly decreased in the NI group (p = 0.045) with large individual variance by contrast to the intervention groups, which maintained their diversity after the intervention (Fig. 2a, NI vs. AED p = 0.011, NI vs. AEx p = 0.007 and NI vs. Diet p = 0.025, respectively, analysis of variance with repeated measures adjusted for change of body weight, baseline value and intervention duration). The changes in microbial diversity may reflect the natural progression of NAFLD in the NI group. Moreover, the principal coordinates analysis (PCoA) based on weighted UniFrac distances shows the...
significant difference between NI and the other groups after the interventions ($p < 0.05$, PERMANOVA, Fig. 2b). These results indicate that microbial diversity deteriorates with the increased HFC, while exercise and dieting may help to maintain the diversity of the gut microbiota. To avoid only including the relative abundance data in the analysis which may introduce the bias, we then measured the absolute abundance of microbiome by qPCR targeting the 16S rRNA gene. We found that there were no significant differences between baseline and after intervention as well as among the groups (Fig. S2a). Moreover, the total bacterial content of each subject in the same group at two time points also did not differ significantly (Fig. S2b–c).

Furthermore, we performed Linear discriminant analysis Effect Size (LDA Effect Size) analysis of ASVs that appeared in more than 20% of the samples. The ASVs showed significant differences between the intervention groups and the NI group after intervention (adjusted $p < 0.05$, log2 fold change > 2) (Fig. 2c). Compared with the NI group, we found that 15 ASVs were enriched and 5 ASVs decreased in the AEx group; and 13 ASVs were enriched in the AEx group, while 9 ASVs were enriched and 6 ASVs were decreased in the Diet group after the intervention. Among these ASVs, ASV2077 (AEx: $p = 0.018$, AEx < 0.01, Diet: $p = 0.032$) and ASV2513 (AEx: $p = 0.012$, AEx: 0.01, Diet: $p = 0.013$) belonged to Bacteroides, and ASV3942 (AED: $p = 0.030$, AEx: $p = 0.024$, Diet: $p < 0.01$), which belongs to Ruminococcus, increased in all intervention groups. In addition, ASV5361, which belongs to Lachnospiraceae, increased in both AEx ($p = 0.027$) and AED ($p = 0.037$) groups. ASV4240, belonging to Bacteroides, increased in both Diet ($p < 0.01$) and AED ($p < 0.01$) groups. Noticeably, some ASVs from the same family or genus showed different behaviors. For example, ASV4432 in Lachnospiraceae was enriched in the AED ($p < 0.01$) group but decreased after the diet ($p = 0.023$) intervention compared with the NI group. To investigate whether changes of microbiome at the genus level were similar to the changes at the genus level, we further performed a LEfSe analysis of the same parameters on the genus level. We found that those observed changed ASVs, if belonging to the same genus, indeed, they do have a similar trend at the genus level (Fig. S3).

We next performed partial Spearman correlation analysis to assess change of ASVs with clinical biomarkers after adjustment for body weight and fat mass ($n = 46$, Fig. 2d). We found that three altered ASVs were significantly negatively associated with the reduction of HFC [ASV2468, belonging to Bacteroides ($r = -0.31$, $p = 0.040$), ASV3307, belonging to Ruminococcaceae ($r = -0.32$, $p = 0.030$), and ASV4538, belonging to Lachnospira ($r = -0.37$, $p = 0.012$)]. In addition, ASV478, which belongs to Phascolarctobacterium ($r = -0.32$, $p = 0.033$); ASV1715, which belongs to Alistipes ($r = -0.35$, $p = 0.022$); and ASV5195 and ASV5305, which belong to Lachnoclostridium ($r = -0.30$, $p = 0.045$, $r = -0.34$, $p = 0.023$, respectively), were negatively correlated with HbA1c. Only ASV776,
Fig. 2 Changes in the structure of gut microbiota before and after intervention and relationship of microbial abundance with clinical parameters.

