Supplementary Information

Individualized Dynamics in the Gut Microbiota Precede Crohn's Disease Flare

Short title: Gut Microbial dynamics Precede Crohn's Disease Flares

Authors: Tzipi Braun, MS[1], Ayelet Di Segni, PhD[1], Amnon Amir, PhD[1,2], Marina BenShoshan, MS[1,2], Sandra Neuman, RN[3], Nina Levhar, MS[3], Michael Bubis, BS[3], Orit Picard, PhD[3], Katya Sosnovski, MS[1,2], Gilat Efroni, PhD[1], Sarit Farage Barhom, PhD[1], Efrat Glick Saar, PhD[1], Avishay Lahad, MD [2,4], Batia Weiss, MD [2,4], Doron Yablecovitch, MD[2,3], Adi Lahat, MD[2,3], Rami Eliakim, MD[2,3], Uri Kopylov, MD[2,3], Shomron Ben-Horin, MD[2,3], Yael Haberman, MD, PhD[2,4,5]*

On behalf of Israeli IBD Research Nucleus (IIRN)

Affiliations:

[1]Cancer Research Center, Sheba Medical Center, Tel-Hashomer, Israel.

[2]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

[3]Department of Gastroenterology, Sheba Medical Center, Tel-Hashomer, Israel

[4]The Pediatric Gastroenterology Unit, The Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel-Hashomer, Israel

[5]Cincinnati Children's Hospital Medical Center, OH, USA.

***Corresponding Author:**

Yael Haberman, MD, PhD

Pediatric Gastroenterology, Hepatology & Nutrition Unit

The Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Israel

Telephone number: 972-3-5302692

Yael.Haberman@sheba.health.gov.il and Yael.haberman@cchmc.org

Supplemental tables

Supplemental Table 1: Microbial taxa associated with CD flare (see additional Excel spreadsheet).

Supplemental Figures

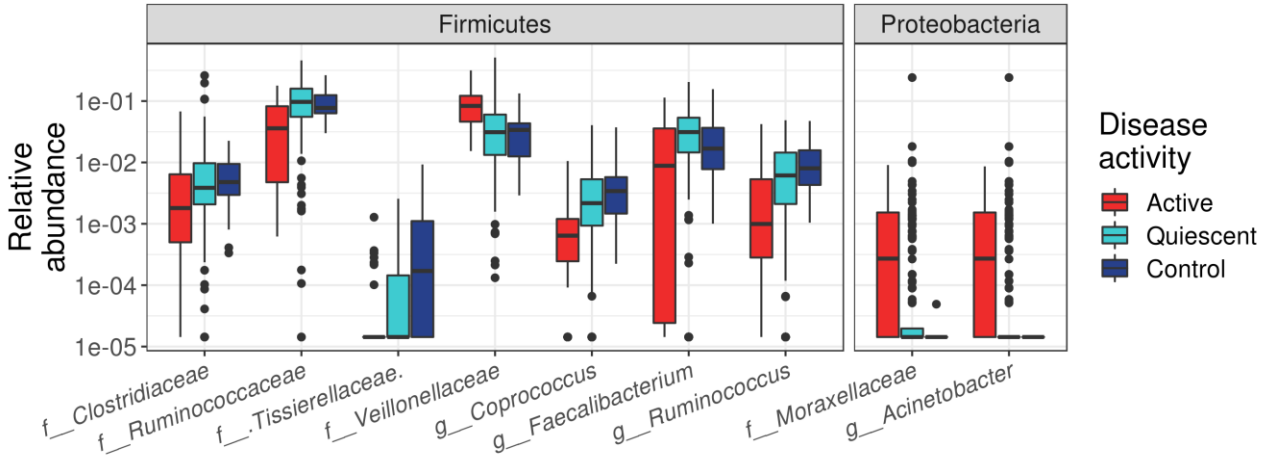

Fig. S1. Significant taxa relative abundance between CD active, remission and control samples. Significance was calculated using MaAslin pipeline (see Methods). Only taxa at the genus (g) and family (f) levels are shown here.

