Metabologenomics identified fecal biomarkers for bowel movement regulation by *Bifidobacterium longum* capsules: an RCT

Y. Nakamura,[1,2]† S. Suzuki,[2,3]† S. Murakami,[1,4]† K. Higashi,[5] N. Watarai,[2] Y. Nishimoto,[1] J. Umetsu,[2] C. Ishii,[4] Y. Ito,[4] Y. Mori,[1] M. Kohno,[6] T. Yamada,[1,2,7]* S. Fukuda [1,4,7-9]*

[1] Metabologenomics Inc., 246-2 Kakuganji, Tsuruoka, Yamagata 997-0052, Japan,
[2] Tokyo Institute of Technology, Department of Life Science and Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan
[3] Education Academy of Computational Life Science (ACLS), 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550 Japan
[4] Keio University, Institute for Advanced Biosciences, 246-2 Kakuganji, Tsuruoka, Yamagata 997-0052, Japan
[5] National Institute of Genetics, Genome Evolution Laboratory, Yata 1111, Mishima 411-8540, Japan
[6] MORISHITA JINTAN CO., LTD., Health Care Product Department, Research & Development Division, 1-2-40 Tamatsukuri, Chuo-ku, Osaka 540-8566, Japan
[7] PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan
[8] Intestinal Microbiota Project, Kanagawa Institute of Industrial Science and Technology, 3-25-13 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa 210-0821, Japan
[9] Transborder Medical Research Center, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan

† These authors contributed equally to this work.
* These authors jointly supervised this work.
Contact:
e-mail: yamada@metagen.co.jp, sfukuda@metagen.co.jp
tel: (+81) 235-64-0330

Abstract

Background: *Bifidobacterium longum* supplementation can be used to regulate bowel movement; however, individuals vary in the response to *B. longum* treatment. One putative factor is the gut microbiota; recent studies have reported that the gut microbiota mediates diet or drug effects. Here, we investigated intestinal features related to *B. longum* effectiveness in increasing bowel movement frequency.

Results: A randomized, double-blind controlled crossover trial was conducted with twenty Japanese subjects selected from 50 participants. The subjects received a two-week dietary intervention consisting of *B. longum* in acid-resistant seamless capsules or similarly encapsulated starch powder. Bowel movement frequency was recorded daily, and time-series fecal collection was conducted for metabologenomic analyses. There were differences among subjects in *B. longum* intake-induced bowel movement frequency. The responders were predictable by machine learning based on the metabologenomic features of the fecal samples collected before *B. longum* intake. Between responders and non-responders, the abundances of nine bacterial genera and of three compounds were significantly different.

Conclusions: Thus, the gut microbiome and metabolome composition have a strong impact on *B. longum* supplementation effectiveness in increasing bowel movement frequency, and gut metabologenomics enables *B. longum* supplementation effect prediction before intake. These findings have implications for the development of personalized probiotic treatments.

Trial registration: UMIN-CTR, UMIN000018924. Registered 07 September 2015, https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000021894

Keywords: bifidobacteria, probiotics, gut microbiota, 16S rRNA gene sequence, CE-TOFMS, responder analysis, machine learning

Main text

Background

Constipation is a globally common gastrointestinal disorder of people of all ages. Constipation has been reported in 5.4% to 26.2% of the general population [1–3] and to predispose individuals to a wide range of diseases, including irritable bowel syndrome (IBS), Parkinson's disease and kidney disease [4–6]. Several diets are used to relieve constipation. For example, dietary fiber and magnesium have been reported to increase bowel movement frequency [3,7–9]. In addition, in recent years, fermented products such as yoghurt have attracted attention because they are useful in relieving constipation. *Bifidobacterium* is a representative microbial genus of probiotics that is frequently used to regulate bowel movements [10]. Multiple randomized controlled trials have shown that bifidobacteria intake increases bowel movement frequency [11–14] or shortens gut transit time [15,16]. However, the response to bifidobacteria supplementation varies from subject to subject. In a study of patients with IBS, a gastrointestinal disorder that causes constipation, 47% of patients exhibited improved symptoms under *Bifidobacterium* supplementation [11].

One factor that could account for these observations is the composition of the gut microbiota. Indeed, individual differences in gut microbiota composition are known to affect probiotic-mediated stimulation [17]. Recent studies based on high-throughput sequencing of the human microbiome have elucidated how the host response to diet is affected by intestinal microbes [18–23]. For example, gut microbial dysbiosis induced by excessive consumption of artificial sweeteners impairs the host's glucose tolerance metabolism [18], and excessive dietary emulsifiers promote colitis and metabolic syndrome via the gut microbiome [19]. In addition, a study has reported that the gut microbiome mediates the effect of drugs, with variation among individuals in the strength of this effect [24]. Anticancer drugs are one of these drugs. The effect of an immune checkpoint inhibitor targeting PD-1/PD-L1 has been found to be related to the relative abundances of *Bifidobacterium longum*, *Collinsella aerofaciens* and *Enterococcus faecium* or *Akkermansia muciniphila* [25,26], and the alpha diversity and relative abundance of bacteria of the *Ruminococcaceae* family in responders have been found to be higher than those in nonresponders [27]. In addition, the inhibitory effect of the immune checkpoint on CTLA-4 has been shown to be related to the abundance of *Bacteroides* [28].

These individual differences can be predicted with some accuracy from omics data for the gut microbiome [20,21]. In inflammatory bowel disease (IBD), which is one of the most well-known gastrointestinal disorders, the response to anti-integrin biologic therapy is predictable from gut microbial function [29]. With prediction of drug or diet effects from the state of the intestinal environment, it might be possible to prescribe according to each case. In other words, the companion diagnosis based on the individual intestinal environment could enable personalized health care and medical cost reduction.

In this study, we conducted a randomized double-blind controlled crossover trial with healthy subjects to determine the effect of *B. longum* BB536 supplementation on bowel movement frequency (Figure 1A). In addition, we conducted 16S rRNA metagenomic and metabolomic analyses of fecal samples collected during the trial to evaluate the influence of *B. longum* BB536 supplementation on the intestinal environment. Furthermore, the bowel movement frequency-increasing effect of *B. longum* BB536 intake for each individual was estimated using a Bayesian statistical model.

Results

The effect of *Bifidobacterium longum* BB536 intake on bowel movement frequency varies among individuals

We confirmed that the bowel movement frequency increased during the test intervention period by the Wilcoxon signed rank test (Figure 1B, Table S1). The bowel movement frequency in the first and second weeks of the test intervention period (4.5 ± 1.9 [/week] and 5.2 ± 2.0 [/week], respectively) increased relative to that in the period before intervention (4.0 ± 1.5 [/week]; $p = 0.0323$ and $p = 0.0012$, respectively). In the second week of the test intervention period, the bowel movement frequency increased relative to that in the second week of the control intervention period (4.3 ± 1.7 [/week], $p = 0.0335$). On the other hand, the bowel movement frequency in the first week of control intervention (5.0 ± 2.2 [/week]) increased relative to that in the period before intervention (4.2 ± 1.8 [/week], $p = 0.0335$).

In this study, individual variation had observed in the increasing effect of *Bifidobacterium longum* supplementation on bowel movement frequency. Therefore, we analyzed the bowel movement frequency of each individual (Figure 1C). We found that the dynamics of bowel movement frequency over the trial period varied greatly among individuals. For example, in MO04, MO10 and MO11, bowel movement frequency tended to increase over the test intervention period, whereas in MO16, MO18, MO19, MO22, MO23 and MO24, there was little change. Similarly, variation was observed in the control intervention, with some individuals exhibiting increases in bowel movement frequency due to the placebo effect and others not.

The intestinal environment of each individual was robust for probiotic diet intervention

To determine how the microbiome composition of the subjects changed over the course of the experiment, the 16S rRNA gene region was amplified from the DNA extracted from the fecal samples and sequenced by high-throughput sequencing. Based on these sequencing data, the relative abundance, alpha diversity and beta diversity at the bacterial genus and species level were calculated. Furthermore, metabolites were extracted from the fecal samples and analyzed by CE-TOFMS. Alpha diversity representing the intra-sample complexity of the microbiome composition and beta diversity representing the inter-sample diversity were calculated from the metabolome data. No overall trends were observed in intestinal microbiome and metabolome composition in either of the two intervention groups (Figure 2A, C). The results of multidimensional scaling (MDS) analysis using beta diversity to analyze similarity between all of the samples showed that the microbiome and metabolome composition of the same subject were similar over the course of the experiment (Figure 2B, D, Figure S1A, B). This result suggests that the effects of individual differences in the intestinal microbiome and metabolome composition were larger than the effect of the intervention.

Subsequently, the Brunner-Munzel test (two-sided test) was carried out to compare the changes in gut microbes and metabolites among the test intervention group (T2, T3), the control intervention group (C2, C3), and the nonintervention group (T1, C1). The results showed that some of the bacterial genera were different in the test intervention group from those in the other groups (p value < 0.05, not corrected). However, when false discovery rate (FDR) correction was applied, no changes were detected (Table S2), and no significant change was detected in alpha diversity (Table S3, Figure S2). The relative abundance of *B. longum* tended to increase during the test intervention period and varied among individuals (Figure S3). Overall, our results demonstrate that the influences of *B. longum* BB536 intake on the intestinal microbiome and metabolome were small relative to the influence of individual differences in intestinal environment and that the intestinal environment of each individual was robust for probiotic diet intervention.

Determining bowel responder by Bayesian Weibull regression

The extent of increase in bowel movement frequency due to *Bifidobacterium longum* supplementation or placebo effect varied among subjects (Figure 1C). To test whether the microbiome or metabolome features of an individual's intestinal environment varied with the magnitude of response, we first quantified the magnitude of the treatment effect in each individual using a Bayesian statistical model. The models were formulated to estimate the sizes of placebo effect and the test supplement effect precisely. In the models, the bowel movement interval of each individual was used based on the fact that bowel movement tends to occur with the lapse of time. The bowel movement intervals were calculated from the daily frequency of bowel movements. The parameters for the models were estimated by Bayesian statistical fitting of bowel movement interval to the models. Subject MO03 was excluded from further analysis because this subject's bowel movements consistently occurred once a day throughout the whole observation period (85 days).