a) Alpha diversity (Shannon index) of gut microbiota was significantly decreased in the NI group by contrast to the intervention groups, which maintained their diversity after intervention duration followed by Sidak correction for multiple comparison between the groups. Contrast results (K Matrix) were used to localize the group differences: *p < 0.05, by contrast to the NI group. AED 8.6 m vs NI 8.6 m, p = 0.011; AEx 8.6 m vs NI 8.6 m, p = 0.007; Diet 8.6 m vs NI 8.6 m, p = 0.025; NI 0 m vs NI 8.6 m, p = 0.031. AED: n = 20, AEx: n = 20, Diet: n = 21 and NI: n = 15. b) Weighted UniFrac distance of samples was calculated from the QiIME 2 ASV level. Ellipsoids represent a 50% confidence interval surrounding each group. c) Differential ASVs between the AED/AEx/Diet groups and NI group (LEfSe analysis, adjusted p < 0.05, log2 fold change > 2). Genus annotations on the left column covers multiple right ASVs. AED: p < 0.016, was positively associated with HBA1c. Although a previous study23 showed the gut microbiota may benefit humans via SCFAs production from carbohydrate fermentation, fecal SCFAs did not change significantly from baseline to follow-up within groups in this experiment. However, we found that some alterations in gut microbes correlated with changes in SCFAs (Fig. 2d). The level of butyric acid significantly correlated with 9 ASVs. Notably, ASV3718 in the genus Faecalibacterium was significantly negatively correlated with 4 out of all 6 SCFAs, while ASV1989 in the genus Bacteroides was positively correlated with 4 SCFAs (Fig. 2d).

To assess microbial functions, we also applied shotgun sequencing in a subset of samples (Fig. 1c) and annotated the data to the KEGG pathway (Fig. 3). By using LEfSe analysis (p < 0.05, LDA score > 2), we found that, in total, 64 specific pathways were significantly different between the AED/AEx/Diet groups and NI group after interventions, most of which fall into the 'carbohydrate metabolism,' 'energy metabolism,' 'glycan biosynthesis and metabolism,' 'lipid metabolism,' 'amino acid metabolism' and 'metabolism of cofactors and vitamins' function pathways. The differences in functional pathways of the three intervention groups relative to the NI group were generally
similar, however, the AED group had more differential pathways than other groups. Notably, within 'lipid metabolism', 'Sphingolipid metabolism' was more vigorously different from that of other pathways (Fig. 3). In addition, most functional pathways were more abundant in the NI group; only 'glycan biosynthesis and metabolism' was more abundant in the intervention groups. Taken together, the above findings indicate that exercise or/and diet intervention(s) distinctly alter the abundance and function of microbiome, which are associated with changes in HFC and SCFAs.

Intervention induced change in gut microbiota co-occurrence network. It has been shown that bacterial species in the human gut may survive, adapt, and decline as interdependent functional groups (guilds) responding to environmental perturbations. To identify bacteria in the gut ecosystem that responded as functional groups to interventions, we adopted "co-abundance groups (CAGs)" to analyze the community structure in the microbial ecosystem. We used the SparCC algorithm to calculate the correlation coefficients among 279 ASVs shared by at least 20% of the samples from all the groups and time points. These 279 ASVs were then clustered into 35 CAGs (Table S2) and their abundance difference between groups were showed in Fig. S4a.

Network analysis can disentangle microbial co-occurrence and provide comprehensive insight into the microbial assembly patterns and community structures. Using Spearman correlation analysis, we constructed a co-occurrence network (Fig. 4a–d) to illustrate the potential interactions among the 35 CAGs in each group with...
baseline and follow-up samples. Each node in the network represents a CAG, and each edge represents a significant correlation ($p < 0.05$) between two CAGs. We found that the number of edges and average degrees that signify the connectivity of CAGs in the gut microbiota network were higher after intervention than at baseline in the AED, AEx and Diet groups, but not in the NI group (Fig. 4e). In addition, the network robustness$^{28}$ was significantly decreased in the NI group but maintained or increased in all intervention groups (Fig. 4f). The greatest increase in robustness score was in the AEx group (+29.3%), followed by the Diet (+12.8%) and AED (+3.1%) groups.

Because the complexity caused by the nonlinear dynamics of the nodes is related to the network architecture, we assessed the distribution of the network degree to judge the topological features of the microbial network (Fig. 4b). We found that most of the connections of the co-occurrence network were
to build a network for each patient (Fig. 5b) and found that responders from the low/non-responders (age: AUC = 0.50, sensitivity = 0.60) were currently connected to the HFC change; however, this was not observed in the Diet group (Fig. S8b). For supervised LASSO classifiers, ROC showed that the edge number of ASV data used to predict the responder exceeded 70% AUC in all three intervention groups (Fig. S8c), which was higher than the metagenomics data. However, the effect of unsupervised classifiers deteriorated and was almost ineffective in the Diet group (Fig. S1d). These results indicated that our Single SparCC networks could be used to construct personalized gut microbial networks for each individual sample and thus differentiate the responders from the low/non-responders, particularly for the exercise intervention. Taken together, these analyses demonstrate that the individual baseline gut microbial network can predict the response to exercise intervention.

