SUPPLEMENTARY MATERIAL

Probiotic-mediated modulation of gut microbiome in students exposed to academic stress: a randomized controlled trial

Jorge Francisco Vázquez-Castellanos[1,2≠], Lucas Ferreira Maciel[1,2≠], Lucas Wauters[3,4≠], Ann Gregory[1,2], Lukas Van Oudenhove[4], Karlien Geboers[3], Kristin Verbeke[4], Tamara Smokvina[5], Jan Tack[3,4], Tim Vanuytsel[3,4*], Muriel Derrien[5*§], Jeroen Raes[1,2*]

[1] Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, KU Leuven, Leuven, Belgium

[2] VIB Center for Microbiology, Leuven, Belgium

[3] Department of Gastroenterology and Hepatology, University Hospitals Leuven, Leuven, Belgium

[4] Translational Research in Gastrointestinal Disorders (TARGID), department of Chronic Diseases and Metabolism (ChroMeta), KU Leuven, Leuven, Belgium

[5] Danone Research & Innovation, Gif-sur-Yvette, France

[§] Present address: Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, KU Leuven, Leuven, Belgium

[≠] These authors contributed equally

[*] These authors jointly supervised this work

Corresponding authors: jeroen.raes@kuleuven.be; tim.vanuytsel@uzleuven.be

Table of content

Supplementary Figure 1. Enterotyping of study cohort with the Dirichlet Multinomial Mixtures (DMM) method. Left: Genus-level microbiota community variation of the Flemish Gut Flora Project (FGFP) represented by Principal Coordinates Analysis (PCoA; Bray-Curtis distance). **Middle**: Study samples are represented by color, samples from the FGFP cohort are represented in gray. All samples are colored by enterotype. **Right**: Identification of the optimal number of Dirichlet components in a matrix of samples from the study cohort and the FGFP. The optimal number of clusters, based on the Bayesian Information Criterion (BIC), was four.

Supplementary Figure 2. Stool characteristics difference between enterotypes. Kruskal–Wallis, * FDR< 0.05, ** FDR<0.01, *** FDR <0.001)

Supplementary Figure 3. *L. rhamnosus* in gut microbiome. A. Strain tracking by StrainPhlan. B. Quantitative abundance of *L. rhamnosus* during trial C. Variation of gut microbiome between subjects by QMP

Supplementary Figure 4. Effect of *L. rhamnosus* CNCM I-3690 on overall gut microbiome composition (genus level) by 16S rRNA gene amplicon sequencing with removal of *Lactobacillus* . **A.** Change of alpha-diversity indices for each subject versus baseline for the placebo and the probiotic group (Wilcoxon test) **B.** Linear mixed-effect models of the effect of the probiotic in the diversity; the model takes as the baseline the “Placebo” and the “Visit 1” as the reference levels for Group and Visit variables. The coefficients and interaction terms are the differences from the baselines. **C.** Within-subject Bray-Curtis dissimilarity to baseline in both groups with lower distance indicating less changes (Wilcoxon test).

Supplementary Figure 5. Effect of *L. rhamnosus* CNCM I-3690 on overall gut microbiome composition (genus level) by 16S rRNA gene amplicon sequencing with exclusion of subjects with baseline calprotectin >100 ug/g. **A.** Change of alpha-diversity indices for each subject versus baseline for the placebo and the probiotic group (Wilcoxon test) **B.** Linear mixed-effect models of the effect of the probiotic in the diversity; the model takes as the baseline the “Placebo” and the “Visit 1” as the reference levels for Group and Visit variables. The coefficients and interaction terms are the differences from the baselines. **C.** Within-subject Bray-Curtis dissimilarity to baseline in both groups with lower distance indicating less changes (Wilcoxon test).

Supplementary Figure 6. Topological metrics of microbial interaction networks. **A**. Differences across timepoints in Degree centrality (number of direct connections per node), Closeness centrality (how easily a node can reach all other nodes), Betweenness centrality (how often a node lies on the shortest path between other nodes, indicating its role as a connector), and Eigenvector centrality (a measure of a node's influence based on the connectivity of its neighbors). Significant differences between the placebo and probiotic groups at each timepoint were assessed using the Wilcoxon test. **B**. Centrality values of *Lactobacillus rhamnosus* over time in the probiotic arm. *L. rhamnosus* shows an increase in centrality between timepoints 2 and 3, suggesting a growing influence within the microbial network.