To investigate individual differences, we prepared several models, some of which assumed a constant treatment effect across individuals and some assuming variation among individuals (Table S4). The models were evaluated by Widely Applicable Information Criterion (WAIC) [30]. The lower the WAIC, the better the model explains the data without overfitting. The WAIC is an index of prediction performance for unknown data. Among the models, the model considering the individual placebo effect and the test supplement effect had the lowest WAIC (Table S4). These results indicated that some subjects had increasing bowel movement frequencies as a result of intake of the test supplement; we defined these subjects as "bowel responders."

Using the model with the best WAIC, the estimated posterior mean of each parameter and its Bayesian credible intervals were calculated for each individual (Table S5, Figure 3). The effects of the test supplement in subjects MO04, MO05 and MO10 were strong as evidenced by the 95% Bayesian credible intervals of the predicted distribution, which did not overlap zero. Therefore, MO04, MO05 and MO10 were defined as strong responders (SRs). In addition, there were 9 other subjects with estimated posterior mean values greater than zero, who were defined as weak responders (WRs). The remaining subjects were defined as nonresponders (NRs).

The abundances of fifteen genera and three metabolites differed between bowel responders and nonresponders as determined by Bayesian statistical model

We then calculated fold changes for each bacterial genus and metabolite abundance for the SR and WR groups and compared them to those for the NR group (FCs and FCw, respectively). To detect features with large or small relative abundances when the effect of the test supplement was strong, the genera and metabolites with fold change values following the order FCs > FCw > 1 or FCs < FCw < 1 were extracted, and the trends were verified using the Jonckheere-Terpstra trend test. Increasing trends were detected for five genera, and decreasing trends were detected for six genera and one metabolite (p value < 0.05) (Figure 4A). For these features, the Brunner-Munzel test was used to compare both responder groups (SRs and WRs) with the NRs group. The relative abundances of g__*Ruminococcus*_E, g__Unassigned85 (g__*Faecalibacterium* related genus), g__CAG-274 (o__Lachnospirales), g__Unassigned207 (o__Clostridiales) and g__Unassigned459 (g__*Roseburia* related genus) were significantly higher in the responder groups than in the NR group (p value < 0.05) (Figure 4B). Conversely, g__*Agathobacter*, g__CAG-24 (f__*Lachnospiraceae*), g__UBA1191 (f__*Anaerovoracaceae*), g__Unassigned66 (g__*Faecalibacterium* related genus), g__Unassigned441 (g__*Lachnoclostridium* related genus) and g__Unassigned390 (g__*Emergencia* related genus) were significantly less abundant in the responder groups than in the NRs group, and the ribulose 5-phosphate level was significantly lower in the responders (p value < 0.05) than in the NRs (Figure 4B, C). Taken together, these results revealed the bowel responders who had a bowel movement frequency-increasing effect due to *B. longum* BB536 intake and clarified that fifteen genera and three metabolites had significantly different abundances in responders compared to those in NRs.

Machine learning enabled us to predict bowel responders from metabologenomic data before probiotic intake

In parallel with our experimental observations of the intestinal environmental factors that determine the response to supplementation with *B. longum* BB536, we also investigated whether responders could be predicted from intestinal environmental features before *B. longum* BB536 intake by a machine learning approach. To test this strategy, we used the machine learning method combined with LASSO regression and logistic regression and predicted whether a subject would be a responder from the metabologenomic data before the test intervention (Figure 5A). The composition of bacterial genera, bacterial species, metabolites or their combinations were used as explanatory variables for responder prediction, and we conducted parameter estimation and performance evaluation in machine learning using a grid search and stratified cross validation. For the stratified cross validation, we chose 7-fold stratified cross validation to use additional training data, considering that the 7 subjects were NRs. As a result, it was possible to predict all responders (SRs and WRs), and the highest performance (area under the receiver operating characteristic curve (AUROC) = 0.810) was observed when using metabolite and bacterial genus data (Figure 5B, Table 1). Responders could also be predicted with high performance (AUROC = 0.766) using metabolite and bacterial species data, which provided more detailed information than did bacterial genera data (Table 1). LASSO regression using bacterial genus and metabolite data with a parameter set for the highest AUROC showed that the g__*Agathobacter*, g__Unassigned66 (g__*Faecalibacterium* related genus), g__*Alistipes*, g__*Mitsuokella*, g__*Parabacteroides*, g__CAG-110 (f__*Ruminococcaceae*), g__*Fusicatenibacter*, g__*Hungatella*, g__UBA1394 (f__*Ruminococcaceae*), g__CAG-65 (f__*Lachnospiraceae*), g__*Sutterella*, g__*Clostridium*_AJ, g__*Lachnospira*, g__*Sellimonas*, g__Unassigned14 (g__*Ruminococcus* related genus), g__*Agathobaculum*, g__Unassigned85 (g__*Faecalibacterium* related genus), g__*Eubacterium*_G, g__Unassigned11 (g__*Ruminococcus* related genus) and g__*Ruminococcus*_E levels were involved in the prediction of responders (Figure 5C). While the exact mechanisms underlying the B. longum-induced effects of the gut microbiota on bowel movement frequency have yet to be elucidated, our machine learning approach allowed us to predict the patients who would respond to *B. longum* BB536 intake with an increased bowel movement frequency.

Discussion

The bowel movement responses to probiotic intake vary among individuals [11], and the gut microbiota is a key factor that can account for these differences [17]. Thus, in this study, we accurately quantified individual differences in the increase in bowel movement frequency in response to intake of *B. longum* BB536 contained in acid-resistant seamless capsules by a statistical model and defined the responders as those whose bowel movement frequency increased noticeably. In addition, the machine learning analysis revealed that responders could be predicted from features of the intestinal environment before intake of the test supplement began, and the predictive performance was improved by using both microbiome and metabolome data. These results suggest that accurate quantification of individual response intensity and machine learning predictions may enable companion diagnostics for response to diet based on the intestinal environment.

Our finding that using both microbiome and metabolome data improved the predictive performance of machine learning shows that a metabologenomic approach combining metagenomics and metabolomics may be important for analyzing the response depending on the individual intestinal environment. In recent years, multiple studies have predicted responses to intervention from metagenomic datasets alone, including datasets of taxonomic composition [22,23,31]; incorporating metabolome data may enhance predictive performance. In addition, previous studies have reported that, in addition to increasing bowel movement frequency, the intake of the *B. longum* BB536 strain encapsulated in an acid-resistant seamless capsule results in a fecal color change from dark to light, a fecal shape change from a hard or watery state to a soft state and a decrease in the amount of ammonia in feces [14,32]. If individual responses with these effects were accurately quantified, then metabologenomic data may be able to predict which subjects will have these responses, as well as which will have increased bowel movement frequency. In addition, several studies have suggested that individual responses to the intake of other dietary factors could be predicted from the intestinal environment [20,21,33]. From these cases, it is expected that it will be possible to predict the effect of probiotics from the intestinal environment before intake in the future, and the development of personalized health care and medical businesses targeting the intestinal environment are expected by using this prediction system.

We established an accurate Bayesian statistical model to estimate the bowel movement frequency-increasing effect due to *B. longum* BB536 intake, and this model revealed which subjects were responders or NRs in this study. This model enabled us to detect shortening of the bowel movement interval due to the intake of *B. longum* BB536, which is consistent with previous research [15,16]. In addition, because Bayesian estimation of the parameters enabled us to calculate a correct interval, this model achieved accurate quantitative evaluation of the effect of *B. longum* BB536 intake on bowel movement frequency, which has been difficult to achieve by threshold-based responder estimation [21].

Few significant differences in the microbiome and metabolome were detected among the baseline, *B. longum* BB536 intake and control supplement samples. This lack of significant differences in the intestinal environment is related to interindividual variation, which masks the effect of the intake of *B. longum* BB536. This finding has also been reported in several studies that have observed changes in the gut microbiome composition by probiotics or diets. Thus, this effect is because the intestinal environment of each individual was robust and individual differences in the intestinal environment override the effect of probiotic intake.

Our observational and machine learning results indicated that a higher or lower relative abundance of 27 genera was associated with whether a subject responded well to *B. longum* BB536 intake, and intriguingly, 23 of these 27 responder-related genera belonged to the class Clostridia. Clostridia is an important bacterial group that produces mainly short-chain fatty acids (SCFAs), including butyrate, in the gut environment [34]. An imbalance of SCFAs may cause IBD and IBS because of the involvement of SCFAs in the maintenance of intestinal homeostasis [35–37]. Clostridial clusters IV and XIVa in Clostridia are reported to produce butyrate, which is related to the promotion of large intestinal peristaltic movement [34]. Among the 23 genera suggested as possibly involved in the bowel responders in this study, nine *Ruminococcaceae* bacteria belong to Clostridial cluster IV, and ten *Lachnospiraceae* bacteria belong to Clostridial cluster XIVa. The bacteria in Clostridial clusters IV and XIVa cross-feed with *Bifidobacterium* and metabolize the inulin-type fructans derived from *Bifidobacterium* to produce butyrate [38–40]. We therefore suggest that the balance of bacteria

belonging to Clostridial clusters IV and XIVa may be related to the increase in bowel movement frequency in response to intake of *B. longum* BB536 via butyrate or other SCFAs and that researching the cross-feeding relationship of these bacteria might lead to a detailed understanding of this relationship.

