Host-Microbiome Determinants of Ready-To-Use Supplemental Food Efficacy in Acute Childhood Malnutrition

---

A

Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Master of Science

by

Gabriel Francesco Hanson

December 2023

APPROVAL SHEET

This

Thesis

is submitted in partial fulfillment of the requirements
for the degree of

Master of Science

Author: Gabriel Francesco Hanson

This Thesis has been read and approved by the examing committee:

Advisor: Sean Moore, MS, MD

Advisor:

Committee Member: Kevin Janes, PhD

Committee Member: Jason Papin, PhD

Committee Member: Sepideh Dolatshahi, PhD

Committee Member:

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, School of Engineering and Applied Science

December 2023

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Dr. Sean Moore, MS, MD, for the mentorship, encouragement, and the opportunities he gave me to be a part of such interesting and impactful work. I've learned a great deal from you about how to do science, how to practice medicine, and how to live a life worth living and I hope that one day I can better the world as you do.

I'm extremely grateful to Dr. Zehra Jamil, MBBS, for allowing me to be a part of this project and for all of the hard work and patience. It's been many long hours of work within and across time zones, but I am proud of the work we have done together and it has been a pleasure to work with you.

I'm deeply indebted to Dr. Brett Moreau, PhD, for all of the mentorship and guidance on this project and on others. This project would not have been possible without all of the weekly meetings and feedback you provided.

I'd like to recognize all of the people who worked on this project who I have never had the chance of meeting. Those who designed this project, who went door to door to distribute materials and collect samples, who performed the sequencing and everything in between. Thank you.

I'm extremely grateful to Dr. Jason Papin, PhD, for providing me a space to learn and grow as a part of his weekly lab meetings and for all of the mentorship he has and continues to give.

I'd like to thank Dr. Kevin Janes, PhD, for his time and mentorship as the chair of my thesis committee and a voice of sharp reason. I've learned a great deal from you both in and outside of the classroom. You've provided excellent and timely advice and guidance.

Many thanks to Dr. Deizi Costa, PhD for your mentorship, patience, and friendship. You have taught me most of what I know about wet lab biology and you remain a constant source of positivity that can brighten any day.

Many thanks to Dr. Vinicios Alves da Silva, PhD. You're a great friend, a brilliant scientist, and I've learned a great deal from you.

Many thanks to Dr. Cassandra Hoffmanm, PhD. Your enthusiasm for science is contagious and I appreciate all of the guidance, in person and remotely.

Special thanks to Dr. Sepideh Dolatshahi, PhD for the mentorship on this project and for providing me with a second chance. You've been an excellent mentor and you continue to push me to be my best.

My deepest gratitude for all my friends at UVA, but particularly Yonathan Aberra and Noah Perry. You both have been a source of betterment in my life since the first day I moved here and I am so grateful for having been friends and roommates for the past 4 years.

Finally, words cannot express my gratitude for my family: Dr Randel Hanson, PhD, Dr. Kathryn Milun, PhD, Tobias Hanson, and Francesco Hanson and all of the aunts, uncles and cousins who support me. So much love, so much support, I am deeply indebted to you all for everything I am and everything I will be.

Note:

This work is an adaptation of a manuscript in development entitled:

“Host-Microbiome Determinants of Ready-To-Use Supplemental Food Efficacy in Acute Childhood Malnutrition” written by

*Zehra Jamil, *Gabriel Hanson, Junaid Iqbal, Brett Moreau, Najeeha Iqbal, Aneeta Hotwani, Furqan Kabir, Sheraz Ahmed, Fayaz Umrani, Kamran Sadiq, Kumail Ahmed, Indika Mallawaarachchi, Jennie Z. Ma, Fatima Aziz, ^S Asad Ali, ^Sean Moore.

* These authors have contributed equally to this work

^ Corresponding authors

Author contributions:

Conceptualization: Sean Moore, Asad Ali, Junaid Iqbal, Zehra Jamil
Methodology: Junaid Iqbal, Zehra Jamil, Gabriel Hanson, Brett Moreau
Investigation: Sheraz Ahmed, Fayaz Umrani, Aneeta Hotwani, Furqan Kabir, Kumail Ahmed, Kamran Sadiq, Fatima Aziz
Visualization: Gabriel Hanson, Zehra Jamil
Funding acquisition: Asad Ali, Sean Moore
Project administration: Najeeha Iqbal, Sheraz Ahmed, Fayaz Umrani, Aneeta Hotwani,
Supervision: Sean Moore, Asad Ali
Writing – original draft: Zehra Jamil, Gabriel Hanson
Writing – review & editing: Sean Moore, Brett Moreau, Najeeha Iqbal, Junaid Iqbal, Asad Ali,

Table of Contents

List of Figures, Schematics, and Tables

Abstract:

Ready-to-use-supplemental foods (RUSF) are an essential intervention against severe and moderate childhood acute malnutrition (SAM and MAM) in community-based settings. However, ponderal growth recovery with RUSF is inconsistent. To test the hypothesis that environmental enteric dysfunction (EED)–and EED-associated perturbations of gut microbiota–undermine RUSF efficacy, we measured EED biomarkers, 16S fecal microbiome composition, and RUSF response in a birth cohort of 416 nutritionally at-risk rural Pakistani children. 9-month-old infants with wasting (weight-for-length Z [WLZ] score <-2, n=187) were supplemented with Acha Mum (a chickpea-based RUSF) for 8 weeks then classified as either RUSF responders (RUSF-R, change in weight-for-age Z [ΔWAZ] score >0, n=75) or RUSF nonresponders (RUSF-NR, ΔWAZ≤0, n=112). Machine learning identified fecal, serum, and urine biomarkers of inflammation, EED, and nutritional status (fecal myeloperoxidase and neopterin; serum prealbumin, glucagon like peptide-2, and C-reactive protein; and urine claudin-15 and creatinine) that predicted RUSF response with 73% accuracy. Remarkably, gut microbiome composition before or after RUSF supplementation predicted response with 93% and 98% accuracy, respectively. RUSF-R harbored an increased relative abundance of Negativicutes (including chickpea-fermenting *Veillonella*) and Clostridia prior to RUSF and decreased Gammaproteobacteria thereafter. Seven RUSF-NR outliers whose microbiome predicted RUSF-R experienced high burdens of inflammation and infection. EED, systemic inflammation, and gut microbial signatures herald RUSF nonresponse in childhood wasting. Effective biomarkers and adjunct therapies targeted to host-microbiome factors underlying wasting are urgently needed to unlock the full potential of ready-to-use supplemental foods in high-risk settings.

Introduction

It is estimated that one in five children worldwide suffer from stunted growth, mostly in low and middle income countries and that nearly 45 percent of deaths among children under 5 years of age are linked to undernutrition[1,2]. Even if not fatal, growth stunting in early life hampers cognitive and physical development and is largely irreversible beyond 2 years of life [1]. As such, there is great need to understand the physiological mechanisms that contribute to growth faltering and to develop strategies and interventions to effectively reverse impaired growth trajectories.