Supplementary Figure 7. Gut-brain modules (GBMs) detected in *L. rhamnosus* CNCMI-3690 genome. In the y-axis is presented the prevalence of each GBM in 533 representative species of the human gut microbiome. The minimum pathway completeness of 50% was required to consider the GBM as present in the genome

Supplementary Figure 8. Effect of *L. rhamnosus* CNCM I-3690 on the GMM composition using a quantitative approach. A. Canberra distance over time in both groups. **B.** Confounding variables indicated as significant and non-redundant according to a distance-based redundancy analysis. R2 (purple) indicates the variance explained by each variable, while Cum_R2 (green) represents the cumulative variance explained. **C.** Negative binomial mixed effect model of the taxa that showed a change in the abundance after the probiotic intake, the model was corrected by the effect of moisture

Supplementary Tables provided as excel

Supplementary Table 1 Significant non-redundant covariates of the gut microbiome composition and metabolic potential

Supplementary Table 2 Significant covariates of the gut microbiome composition and metabolic potential

Supplementary Table 3 Effect of intervention on alpha-diversity (genus and mOTU level)

Supplementary Table 4 Baseline comparison between groups at species level

Supplementary Table 5 Baseline association between alpha-diversity and microbiome / host variables

Supplementary Table 6 Baseline association between species and host variables

Supplementary Table 7 Differential abundance of species between enterotypes

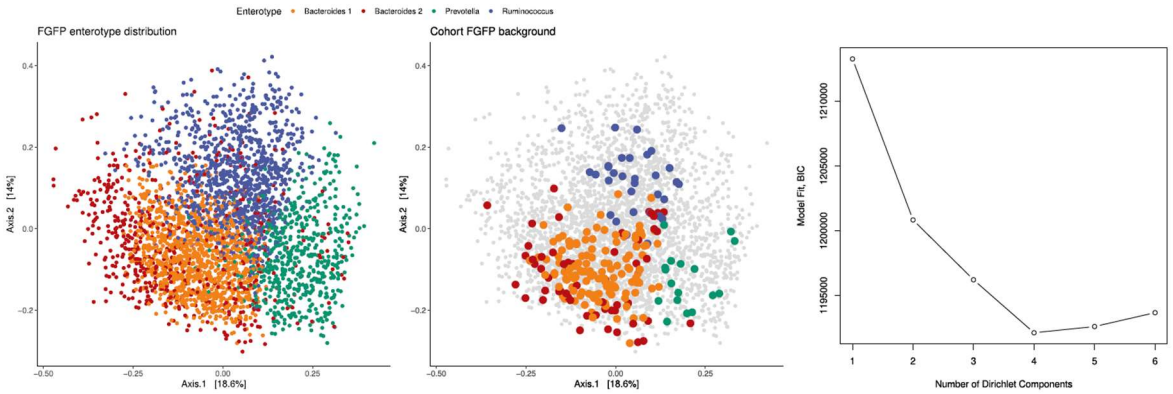
Supplementary Table 8 Differences between microbiome and host variables between enterotypes