Discussion
In the current study, in patients with NAFLD and pre-diabetes, we identified the characteristics of the gut microbiota responding to an 8.6-month aerobic exercise and/or low-carbohydrate dietary intervention. We showed that after combined exercise and diet intervention, changes are prominent in gut microbial composition in which keystone taxa became divergent but their connections remained stable. By contrast, with exercise or diet intervention alone, effects were significant regarding network connectivity and robustness between taxa. Moreover, the personalized microbial network was able to predict the intervention efficiency of individual patients.

Members from Ruminococcus have been reported to produce SCFAs from complex carbohydrates \cite{31,32}. A recent study \cite{33} showed that Ruminococcaceae was negatively correlated with the fibrosis severity. In agreement with this, members in Ruminococcaceae was found negatively correlation with HFC in our study. In addition, some members of Bacteroides contribute to the release of energy from dietary fiber and starch, and they are likely to be a major source of propionate \cite{34}. Accordingly, we found that members in Bacteroides was positively correlated with propionic acid, isobutyric acid and isovaleric acid. A previous study showed that compared to western countries, the abundance of Bacteroides in NAFLD is much lower in Chinese individuals \cite{35}. However, our result suggested an important energy-extracting role of members in Bacteroides, although their abundance is relatively low in Chinese population. Of note, there are currently conflicting results about the shifts of microbiota at the taxon level in NAFLD-related studies. For example, some studies showed that Bacteroides abundance was higher in NASH (nonalcoholic steatohepatitis, a stage of NAFLD progression) than in non-NASH \cite{36}. Another article showed the phylum Bacteroidetes (44.63%) tended to be more abundant in healthy subjects than in NAFLD \cite{35}. Our results showed that members in Bacteroides abundance was increased after both diet and exercise intervention, and this increase was correlated with decreased HFC. However, previous studies showed higher fiber intake was marginally associated with lower abundance of Bacteroides uniformis \cite{37}, and Bacteroides was increased in participants with obesity but decreased in lean participants after exercise intervention \cite{38}. These inconclusive results may be ascribed to differences in ethnicity, living environment and lifestyle among the cohorts. Thus, the mechanism of action of the key microbiota needs to be investigated in future studies.

The long-term adaptation after intervention may account for maintenance of gut microbiota alpha diversity in all intervention groups in contrast to the NI group in which alpha diversity decreased. However, the connectivity and robustness of the microbiota co-occurrence networks improved in three
Various ecological microbiome studies have shown that healthy people have a higher average network degree and hence greater connectivity of their gut microbiota network. A poorly developed microbial network usually has lower functionality due to fewer taxa present that can support their function in ecosystems. In combination with these findings, our results indicate that microbiota as an ecosystem were more healthy and stable after intervention. Interestingly, when

**Fig. 5 Intervention efficiency predicted by individual baseline gut microbial networks on metagenomics species data.** a Prediction of responders using our single SparCC network method. b The gut microbial network for each individual before intervention based on metagenomics species data by the Single SparCC network method. Three numbers under every network represent subject ID, HFC change and edge numbers. c Linear regressions of edge number (Single SparCC network) with the change of HFC after intervention. AED: n = 12, AEx: n = 12, Diet: n = 10. d ROC curve of edge number for differential responders from low/non-responders within each intervention group by unsupervised classification. AED: n = 12, AEx: n = 12, Diet: n = 10. e ROC curve by supervised Lasso model performance. AED: n = 12, AEx: n = 12, Diet: n = 10.

intervention groups. Various ecological microbiome studies have shown that healthy people have a higher average network degree and hence greater connectivity of their gut microbiota network. A poorly developed microbial network usually has lower functionality due to fewer taxa present that can support their function in ecosystems. In combination with these findings, our results indicate that microbiota as an ecosystem were more healthy and stable after intervention. Interestingly, when
ASVs abundance is used to predict the change of HFC, it has the best predictive power in the diet group, but is not predictive in the exercise group. However, when assessed the robustness of co-occurrence networks, the exercise intervention showed the highest increment of robustness score. A possible explanation is that diet intervention directly affects the structure and composition of gut microbiota through changes in nutrients, while exercise intervention influences the interaction between bacteria through the regulation of host metabolism. In addition, this study confirmed that co-occurrence networks of microbiota are scale-free networks, and further defined four keystone taxa, which have an prominently high degree in each network. A recent study found that keystone taxa that are highly connected in the microbiome can explain microbiome compositional turnover better than all taxa combined. Therefore, we speculate that what exercise intervention improves not the gut microbiota diversity, but the stability of the interaction network of the gut microbiota ecosystem, thereby improving the metabolism of the host. The keystone taxa determined by the microbiota population network topology may be a new perspective beyond finding key bacteria by abundance and diversity comparison.