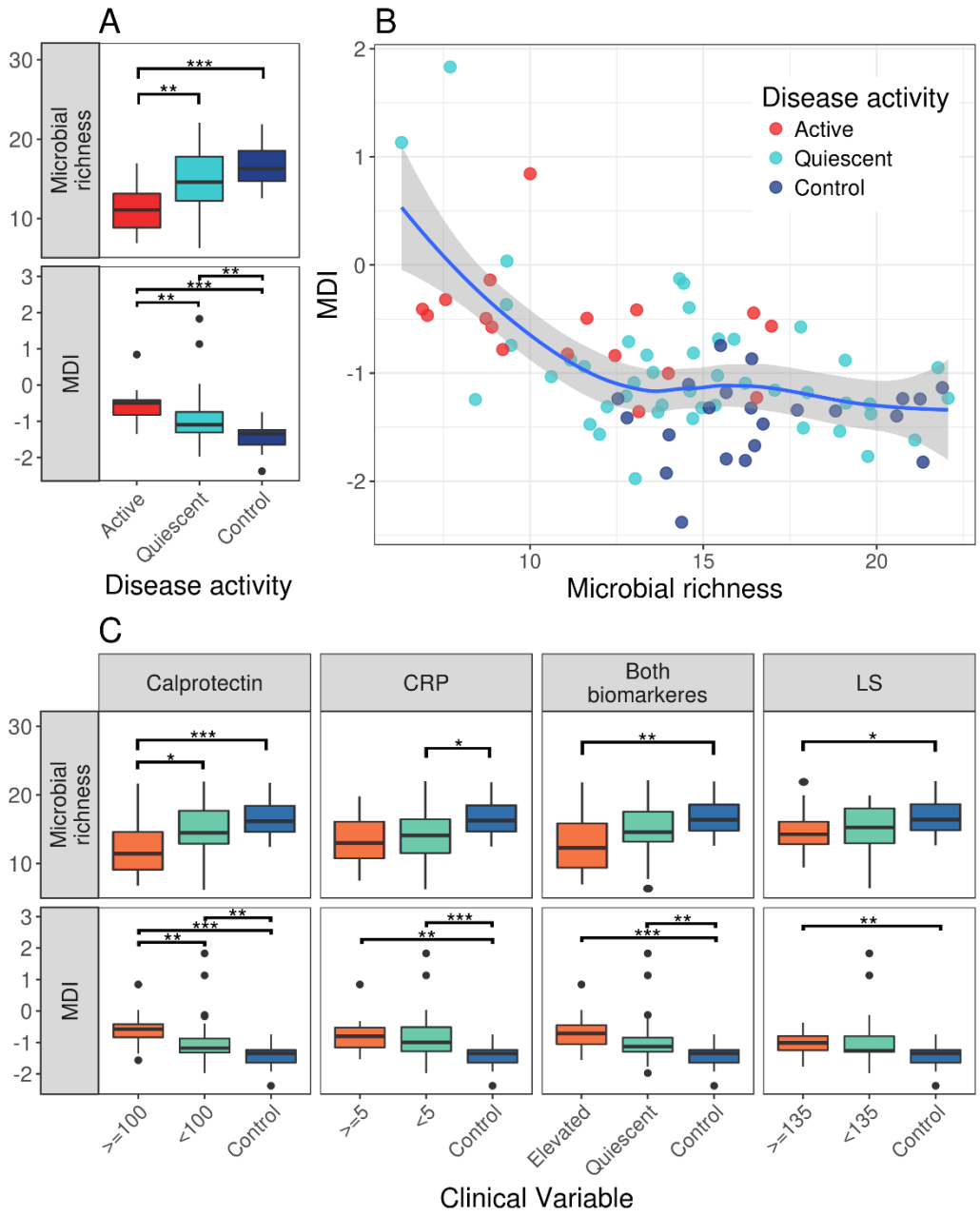

Fig. S2. CD patients in clinical, biomarkers, and mucosal remission show microbial richness and dysbiosis in between active CD and control, when analyzing one sample per patient. To avoid a bias for specific patients that have many samples, one sample per individual was used **(A)** Microbial richness (upper) and microbial dysbiosis index (lower) of controls (n=22), CD active (n=17) and remission (n=45) patients, as defined by PGA. **(B)** Scatter plot of samples microbial richness plotted against microbial dysbiosis index. One sample per patient is plotted (Spearman's rank correlation rho = -0.46, P < 1.3e-5, n=84), with a local polynomial regression fitting, and a 0.95 confidence interval. **(C)** Microbial richness (upper) and microbial dysbiosis index (lower) of controls (n=22) and, from left to right: fecal calprotectin lower (n=37) or higher (n=25) than 100 µg/g. CRP lower (n=46) or higher (n=12) than 5 mg/l. Both calprotectin and CRP lower than the designated threshold value (n=30) or at least one is higher (n=30). Lewis score lower (n=16) or higher (n=24) than 135. Asterisks indicate significant differences (Bonferroni corrected Mann Whitney U test: ***q < 0.001, **q < 0.01, *q < 0.05).