The gut microbiome has been increasingly recognized as a major mediator of systemic health particularly in early life. Indeed the human gut microbiome, composed of countless bacteria, but also fungi and viruses, has been shown to modulate systemic metabolism, prime the immune system, and influence brain activity [3]. Food and nutrients that enter the gut are processed and broken down and provide the basis for a complex ecosystem of dense bacterial growth. During birth, an infant is colonized with the microbes from maternal skin, oral, and vaginal communities as well as through interactions with the external environments. Breastfeeding selects for and supports a distinct community of commensal microbes, setting the stage for an infant's healthy gut development [1].

In recent years, scientists and doctors have begun to develop interventions and therapies that target the gut microbiome [4,5,6]. In particular, several large studies of malnutrition have sought to characterize the gut microbiome and develop nutritional therapies that alter gut microbial composition and support infant growth. Indeed, while initially hampered by technological limitations, our understanding of the microbiome and its effects on human health has grown rapidly. With the introduction of high throughput sequencing methods to detect and quantify taxa, and the development of computational techniques to make sense of the data structures and patterns, we are now better able to understand and appreciate how the microbiome exerts its influence on health and can begin to develop interventions that strategically support the colonization of commensal microbes that support growth [7].

In this thesis, I seek to add to the existing literature an analysis of biomarkers, anthropometry, and gut microbiomes in a cohort of stunted Pakistani children who received a chickpea based nutritional intervention. I seek to identify the factors that may impact how an infant responds to a nutritional intervention, biomarkers that could indicate whether an infant will respond, and microbial taxa that may contribute to the response. I hope this work will improve our understanding of how the gut microbiome influences growth and health in early life and help develop more efficient and effective nutritional interventions in the future.

Background

Numerous efforts have been made to achieve the WHO Global Nutrition Targets of reducing wasting to less than 5% and stunting by 40% by 2025 [8]. However, limited success has been achieved due to considerable disparities in socioeconomic and demographic factors in low- and middle-income countries (LMIC) [9]. Sub-optimal growth is multifactorial and to achieve these targets, it is necessary to further explore underpinning mechanistic pathways to improve the results obtained by relevant interventions. To eradicate childhood malnutrition and treat SAM using effective, locally produced ready-to-use therapeutic foods (RUTFs), continuous efforts are made to improve the formulation. Locally manufactured RUSF has been adopted to improve compliance and minimize intervention related gastro-intestinal side effects [10]. However, complementary feeding interventions have shown inconsistent growth improvements in children living in LMIC. In our study conducted on Pakistani children living in Matiari, 187 children received RUSF out of which only 40% cases with SAM or MAM responded to the eight week's nutritional intervention.

In recent years, development of the gut microbial community in young children has been explored to gain better insight on the causal link between the microbiome and undernutrition [11]. Using culture independent methods and machine learning, studies have associated changes in the relative abundance of certain bacterial strains associated with normal microbiota development [12]. In a subsequent study, fecal microbiota of children with severe acute malnutrition (SAM) was found to resemble younger healthy children, hinting at an "immaturity" of the development of the microbiome as a contributing factor to undernutrition [13]. In another study, legume supplementation did not affect structure of gut microbial communities in a clinical trial highlighting the need for a better understanding of the mechanisms of malnutrition and response to interventions [14].

Towards this end, studies where transplantation of microbial communities isolated from undernourished children into germ-free mice have shown reduction in lean body mass gain, alteration in metabolism and varied bone growth [12]. Furthermore, this impaired growth phenotype was reversed by administration of cultured consortium of age and growth discriminatory taxa to the gnotobiotic mice with compromised growth due to colonization with immature microbiota. Based on these lab findings, a clinical study comparing microbiota-directed complementary food (MDCF) based on chickpea along with other combination and rice-lentil-based RUSF as a dietary intervention for children with moderately acute malnutrition (MAM) identified a better response to MDCF [15]. In addition to abundance of bacterial taxa, they reported change in plasma biomarkers, mediators of growth, bone formation, neurodevelopment, and immune function. One of the selected MDCF diets was shown to manipulate the abundance of microbes associated with growth in children in response to the intervention [16]. In another study, Chen et al. harvested bacterial strains isolated from the aspirate of wasted children refractory to nutritional intervention and found negative correlations with linear growth and positive correlations with duodenal proteins involved in immunoinflammatory responses. Colonization of cultures from harvested duodenal strains obtained from children with EED led to development of enteropathy in gnotobiotic mice, further supporting the link between growth stunting, microbiota, and enteropathy [17].

From these observations we set out to explore features due to which only some of the undernourished children responded to conventional supplementary food in clinical studies. We inquired if there could be any difference in their baseline microbiome that might have catalyzed the success of the dietary intervention on growth and health outcomes. And secondly, how does dietary intervention itself modify the gut microbiota. To answer this, we compared the biomarkers of inflammation and fecal microbiome profiles of a subset of responders and nonresponders to RUSF in the SEEM study and their associations with children's anthropometrics. We further evaluated if the pre-intervention fecal microbial community along with serum, fecal and urinary biomarkers may be helpful in predicting the probability of a child to respond to a nutritional intervention. Based on change in weight-for-age (WAZ) scores as a measure of response, we observed the change in the fecal microbial composition with administration of Acha-Mum in a longitudinal follow-up study. We hope the results of this study may lead to enduring benefits from such interventions.

Results

Participant characteristics

In the parent SEEM study, newborns were monitored for six months and enrolled as either reference controls (WLZ ≥0, LAZ ≥-1, n = 51) or cases (WLZ ≤-2, n = 365) based on two consecutive months of anthropometry (Fig. 1A). Of these cases, a subset with WLZ < -2 (n = 187) were selected to participate in an eight-week RUSF intervention at nine months of age. Blood, fecal, and urine samples were collected at nine months for baseline biomarker evaluation (Fig. 1B). Overall, 41% of cases in the SEEM study showed improvement in growth, as defined by reaching WLZ ≥ -2 by one week post-intervention [18].

This study was designed using a subset from the SEEM intervention cohort (n = 60) to define the role of the microbiome in response to nutritional intervention. Selected cases were divided evenly between the responder group (RG), which comprised the 30 best responders, and the nonresponder group (NRG), which consisted of the 30 worst responders. Best and worst responders were selected to maximize the power of this analysis with a smaller cohort size. Response in this sub-study was based on change in WAZ over the course of nutritional intervention rather than WLZ, as we reasoned that weight relative to the expected normal for a specific age was a better marker of an infant's growth trajectory. In addition to cases, a subset of controls (n = 28), who did not receive any intervention, were selected as a reference control group (CG). Fecal samples were collected at four-time points for 16S rRNA sequencing: pre-intervention (nine months), one-week post-intervention (12 months), one month post-intervention (13 months), and six months post-intervention (18 months).

The baseline clinical characteristics of the cases (including both responders and nonresponders) and controls with 16S data are summarized in Table 1. As expected, children in the control group had better anthropometrics at birth, and this difference was even more pronounced by nine months of age. Both maternal and paternal BMI were also significantly higher in the control group. Breastfeeding patterns were similar between both groups over the first six months of life. Each group in this sub-study is broadly representative of the larger SEEM cohort, as shown in **Supplementary Table 1**.