No disease-specific drug/medical therapy for NAFLD is available due to its complexly pathogenetic mechanism, and management of lifestyle (such as exercise and diet) in patients is the only efficient intervention in clinical practices. However, studies have shown that not all participants respond to interventions of exercise or diet in a similar way, which may be attributable to the individual-variation and response of gut microbiota in patients. In some studies, 50% of the study participants show no change in their gut microbiota composition after an exercise intervention. Then prediction of improvement of clinical phenotype is of great clinical significance. Using the network features of gut microbiota to predict individual response to exercise and diet interventions may help to choose more precise and effective treatment for different NAFLD phenotypes.

Comparing with the conventional machine learning methods, although the reference samples can be viewed as the training samples, our method can further extract the second-order statistical information (e.g., correlation or association or network) of each individual approximately, thus providing system-wide features of each test sample. In general, our results are important in terms of how to assess the intervention effect on physiology and pathophysiology and to develop personalized lifestyle treatment for such patients through their individual gut microbiota network.

Our study provides novel information on how exercise and diet intervention affect the gut microbiome for patients with comorbidity conditions at both the population and individual levels. However, there are some limits in this study. Our sample size for metagenomics is relatively small but it is comparable with most of the studies. The dietary and exercise intervention methods are often different from trial to trial, and even if the intervention methods are the same, the intervention contents and intensity are different which makes the comparison with other trials challenging. In addition, since the trial was not originally designed to predict response efficiency of personalized microbiota network, our predictions were not validated by independent samples. Nevertheless, this study provides a new perspective in the research of gut microbiota, that is, in addition to focusing on the differential strains themselves, the topological properties of the microbial ecosystem community network are also of great significance. Using the network features of gut microbiota to predict individual response to exercise and diet interventions may help to choose more precise and effective treatment for different NAFLD phenotypes.

**Methods**

The study was approved by the Ethics Committee of Shanghai Institute of Nutrition (06.01.2013) and has been retrospectively registered in the International Standard Randomized Controlled Trial Number Register (ISRCTN 42622771). We acknowledge that one limitation of the study is that the trial was retrospectively registered. Two of the primary outcomes (Glucose tolerance test and Liver fat content) have been previously reported. All participants signed informed consent forms before beginning the study. The study was conducted in accordance with the principles of the Declaration of Helsinki.

**Subjects and design.** This study was a randomized controlled trial. Detailed information was given in our previous publication. The enrollment of participants was between January 6 and December 25, 2013. The eligibility criteria for intervention groups were as follows: men or women aged 50-65 years with impaired fasting glucose (IFG between 5.6 to 6.9 mmol/L) or impaired glucose tolerance (IGT between 7.8 to 11.0 mmol/L 2 hour after the intake of 75 g glucose), and diagnosed as NAFLD by ‘H MRS (liver fat > 5%) and by questionnaire that ongoing or recent alcohol consumption is <21 drinks on average per week in men and <14 drinks on average per week in women; no chronic cardiovascular, serious musculoskeletal or gastrointestinal problems and not on extreme diets; and for women, serum follicle-stimulating hormone level greater than 30 IU/L and last menstruation more than 6 months ago but within 10 years. The exclusion criteria included body mass index (BMI) >34 kg/m²; serious cardiovascular or musculoskeletal problems; diagnosed with Type 1 diabetes and T2D; and mental illness.

In the initial study design, we estimated the sample size based on previous literature of liver fat and microbiota as primary outcome variable. We estimated that 34 subjects in each group would have 85% power for mean comparison between the groups and therefore we targeted to have 50 subjects in each group. However, during the recruitment period, we found that only about 20% of the subjects met the inclusion criteria and due to the limited funding, we re-calculated the sample size based on our previous similar type of study. Thus, for this report, sample size estimation for the primary outcome HFC with 29 individuals has 95% power to test against the null hypothesis that there is no change in any group. Further, when intervention groups compared with NI group having 17 subjects, the power for the HFC was 84% with < 0.05 two-sided significance level. After every subject reached minimal 6 months intervention, the study did not continue to follow-up all the subjects and consequently, the follow-up study was terminated. There were no sex distribution differences between the groups and the male/female ratio for each group was n = 6/23 for AEX, n = 6/22 for Diet, n = 7/22 for AED, and n = 7/22 for NI group, respectively.