Our results also suggest that butyrate may play a role in the response to *B. longum* BB536 intake. Compared with NRs, responders tended to show an increase in the abundance of several butyrate-producing bacteria during the test intervention period (Figure 4). This result also supports a relationship among the gut microbiota, butyrate and an increase in bowel movement. Furthermore, it has been reported that the efficiency of butyrate production in the sugar metabolism of these bacteria varies among bacterial species and habitat. Previous studies have indicated that the sugar-metabolizing abilities of *Ruminococcus bromii*, *Eubacterium rectale* (*Agathobactor rectale*), *Bacteroides thetaiotaomicron*, *Bifidobacterium adolescentis*, *Anaerostipes caccae* and *Eubacterium hallii* vary with habitat [39,41]. In addition, *Ruminococcus* and *Bifidobacterium* are localized in the intestinal lumen, and *Clostridium* is localized in the intestinal lumen and on the intestinal epithelial cell surface, whereas *E. rectale* preferentially colonizes the mucous layer to produce and promote the utilization of butyrate by colonic epithelial cells [34,42]. Taken together, these findings suggest that variation in butyrate productivity among bacteria and their locations in the large intestine may influence whether a subject responds to *B. longum* BB536 intake.

In our study, the most valuable points are the clinical trial with a double-blind placebo-controlled design and multi-omics features to define responsiveness to probiotic intervention; however, two limitation points should be considered. First, 16S metagenomic data were used to evaluate the gut microbiome. Since 16S metagenomic data reveal only the taxonomic composition and cannot be used for detailed functional analysis, it is necessary to conduct a future study with shotgun metagenomics to clarify the hypothesis proposed in this study. The other limitation is that the final number of subjects was 20. In this study, acquiring the time-series data of each individual compensates for that; however, in the future, the results need to be validated by a larger cohort study. Although no such large cohort has been conducted for the effect of probiotics on bowel movement frequency, the expectations and demands for such a study have risen considerably.

Conclusions

In summary, our results show that the gut microbiota composition has a strong impact on the effectiveness of *B. longum* BB536 supplementation in increasing bowel movement frequency, and the gut metabologenomic data enable us to predict the effect of *B. longum* BB536 supplementation before intake. These findings have interesting implications for personalized treatment of chronic constipation. In addition, these findings suggest that the influence of the intestinal environment on the response to intervention should be considered not only for probiotics but also for all health effects, adverse effects and side effects for all oral intake interventions. When the effect of probiotics, diets or drugs is dependent on the intestinal environment before intake, as it is in the present study, the demand for intestinal environment alterations to enhance the effect may also increase. Furthermore, the development of personalized health care and medical businesses targeting the intestinal environment are expected to use this intervention effect quantification and responder prediction system.

Methods

Trial design and recruitment

In this study, we conducted a randomized double-blind controlled crossover trial with healthy subjects from September to December 2015 to quantify the increase in bowel movement frequency in response to intake of acid-resistant seamless capsules containing *Bifidobacterium longum* BB536. In addition, we analyzed the microbiome and metabolome profiles of fecal samples collected during the trial period (Figure 1A, Figure S4, Table S6) (see also Supplementary Protocol).

The sample size was estimated based on a previous randomized controlled trial using the probiotic strain *B. longum* BB536 as the test supplement [43]. In this previous trial, the bowel movement frequency during the test intervention period was 9.6 ± 0.6 [/ 2 weeks], and that during the control intervention period was 8.4 ± 0.4 [/ 2 weeks]. We applied power analysis to estimate the number of subjects with a significance level of 5%. The statistical power was 80%, with the expectation that 20% of subjects would drop out in this trial. From these conditions, the number of participants required for this trial was estimated to be eight.

A preliminary bowel movement test was conducted for 50 Japanese participants, and 24 subjects who fulfilled the criteria for age, male-female ratio and bowel movement status were selected before the main trial from the preliminary results. Four out of 24 subjects were not used in the subsequent analysis, as they violated the protocols (by taking medication or failing to collect fecal samples). In these twenty subjects, the average age was 47.7 ± 5.8, and the male/female ratio was 8:12. The remaining baseline demographic and clinical characteristics did not significantly differ between the control and treatment groups (Table S7).

Subject characteristics and exclusion criteria

This clinical study was approved by the ethics committee and then requested to CPCC Co., Ltd. All of the participants consented to participate in this study after the study was fully explained to them.

The subject inclusion criteria were as follows: 1) aged 40 to 60 years at the time of providing consent and 2) stool frequency of three to five times per week or more than seven times per week. The exclusion criteria were as follows: 1) laparotomy surgery performed within six months before the start of the study, 2) intake of antibiotics for more than one week within six months before the start of the study, 3) allergies to the test supplement, 4) significant lifestyle changes intended during the examination period, 5) susceptibility to chronic diarrhea, 6) history of significant liver dysfunction, gastric dysfunction, or cardiovascular disease, 7) suspected chronic or acute infection, 8) known or potential pregnancy or breastfeeding, 9) participation in other trials within the month before the start of the study and 10) otherwise assessed to be inappropriate for the study by the physician involved in the study. All subjects were monitored for defecation frequency for two weeks before the start of the study, and it was confirmed that the defecation frequency was either three to five times per week or seven times per week or more. Among the subjects who completed the study, those who met any of the following conditions were excluded from further analysis: 1) considered to have met one or more of the exclusion criteria, 2) an inoculation rate of the test supplement less than 80%, 3) large variation in the diet record or lack of confirmation of no change from the record, 4) large variation in the content of the lifestyle record or lack of confirmation of no change from the record, 5) continued or repeated ingestion of medicines, foods or supplements that might influence the outcomes of the trial and 6) otherwise assessed to be inappropriate for the study.

Trial intervention: dietary information

In this trial, we used acid-resistant seamless capsules containing approximately $5.0 * 10^9$ *B. longum* BB536 (0.53 g per serving) as the test supplement (MORISHITA JINTAN CO., LTD, Japan). These capsules are spherical with a diameter of approximately 2.4 mm and consist of two types of acid-resistant layers. It has

been found that approximately 90% of encapsulated *B. longum* BB536 survive for two hours in artificial gastric fluid adjusted to pH 1.2 [12]. The acid-resistant pH-dependent disintegrating membrane, which is the outermost layer of the capsule, passes through the stomach and collapses as the surrounding pH becomes neutral in the small intestine. Subsequently, the hardened fat layer, which is an intermediate layer, is dissolved by the surface-active action of bile acid, the digestive action of lipase and physical stimulation by intestinal motility, and then the internal bacteria are released. The released bacteria have been reported to grow in the intestine [44,45]. For the control supplement, the same capsule containing only starch was used (MORISHITA JINTAN CO., LTD, Japan). The energy, protein, lipid, carbohydrate, and sodium contents were adjusted to be equivalent between the test supplement and the control supplement. In addition, the appearances of the test and control supplements were processed to be indistinguishable.

Trial intervention: randomization and blinding

Randomization in the trial was performed by a blocked stratified randomization method. First, the 24 subjects who passed the selection criteria were assigned to two groups (group A, group B) by stratification, taking into consideration age, male-female ratio and bowel movement status before the trial period. Subsequently, the symbol “A” or “1”, representing either the test or control supplement, respectively, was randomly assigned to each group of subjects, and a test supplement assignment table with the test supplement symbol and the subject identification code was prepared. Immediately after the test supplement was assigned to one group, the table was sealed and maintained secure until the end of the study such that most of the investigators were blinded to the group assignments. Each subject orally took one serving of the test or control supplement per day with as much water as they desired. One kind of intervention was taken for two weeks, and then the other kind was taken for two weeks after a four-week washout period. The capsules were stored at room temperature by the subjects until intake.

Trial outcomes and sample collection

The primary outcomes in this study were bowel movement frequency and the features of the intestinal environment (intestinal microbiome and metabolome). The bowel movement frequency data as well as data on other activities were acquired by a questionnaire that was completed once a day (Table S8). The questionnaire included life status, diet and bowel movement status. All subjects were prohibited from consuming 1) beverages or foods containing lactic acid bacteria or bifidobacteria in large amounts, 2) beverages or foods rich in dietary fiber or oligosaccharides, 3) dietary supplements, 4) functional yoghurt, and 5) fermented soybeans (natto), and each participant noted on the questionnaire if any of these items had been consumed. The features of the intestinal environment were obtained from the stool specimen collected one week before the intervention, during the second week of the intervention period, and one week after the intervention. Feces were collected by each subject at home using a collection sheet (Nagaseru; OZAX, Japan) laid in the toilet bowl and a feces collection tube (feces container 54 x 28 mm; SARSTEDT AG & Co, Germany). After collection, the stool samples were stored immediately in the household freezer and then collected later and transported to the laboratory by refrigerated transport.

DNA extraction and 16S rRNA gene analysis

We used the protocol described in a previous report [46] for DNA extraction from stool samples. From the extracted DNA samples, the 16S rRNA gene region was amplified using the universal primers 27Fmod and 338R for the V1-V2 region of the bacterial 16S rRNA gene [47]. For sequencing of the amplicon DNA, a MiSeq platform (Illumina, USA) was used in paired-end mode with 600 cycles. The obtained 16S rRNA gene sequence data are available in the DDBJ DRA (DRA accession number: DRA006874).

For 16S rRNA gene analysis, the sequenced forward and reverse reads of each sample (average number of reads: 44,441 ± 5,959) were merged using vsearch version 1.9.3 (options: --fastq_maxee 9.0 --fastq_truncqual 7 --fastq_maxdiffs 300 --fastq_maxmergelen 330 --fastq_minmergelen 280) [48]. Subsequently, we obtained high-quality reads by quality filtering using fastp with the default parameters [49]. High-quality read FASTQ files were converted to FASTA format using SeqKit [50]. The final set of high-quality reads was used for

identifying exact sequence variants (ESVs) using the Deblur pipeline [51]. In the pipeline, artefact sequences including PhiX were removed, and de novo chimaera sequences were analyzed and removed. All ESVs were aligned to the Genome Taxonomy Database (GTDB) version 86 [52] using BLAST version 2.8.1 [53], and 99% identity was used as the threshold for the species level [54], whereas 94% identity was used for the genus level [55]. The correspondence table of ESVs and phylogenetic information and the number of reads in each step are shown in Table S9 and S10.