Supplementary Table 9 Differences in fecal parameters between groups at baseline

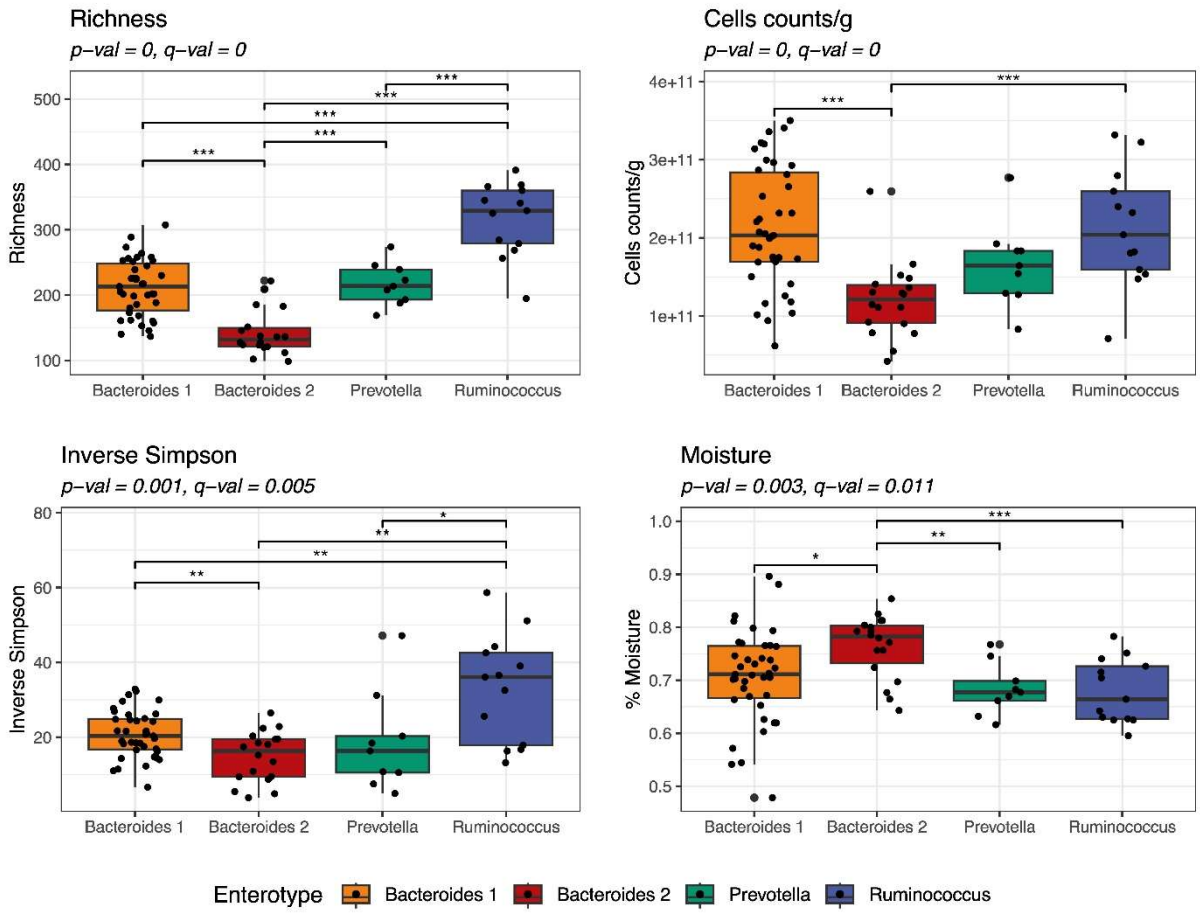
Supplementary Table 10 Shotgun-based alpha-diversity analysis

Supplementary Table 11 ANOVA model comparisons of the delta microbiome confounders and clinical variables

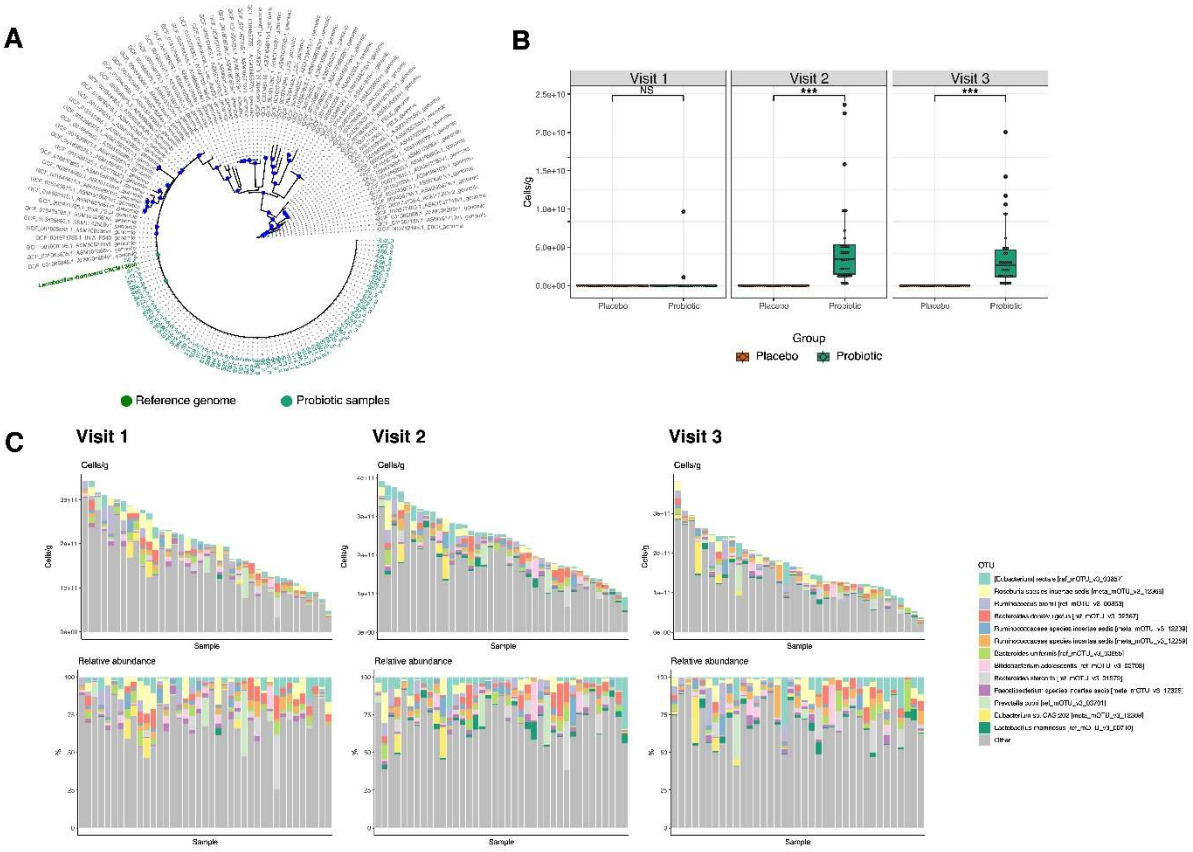
Supplementary Table 12 Delta association between species/GMM and host variables