**Interventions.** The duration of the intervention ranged from 6.5 months to 11.1 months and on average 8.6 months. There were no significant differences in
Subsequent purification of bacterial DNA extraction was performed following the 16S rRNA Sequence analysis. The extracted DNA was sequenced using an Illumina Miseq (Illumina, San Diego, CA, USA) sequencing platform with a 6-cycle index reads method to eliminate spurious correlation. The raw sequencing data were analyzed by QIIME 2. In the process of running the DADA2 pipeline, based on the quality profile of the data, forward and reverse reads were trimmed accordingly to ensure that the median quality score for each position was above 32. The taxonomy of all ASVs was annotated by the SILVA (v122) reference database. All samples were rarefied to 10,000 per sample for downstream analysis.

**Metagenomics analysis.** Forty-two paired samples were sequenced at the Novogene Bioinformatics Institute (Novogene, Beijing, China). After construction of 350 bp insert libraries, metagenomic samples were pair-end sequenced on the HiSeq X Ten System. Cluster generation, template hybridization, isothermal amplification, linearization, and blocking denaturing and hybridization of the sequencing primers were performed according to the workflow specified by the service provider. The pipeline was consistent with previous studies. We removed contamination of host DNA and assembled the metagenome by using Bowtie (Default parameter Settings:—end-to-end,—sensitive, -l 200, -x 400) and SOAP denovo (K-mers = 53). To obtain the proportion of spurious correlation, SparCC was used to detect the association between the characteristics of each sample. Gut microbiota 16 S sequencing and metagenomics data suffered from spurious correlation. SparCC is the most popular method to detect the association between species. The correlation coefficient of each sample pair was calculated as the product of the log-transformed relative abundance of each species. The correlation coefficient was calculated using the SparCC package.

**Short chain fatty acid measurement.** Slice of human feces were homogenized with 1.2 mL of phosphate buffer solution in weighed 2 mL centrifuge tubes (marker as W1). The mixture was centrifuged completely and centrifuged twice at 4 °C at 16000 rpm for 15 min. The supernatants were filtered through a 0.22 µm filter (EMD Millipore). Feces residue and tube were dried in 80 °C drying apparatus overnight and weighed (marked as W2). 200 µL of the supernatant was acidified with 100 µL of 50% (v/v) sulfuric acid. After vortexing and standing for 2 min, the organic phase was extracted by adding 400 µL of diethyl ether. The extracted solution was measured by gas chromatography (GC) on an Agilent 6890 (Agilent Technologies, CA, USA). The concentrations of 6 short chain fatty acids (acetic acid, propionic acid, butyric acid, isobutyric acid, valeric acid and isovaleric acid) were calculated by each peak area and feces dry weight (W2-W1).

**Co-abundance gene networks (CAGNs).** The ASVs shared by at least 20% among all samples were considered key ASVs. The correlation among 279 key ASVs was calculated by the SparCC algorithm (with a bootstrap procedure repeated 100 times, and then correlation matrices were computed from the resampled data matrices. Once the bootstrapped correlation scores have been computed, only ASVs with correlation scores greater than 0.4 were classified into CAGNs. The p value at the desired cut-off < 0.05.

**Network fragility.** Given the N node (CAG) number in a network, the N nodes were removed one by one in a random order. The fragility statistics (y axis) were the ratio of the node number in the largest connected component to the node number N, and were calculated when each node was removed, while the percentage of removed nodes was taken as the x axis. The robustness value R for a network was related to a function σ, which was dependent on the decreasing order of fragility statistics as nodes were removed, i.e.,

\[ R = \frac{1 - N_i}{N - i} \]

where i represents the ith CAG. To ensure the results were stable, we repeatedly remove the nodes in each network for 1000 times. The median fragility statistics of the 1000 curves were taken as the coordinates for the final curve to plot the robustness curve.

**Single SparCC network analysis.** The traditional correlation coefficient mainly exploits the general information in a group of samples. Rather than focus on a ‘mean’ value within a group, our Single SparCC network is proposed to investigate the co-occurrence features of each sample pair. The beta 16S sequencing and metagenomics data suffered from spurious correlation. SparCC is the most popular method to detect the association between species. The correlation coefficient of each sample pair was calculated as the product of the log-transformed relative abundance of each species. The correlation coefficient was calculated using the SparCC package.