Response to nutritional intervention as measured by change in WAZ

Of the 187 intervention cases from the original SEEM cohort, 75 (40.1%) showed a growth response whereas 112 (59.9%) did not. In this sub-study, the 30 cases classified as responders demonstrated a mean monthly improvement in WAZ of 0.413 over the course of the nutritional intervention (**Fig. 2A, Table 2**). In contrast, the 30 cases classified as nonresponders either experienced no change in WAZ score or saw a decline, with a mean monthly change over intervention of -0.081. Response was associated with significant improvement in malnutrition status in the response group, as children with severe acute malnutrition (SAM) went from 33% of the group to 0% and children with moderate malnutrition (MAM) went from 63% to 20% (**Fig. 2B**). In contrast, rates of SAM increased in nonresponders from 27% at baseline to 57% post-intervention. WAZ and WLZ scores were lower in responders at the time of intervention, indicating that response was not due to improved anthropometry at baseline, although responders did have higher WLZ scores at birth and improved maternal and paternal BMI (**Sup. Table 2**). All other baseline anthropometric measures were similar between groups, including adherence to the Acha mum (chickpea-derived nutritional supplementation) and GI complications such as episodes of diarrhea or vomiting (**Fig. S1**).

We also examined the growth trajectory in these 60 cases after nutritional intervention was concluded. Improvement in WAZ and WLZ scores was seen upon introduction of nutritional intervention in the responder group (**Fig. 2C**). Improved WAZ and WLZ scores continued in this group after intervention was completed, persisting until the end of follow-up at 24 months. The nonresponder group did not see immediate improvement in WAZ or WLZ score upon nutritional intervention, but did see a delayed improvement starting around 13 months. This finding is supported by the significant difference in

monthly change in WAZ and WLZ score over the course of intervention between responders and nonresponders, but the minimal change in these scores between groups once intervention has ended (**Table 2**). Together, these results suggest that the effects of nutritional intervention are more significant and immediate in this subset of responders, and while nonresponders do not see immediate changes in WAZ or WLZ, there is a slight improvement that is delayed. LAZ scores were largely unchanged in both responders and nonresponders over the course of intervention and during follow-up, although by six months post-intervention the change in LAZ was 0.033 in responders and -0.055 in nonresponders for a between-group difference of 0.088. All anthropometric measures gradually declined over the course of follow-up in controls who did not receive nutritional intervention, suggesting that nutritional intervention was still beneficial, even in nonresponders.

Pre-intervention biomarker profiles are distinct between wasted children and controls

We first sought to characterize how biomarkers of growth, inflammation, and gut health differed pre-intervention between case and control groups. To maximize the power of this analysis, biomarker data from the entire SEEM cohort (n = 235), including the 88 participants from this sub-study, were compared. Principal Component Analysis (PCA) was used to identify global differences in biomarker profiles between samples (**Fig. 3A**), and this model was able to separate cases and controls along Principal Component 2 (PC2). Most of the variation along this axis was driven by biomarkers of growth and nutritional status (IGF-1, prealbumin, GLP-2, and leptin) as well as the acute phase proteins CRP and AGP (**Fig. S2**). While control samples clustered fairly tightly, cases had a more heterogeneous distribution along PC1, with a subset of samples designated as cases regardless of nutritional status. PC1 represents most of the variation within the model and was primarily driven by inflammatory cytokines such as IL-1b, IFN-γ, and TNFa, indicating that this subset of cases were defined by altered inflammatory cytokine profiles. PCA for the select 88 study participants in this study (60 wasted children and 28 controls) showed similar clustering as the larger group (**Fig. S3)**, indicating that biomarker profiles from the individuals from this sub-study are broadly representative of the total cohort.

These PCA findings were confirmed by univariate analysis between groups, which identified that biomarkers of growth and nutritional status, including IGF-1, prealbumin, GLP-2, and leptin, were significantly elevated in control samples compared to cases (**Fig. 3B-D**). Cases displayed evidence of systemic inflammation, with significantly higher concentrations of inflammatory biomarkers such as C-reactive protein (CRP), IL-6, IFN-γ, TNF-α, and IL-1β. Interestingly, the anti-inflammatory cytokine IL-10 was also significantly increased in cases compared to controls, which may represent feedback regulation of these inflammatory pathways. Additionally, the urinary biomarker claudin-15 was significantly higher in cases, consistent with previous associations with decreased barrier function observed in EED cases [19,20]. Interestingly, there was no evidence of increased intestinal inflammation in cases compared to controls: myeloperoxidase (MPO) and neopterin (NEO) levels were similar between groups, and lipocalin-2 was elevated in controls compared to cases.

Pre-Intervention Biomarkers are Predictive of Nutritional Intervention Response

Because the heterogeneous biomarker profiles observed by PCA in case samples suggested a mixed population in this group, we hypothesized that the subset of cases with distinct biomarker profiles would be less likely to respond to nutritional intervention. Relabeling of cases by nutritional response confirmed this hypothesis, with most samples that failed to cluster with control samples belonging to the nonresponder group (**Fig. 4A**). This trend was also observed in the 88 children in this 16S subset (**Fig. S4: PCA 16S**). While heterogeneous samples were more likely to be nonresponders, the responder and nonresponder groups as a whole did not clearly separate,, indicating that biomarker profiles were broadly similar between these groups. This finding was further supported by univariate analysis of biomarkers, which found fewer significant differences between responders and nonresponders than observed between cases and controls (**Fig. 4B-D**. Therefore, we took a machine learning approach to identify biomarkers that most accurately predicted response to nutritional intervention.

Biomarkers were incorporated into an ensemble of Random Forest (RF) models with 20-fold cross-validation. Using this approach, the biomarker data collected was able to correctly predict response to nutritional intervention with 60% accuracy. A distinct set of nine features (GLP-2, MPO, prealbumin, NEO, claudin-15, creatinine, CRP, IL-10, and MCP-1) representing a mix of growth and inflammatory status biomarkers were identified as most important for distinguishing responders from nonresponders. There was a trend towards increased markers of growth and nutritional status (GLP-2, prealbumin, creatinine) in responders, which is consistent with the elevation of these markers being beneficial in control samples. Interestingly, there was a trend towards increased intestinal inflammation (MPO, NEO) in nonresponders, a change that was not observed between cases and controls. IL-10 was significantly higher in responders, while CRP was significantly higher in nonresponders, highlighting elevated systemic inflammation as predictive of failure to respond. To interrogate the importance of individual biomarkers, a logistic regression model was trained based on these nine biomarkers and correctly predicted response with 68% accuracy on training set data and 65% accuracy on withheld test set data (**Fig. 4E)**. Regression models built using individual biomarkers or the top 4 biomarkers according to random forest performed more poorly, with accuracy ranging from 40-60%, highlighting the additional predictive capacity gained from incorporating all RF-identified biomarkers **(Fig. 4F)**. Broadly, these results found that biomarkers of growth as well as intestinal and systemic inflammation could distinguish between responder and nonresponder groups.