Metabolite extraction and analysis

Metabolome analysis was conducted as described previously with some modifications (PMID: 25525179, 26798400). To extract metabolites from feces, samples were lyophilized using a VD-800R lyophilizer (TAITEC) for at least 24 h. Freeze-dried feces were disrupted with 3.0 mm zirconia beads by vigorous shaking (1,500 rpm for 10 min) using a Shake Master NEO homogenizer (Biomedical Science). A total of 500 µl of methanol containing the internal standards (20 µM each of methionine sulfone and D-camphor-10-sulfonic acid (CSA)) was added to 10 mg of disrupted feces. Samples were further disrupted with 0.1 mm zirconia/silica beads by vigorous shaking (1,500 rpm for 5 min) using the Shake Master neo. Next, 200 µl of ultrapure water and 500 µl of chloroform were added before centrifuging at 4,600×g for 15 min at 20°C. Subsequently, 150 µl of the aqueous layer was transferred to a centrifugal filter tube (UltrafreeMC-PLHCC 250/pk for Metabolome Analysis, Human Metabolome Technologies) to remove protein and lipid molecules. The filtrate was centrifugally concentrated and dissolved in 50 µl of ultrapure water immediately before CE-TOFMS analysis. The measurement of extracted metabolites in both positive and negative modes was performed by CE-TOFMS. All CE-TOFMS experiments were performed using the Agilent CE capillary electrophoresis system (Agilent Technologies). Annotation tables were produced from measurements of standard compounds and were aligned with the datasets according to similar values and normalized migration times. Then, peak areas were normalized against those of the internal standards, which were methionine sulfone and CSA for cationic and anionic metabolites, respectively. Concentrations of each metabolite were calculated based on the relative peak areas and the concentrations of standard compounds. The obtained metabolome data are available in Table S11 and 12.

Bioinformatics and statistical analyses

All statistical analyses were performed using Python. The Shannon diversity index was used for alpha diversity, and the Spearman correlation distance was used for beta diversity. MDS was calculated using beta diversity. A two-sided Wilcoxon signed rank test was used to compare bowel movement frequency data between the two groups. A two-sided Brunner-Munzel test was used for comparing metabologenomic data between the two groups. A two-sided Jonckheere-Terpstra trend test was used to verify the trend in metabologenomic data among the groups. The Benjamini-Hochberg false discovery rate correction (FDR) method was used for multiple test correction. For gut microbiome analysis, taxonomic composition data in the genus, species, and ESV levels were used. Of these, ESV-level data were used only to calculate the alpha diversity. The genus-level data were used for the other statistical analyses. The species-level data were used only for the prediction of bowel responders. For gut metabolome analysis, relative area compared with that of an internal standard and content of metabolite data was used. The metabolite relative area data were used only to calculate the alpha diversity. The metabolite content data were used for the other statistical analyses.

Bowel responder definition based on Bayesian Weibull regression

Using the data from the questionnaires recorded by each subject over the course the study, subjects who strongly responded to treatment, i.e., who showed an increased bowel movement frequency, were estimated using a statistical model. An increasing effect of the test supplement on bowel movement frequency was defined as a shortened interval between bowel movements; however, other measures of bowel movement frequency or bowel movement probability were also evaluated.

Paying attention to the fact that bowel movement tends to occur with the lapse of time, the bowel movement interval of each individual was modeled by the Weibull distribution. The Weibull distribution is widely used to investigate the effects of drugs on survival time, and it has a shape parameter and a scale parameter [56]. The effect of intake itself was estimated based on a proportional hazards model, assuming that intake influences the scale parameter. The corresponding mathematical formulas are shown below.

$$I_{s,i} \sim \mathrm{Weibull}(k_s, \lambda_{s,i})$$

$$k_s \sim \mathrm{HalfNormal}(0, \sigma_k)$$

$$\lambda_{s,i} = \begin{cases} \exp(gut_state_s + \beta_{s,\mathrm{placebo}} + \beta_{s,\mathrm{test}}) & \text{(if eating } \textit{Bifidobacterium}\text{)} \\ \exp(gut_state_s + \beta_{s,\mathrm{placebo}}) & \text{(if eating placebo supplement)} \\ \exp(gut_state_s) & \text{(else)} \end{cases}$$

$$gut_state_s \sim \mathrm{Normal}(0, \sigma_{gut_state})$$

$$\beta_{s,\mathrm{placebo}} \sim \mathrm{Normal}(0, \sigma_{\mathrm{placebo}})$$

$$\beta_{s,\mathrm{test}} \sim \mathrm{Normal}(0, \sigma_{\mathrm{test}})$$

s : subject index

i : stool index

$I_{s,i}$: time interval of ith stool of subject s

k_s : shape parameter of subject s

$\lambda_{s,i}$: scale parameter of ith stool of subject s

σ_k : scale parameter of shape parameter k

gut_state_s : personal variable parameter to explain normal gut status

$\beta_{s,\mathrm{placebo}}$: personal variable parameter to explain intake effect occur in both *Bifidobacterium* and controll suppleme

$\beta_{s,\mathrm{test}}$: personal variable parameter to explain intake effect occur in only *Bifidobacterium* supplement

σ_{gut_state} : scale paramter of gut_state

$\sigma_{\mathrm{placebo}}$: scale paramter of $\beta_{s,\mathrm{placebo}}$

σ_{test} : scale paramter of $\beta_{s,\mathrm{test}}$

In this model, it is assumed that the effect of each variable at time t occurs as a product of the hazard function in the basis state. The hazard function is a function that expresses the probability of occurrence of an event after time lapse Δt from time t. In this study, it was assumed that the effect of the variable accompanying the bowel movement time, such as the influence of the intake of the test supplement, occurs as a product of the hazard function in the normal state and that the occurrence probability of bowel movement changes. This assumption is equivalent to the assumption that factors other than intake do not influence the hazard function throughout the observation period. This assumption is supported by the empirical observation that the condition in the intestine is stable in a state where food intake is strictly restricted, as it is in the cohort of the present study. Moreover, the parameters in the Weibull distribution in the model and the effect of intake are unknown. Such values were estimated from the observed data. The record of bowel movements acquired in this study did not include the time of each movement but reported the number of bowel movements per day. Therefore, the interval between successive bowel movements was estimated by dividing 24 h by the frequency. One subject (MO03) consistently had one bowel movement a day throughout the entire observation period of 85 days; thus, the bowel movement interval was always 24 h. Since a variable with zero variance is

inappropriate for model estimation, this subject was excluded from subsequent analysis. Furthermore, multiple models were constructed and compared using the Widely Applicable Information Criterion (WAIC) [30]. WAIC approximates the prediction error (generalization error) for unknown data of the model; WAIC can also be used for a model in which the posterior distribution of parameters cannot be approximated by a normal distribution [57].

Parameter estimation was performed by a Hamiltonian Monte Carlo method based on NUTS algorithm using Python and Stan [58]. A generalized linear model of the Weibull distribution using Markov Chain Monte Carlo (MCMC) was applied following a previous study [59]. The number MCMC iterations was 20,000, the number of chains was 8, and the first 1,000 steps in each iteration were discarded as burn-in. In a previous study, MCMC convergence was confirmed by a potential scale reduction factor (PSRF or Rhat) less than 1.1 for all estimates [60]. Similarly, in the present study, Rhat was calculated based on the variance of multiple MCMC chains and was used for convergence determination. Initial values were used for the other parameters. The implemented software is available from https://github.com/metagen/EEBIIC.

Other defecation responder estimation models

The defecation dynamics were modeled as defecation frequency and defecation probability. Regarding defecation frequency, it was assumed that the number of defecations per day for each individual followed a Poisson distribution. The Poisson distribution is used when modeling count data and has one ratio parameter. The effect of ingestion was estimated assuming that the ratio parameter changes according to ingestion. For this method, the questionnaire-based defecation frequencies obtained from each subject were used.

Regarding defecation probability, it was assumed that occurrence of defecation per day of each individual can be modeled by the Bernoulli distribution. The Bernoulli distribution is used in expressing the occurrence probability of alternative choices, such as the side of a coin, and has one probability parameter. The effect of ingestion was estimated assuming that the probability parameter changes through a logistic function according to ingestion. For this method, the defecation frequency data were converted to presence/absence data of defecation each day, with 1 recorded if a bowel movement occurred and 0 otherwise.

Prediction of responders by machine learning

We used the relative abundance of bacterial species or genera and the fecal metabolite content before the test intervention (T1) of 19 subjects, i.e., all subjects except MO03, as feature values. MO03 was excluded from this analysis because the bowel movements of this subject always occurred once a day throughout the whole observation period (85 days), which was not suitable for determining responders by the Bayesian statistical model. Of all feature values, those of the relative abundance or the fecal content exceeding a certain threshold were selected. For the threshold value, a grid search was conducted to search for the one with the maximum predicted AUROC. Subsequently, each feature value was standardized by centering to a mean of 0 and dividing by the standard deviation (z-score) of each feature, considering the following potential problems: 1) it is possible that a feature value with a large range of possible values will have a large influence on the prediction because the range of values that can be taken with the microbial relative abundance and metabolite contents are different; and 2) the feature importance of the microbiome and metabolome cannot be directly compared. Nineteen subjects were classified into two groups, responders (SRs and WRs) and NRs. To identify metabologenomic markers that can distinguish responders from NRs, we constructed classification models based on metabologenomic data using LASSO regression and logistic regression algorithms. The model was validated by 7-fold stratified cross-validation testing (we resampled dataset partitions 100 times). A LASSO regression model was used for feature selection from the training dataset, which was fitted to the mean value of the increase in bowel movement frequency with the test supplement for each individual estimated by a statistical model. Only the features with nonzero coefficients were extracted. Regarding the LASSO regression hyperparameters, a grid search was performed to search for the parameter set with the maximum predicted AUROC. Test data were predicted by a logistic regression model that learned from feature-selected

training data. For a hyperparameter in logistic regression, we used the default parameters in sklearn.linear_model.LogisticRegression (version 0.19.1).