The correlation of species i and j can be inferred from their ratio. Even though high-throughput sequencing only produce relative quantification of microbiome, the ratio of those relative value is the same with real abundance. Define \( t_{ij} \) to describe variance of log transformed relative value of component \( x_i \) and \( x_j \).

\[ t_{ij} = \frac{\log(\omega_j) - \log(\omega_i)}{\log(\omega_j) - \log(\omega_i)} = \frac{\log(x_j) - \log(x_i)}{\log(x_j) - \log(x_i)} \]

where \( x_i \) and \( x_j \) are the relative abundances of microbes \( i \) and \( j \), while \( \omega_i \) and \( \omega_j \) are the absolute abundances of those two microbes.
In SparCC method, \( t_i \) was related to Pearson correlation coefficient \( \rho_{ij} \) by equation:

\[
t_i = \frac{1}{2\omega_{i+1}} \left( b(a+1) + \frac{1}{\frac{1}{2} \sum_{j \neq i} \log \frac{x_i^{(j)}}{x_j^{(i)}} - \mu_{	ext{bgps}}} \times \right) - 2b \times \frac{1}{\frac{1}{2} \sum_{j \neq i} \log \frac{x_i^{(j)}}{x_j^{(i)}} - \mu_{	ext{bgps}}})
\]

where \( \omega_{ij} \) and \( \omega_{ij} \) are the variances of the log-transformed real abundance of species \( i \) and \( j \). Because of the sparse nature of microbial association network, we can assume that the average of all \( \rho_{ij} \) correlation is small enough to ignore, which allow us to solve for \( \omega_{ij} \) and \( \omega_{ij} \). Then put in the values of \( \omega_{ij} \) and \( \omega_{ij} \) back to Eq. (2), and end up with the approximated correlation \( \rho_{ij} \).

Letting \( \rho_{ij} = - \frac{1}{2\omega_{ij}} b(a+1) \), we achieve an analytical solution for \( \rho_{ij} \):

\[
\rho_{ij} = - \frac{1}{2\omega_{ij}} b(a+1) \sum_{j \neq i} \log \frac{x_i^{(j)}}{x_j^{(i)}} - \mu_{	ext{bgps}}\]

and the number of nodes \( \mu_{	ext{bgps}} \) and the number of edges \( \mu_{	ext{bgps}} \) are used for network construction. The networks in each intervention group (before and after intervention) were merged together to facilitate direct comparison and were further visualized in Cytoscape software (v3.8.1). To detect differential taxonomy, KEGG pathway analysis and linear discriminant analysis effect size (LDAseq v1.0), which is an algorithm for high-dimensional biomarker discovery and identifies genomic features such as genes, pathways or taxa as well as characterizing the differences between two or more biological conditions or classes) was performed with cut-offs (LDA Score > 2.0, \( p < 0.05 \)).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The study protocol has been published previously (https://doi.org/10.1186/s1471-2458-14-48). The raw illumina sequence data including 16s rRNA sequence and metagenomics generated in this study have been deposited to the NCBI database under BioProject accession no. PRJNA575939 (https://www.ncbi.nlm.nih.gov/bioproject?term=PRJNA575939). The data is available for academic use under controlled access in compliance with the regulation of the Ministry of Science and Technology (MOST) of China for the deposit and use of human genomic data. All individual de-identified participant data that support the findings of this study are available from the corresponding author upon reasonable request based on the rule of data protection and consent. The access to the controlled data will be valid for one year from the time of the data accessibility approved. The processed data are available within the Source Data file.

**Code availability.** Codes used for computing population and individual networks are available on GitHub (https://github.com/crtsjtu/AELC).

Received: 21 February 2021; Accepted: 11 April 2022; Published online: 10 May 2022