Pre-intervention fecal microbiome diversity differs between wasted children and controls

Having investigated the relationships between biomarkers and response to nutritional intervention, we next explored the microbiome community profiles of cases (including both responders and nonresponders) and controls. 16S rRNA sequencing was performed on stool samples collected at baseline pre-intervention and at several time points post-intervention, and amplicon sequence variants (ASVs), which represent unique biological sequences that can be assigned a taxonomy, were identified. We first looked broadly at differences in both alpha (within sample) and beta (between sample) diversity between groups. Alpha diversity was significantly lower in controls compared to cases (**Fig. 5A**). This was true both when measuring observed ASVs as a measure of richness and when measuring Simpson's Diversity, which accounts for both richness and evenness. Cases and controls also had significantly different beta diversity, as visualized using nonparametric multidimensional scaling (NMDS) on a Bray-Curtis Dissimilarity matrix (**Fig. 5B**). These results together suggest that cases have a distinct microbiome from controls, characterized by increased diversity and greater evenness of taxa.

Next, we examined community composition to determine which specific phylogenetic groups were different between cases and controls. Consistent with the decreased diversity observed in controls, community composition of control samples was largely dominated by Actinobacteriota, while this phylum made up a significantly smaller proportion of the total community in cases (**Fig. 5C**). Decreased Actinobacteriota relative abundance in cases was compensated by significantly increased abundance of Proteobacteria, as well as trends towards increased Firmicutes and Bacteroidota, when compared to controls. These changes were also reflected at the class level (**Fig. 5D**), with a relative decrease in *Actinobacteria* in cases along with increased *Gammaproteobacteria* and trends towards increased *Bacteroidia*, *Negativicutes*, and *Clostridia*.

Finally, a LASSO regularized, orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA) model was constructed to discriminate between pre-intervention case and control samples and identify ASVs that best discriminate between groups (**Fig 5E-F**). The model outperformed all 1000 randomly permuted models ($p < 0.001$) and achieved a cross validation accuracy of 78%. Variable Importance in Projection (VIP) scores were plotted for top features and reflect ASVs important for discriminating between groups (absolute value of VIP >1 is an important feature, with the sign of VIP indicating which group it is related to) **(Fig S5A)**. This analysis identified ASV1620, a member of the *Bifidobacterium* genus, as particularly important in discriminating between cases and controls. Univariate analysis showed

ASV1620 was significantly enriched in control samples (**Fig. S5b**). This ASV is a member of the Actinobacteriota phylum and *Bifidobacterium* genus, which are known to be important in early life [21]. As LASSO hides features that are linearly correlated to key driver features, ASVs with high correlation to those identified in the model may also be of significance and are shown in a correlation network (**Fig. S5c**). Taken together, these analyses show greater microbiome diversity at the expense of *Bifidobacteria* species is a signature feature of undernourished cases prior to RUSF.

Pre-intervention microbiome profiles are broadly similar between responders and nonresponders

Changes in the microbiome between responder and nonresponder groups were also examined at both a high-level view with microbiome diversity as well as at a more granular level with ASVs associated with response to nutritional intervention. We first focused on baseline differences between responders and nonresponders pre-intervention before looking at longitudinal changes post-intervention. No significant differences in either alpha or beta diversity measures were observed at baseline between responder and nonresponder samples **(Fig 6A-B**), indicating that the microbiomes from these groups are broadly more similar to each other than cases and controls. Community composition in nonresponders mostly consisted of enrichment in Actinobacteriota, while responders had significantly lower Actinobacteriota as a proportion of the total along with increased abundance of Proteobacteria (**Fig. 6C-D**), similar to the changes observed in cases relative to controls. Similar trends were also observed at the class level, where the relative abundance of Actinobacteria and Gammaproteobacteria were significantly decreased and increased, respectively, in responders. In addition to these significant changes, there was a trend towards higher relative abundance of Firmicutes at the phylum level and Negativicutes at the class level in responders at baseline.

We then generated a LASSO regularized, orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA) model to distinguish responders from nonresponders based on gut microbiome composition prior to RUSF. Remarkably, this model outperformed all 1000 randomly permuted models ($p < 0.001$) and achieved a cross validation accuracy of 93%, indicating distinct differences at the ASV level between responders and nonresponders (**Fig. S6A)**. VIP scores were plotted and used to determine ASVs important for discriminating between groups (**Fig. 6E-F).** Univariate analysis showed significant enrichment in several of the important features (**Fig. S6B).** ASVs with high correlation (70%) to those identified in the model that may also be of significance are shown in a heatmap (**Fig. S6c**). Together, this analysis points to a distinct pre-RUSF microbiome state characterized by increased colonization of *Proteobacteria* in children whose ponderal growth markedly improved after RUSF.

Responders exhibit changes in microbiome composition immediately post-intervention, but composition eventually converges in both responders and nonresponders

We next turned our attention to longitudinal changes in the microbiome after nutritional intervention (12 months of age). Alpha diversity increased during the intervention in both responder and nonresponder groups (**Fig. 7A**), a change that was also observed in controls over this time point (**Fig S7a**). No significant changes were observed in the structure microbiome of the nonresponders at the phylum and class levels. However, response to the nutritional intervention was associated with a significant decrease in *Proteobacteria* relative abundance from 9 to 12 months in the responder group, and members of the *Firmicutes* phylum, particularly the classes *Clostridia* and *Negativicutes,* made up a larger proportion of the microbiome as a whole (**Fig. 7B**). While there was also a trend towards higher Clostridia and Negativicutes in this group post-intervention, Negativicutes relative abundance was higher in responders post-intervention than in nonresponders.

In the responders, a significant decline in the relative abundance of Proteobacteria were observed, driven by class Gammaproteobacteria. Importantly, a significant increase in the relative abundance of Negativicutes was seen in the responders and while exploring this trend in the nonresponders, we saw a similar significant trend (**Fig. S7B**). While comparing these two transformations, the relative abundance

of Negativicutes was significantly higher in responders pre and post intervention (**Fig. 7B, Fig. S7c**). Furthermore, a significant increase was also observed in members of the Clostridia class over the course of the intervention in responders. In the nonresponders no taxa was significantly different pre and post intervention. At the family level, the observed trends in the Negativicutes class were driven by members of *Veillonellaceae*, the trends in the Proteobacteria phyla were driven by members of *Enterobacteriaceae,* and the trends in the Actinobacteria phyla were driven by members of the *Bifidobacteriaceae* family (**Fig. 7C**).

To identify specific ASVs associated with RUSF response, we constructed a LASSO regularized OPLS-DA model to discriminate between RUSF-R and RUSF-NR samples immediately post-intervention (12 months of age). The model outperformed all of 1000 randomly permuted models ($p < 0.001$) and achieved a cross validation accuracy of 98% **(Fig S8A)**. A plot of VIP scores showed several ASVs as being important in discriminating between R and NR samples post intervention. ASV2304, a member of the *Lachnoclostridium* genus, was the most important feature of RUSF-NR samples. Univariate analysis showed this ASV to be present significantly more often in nonresponder samples than responder samples. The most important microbial feature relevant to responder samples was ASV578. ASV578, a member of the *Veillonella* genus and Negativicutes class , was present more frequently and more abundantly in responders (**Fig S8b**). ASVs with high correlation (>70%) to those identified in the model that may also be of significance are shown in a heatmap (**Fig. S8c**). Together, this analysis shows an RUSF-driven transformation of the gut microbiome in RUSF-R distinct from that observed in RUSF-NR over the course of the nutritional intervention. This transformation is characterized by a robust decrease in Gammaproteobacteria and an increase in Negativicutes.