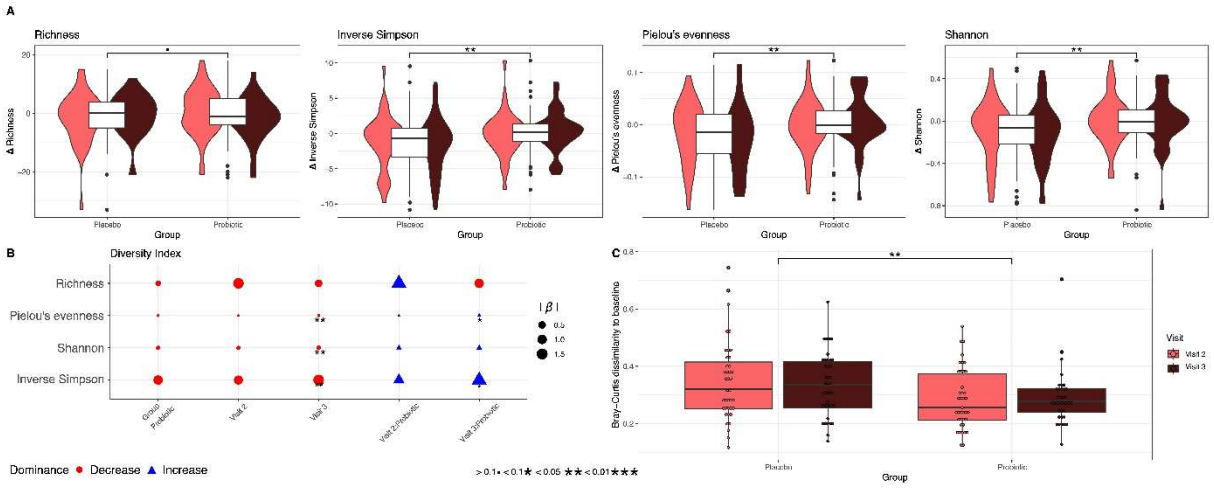
Supplementary Figure 1. Enterotyping of study cohort with the Dirichlet Multinomial Mixtures (DMM) method. Left: Genus-level microbiota community variation of the Flemish Gut Flora Project (FGFP) represented by Principal Coordinates Analysis (PCoA; Bray-Curtis distance). **Middle**: Study samples are represented by color, samples from the FGFP cohort are represented in gray. All samples are colored by enterotype. **Right**: Identification of the optimal number of Dirichlet components in a matrix of samples from the study cohort and the FGFP. The optimal number of clusters, based on the Bayesian Information Criterion (BIC), was four.