**References.**

1. Rinel, M. E. Nonalcoholic fatty liver disease: a systematic review. *JAMA* 2313, 2623–2627 (2015).
2. Williams, C. D. et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* 140, 124–131 (2011).
3. Leite, N. C., Salles, G. F., Araujo, A. L., Villela-Nogueira, C. A. & Cardoso, C. R. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. *Liver Int. 29*, 113–119 (2009).
4. Lonardo, A., Ballestri, S., Marchesini, G., Angulo, P. & Loria, P. Nonalcoholic fatty liver disease: a precursor of the metabolic syndrome. *Dig. Liver Dis. 47*, 181–190 (2015).
5. Betrapally, N. S., Gillevent, P. M. & Bajaj, J. S. Gut microbiome and liver disease. *Transl. Res* 179, 49–59 (2017).
6. Qin, N. et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature* 513, 59–64 (2014).
7. Betrapally, N. S., Gillevent, P. M. & Bajaj, J. S. Changes in the intestinal microbiome and nonalcoholic liver diseases: causes or effects? *Gastroenterology* 150, 1745–1755.e1743 (2016).
8. Haas, J. T., Francque, S. & Staelens, B. Pathophysiology and mechanisms of nonalcoholic fatty liver disease. *Annu. Rev. Physiol. 78*, 181–205 (2016).
9. Ratziu, V., Bellentani, S., Cortez-Pinto, H., Day, C. & Marchesini, G. A position statement on NAFLD/NASH based on the EASL 2009 special conference. *J. Hepatol.* 53, 372–384 (2010).
10. Mardiniouga, A. et al. An integrated understanding of the rapid metabolic benefits of a carbohydrate-restricted diet on hepatic steatosis in humans. *Cell Metab.* 27, 559–571.e555 (2018).
11. Johnson, N. A. et al. Aerobic exercise training reduces hepatic and visceral lipids in obese individuals without weight loss. *Hepatology* 50, 1105–1112 (2009).
12. Oh, S. et al. Moderate to vigorous physical activity volume is an important factor for managing nonalcoholic fatty liver disease: a retrospective study. *Hepatology* 61, 1205–1215 (2015).
13. Clarke, S. F. et al. Exercise and associated dietary extremes impact on gut microbial diversity. Gut 63, 1913–1920 (2014).
14. Mailing, L. J., Allen, J. M., Buford, T. W., Fields, C. J. & Woods, J. A. Exercise and the gut microbiome: a review of the evidence, potential mechanisms, and implications for human health. Exerc Sport Sci Rev. 47, 75–85 (2019).
15. Codella, R., Luzi, L. & Terruzzi, I. Exercise has the guts: how physical activity may positively modulate gut microbiota in chronic and immune-based diseases. Dig Dis 50, 331–341 (2018).
16. Anjo, D. et al. Who are the non-responder patients to cardiac rehabilitation? Eur. Heart J. 34, https://doi.org/10.1093/eurheartj/ehs358 (2013).
17. Mohr, A. E. et al. The athletic gut microbiota. J. Int. Soc. Sports Nutr. 17, 24 (2020).
18. Kirwan, L. et al. Phenotypic factors influencing the variation in response of circulating cholesterol level to personalised dietary advice in the Food4Me study. Br. J. Nutr. 116, 2011–2019 (2016).
19. Sparks, I. M. Exercise training response heterogeneity: physiological and molecular insights. Diabetologia 60, 2329–2336 (2017).
20. Banerjee, S., Schlaeppe, K. & van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. Nat. Rev. Microbiol. 16, 567–576 (2018).
21. Cheng, S. et al. Effect of aerobic exercise and diet on liver fat in pre-diabetic patients with non-alcoholic-fatty-liver-disease: A randomized controlled trial. Sci. Rep. 7, 15952 (2017).
22. Segata, N. et al. Metagenomic biomarker discovery and explanation. Genome Biol. 13, R62 (2012).
23. Koh, A., De Vadder, F., Kovatcheva-Datchary, P. & Bäckhed, F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell 165, 1332–1345 (2016).
24. Nielsen, H. B. et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat. Biotechnol. 32, 822–828 (2014).
25. Zhang, C. et al. Dietary modulation of gut microbiota contributes to alleviation of both genetic and simple obesity in children. eBiologyMedicine 2, 968–984 (2015).
26. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. PLoS Comput. Biol. 8, e1002687 (2012).
27. Faust, K. & Raes, J. Microbial interactions from networks to models. Nat. Rev. Microbiol. 10, 538–550 (2012).
28. Iyer, S., Killingback, T., Sundaram, B. & Wang, Z. Attack robustness and centrality of complex networks. PLoS ONE 8, e59613 (2013).
29. Strogatz, S. H. Exploring complex networks. Nature 410, 268–276 (2001).
30. Liu, X., Wang, Y., Ji, H., Ahara, K. & Chen, L. Personalized characterization of diseases using sample-specific networks. Nucleic Acids Res. 44, e164 (2016).
31. Christopherson, M. R. et al. Unique aspects of fiber degradation by the ruminal ethanologen Ruminococcus albus 7 revealed by physiological and transcriptomic analysis. BMC Genomics 15, 1066 (2014).
32. Crost, E. H. et al. Utilisation of mucin glycans by the human gut symbiont Clostridium butyricum. Proc. Natl. Acad. Sci. USA 110, 15952 (2013).
33. Lee, G. et al. Distinct signatures of gut microbiome and metabolites associated with significant fibrosis in non-obese NAFLD. Nat. Commun. 11, 4982 (2020).
34. Flint, H. J. & Duncan, S. H. In Encyclopedia of Food Microbiology (Second Edition) (eds Carl A. Batt & Mary Lou Tortorelli) 203–208 (Academic Press, 2014).
35. Shen, F. et al. Gut microbiota dysbiosis in patients with non-alcoholic fatty liver disease. Hepatobiliary Pancreat. Dis. Sci. 16, 375–381 (2017).
36. Bourrier, J. et al. The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. Hepatology 63, 764–775 (2016).
37. Lin, D. et al. Association of dietary fibre intake and gut microbiota in adults. Br. J. Nutr. 120, 1014–1022 (2018).
38. Allen, J. M. et al. Exercise alters gut microbiota composition and function in lean and obese humans. Med. Sci. Sports Exerc 50, 747–757 (2018).
39. Wagg, C., Schlaeppe, K., Banerjee, S., Kuramai, E. E. & van der Heijden, M. G. A. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. Nat. Commun. 10, 4841 (2019).
40. Strona, G. & Lafferty, K. D. Environmental change makes robust ecological networks fragile. Nat. Commun. 7, 12462 (2016).
41. Kim, H.-J. et al. Fragile skin microbiomes in megacities are assembled by a predominantly niche-based process. Sci. Adv. 4, e1701581 (2018).
42. Herren, C. M. & McMahon, K. D. Keystone taxa predict compositional change in microbial communities. Environ. Microbiol. 20, 2207–2217 (2018).
43. Kurtz, Z. D. et al. Sparse and compositionally robust inference of microbial ecological networks. PLoS Comput. Biol. 11, e1004226 (2015).
44. Liu, Y. et al. Gut microbiome fermentation determines the efficacy of exercise for diabetes prevention. Cell Metab. 31, 77–91 e75 (2020).
45. Tan, X., Alén, M., Wilkund, P., Partinen, M. & Cheng, S. Effects of aerobic exercise on home-based sleep among overweight and obese men with chronic insomnia symptoms: a randomized controlled trial. Sleep. Med. 25, 113–121 (2016).
46. Tan, X. et al. Effect of six-month diet intervention on sleep among overweight and obese men with chronic insomnia symptoms: a randomized controlled trial. Nutrients 8, (2016).
47. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–596 (2013).
48. Xing, J. et al. Hypoxia induces senescence of bone marrow mesenchymal stem cells via altered gut microbiota. Nat. Commun. 9, 2020–2020 (2018).
49. Li, J. et al. Gut microbiota dysbiosis contributes to the development of hypertension. Microbiome 5, 14 (2017).
50. Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1, 18 (2012).
51. Wishart, D. S. et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res. 46, D608–d617 (2018).