Z-score-standardized features, which were bacterial genera of relative abundance > 0.001 or metabolites present at > 10,000 [nmol/g] in fecal content before test supplement intake and were the conditions with the highest prediction AUROC values in the cross validation, were used for LASSO regression (alpha = 0.001). As a result, the regression coefficients of each bacterial genus and metabolite were calculated, and the features contributing to the responder prediction were extracted.

List of abbreviations

IBS: irritable bowel syndrome
IBD: inflammatory bowel disease
MDS: multidimensional scaling
FDR: false discovery rate
WAIC: Widely Applicable Information Criterion
SRs: strong responders
WRs: weak responders
NRs: nonresponders
AUROC: area under the receiver operating characteristic curve
SCFAs: short-chain fatty acids
ESVs: exact sequence variants
CSA: D-camphor-10-sulfonic acid
MCMC: Markov Chain Monte Carlo
PSRF: potential scale reduction factor

Declations

Ethics approval of research: The human rights of the subjects who participated in this study were protected at all times, and the study observed the Helsinki Declaration and Ethical Guidelines on Epidemiological Research in Japan referring to cases concerning standards for clinical trials of drugs. This randomized controlled trial was conducted with the approval of the clinical trial ethics review committee of Chiyoda Paramedical Care Clinic.
Patient consent: Obtained.
Clinical trial registration: UMIN-CTR, UMIN000018924. Registered 07 September 2015, https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000021894
Data and materials availability: The raw sequencing dataset reported in this paper is available in the DDBJ Sequence Read Archive (DRA) repository (DRA006874; https://ddbj.nig.ac.jp/DRASearch/submission?acc=DRA006874). The metabolomic datasets by CE-TOFMS are included within the article and its additional files.
Competing interests: The authors declare no conflict of interests.
Funding: None declared.
Author contributions: S.F. and T.Y. conceived and designed the project. M.K., T.Y. and S.F. designed the clinical trial. M.K. prepared the test supplement used for intervention. S.M., C.I. and Y.I. extracted DNA from fecal samples. C.I. and Y.I. performed the microbiome analysis. S.M., C.I. and Y.I. performed the CE-TOFMS-based metabolome analysis. Yuya N. conducted statistical analysis. All analyses related to metabologenomic data were performed by Yuya N. Shinya S. performed responder classification by Bayesian modeling. Yuya N. predicted responders from metabologenomic data by machine learning. Y.M. performed the PCR amplification of the 16S rRNA region of *Bifidobacterium*. Yuichiro N., N.W., K.H. and J.U. contributed to the study design. Yuya N. and Shinya S. constructed figures and described the thesis. All authors reviewed and approved the final version of the paper.
Acknowledgments: We acknowledge Shuji Suzuki for his insight and feedback regarding responder prediction by machine learning. We would like to thank MORISHITA JINTAN CO., LTD, and CPCC Co., Ltd., who conducted the clinical trial. We also thank the Tokyo Institute of Technology Yamada laboratory and Metabologenomics, Inc., staff, who participated in discussion of this research.
Authors' information:

 Name Dr. Takuji Yamada	**Biography** Dr.Takuji Yamada received his doctorate (Doctor of Science) in 2007 from Kyoto University. That year, he was appointed Assistant Professor at the Institute for Chemical Research, Kyoto University. He was then a postdoctoral fellow (2008-2010) and a technical officer (2010-2012) at the European Molecular Biology Laboratory (EMBL). In 2012, he joined Tokyo Institute of Technology as Associate Professor in the Department of Biological Information. Since 2016, he has been Associate Professor in the Department of Life Science and Technology, School of Life Science and Technology, Tokyo Institute of Technology.
 Name Dr. Shinji Fukuda	**Biography** After obtaining his Ph. D. (Gut Microbiology) at Graduate School of Agriculture, Meiji University in 2006, Dr. Shinji Fukuda undertook several research positions at RIKEN. He was then appointed to Project Professor of the Institute for Advanced Biosciences, Keio University in 2019. Currently, he is also holding other posts as visiting professor in University of Tsukuba School of Medicine from 2016; group leader of Intestinal Ecosystem Regulation Group in KISTEC-KAST from 2017; Co-PI of ERATO in JST from 2019; and visiting professor in University of Technology Malaysia from 2019. He founded and is President CEO of Metabologenomics, Inc. from 2015.

References

1. Chu H, Zhong L, Li H, Zhang X, Zhang J, Hou X. Epidemiology Characteristics of Constipation for General Population, Pediatric Population, and Elderly Population in China [Internet]. Gastroenterology Research and Practice. 2014 [cited 2019 Mar 26]. Available from: https://www.hindawi.com/journals/grp/2014/532734/

2. Suares NC. Prevalence of, and risk factors for, chronic idiopathic constipation in the community: systematic review and meta-analysis. The American journal of gastroenterology [Internet]. 2011 [cited 2019 Mar 26];106. Available from: insights.ovid.com

3. Murakami K, Sasaki S, Okubo H, Takahashi Y, Hosoi Y, Itabashi M. Association between dietary fiber, water and magnesium intake and functional constipation among young Japanese women. European Journal of Clinical Nutrition [Internet]. 2007 [cited 2019 Mar 26];61:616–22. Available from: https://www.nature.com/articles/1602573

4. Forootan M, Bagheri N, Darvishi M. Chronic constipation. Medicine [Internet]. 2018 [cited 2019 May 10];97. Available from: insights.ovid.com

5. Stirpe P, Hoffman M, Badiali D, Colosimo C. Constipation: an emerging risk factor for Parkinson's disease? European Journal of Neurology [Internet]. 2016 [cited 2019 May 10];23:1606–13. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ene.13082

6. Sumida K, Molnar MZ, Potukuchi PK, Thomas F, Lu JL, Matsushita K, et al. Constipation and Incident CKD. JASN [Internet]. 2017 [cited 2019 May 10];28:1248–58. Available from: https://jasn.asnjournals.org/content/28/4/1248

7. Portalatin M, Winstead N. Medical Management of Constipation. Clinics in Colon and Rectal Surgery [Internet]. 2012 [cited 2019 Mar 26];25:012–9. Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0032-1301754

8. Bothe G, Coh A, Auinger A. Efficacy and safety of a natural mineral water rich in magnesium and sulphate for bowel function: a double-blind, randomized, placebo-controlled study. Eur J Nutr [Internet]. 2017 [cited 2019 Mar 26];56:491–9. Available from: https://doi.org/10.1007/s00394-015-1094-8

9. Abdullah MMH, Gyles CL, Marinangeli CPF, Carlberg JG, Jones PJH. Dietary fibre intakes and reduction in functional constipation rates among Canadian adults: a cost-of-illness analysis. 1 [Internet]. 2015 [cited 2019 Mar 26];59. Available from: https://foodandnutritionresearch.net/index.php/fnr/article/view/773

10. Simrén M, Barbara G, Flint HJ, Spiegel BMR, Spiller RC, Vanner S, et al. Intestinal microbiota in functional bowel disorders: a Rome foundation report. Gut [Internet]. 2013 [cited 2018 Dec 5];62:159–76. Available from: https://gut.bmj.com/content/62/1/159

11. Guglielmetti S, Mora D, Gschwender M, Popp K. Randomised clinical trial: Bifidobacterium bifidum MIMBb75 significantly alleviates irritable bowel syndrome and improves quality of life — a double-blind, placebo-controlled study. Alimentary Pharmacology & Therapeutics [Internet]. 2011 [cited 2018 Dec 5];33:1123–32. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2036.2011.04633.x

12. Kohno M, Tanaka-Azuma Y, Kamaguchi R, Tagawa D. Application of enteric seamless capsules containing Bifidobacterium longum to functional foods. Nippon Yakurigaku Zasshi. 2016;148:310–4.

13. Wong CB, Odamaki T, Xiao J. Beneficial effects of Bifidobacterium longum subsp. longum BB536 on human health: Modulation of gut microbiome as the principal action. Journal of Functional Foods [Internet]. 2019 [cited 2019 Mar 26];54:506–19. Available from: http://www.sciencedirect.com/science/article/pii/S1756464619300684

14. Mamiko K, Tomoe Y, Youichi M, Masanori A, Yuzo K. Effects of Enteric Capsules Containing Bifidobacteria and Lactic Acid Bacteria on the Fecal Properties in Healthy Volunteers. Journal of Intestinal Microbiology [Internet]. 2004 [cited 2018 Dec 5];18:87–92. Available from: https://www.jstage.jst.go.jp/article/jim/18/2/18_2_87/_article/-char/en/

15. Ibarra A, Latreille-Barbier M, Donazzolo Y, Pelletier X, Ouwehand AC. Effects of 28-day Bifidobacterium animalis subsp. lactis HN019 supplementation on colonic transit time and gastrointestinal symptoms in adults with functional constipation: A double-blind, randomized, placebo-controlled, and dose-ranging trial. Gut Microbes [Internet]. 2018 [cited 2019 Mar 26];9:236–51. Available from: https://doi.org/10.1080/19490976.2017.1412908

16. Dimidi E, Zdanaviciene A, Christodoulides S, Taheri S, Louis P, Duncan PI, et al. Randomised clinical trial: Bifidobacterium lactis NCC2818 probiotic vs placebo, and impact on gut transit time, symptoms, and gut microbiology in chronic constipation. Alimentary Pharmacology & Therapeutics [Internet]. 2019 [cited 2019 Mar 26];49:251–64. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/apt.15073

17. Reid G, Gaudier E, Guarner F, Huffnagle GB, Macklaim JM, Munoz AM, et al. Responders and non-responders to probiotic interventions. Gut Microbes [Internet]. 2010 [cited 2018 Dec 5];1:200–4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3023600/

18. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, et al. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature [Internet]. 2014 [cited 2018 Dec 5];514:181–6. Available from: https://www.nature.com/articles/nature13793

19. Chassaing B, Koren O, Goodrich JK, Poole AC, Srinivasan S, Ley RE, et al. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. Nature [Internet]. 2015 [cited 2018 Dec 5];519:92–6. Available from: https://www.nature.com/articles/nature14232

20. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell [Internet]. 2015 [cited 2018 Dec 5];163:1079–94. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0092867415014816

21. Korem T, Zeevi D, Zmora N, Weissbrod O, Bar N, Lotan-Pompan M, et al. Bread Affects Clinical Parameters and Induces Gut Microbiome-Associated Personal Glycemic Responses. Cell Metabolism [Internet]. 2017 [cited 2018 Dec 5];25:1243-1253.e5. Available from: http://www.sciencedirect.com/science/article/pii/S1550413117302887

22. Doherty MK, Ding T, Koumpouras C, Telesco SE, Monast C, Das A, et al. Fecal Microbiota Signatures Are Associated with Response to Ustekinumab Therapy among Crohn's Disease Patients. mBio [Internet]. 2018 [cited 2019 May 10];9:e02120-17. Available from: https://mbio.asm.org/content/9/2/e02120-17

23. Bazin T, Hooks KB, Barnetche T, Truchetet M-E, Enaud R, Richez C, et al. Microbiota Composition May Predict Anti-Tnf Alpha Response in Spondyloarthritis Patients: an Exploratory Study. Scientific Reports [Internet]. 2018 [cited 2019 May 10];8:5446. Available from: https://www.nature.com/articles/s41598-018-23571-4

24. Wilkinson EM, Ilhan ZE, Herbst-Kralovetz MM. Microbiota–drug interactions: Impact on metabolism and efficacy of therapeutics. Maturitas [Internet]. 2018 [cited 2019 Mar 26];112:53–63. Available from: http://www.sciencedirect.com/science/article/pii/S0378512218301348

25. Matson V, Fessler J, Bao R, Chongsuwat T, Zha Y, Alegre M-L, et al. The commensal microbiome is associated with anti–PD-1 efficacy in metastatic melanoma patients. Science [Internet]. 2018 [cited 2018 Dec 5];359:104–8. Available from: http://science.sciencemag.org/content/359/6371/104

26. Routy B, Chatelier EL, Derosa L, Duong CPM, Alou MT, Daillère R, et al. Gut microbiome influences efficacy of PD-1–based immunotherapy against epithelial tumors. Science [Internet]. 2018 [cited 2018 Dec 5];359:91–7. Available from: http://science.sciencemag.org/content/359/6371/91

27. Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinets TV, et al. Gut microbiome modulates response to anti–PD-1 immunotherapy in melanoma patients. Science [Internet]. 2018 [cited 2019 Feb 19];359:97–103. Available from: http://science.sciencemag.org/content/359/6371/97

28. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science [Internet]. 2015 [cited 2018 Dec 5];350:1079–84. Available from: http://science.sciencemag.org/content/350/6264/1079

29. Ananthakrishnan AN, Luo C, Yajnik V, Khalili H, Garber JJ, Stevens BW, et al. Gut Microbiome Function Predicts Response to Anti-integrin Biologic Therapy in Inflammatory Bowel Diseases. Cell Host & Microbe [Internet]. 2017 [cited 2019 Mar 26];21:603-610.e3. Available from: http://www.sciencedirect.com/science/article/pii/S1931312817301567

30. Watanabe S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. Journal of Machine Learning Research [Internet]. 2010 [cited 2018 Dec 5];11:3571–94. Available from: http://www.jmlr.org/papers/v11/watanabe10a.html

31. Korpela K, Flint HJ, Johnstone AM, Lappi J, Poutanen K, Dewulf E, et al. Gut Microbiota Signatures Predict Host and Microbiota Responses to Dietary Interventions in Obese Individuals. PLOS ONE [Internet]. 2014 [cited 2019 May 10];9:e90702. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0090702

32. Ogata T, Iino H, Kingaku M, Yaesima T, Teraguchi S, Fukuwatari Y, et al. Effect of Bifidobacterium longum BB536 yogurt administration on the intestinal environment of healthy adults. Microbial Ecology in Health and Disease [Internet]. 1999 [cited 2018 Dec 5];11:41–6. Available from: https://doi.org/10.1080/089106099435916

33. Bennet SMP, Böhn L, Störsrud S, Liljebo T, Collin L, Lindfors P, et al. Multivariate modelling of faecal bacterial profiles of patients with IBS predicts responsiveness to a diet low in FODMAPs. Gut [Internet]. 2018 [cited 2018 Dec 5];67:872–81. Available from: https://gut.bmj.com/content/67/5/872

34. Rivière A, Selak M, Lantin D, Leroy F, De Vuyst L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. Front Microbiol [Internet]. 2016 [cited 2018 Dec 5];7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4923077/

35. Lopetuso LR, Scaldaferri F, Petito V, Gasbarrini A. Commensal Clostridia: leading players in the maintenance of gut homeostasis. Gut Pathogens [Internet]. 2013 [cited 2018 Dec 5];5:23. Available from: https://doi.org/10.1186/1757-4749-5-23

36. Jalanka-Tuovinen J, Salojärvi J, Salonen A, Immonen O, Garsed K, Kelly FM, et al. Faecal microbiota composition and host–microbe cross-talk following gastroenteritis and in postinfectious irritable bowel syndrome. Gut [Internet]. 2014 [cited 2019 Feb 19];63:1737–45. Available from: https://gut.bmj.com/content/63/11/1737

37. Tamboli CP, Neut C, Desreumaux P, Colombel JF. Dysbiosis in inflammatory bowel disease. Gut [Internet]. 2004 [cited 2019 Feb 19];53:1–4. Available from: https://gut.bmj.com/content/53/1/1.1

38. Rivière A, Gagnon M, Weckx S, Roy D, Vuyst LD. Mutual Cross-Feeding Interactions between Bifidobacterium longum subsp. longum NCC2705 and Eubacterium rectale ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides. Appl Environ Microbiol [Internet]. 2015 [cited 2018 Dec 5];81:7767–81. Available from: https://aem.asm.org/content/81/22/7767

39. Moens F, Verce M, De Vuyst L. Lactate- and acetate-based cross-feeding interactions between selected strains of lactobacilli, bifidobacteria and colon bacteria in the presence of inulin-type fructans. International Journal of Food Microbiology [Internet]. 2017 [cited 2018 Dec 5];241:225–36. Available from: http://www.sciencedirect.com/science/article/pii/S0168160516305530

40. Falony G, Vlachou A, Verbrugghe K, Vuyst LD. Cross-Feeding between Bifidobacterium longum BB536 and Acetate-Converting, Butyrate-Producing Colon Bacteria during Growth on Oligofructose. Appl Environ Microbiol [Internet]. 2006 [cited 2018 Dec 5];72:7835–41. Available from: https://aem.asm.org/content/72/12/7835

41. Ze X, Duncan SH, Louis P, Flint HJ. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. The ISME Journal [Internet]. 2012 [cited 2018 Dec 5];6:1535–43. Available from: https://www.nature.com/articles/ismej20124

42. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. Role of the normal gut microbiota. World J Gastroenterol [Internet]. 2015 [cited 2018 Dec 5];21:8787–803. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528021/

43. Jin-zhong X, Shizuki K, Toshitaka O, Kazuhiro M, Tomoko Y, Keiji I, et al. Effect of yogurt containing Bifidobacterium longum BB536 on the defecation frequency and fecal characteristics of healthy adults : A double-blind cross over study. Japanese Journal of Lactic Acid Bacteria [Internet]. 2007 [cited 2018 Dec 5];18:31–6. Available from: https://www.jstage.jst.go.jp/article/jslab/18/1/18_1_31/_article/-char/en

44. Asada M, Kamaguchi R, Nakatsuji M, Takadera T, Yoshikado M. Seamless capsule [Internet]. 2009 [cited 2018 Dec 5]. Available from: https://patents.google.com/patent/WO2009142170A1/en

45. Kamaguchi R, Mizutani M. Capsule disintegrable in large-intestine-specific manner [Internet]. 2013 [cited 2018 Dec 5]. Available from: https://patents.google.com/patent/WO2013145379A1/en

46. Murakami S, Goto Y, Ito K, Hayasaka S, Kurihara S, Soga T, et al. The Consumption of Bicarbonate-Rich Mineral Water Improves Glycemic Control. Evid Based Complement Alternat Med. 2015;2015:824395.

47. Kim S-W, Suda W, Kim S, Oshima K, Fukuda S, Ohno H, et al. Robustness of Gut Microbiota of Healthy Adults in Response to Probiotic Intervention Revealed by High-Throughput Pyrosequencing. DNA Res [Internet]. 2013 [cited 2018 Dec 5];20:241–53. Available from: https://academic.oup.com/dnaresearch/article/20/3/241/368334

48. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ [Internet]. 2016 [cited 2018 Dec 5];4:e2584. Available from: https://peerj.com/articles/2584

49. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics [Internet]. 2018 [cited 2019 Jul 10];34:i884–90. Available from: https://academic.oup.com/bioinformatics/article/34/17/i884/5093234

50. Shen W, Le S, Li Y, Hu F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLOS ONE [Internet]. 2016 [cited 2019 Jul 10];11:e0163962. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0163962

51. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Xu ZZ, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems [Internet]. 2017 [cited 2019 Jul 10];2:e00191-16. Available from: https://msystems.asm.org/content/2/2/e00191-16

52. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nature Biotechnology [Internet]. 2018 [cited 2019 Jul 10];36:996–1004. Available from: https://www.nature.com/articles/nbt.4229

53. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.