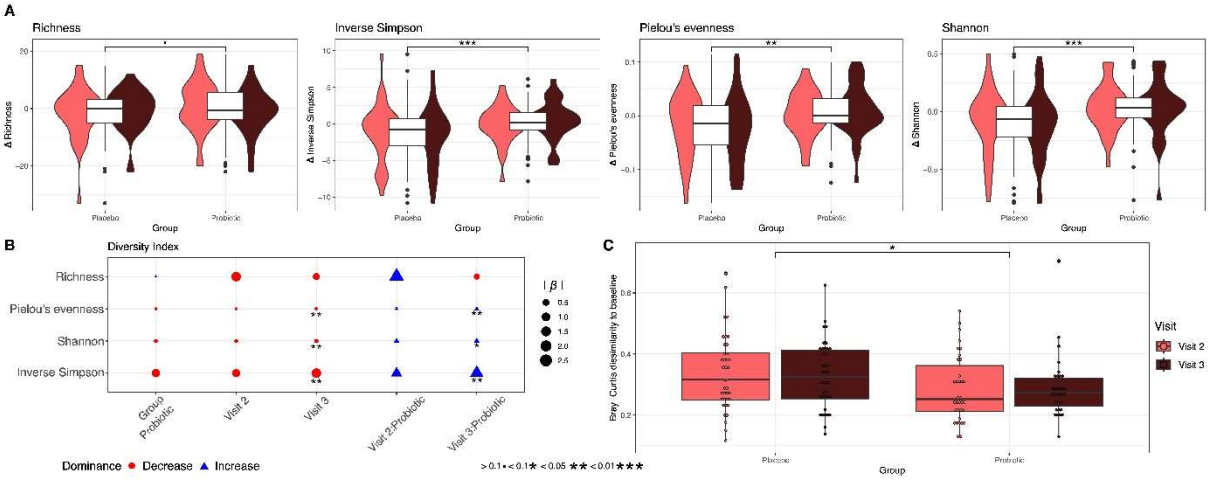
Supplementary Figure 2. Stool characteristics difference between enterotypes. Kruskal–Wallis, * FDR< 0.05, ** FDR<0.01, *** FDR <0.001)



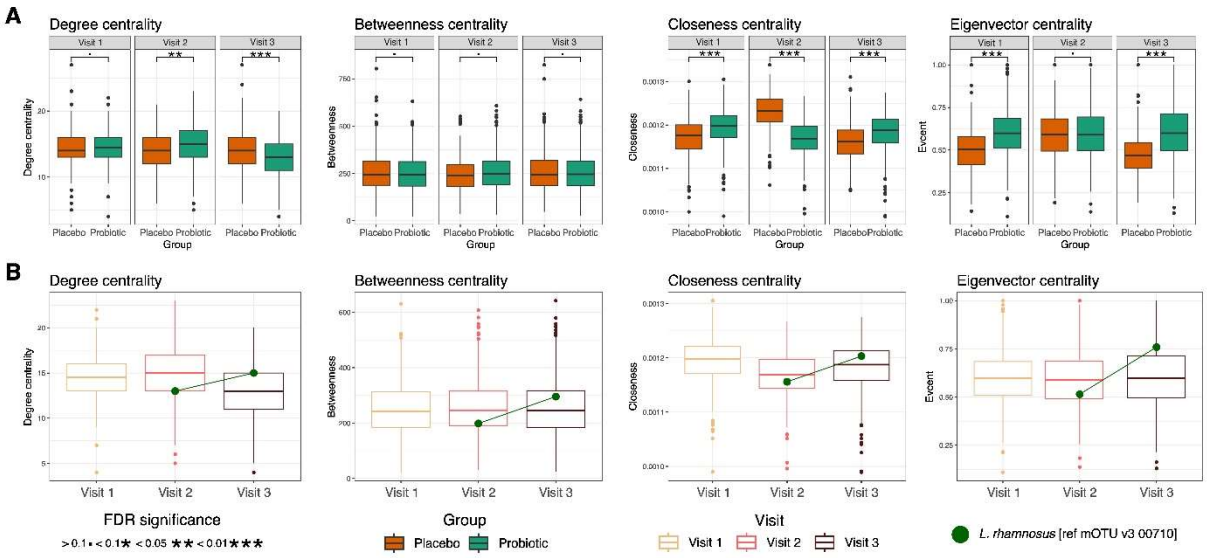
Supplementary Figure 3. *L. rhamnosus* in gut microbiome. A. Strain tracking by StrainPhlan. B. Quantitative abundance of *L. rhamnosus* during trial C. Variation of gut microbiome between subjects by QMP