Pre-intervention raised biomarkers led to nonresponsiveness despite fecal microbiome similarity with responders

We previously identified several nonresponders as having distinct biomarker profiles compared to control and responder samples, and these profiles were primarily driven by a significant increase in biomarkers of systemic inflammation (**Fig. 8A**). We next asked whether these individuals had a distinct microbiome profile compared to nonresponders with similar immune profiles to responders. The microbiomes of nonresponders with a more inflammatory biomarker profile were more similar to responders than to non-inflammed nonresponders (**Fig. 8B**). This was most evident by lower levels of Actinobacteriota and increased levels of *Proteobacteria* when compared to non-inflammed nonresponders (**Fig. 8C**). Follow up analysis showed these patients reported higher rates of diarrheal episodes and acute respiratory infections (**Supp table 3**). Together, this seems to indicate that while microbiome structure pre-intervention plays a role in determining response, systemic inflammation and acute illness play an orthogonal role.

Discussion:

We present results of a study conducted on Pakistani children aged 9-12 months with moderate and severe acute malnutrition who received nutritional intervention. We assessed inflammatory biomarkers before intervention and monitored response to intervention as an improvement in WAZ. The fecal microbiome was explored longitudinally while we measured anthropometrics up to 24 months of age. We found that pre-intervention biomarker profiles differentiated the two response groups; children with heightened inflammation being less likely to respond to supplementation. Despite differences in gut microbiomes between responders and nonresponders prior to the intervention, their microbiome profiles converged to share a similar composition over the course of six months post-intervention.

Responders were born with significantly higher birth weights, supporting the role of antenatal care and birth dynamics in achieving optimum growth later in life. Similarly, a trend of higher maternal or paternal BMI was seen in association with the response but was insignificant. We identified a set of nine biomarkers with a decent predictive accuracy for children at risk of failing to respond in our cohort. This included both; the biomarkers of good growth (prealbumin, leptin, GLP-2 and IGF-1) as well as markers of inflammation (fecal MPO, serum AGP and CRP). These potential biomarkers of response have been explored in association with microbial infections in the context of environmental enteropathy [22]. These findings provide valuable insight into the role of baseline biomarkers to predict those cases that may require additional therapeutic interventions to respond effectively to supplementation in the clinical settings.

Differences in the fecal microbiome profile at pre-intervention were more pronounced between the cases and controls than between responder and nonresponder cases highlighting a consistent stunted phenotype. With a greater alpha diversity seen in the cases, we saw a more diverse microbiome found in their fecal samples compared to controls. The phylum Actinobacteria which includes the genus *Bifidobacterium* was higher in our controls while phylum Proteobacteria and Bacteroidia in the undernourished cases. However, within the undernourished cases, Actinobacteria had no association with change in WAZ in response to intervention. *Bifidobacteria* are important members of the microbiome in early life and dominate the gut of breastfed children, however their abundance declines by the end of infancy [23]. This association has been confirmed by Chen et al which observed a negative association between a *Bifidobacterium* species and microbiota-directed complementary food intervention associated change in WLZ [16]. The higher abundance of *Actinobacteria* in controls might be explained by the fact that these children may not have served as true controls with a WLZ just above the borderline yet lower than the cutoffs to be labeled as healthy controls (median WAZ of -0.11 at 9 months).

Changes in the gut microbiota were observed in both responders and nonresponders to the intervention. The abundance of Gammaproteobacteria and Negativicutes was found to be linked to the response to intervention. Specifically, we saw Negativicutes increased over time, while Gammaproteobacteria decreased. Gammaproteobacteria, a large class of bacteria that includes harmful species like *Klebsiella*, *Salmonella*, and *Shigella*, were found in higher numbers before the intervention and decreased significantly afterwards in the responders. We anticipate that the high prevalence of these facultative anaerobes prior to intervention may have contributed to the observed growth faltering in these children. However, the nutritional intervention appeared to be sufficient in restructuring their microbiome towards a composition more conducive to growth, such as an increase in Negativicutes including *Megasphera* and *Veillonella*. Studies have shown these species to ferment chickpeas and release beneficial phenolic acids [24]. This could explain the link between the increased presence of these bacteria and the use of chickpeas in the Acha-mum used as supplementation in our study. Furthermore, these bacteria have been identified in other nutritional studies as beneficial in preventing necrotizing enterocolitis, suggesting a positive role for them in the gut [14]. Interestingly, we also saw several species of the genus *Prevotella* in our analysis. For example the abundance of *Prevotella copri* increased with age in all the children. Several *Prevotella* species inhabit the human gut, among which *P. copri* is the most prevalent in non-industrialized settings likely due to its association with high-fiber low-fat diets [22]. Increases in *P. copri* accompanied by

reductions in *B. longum* in association to microbiota-directed complementary food has been reported in a Bangladeshi study [16]. However, conflicting observations related to P. copri and human health suggest the need to further characterize human microbial diversity for a better understanding of its function [25,26].

This study possesses several strengths. Firstly, it included children with varying degrees of wasting, from moderate to severe, and involved meticulous sample collection. Secondly, we included controls from the same rural settings who did not receive the nutritional intervention, providing a set of age-matched controls unaffected by the intervention yet living in the same environment and consuming similar diets. Thirdly, longitudinal collection of the fecal samples enabled us to observe the restructuring of the microbiome and its correlation with response. Lastly, by monitoring anthropometrics up to 24 months of age, we explored the long-term effects of the intervention, particularly on LAZ scores, as delayed improvement was observed in subsequent months. However, the study also had limitations. These included exploration of the 16S rRNA fecal analysis of 30 best and 30 worst responders, which we assumed represented the larger group. We focused on the extreme responders to study WAZ-associated microbial taxa that might not have been evident in cases with minimal weight change. Another limitation was the lack of fecal sample collection during the intervention, which could have better captured the dynamics of microbiome transformation. Because species-level resolution was not consistently achieved with this 16S approach, any potential probiotic candidates need to be confirmed in gnotobiotic mice before clinical trials. Lastly, while machine learning tools were useful for identifying associated taxa, they may assign strong weights to taxa found in extremely low abundance in one group, so results should be interpreted with caution.

For future studies, we suggest exploring longer intervention durations, especially for nonresponders, as we observed late catch-up in the transformation of a few ASVs. Profound growth improvements during the supplementation duration cease afterwards, suggesting the need for follow-up support during critical growth periods, such as the first 1000 days of life. Efforts to minimize oro-fecal contamination, use clean drinking water, and treat enteropathogen infections with anti-inflammatory drugs or antibiotics may enhance the nutritional intervention's impact. Our results underscore the importance of healthy birth anthropometrics, emphasizing the role of maternal health in optimal childhood development. Probiotic co-administration based on the intervention's composition may assist in achieving desirable results, hence further studies should focus on this area. Finally, strategies to incorporate pre-screening the inflammatory state of a child through biomarker quantification and potentially delaying the intervention until acute inflammation is reduced may increase intervention efficiency. To conclude, baseline characteristics including host factors, gut inflammation and birth dynamics have an equally important effect on the response to intervention as of baseline fecal microbiome or its transformation across the intervention. Multi-centered studies are needed to explore microbiota with an added potential as biomarkers. Efforts are needed to harmonize methods and interpretation of findings and compile microbiome data from diverse sites with a high prevalence of malnutrition as it is critical for an informed and meaningful clinical and public health practice.