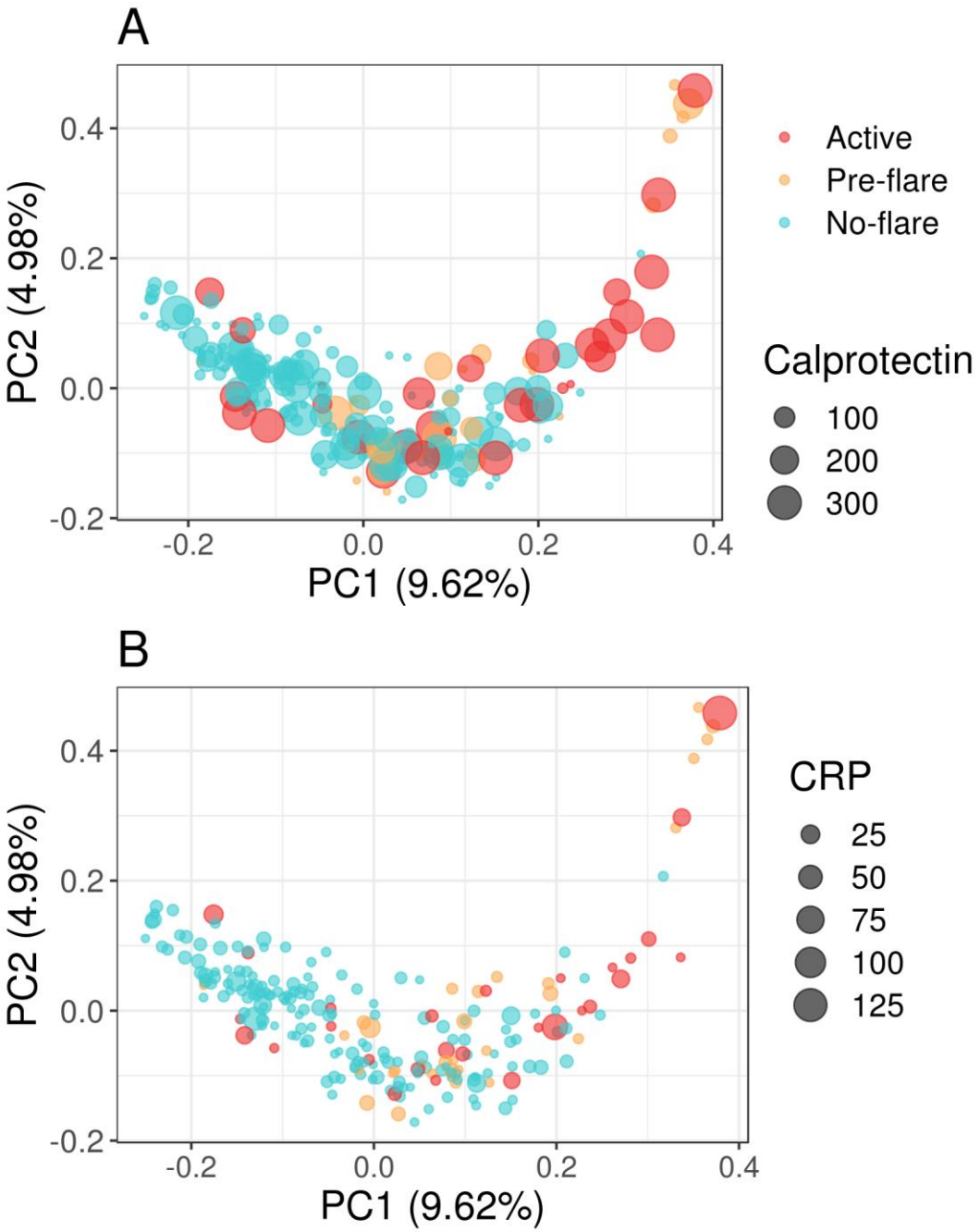

Fig. S3. CRP, fecal calprotectin, and gut microbiome variations in our CD cohort. Unweighted UniFrac PCoA plots colored as indicated. Circle size represents (A) Calprotectin or (B) CRP values. Fecal calprotectin is shown for all CD samples. CRP values are shown only for CD samples for which values were available. Summarized calprotectin and CRP values in each group are also noted in Table 1.