54. Edgar RC. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. Bioinformatics [Internet]. 2018 [cited 2019 Jul 10];34:2371–5. Available from: https://academic.oup.com/bioinformatics/article/34/14/2371/4913809

55. Mori H, Maruyama T, Yano M, Yamada T, Kurokawa K. VITCOMIC2: visualization tool for the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing. BMC Systems Biology [Internet]. 2018 [cited 2019 Jul 10];12:30. Available from: https://doi.org/10.1186/s12918-018-0545-2

56. Mudholkar GS, Srivastava DK, Kollia GD. A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. Journal of the American Statistical Association [Internet]. 1996 [cited 2018 Dec 5];91:1575–83. Available from: https://www.jstor.org/stable/2291583

57. Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. Stat Comput [Internet]. 2014 [cited 2018 Dec 5];24:997–1016. Available from: https://doi.org/10.1007/s11222-013-9416-2

58. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software [Internet]. 2017 [cited 2018 Dec 5];76:1–32. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v076i01

59. Dellaportas P, Smith AFM. Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling. Journal of the Royal Statistical Society Series C (Applied Statistics) [Internet]. 1993 [cited 2018 Dec 5];42:443–59. Available from: https://www.jstor.org/stable/2986324

60. Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics [Internet]. 1998 [cited 2018 Dec 5];7:434–55. Available from: https://www.tandfonline.com/doi/abs/10.1080/10618600.1998.10474787

Figures

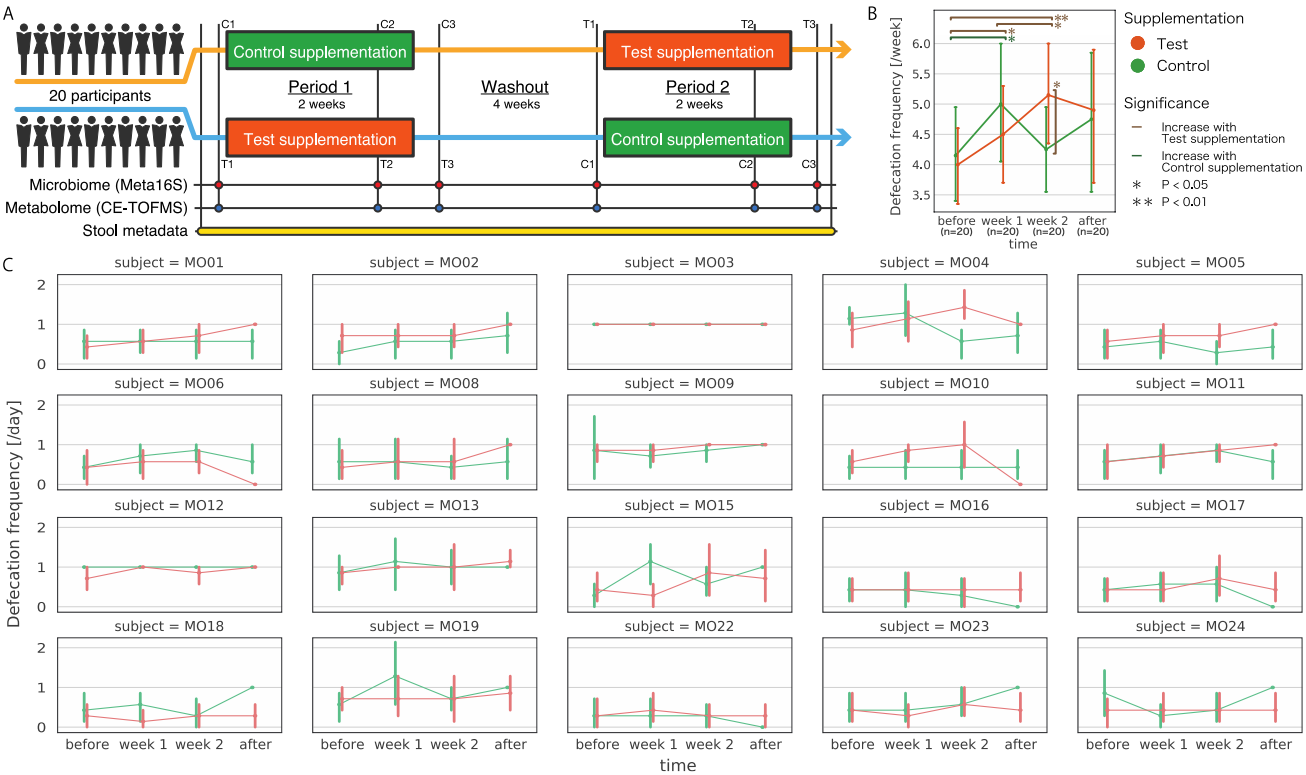

Figure 1. Schematic of the randomized double-blind controlled trial to investigate the increase in bowel movement frequency induced by the intake of *Bifidobacterium longum* BB536 encapsulated in an acid-resistant seamless capsule. A: The illustration indicates a study design of a randomized double-blind controlled trial to investigate the increase in bowel movement frequency due to the intake of *B. longum* BB536 encapsulated in an acid-resistant seamless capsule. B, C: The line graphs indicate the bowel movement frequency per week (B) and the bowel movement frequency per day for each subject (C) throughout the trial period. A Wilcoxon signed rank test (two-sided) was carried out for comparisons between periods (before (n=20); week 1 (n=20); week 2 (n=20); after (n=20)). The points represent the mean bowel movement frequency per week (or day), and the error bars represent the corresponding 95% confidence intervals.

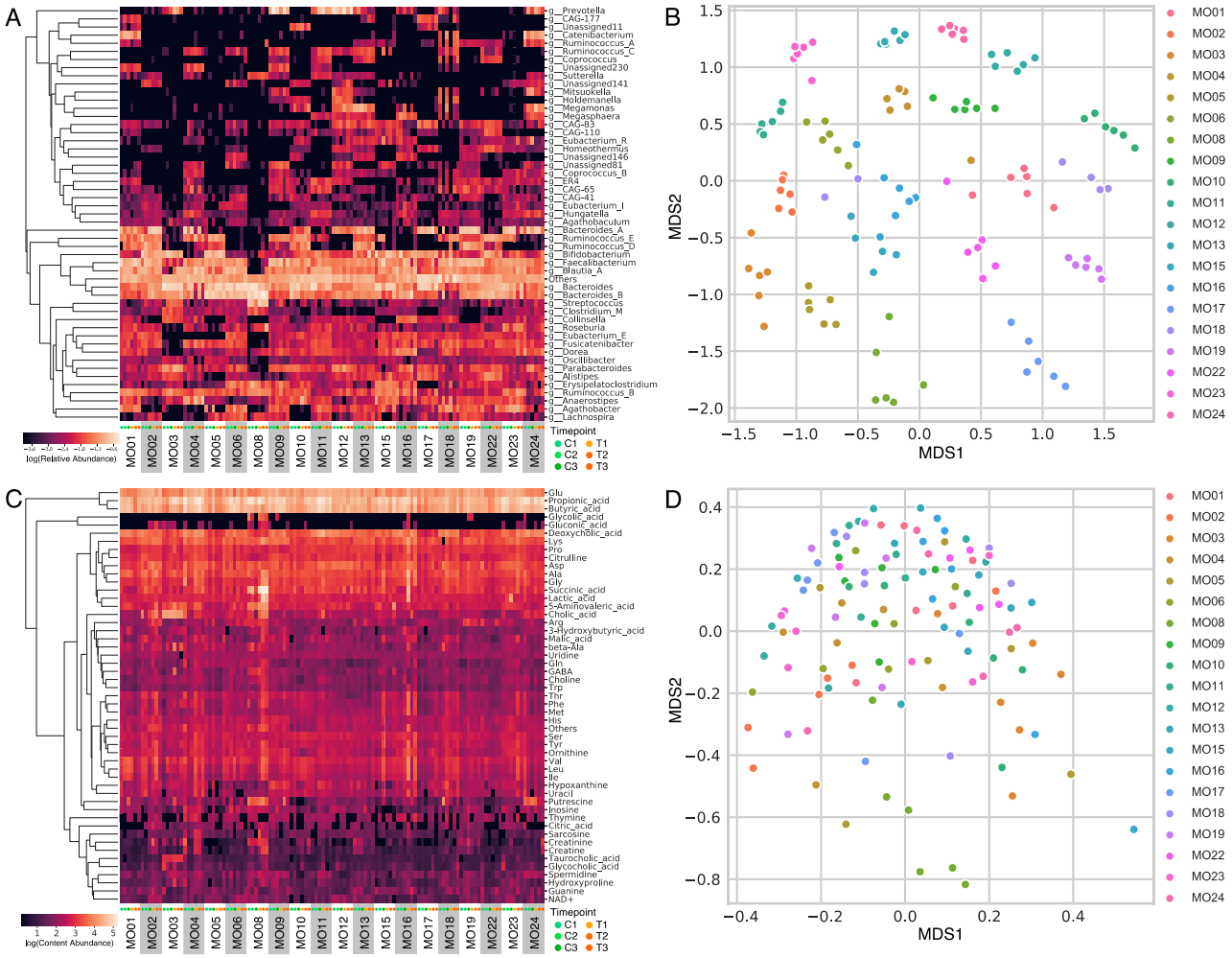

Figure 2. Overall view of the metabologenomic data. A: The heatmap represents the relative abundance of the top 50 genera. B: The scatter plot shows the results of MDS using beta diversity (Spearman correlation coefficient distance) calculated from the microbiome composition. C: The heatmap represents the abundance of the top 50 metabolites in the fecal content. D: A scatter plot showing the results of MDS using beta diversity (Spearman correlation coefficient distance) calculated from the metabolite composition. The same color was used for each subject.