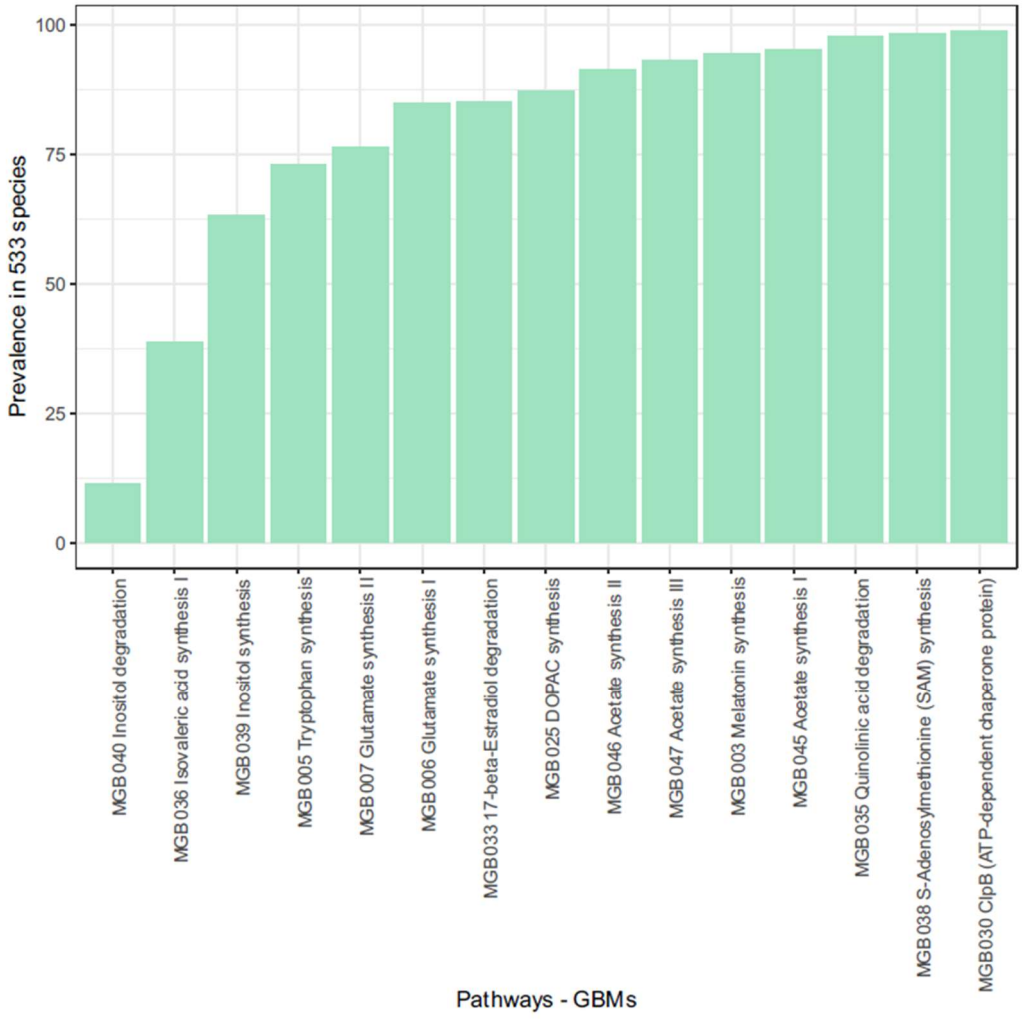
Supplementary Figure 4. Effect of *L. rhamnosus* CNCM I-3690 on overall gut microbiome composition (genus level) by 16S rRNA gene amplicon sequencing with removal of *Lactobacillus* . **A.** Change of alpha-diversity indices for each subject versus baseline for the placebo and the probiotic group (Wilcoxon test) **B.** Linear mixed-effect models of the effect of the probiotic in the diversity; the model takes as the baseline the “Placebo” and the “Visit 1” as the reference levels for Group and Visit variables. The coefficients and interaction terms are the differences from the baselines. **C.** Within-subject Bray-Curtis dissimilarity to baseline in both groups with lower distance indicating less changes (Wilcoxon test).