Material and Methods

Study participant details: Study participants were selected based on the availability of fecal samples for longitudinal analysis of gut microbiota from a larger SEEM study in which children were recruited as controls and undernourished cases at risk of EED from a rural setting (Matiari) in Pakistan. The SEEM study explored the gut tissue histology, transcriptomics and intestinal permeability in undernourished cases refractory to nutritional intervention. The study was approved by Aga Khan University Ethical Review Committee (ERC # 2021-0535-19973) while written informed consent was obtained from the participant's parents. The study protocol is already published [27].

For enrollment in the SEEM study, newborns were followed up to six months of age and on the basis of two consecutive months of anthropometric measurements they were enrolled either as controls (WLZ $\geq$0, LAZ $\geq$-1) or cases (WLZ $\leq$-2). Out of the 350 cases, 187 were selected for RUSF nutritional intervention. The response to this intervention was monitored as improvement in WLZ $\geq$ -2 at one week after post intervention irrespective of their WLZ at pre-intervention time point. 41% of cases responded in the SEEM study. For this sub-study, those children with a maximum gain in WAZ (change > 0.5) were selected as the responder group (n=30) while children who experienced a decline in WAZ or no change were selected as nonresponder group (n=30). Regarding nutritional status of these 60 children, at the beginning of the intervention two were at risk of wasting, 40 had MAM while 18 had SAM. A third group was included in the analysis as a control group (CG). These children were enrolled in the study as controls and did not receive any intervention (n=28), however, they reported a gradual decline in their anthropometrics with age.

Data collection: Information on general demographic characteristics was collected by trained community health workers through interviews with the mothers, which included the child's birth dynamics, breastfeeding history, and maternal parameters. Anthropometric measurements were performed using standard procedure and equipment, with weight measured to the nearest 20g precision electronic scale (TANITA 1584) and length to the nearest 1 mm using rigid length board with a movable foot piece. These measurements were collected on a monthly basis up to the age of 24 months. Based on established cut-offs for nutritional indicators, the participants were categorized as stunted (LAZ< -2), underweight (WAZ < -2), at risk of wasting (WLZ <-1 and $\geq$ -2), moderately wasted (WLZ < -2) and severely wasted (WLZ < -3), respectively.

Sample collection for inflammatory biomarkers: Blood, fecal, and urinary samples were collected from the participants at nine months of age (**Fig. 1A**). 1 to 2 ml blood was collected from which serum was aliquoted in small volumes to avoid freeze-thawing and transported at 4°C from the field site lab to the Infectious Disease Research Laboratory (IDRL) where they were stored at -80°C until processed. Commercial ELISA kits were used for estimation of GLP-2 (USCN, Life Sciences Inc, Wuhan, China) while CRP, ferritin and AGP were analyzed using a Hitachi 902 analyzer (Roche Diagnostics, Holliston, MA) and IGF-1 was measured using a LIAISON (Diasorin Saluggia (VC) Italy). All assays were performed following manufacturer protocols. For fecal samples, the caretaker of the child collected the sample using a wooden spatula into a clean container provided in the stool collection kit. The fecal sample was transferred into multiple cryogenic vials for long-term storage.Commercial ELISA kits were used for the estimation of MPO (Immunodiagnostic AG, Stubenwald-Allee, and Bensheim) and NEO

(GenWay Biotech, San Diego. CA). Fecal lipocalin (LCN) was measured by DuoSet ELISA DY1757. All plates were read on Biorad iMark (Hercules, CA) plate reader.

For the evaluation of serum cytokines, commercially available MILLIPLEX MAP Human Cytokine/Chemokine (MERCK) kit was used. The screening panel including IFN-γ (interferon- γ), IL-10, IL-12 (IL-12p70), IL-1β, IL-6 ,IL-8 ,IP-10 (Interferon-gamma-induced protein 10, also called as CXCL10), MCP-1((Monocyte chemoattractant protein-1, also called as CCL2) and TNF-α (Tumor necrosis factor) as per manufacturers' instructions using Bioplex- 200 instrument. The data was analyzed using Bioplex Manager 6.1.

Administration of nutritional intervention: Enrolled cases with WLZ < -2 at nine months were selected for nutritional intervention. Acha-mum which is a Ready to use supplementary food (RUSF), was given at a dose of one sachet per day to cases with MAM (WLZ scores between -2 and -3) while children with MAM (WLZ < -3) were administered sachets as per child's weight (200 kcal/kg/day) (Organization WHO, 2019. Meeting report: WHO technical consultation: nutrition-related health products and the World Health Organization model list of essential medicines–practical considerations and feasibility: Geneva, Switzerland, 20–21 September 2018). Its composition is provided in the supplementary material. This eight-week intervention was monitored through weekly visits by the local team to document compliance, side effects, needs for medical assistance, and other details. Compliance was calculated (weekly) based on the empty wrappers returned by mothers [Compliance = (Total packet used/total packet given) *100]. One week after completion of the intervention, the response was measured as the overall change in the WAZ of the children.

Longitudinal fecal sample collection for 16S ribosomal RNA gene sequencing: Mothers of the study participants were trained to collect the fecal samples and provided with mobile cards for timely communication with the collection team in the field. As described above, the fecal sample was collected by the mothers who sent an immediate message to the field team to pick up the sample within 30 minutes. Overall four samples were collected from each child; one sample at nine months to capture microbiota at the pre-intervention stage while three samples were collected at twelve, thirteen, and eighteen months to evaluate post-intervention taxa (**Fig. 1b)**. Fresh samples were transferred into a pre-chilled cryovial, snapped to the aluminum cryo cone and placed into a freezing container (Coleman) at 2° to 8°C. The time taken from the passage of fresh stools to being snap frozen in a dry shipper was less than 30 minutes. The Coleman was carried in liquid nitrogen dry shipper to the local laboratory and later shifted to IDRL, AKU on dry ice where samples were stored at -80C until shipped to BGI Genomics (formerly Beijing Genomics Institute). 30ng qualified DNA template was tested for sample integrity by agarose gel electrophoresis and concentration by a microplate reader (Qubit fluorometer). The quantified samples (6 -100ng/uL) were normalized to 30ng DNA per reaction. fusion primers were designed to include Illumina adapter sequences, an 8-nucleotide index sequence and gene-specific primer and added to the Polymerase chain reaction (PCR) reaction system. All PCR products were purified by Agencourt AMPure XP beads, dissolved in Elution Buffer and labeled for library construction. Library size and concentration were detected by Agilent 2100 Bioanalyzer. Qualified libraries were sequenced pair-end on the HiSeq 2500 platform according to their insert size. The raw reads were filtered to remove the adapter and low-quality bases. Paired-end reads were added to the tags by Fast length Adjustment of Short read program (FLASH, v12.11)The sequence data are being deposited in NCBI Sequence Read Archive (SRA). Alpha rarefaction curves were generated to assess the effect of sampling depth on ASV abundance. The plots indicate the

detection of ASVs had already attained a plateau at 60000 reads. This trend was separately confirmed in the three groups.