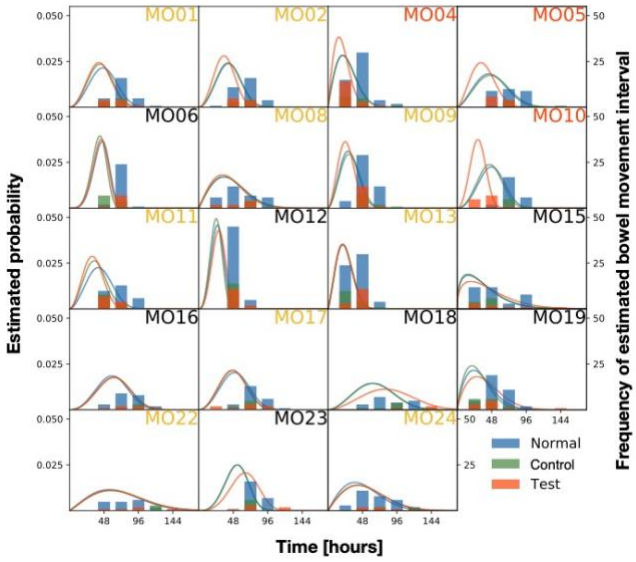

Figure 3. Histogram of the observed bowel movement time interval and estimated probability density function of the Weibull distribution. The vertical axis on the left shows the value of the probability density function, and the vertical axis on the right shows the frequency of the estimated bowel movement interval in the whole observation period (85 days). The responder status of the subject was defined based on the Bayesian posterior distribution: SRs are indicated in orange labels, WRs are indicated in yellow labels, and NRs are indicated in black labels.

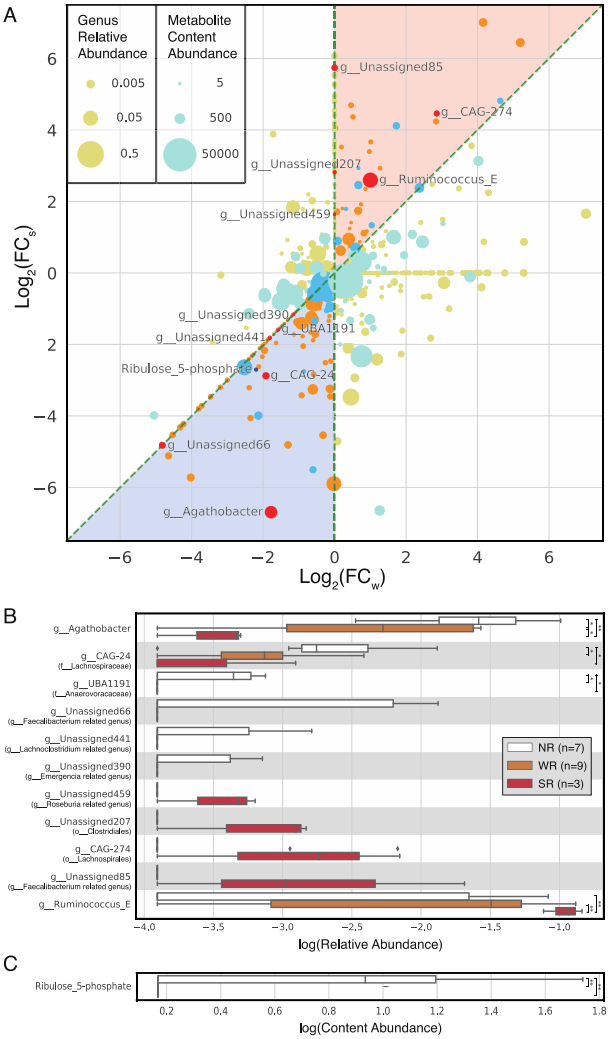

Figure 4. Intestinal environmental features related to the increasing bowel movement frequency due to *B. longum* BB536 intake. A: A scatter plot showing the fold change in microbiota and metabolite abundances in responders compared to those in NRs. The vertical axis represents the log 2 value of the SRs' fold change compared to that of the NRs (FCs), and the horizontal axis represents the log 2 value of the WRs' fold change compared to that of the NRs (FCw). The warm plot corresponds to the microbiota, and its size represents the relative abundance. Red represents bacterial genera that were subjected to the Jonckheere-Terpstra trend test and had a p value < 0.05. The cold plot corresponds to metabolites, and its size represents the fecal content. Blue represents metabolites that were subjected to the Jonckheere-Terpstra trend test and had a p value < 0.05. The light orange region is FCw < FCs, that is, NR < WR < SR, and the light blue region is FCw> FCs, that is, NR > WR > SR. B, C: The relative abundance of any bacterial genus and metabolite content of fecal samples that had a p value < 0.05. Lineage names in parentheses represent the most closely related GTDB taxonomy. * indicates a p value < 0.05, according to the Brunner-Munzel test, and ** indicates a p value < 0.01. The error bars indicate the 95% confidence interval.

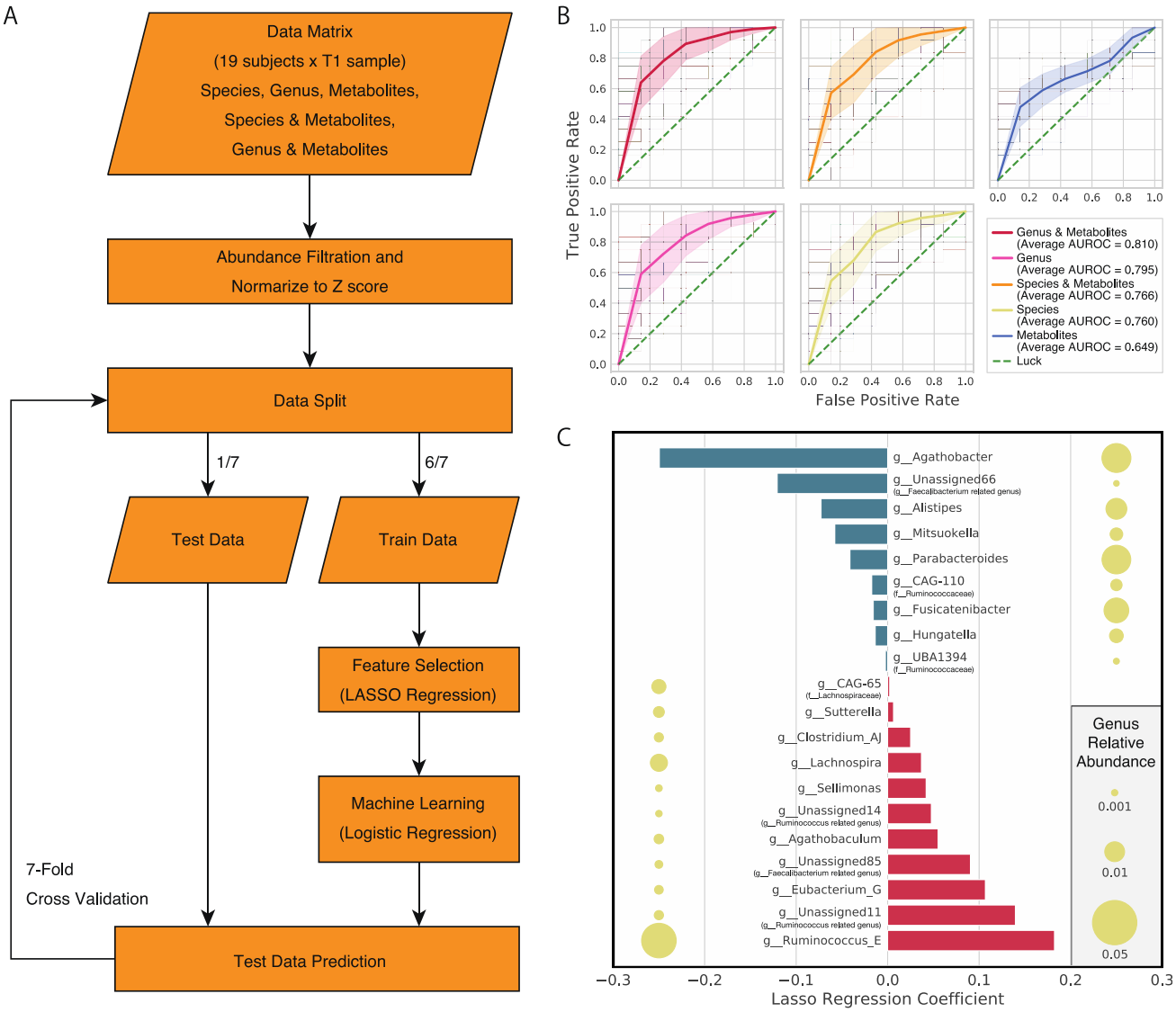

Figure 5. Machine learning-based responder prediction from intestinal environmental features before the intake of the test supplement. A: A flow chart of the machine learning algorithm for responder prediction. Feature selection by LASSO regression and machine learning by logistic regression algorithm were used. B: A receiver operating characteristic (ROC) curve when predicting, from intestinal omics data collected before intake of the test supplement, who will be a responder (SR and WR) (i.e., experience a significant increase in bowel movement frequency). C: Regression coefficient of the LASSO regression with respect to the responder score in the parameter with the highest accuracy in the 7-fold stratified validation, using bacterial genus and metabolite data collected before the test supplement intake began. The size of the circle adjacent to the bacterial genus name represents the average relative abundance or content (no metabolite features were detected). Lineage names in parentheses represent the most closely related GTDB taxonomy.

Tables

Table 1. Summary of the responder predictions from machine learning. *: AUROC; area under the receiver operating characteristic curve. **: For each data set, the parameters, AUROC, accuracy and F-measure at the highest performance are shown.

Data	Microbiome relative abundance threshold**	Metabolome content abundance threshold**	LASSO regression parameter alpha**	Average AUROC*	Accuracy	F-measure
Genus and metabolite	0.001	10000	0.001	0.810	0.746	0.797
Genus	0.001	-	0.005	0.795	0.739	0.797
Species and metabolite	0.005	10000	0.1	0.766	0.724	0.788
Species	0.005	-	0.1	0.760	0.721	0.787
Metabolite	-	100	0.001	0.649	0.617	0.675