**Data processing.** Marker-gene sequencing data was processed using the Quantitative Insights Into Microbial Ecology (QIIME) package and mapped to 97% operational taxonomic units (OTUs) using the Greengenes reference database and USEARCH. Within-sample biodiversity was measured using whole-tree phylogenetic diversity. To quantify differences in microbial composition between subjects over time, UniFrac distance between all pairs of samples was calculated. Functional repertoire of the microbiome was predicted from the marker gene sequences using the PICRUSt tool.

**Statisical analysis of differential taxa by treatment group.** Differences in microbial biodiversity between both within subjects before and after the dietary intervention, and between subjects in the two treatment groups using a linear model in the statistical package R, controlling for age, BMI, and gender. Significance of associations was corrected for the total number of comparisons using false discovery rate correction, where significance is determined by q-values < 0.25. This analysis was conducted with bacterial taxa summarized at the levels of genus and phylum, and was conducted with relative abundances of taxa at both timepoints individually, as well as with the change in relative abundance of each taxon.

**Comparison of beta-diversity between groups.** Weighted and unweighted UniFrac measures of between-sample ecological variation was used to test the hypothesis that subjects are more similar to other subjects within the same treatment group. This test was performed using the Adonis function in the *vegan* package in *R*, similar to the PERMANOVA analysis.

**Comparison of alpha-diversity between groups.**  A two-tailed test of differentiation of within-sample biodiversity or alpha diversity, as measured by whole-tree phylogenetic diversity, between treatment groups. Sequences from all samples were subsampled at the sample sequencing depth of 19,490 sequences to control for differential sequencing effort. Alpha diversity values were verified to be normally distributed by a Shapiro-Wilk test then compared using a t-test. Additionally, alpha diversity was compared using a paired t-test to check whether diversity was higher in one group versus another after the treatment.

**Statisical analysis of associations between taxa and metadata.** Linear associations between relative abundances of common taxa (those present in at least .1% of samples) and clinical metadata covariates of interest, as provided by the GM-Tufts team, was performed. For this statistical testing we performed a generalized linear regression in the statistical package R, while controlling for age, BMI, and gender. Significance of associations was corrected for the total number of comparisons using false discovery rate correction. Additionally, correlation tests (using Spearman’s correlation) were conducted to control for outliers, and hence associations were identified as significant if they had both q-values < .25 and correlation test p-values < .05. These tests were performed within the control subjects only by regressing clinical covariates on relative abundance of various taxa at baseline, and in control and treatment subjects together stratified by treatment group. Additionally, changes in relative abundances of taxa and changes in clinical variables of interest were also used for similar tests.

**Predictive modeling of clinical covariates by microbiome composition.** Changes in clinical covariate values, as provided by the GM-Tufts team, in response to the fiber intervention, was tested for predictive capability from the configuration of the pre-intervention microbiome as well as the configuration of the change in the microbiome. These tests were performed using the randomForests package in R.