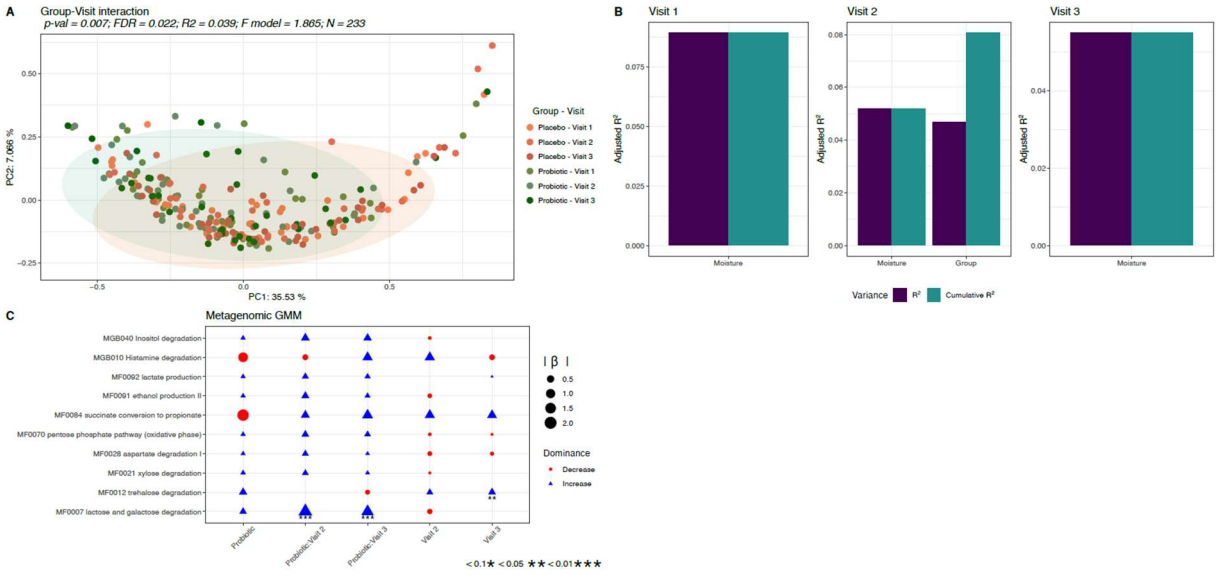
Supplementary Figure 5. Effect of *L. rhamnosus* CNCM I-3690 on overall gut microbiome composition (genus level) by 16S rRNA gene amplicon sequencing with exclusion of subjects with baseline calprotectin >100 ug/g. **A.** Change of alpha-diversity indices for each subject versus baseline for the placebo and the probiotic group (Wilcoxon test) **B.** Linear mixed-effect models of the effect of the probiotic in the diversity; the model takes as the baseline the “Placebo” and the “Visit 1” as the reference levels for Group and Visit variables. The coefficients and interaction terms are the differences from the baselines. **C.** Within-subject Bray-Curtis dissimilarity to baseline in both groups with lower distance indicating less changes (Wilcoxon test).



Supplementary Figure 6. Topological metrics of microbial interaction networks. **A**. Differences across timepoints in Degree centrality (number of direct connections per node), Closeness centrality (how easily a node can reach all other nodes), Betweenness centrality (how often a node lies on the shortest path between other nodes, indicating its role as a connector), and Eigenvector centrality (a measure of a node's influence based on the connectivity of its neighbors). Significant differences between the placebo and probiotic groups at each timepoint were assessed using the Wilcoxon test. **B**. Centrality values of *Lactobacillus rhamnosus* over time in the probiotic arm. *L. rhamnosus* shows an increase in centrality between timepoints 2 and 3, suggesting a growing influence within the microbial network.



Supplementary Figure 7. Gut-brain modules (GBMs) detected in *L. rhamnosus* CNCMI-3690 genome. In the y-axis is presented the prevalence of each GBM in 533 representative species of the human gut microbiome. The minimum pathway completeness of 50% was required to consider the GBM as present in the genome



Supplementary Figure 8. Effect of *L. rhamnosus* CNCM I-3690 on the GMM composition using a quantitative approach. A. Canberra distance over time in both groups. **B.** Confounding variables indicated as significant and non-redundant according to a distance-based redundancy analysis. R2 (purple) indicates the variance explained by each variable, while Cum_R2 (green) represents the cumulative variance explained. **C.** Negative binomial mixed effect model of the taxa that showed a change in the abundance after the probiotic intake, the model was corrected by the effect of moisture