16S ribosomal RNA gene analysis and statistical analysis: Sequence analysis was performed in R using DADA2 (version 1.22.0). The forward read was truncated to 200 base pairs and reads with ambiguous 'N' bases, and >2 expected errors were removed. Chimeras were removed. Forward and reverse reads were aligned and resulting sequence variant counts (SVs) and taxonomic calls were assigned using the Greengenes 16S rRNA gene database. Models were created and data were analyzed in R using the following packages: Phyloseq (v. 1.38.0), caret (v. 6.0.92), GGplot2 (v. 3.3.6), mixOmics (6.18.1) and in python using Jupyter notebooks, pandas, scipy, numpy, matplotlib, and seaborn. Venn diagrams were created using the R packages VennDiagram (1.7.3), and ggvenn (0.1.9). Radar plots were created using the R package fsmb (0.7.3). Wilcoxon rank sum tests were used to compare the alpha diversity and genus relative abundances.

Model generation: The R package randomForest (4.7-1.1) was used to generate a Random Forest (RF) model to classify the samples according to the response to the nutritional intervention and determine important features. Relative SV counts were used to train the RF. The forest used has 1000 trees with a node split (mtry) set to the default of the square root of the number of samples. Biomarkers that were important for classifying patient response to the nutritional intervention at 9 months were also identified using an RF model. Patients with more than one missing test and biomarkers with more than 3 missing patients were removed. After this filtering, we were left with 148 observations (children) and 23 variables (biomarkers listed in the supplementary file).

OPLSDA models were created using MATLAB using a custom pipeline initially developed by Remziye Erdogan. Lasso regularization was used to better guide the feature selection and model fitting process and achieves improved classification by allowing the selection of a subset of the covariates instead of using all of them. Here, a 5-fold cross-validation was repeated 1000 times to calibrate the performance of the model. Permutation testing results shuffling the labels of the samples show goodness of fit of the model vs a null distribution a and the cross validation accuracy showing the stability of the model are reported in supplemental figures. Principal component analysis was performed in R and plotted using ggbiplot (0.55). Correlations between SV abundance and biomarker levels were plotted using corrplot (0.92).

All code and data are available on github (https://github.com/gabehanson/SEEM_microbiome_analysis)

Bibliography:

1. Cowardin CA, Syed S, Iqbal N, Jamil Z, Sadiq K, Iqbal J, et al. Environmental enteric dysfunction: gut and microbiota adaptation in pregnancy and infancy. Nat Rev Gastroenterol Hepatol. 2023;20(4):223–37.
2. World Health Organization. (2021, June 9). "Malnutrition" Retrieved from https://www.who.int/news-room/fact-sheets/detail/malnutrition
3. Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. Genome Med 2016;8(1):1–11.
4. Mahfuz M, Das S, Mazumder RN, Masudur Rahman M, Haque R, Bhuiyan MMR, et al. Bangladesh Environmental Enteric Dysfunction (BEED) study: Protocol for a community-based intervention study to validate non-invasive biomarkers of environmental enteric dysfunction. BMJ Open. 2017;7(8).
5. Murray-Kolb LE, Rasmussen ZA, Scharf RJ, Rasheed MA, Svensen E, Seidman JC, et al. The MAL-ED Cohort Study: Methods and Lessons Learned When Assessing Early Child Development and Caregiving Mediators in Infants and Young Children in 8 Low-and Middle-Income Countries. 2014;
6. Chen RY, Mostafa I, Hibberd MC, Das S, Mahfuz M, Naila NN, Islam MM, Huq S, Alam MA, Zaman MU, Raman AS, Webber D, Zhou C, Sundaresan V, Ahsan K, Meier MF, Barratt MJ, Ahmed T, Gordon JI. A Microbiota-Directed Food Intervention for Undernourished Children. N Engl J Med. 2021 Apr 22;384(16):1517-1528. doi: 10.1056/NEJMoa2023294. Epub 2021 Apr 7. PMID: 33826814; PMCID: PMC7993600.
7. Hernández Medina R, Kutuzova S, Nielsen KN, Johansen J, Hansen LH, Nielsen M, et al. Machine learning and deep learning applications in microbiome research. ISME Commun. 2022;2(1):1–7.
8. Organization WH. Global Nutrition Monitoring Framework: operational guidance for tracking progress in meeting targets for 2025. 2017.
9. Verma P, Prasad JB. Stunting, wasting and underweight as indicators of under-nutrition in under five children from developing Countries: A systematic review. Diabetes & metabolic syndrome. 2021;15:102243.
10. Das JK, Salam RA, Saeed M, et al. Effectiveness of interventions for managing acute malnutrition in children under five years of age in low-income and middle-income countries: a systematic review and meta-analysis. Nutrients. 2020;12:116.
11. Monira S, Nakamura S, Gotoh K, et al. Gut microbiota of healthy and malnourished children in Bangladesh. Frontiers in microbiology. 2011;2:228.
12. Blanton LV, Charbonneau MR, Salih T, et al. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. Science. 2016;351:aad3311.
13. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature. 2014;510:417.
14. Ordiz MI, Janssen S, Humphrey G, et al. The effect of legume supplementation on the gut microbiota in rural Malawian infants aged 6 to 12 months. The American journal of clinical nutrition. 2020;111:884-92.
15. Gehrig JL, Venkatesh S, Chang H-W, et al. Effects of microbiota-directed foods in gnotobiotic animals and undernourished children. Science. 2019;365:eaau4732.
16. Chen RY, Mostafa I, Hibberd MC, et al. A microbiota-directed food intervention for undernourished children. New England Journal of Medicine. 2021;384:1517-28.
17. Chen RY, Kung VL, Das S, et al. Duodenal microbiota in stunted undernourished children with enteropathy. New England Journal of Medicine. 2020;383:321-33.