Acknowledgements

Funding to S.C. was provided by the China State Sport General Administration (2013B040, 2015B039), the National Natural Science Foundation of China (Nos. 31571219), the Shanghai Jiao Tong University Zhiyuan Foundation (CP201403), and to S.L. the Inter-disciplinary Program of Shanghai Jiao Tong University (ZH2018QNB05). Funding to L.C. was supported by the National Key R&D Program of China (No. 2017YFA0505500), Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB18000000), National Natural Science Foundation of China (Nos. 31930022, 31717147, 11202608). Funding to C.Z. was supported by the National Natural Science Foundation of China (Nos. 31930022 and 81871091). In addition to the authors, we thank Mengra Lin for contribution to the absolute microbial abundance measurement. We would like to thank the study team members and participants for their contributions to the success of this trial.

Author contributions

S.C. designed the study and has full access to all of the data in the study and takes full responsibility for the integrity of the data and for the accuracy of the data analysis. S.C., C.Z., S.L., J.G., D.L. participated in subject recruitment, measurements and intervention process. Y.Y., R.C., S.L. and X.Y. processed fecal samples and the 16 s rRNA sequence analysis. L.W. and L.C. involved in the metagenomics data analysis. X.Z. performed SCFA analysis. R.C., L.W., Y.Y. and T.X. performed the statistical analyses. L.X., P.W., C.Z., L.C. and S.C. checked data analyses and interpretation of the results. R.C., L.W. and S.C. drafted the manuscript. L.X., P.W., C.Z., L.C. and S.C. performed critical revision of the manuscript. All authors have approved the final submission version of the manuscript.

Competing interests

The authors declared no conflict of interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-29968-0.

Correspondence and requests for materials should be addressed to Chenhong Zhang, Luonan Chen or Sulin Cheng.

Peer review information Nature Communications thanks Yinghui Wei and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2022
These authors contributed equally: Runtan Cheng, Lu Wang. 
email: zhangchenhong@sjtu.edu.cn; lnchen@sibs.ac.cn; sulin.cheng@jyu.fi