Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms, Herfindal AM *et al.*

Supplementary method descriptions

Bacterial α-diversity and β-diversity

β-diversities were calculated through QIIME default script (core_diversity_analyses.py) based on one single normalized OTU table, while α-diversities were calculated as mean diversity based on OTU tables from ten rarefying iterations.

NMDS

Non-metric multidimensional scaling (NMDS) of Jaccard, Bray-Curtis, and weighted UniFrac distances was conducted in R (1) using the function metaMDS() from the *vegan* package (k=3, autotransform=FALSE, try=100) (2). To the previous.best argument, ordination output from the cmdscale() function (k=3) from the *stats* package (1) was used. The goodness of fit of the NMDS ordinations (compliance between the Euclidian ordination distances and the original pairwise distances) were evaluated by the stress value and Shepherd's plot. Shepherd's plots were made using the stressplot() function from the *vegan* package. Stress >0.2 was considered poor, 0.1-0.2 was considered fair, 0.05-0.1 was considered good and stress < 0.05 was considered excellent.

Multiple linear regression (ANCOVA)

Multiple linear regressions models, referred to as ANCOVA models, were first fitted using the lm() function from the *stats* package (1), with

formula = 4-week follow-up values ~ baseline values * diet group

and defining options(contrasts=c("contr.treatment", "contr.poly")) to obtain the desired contrast estimates (illustrated in A in figure below). The control group was always used as reference condition. The diet group baseline interaction (β_3 estimate in A in figure below) was excluded when not significant ($P > 0.05$), and the models were then fitted with

Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms, Herfindal AM *et al.*

formula = 4-week follow-up values ~ baseline values + diet group

(illustrated in B in figure below).

Reported estimates and corresponding *P* values were obtained by using the base function summary() on the fitted model object. In the main text/tables of the paper, only the diet group baseline interaction (β_3 estimate in A in figure below) or diet group (β_2 estimate in B in figure below) effect with corresponding *P* values are reported. An β_2 estimate of for example 10 in ANCOVA model without interaction, means that the low FODMAP diet group had a value of 10 higher than the control group at 4-week follow-up when adjusting for baseline values.
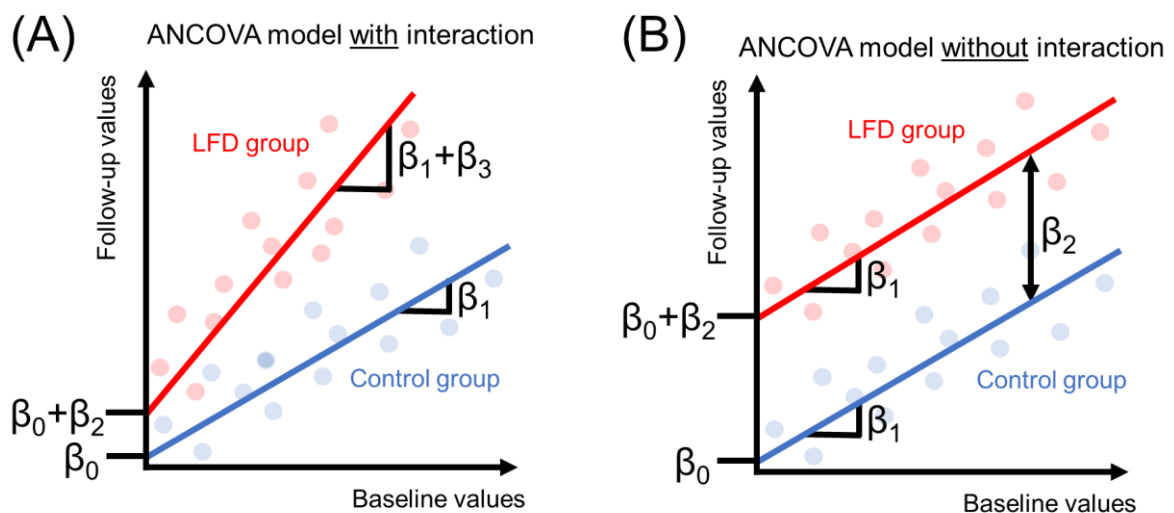
In ANCOVA model plots in the Supplemental Figure 5, 6, 7, 8 and 9, the estimated model equation for each diet group are reported:

Control group (without interaction): $F = \beta_0 + (\beta_1 * B)$

LFD group (without interaction): $F = (\beta_0 + \beta_2) + (\beta_1 * B)$

Control group (with interaction): $F = \beta_0 + (\beta_1 * B)$

LFD group (with interaction): $F = (\beta_0 + \beta_2) + ((\beta_1 + \beta_3) * B)$

Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms, Herfindal AM *et al.*

Assumptions of described linear regression models were evaluated in the following manner:

(1) The linear relationship between 4-week follow-up and baseline values was assessed by fitting a linear regression model (lm() function, formula = 4-week follow-up values ~ baseline values) separately for each diet group.

(2) Homogeneity of regression slopes was assessed by fitting a linear regression model where the interaction term between diet group and baseline values was included (lm() function, formula = 4-week follow-up values ~ baseline values * diet group).

(3) The normality of the residuals was assessed using normal QQ plot together with histogram, Shapiro-Wilk normality test, kurtosis and skewness statistics. Normal QQ plot was made using plot(model, 2). Histogram was made using the function plotNormalHistogram() from the *rcompanion* package (3). The function shapiro_test()from the *rstatix* package (4) was used to perform Shapiro-Wilk normality test. Skewness and kurtosis were calculated using the skewness() and kurtosis() functions, respectively, from the *e1071* packages (5).

(4) Constant variance of the residuals was assessed by plotting the residuals against fitted values, and by applying a Levene's test on the residuals between diet groups. Levene's test was performed using the levene_test() function from the *rstatix* package (4).

(5) Presence of influential values was assessed using "Residuals vs Leverage" plot obtained by plot(model, 5).

Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms, Herfindal AM *et al.*

Other methods related to data analyses

Method	Function/package description
Benjamini-Hochberg *P* value adjustment	p.adjust(method="BH") function from *stats* package.
Independent two-sample t-test	t.test() function from the *stats* package (1). Value "p.value" from t.test () object is reported.
Dependent (paired) two-sample t-test	t.test(paired=TRUE) function from the *stats* package (1). Value "p.value" from t.test () object is reported.
Wilcoxon rank-sum test for unpaired samples	wilcox.test() function from the *stats* package (1). Value "p.value" from wilcox.test() object is reported.
Wilcoxon signed-rank exact test or Wilcoxon signed-rank test with continuity correction for paired samples	wilcox.test(paired=TRUE) function from the *stats* package (1). Value "p.value" from wilcox.test() object is reported.
Spearman's correlation	cor.test(method="spearman") function from *stats* package (1). Value "estimate" (Spearman's rank correlation coefficients, Rho) and "p.value" from cor.test() object reported.
Figures	Almost all plots were made by applying various functions from the *ggplot2* package (6). Correlogram was made using the corrplot() function from the *corrplot* package (7).

Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms, Herfindal AM *et al.*

References

1. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2022. https://www.R-project.org/ (version 4.2.1).
2. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, et al. vegan: Community Ecology Package. 2020. https://CRAN.R-project.org/package=vegan (version 2.5-7).
3. Mangiafico S. rcompanion: Functions to Support Extension Education Program Evaluation. 2021. https://CRAN.R-project.org/package=rcompanion (version 2.4.1).
4. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. 2021. https://CRAN.R-project.org/package=rstatix (version 0.7.0).
5. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071). TU Wien, 2021. https://CRAN.R-project.org/package=e1071 (version 1.7-9).
6. Wickham H. ggplot2: Elegant Graphics for Data Analysis Springer-Verlag New York, 2016. https://ggplot2.tidyverse.org (version 3.3.5).
7. Wei T, Simko V. R package "corrplot": Visualization of a Correlation Matrix. 2021. https://github.com/taiyun/corrplot (version 0.88).