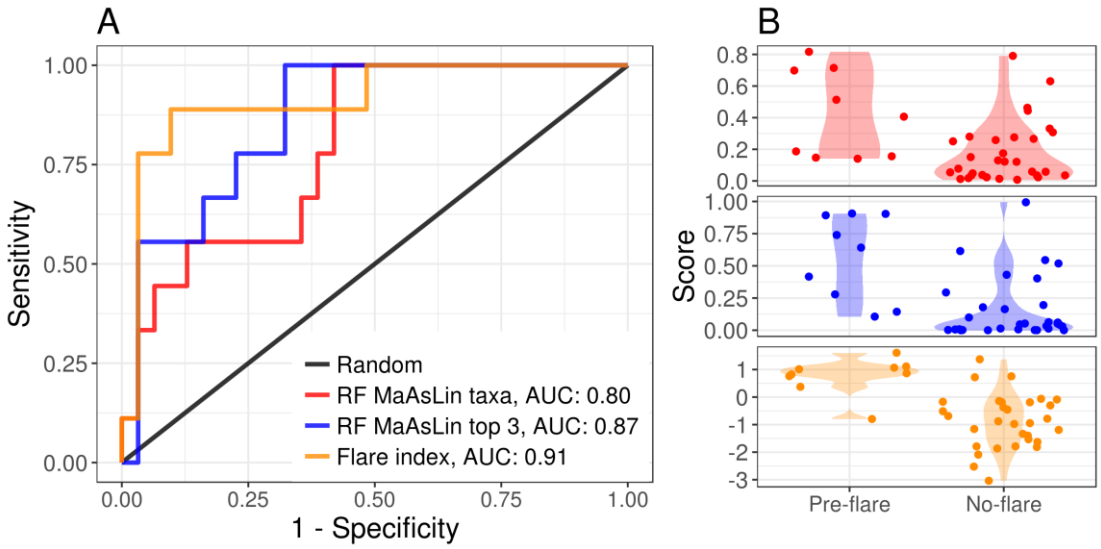

Fig. S4. Classification of pre-flare and no-flare samples using random forest machine learning algorithm. (A) ROC analysis of the Random forest results classifying pre-flare and no-flare samples using the taxa significantly different between the groups in the MaAsLin analysis (red), Random forest results using the top three significant taxa in the classification, as calculated by random forest mean decreased gini (blue) and flare index (orange), as noted in Fig 4A. The area under curve (AUC) for each is noted in the legend. **(B)** Violin plots of pred score for the random forest results, and flare index, of the noted Classification between pre-flare and no-flare samples.

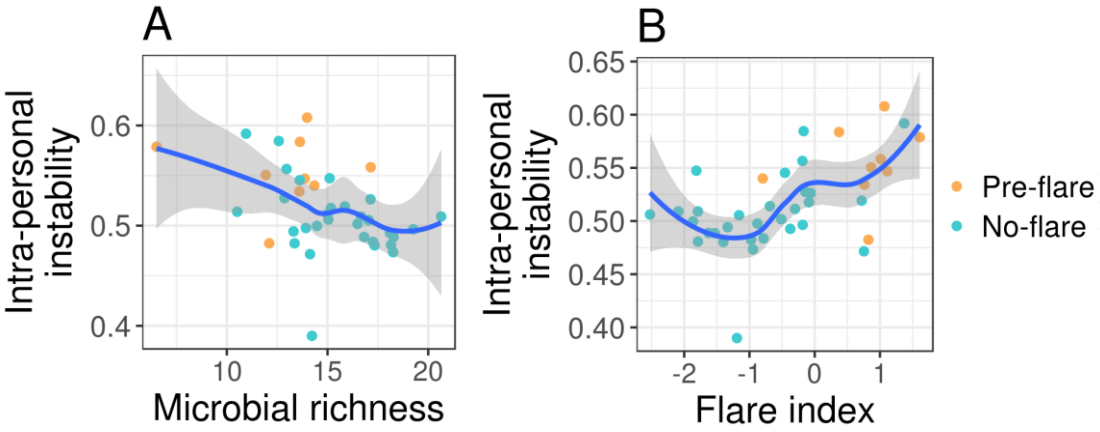

Fig. S5. Intra-personal variation (instability) in relation to other microbial variables. Scatterplots of Intra-personal instability against average values per patient of **(A)** Microbial richness **(B)** Flare index, with a local polynomial regression fitting, and a 0.95 confidence interval.