18. Jamil Z, Iqbal NT, Idress R, et al. Gut integrity and duodenal enteropathogen burden in undernourished children with environmental enteric dysfunction. PLoS neglected tropical diseases. 2021;15:e0009584.
19. Marie, C., Ali, A., Chandwe, K., Petri, W. A., & Kelly, P. (2018). Pathophysiology of environmental enteric dysfunction and its impact on oral vaccine efficacy. *Mucosal Immunology*, *11*(5), 1290–1298.
20. Harper KM, Mutasa M, Prendergast AJ, Humphrey J, Manges AR (2018) Environmental enteric dysfunction pathways and child stunting: A systematic review. PLOS Neglected Tropical Diseases 12(1): e0006205.
21. 1. Underwood MA, German JB, Lebrilla CB, Mills DA. Bifidobacterium longum subspecies infantis: Champion colonizer of the infant gut. Pediatr Res. 2015;77(1):229–35.
22. Uddin MI, Hossain M, Islam S, et al. An assessment of potential biomarkers of environment enteropathy and its association with age and microbial infections among children in Bangladesh. Plos one. 2021;16:e0250446.
23. O'Brien, C. E., Meier, A. K., Cernioglo, K., Mitchell, R. D., Casaburi, G., Frese, S. A., Henrick, B. M., Underwood, M. A., & Smilowitz, J. T. (2022). Early probiotic supplementation with B. infantis in breastfed infants leads to persistent colonization at 1 year. *Pediatric Research*, *91*(3), 627–636.
24. Perez-Perez LM, Huerta-Ocampo JÁ, Ramos-Enríquez JR, et al. Interaction of the human intestinal microbiota with the release of bound phenolic compounds in chickpea (Cicer arietinum L.). International Journal of Food Science & Technology. 2021;56:6497-506.
25. Yeoh YK, Sun Y, Ip LYT, et al. Prevotella species in the human gut is primarily comprised of Prevotella copri, Prevotella stercorea and related lineages. Scientific Reports. 2022;12:9055.
26. De Filippis F, Pasolli E, Tett A, et al. Distinct genetic and functional traits of human intestinal Prevotella copri strains are associated with different habitual diets. Cell host & microbe. 2019;25:444-53. E3.
27. Iqbal NT, Syed S, Sadiq K, Khan MN, Iqbal J, Ma JZ, et al. Study of Environmental Enteropathy and Malnutrition (SEEM) in Pakistan: Protocols for biopsy based biomarker discovery and validation. BMC Pediatr. 2019;19(1):1–17.

Tables:

Table 1: Baseline Clinical Characteristics of cases and controls with 16S data

Variables	All cases with 16S	Controls	P-values
Total Number	60	28	
Demographics			
Gender (# of males)	39	18	0.940
Birthplace (hospital vs home)	43 vs 17	24 vs 3	0.120
Gestational age- median (q25, q75)	39 (38, 39)	39 (39, 39)	0.054
Birth anthropometrics			
WAZ	-1.99 (-2.69, -1.32)	-1.03 (-1.39, -0.31)	**0.000**
LAZ	-1.59 (-2.51, -1.06)	-0.88 (-1.54, -0.16)	**0.001**
WLZ	-1.39 (-1.99, -0.72)	-0.45 (-1.16, -0.09)	**0.002**
Maternal factors			
age (at the time of delivery)	27 (23.5, 35)	27 (25, 30)	0.263
weight (at the time /after	46.0 (40.75, 50.5)	49.25 (44.5, 53.60)	**0.050**
height	153.25 (150.39, 156.49)	151.95 (148.25, 55.55)	0.198
BMI	19.44 (17.63, 21.89)	20.92 (19.76, 24.46)	**0.045**
Paternal BMI	19.39 (18.47, 21.89)	22.46 (20.27, 24.32)	**0.012**
Breast feeding status (N)	60	28	
exclusive (%)	36.67	35.71	
partial (%)	55	53.57	
none (%)	8.34	3.34	
N/A	0	1	
Pre-intervention anthropometrics			
WAZ 9mo	-3.465	-0.11	**0.000**
LAZ 9mo	-2.585	-0.88	**0.000**
WLZ 9mo	-2.555	0.53	**0.000**
Pre-intervention status n(%)			
Underweight (≤-2)	57	0	
Stunted (≤-2)	39	5	
Wasted (≤-2)	53	0	
Pre-intervention markers			
Hemoglobin	10.45 (9.2, 11.4)	10.45 (9.5, 11.3)	0.267
Pre-albumin	14.2 (11.55, 16.15)	15.2 (13.6, 16.35)	0.322
Serum AGP mg/dl	116.445 (91, 138.05)	115 (85, 134)	0.207
Serum IGF-1 ng/ml	17.67 (10.41, 31.73)	24.46 (17.57, 36.48)	0.055
Serum CRP	0.202 (0.061, 0.56)	0.13 (0.07, 0.37)	**0.030**
Serum Ferritin	21.8 (11, 55)	16.95 (5.8, 34.8)	0.212
Serum Leptin	145.06 (94.96, 260.20)	258.17 (179.39, 345.51)	0.071
Serum GLP-2	1156.45 (863.85, 1532.66)	1260.87 (649.51, 2286.14)	0.597
Fecal MPO ng/ml	3550 (1575, 8829.75)	4672.25 (2640.25, 9275)	0.396
Fecal NEO nmol/L	1925 (1042.15, 2875)	1341.5 (485.95, 2600)	0.914
Fecal Lipocalin	27422.4 (16750.4, 41399.6)	40097.9 (20119.5, 94775)	**0.024**
Urine claudin-15	2.06 (0.86, 2.87)	0.92 (0.69, 1.15)	**0.000**
Urine creatinine	125.63 (91.94, 263.56)	176.57 (129.65, 212.41)	0.841

Table 2: Clinical response to the RUSF supplementation in responders and nonresponders

Anthropometric features	Responders	Nonresponders	Difference (95% CI) †
At baseline (mean ± SD)			
Weight-for-age z score	-3.740 ± 1.140	-3.348 ± 1.075	-0.392 (-0.989, 0.204)
Length-for-age z score	-2.774 ± 1.393	-2.487 ± 1.260	-0.287 (-1.002, 0.428)
Weight-for-length z score	-2.803 ± 0.702	-2.564 ± 0.858	-0.239 (-0.661, 0.183)
Mean rate of growth per month during intervention (95% CI) †			
Weight-for-age z score	0.413 (0.360, 0.466)	-0.081 (-0.133, -0.028)	0.494 (0.401, 0.588)
Length-for-age z score	-0.006 (-0.048, 0.037)	-0.082 (-0.124, -0.039)	0.076 (-0.005, 0.157)
Weight-for-length z score	0.481 (0.398, 0.563)	-0.130 (-0.212, -0.048)	0.611 (0.500, 0.722)
Mean rate of growth during six months of follow-up (95% CI) ‡			
Weight-for-age z score	0.046 (0.011, 0.081)	0.022 (-0.013, 0.058)	0.024 (-0.035, 0.083)
Length-for-age z score	0.033 (-0.001, 0.066)	-0.055 (-0.085, -0.026)	0.088 (0.044, 0.133)
Weight-for-length z score	0.009 (-0.041, 0.058)	0.053 (0.008, 0.097)	-0.044 (-0.111, 0.023)

†Values for the between-group difference at baseline were derived from a linear model predicting anthropometric features at the start of treatment as a function of the treatment group after adjustment for age and sex. Values for the between-group difference in the growth rate per month during the treatment period and during follow-up were derived from a mixed-effects linear model predicting anthropometric features as a function of the interaction between treatment group and the number of weeks since the initiation of nutritional supplementation after adjustment for the baseline variables plus the number of weeks of treatment, the treatment group, and a random intercept for each participant to account for the within-participant correlation. Positive values indicate a faster growth rate in responders.
‡ Values for the rate of growth per month during the six months of follow-up were derived from a mixed-effects linear model. This model predicts anthropometric features as a function of the number of weeks since the initiation of nutritional supplementation after adjustment for the baseline variables plus a random intercept for each participant to account for the within-participant correlation.

Figures:

Figure 1

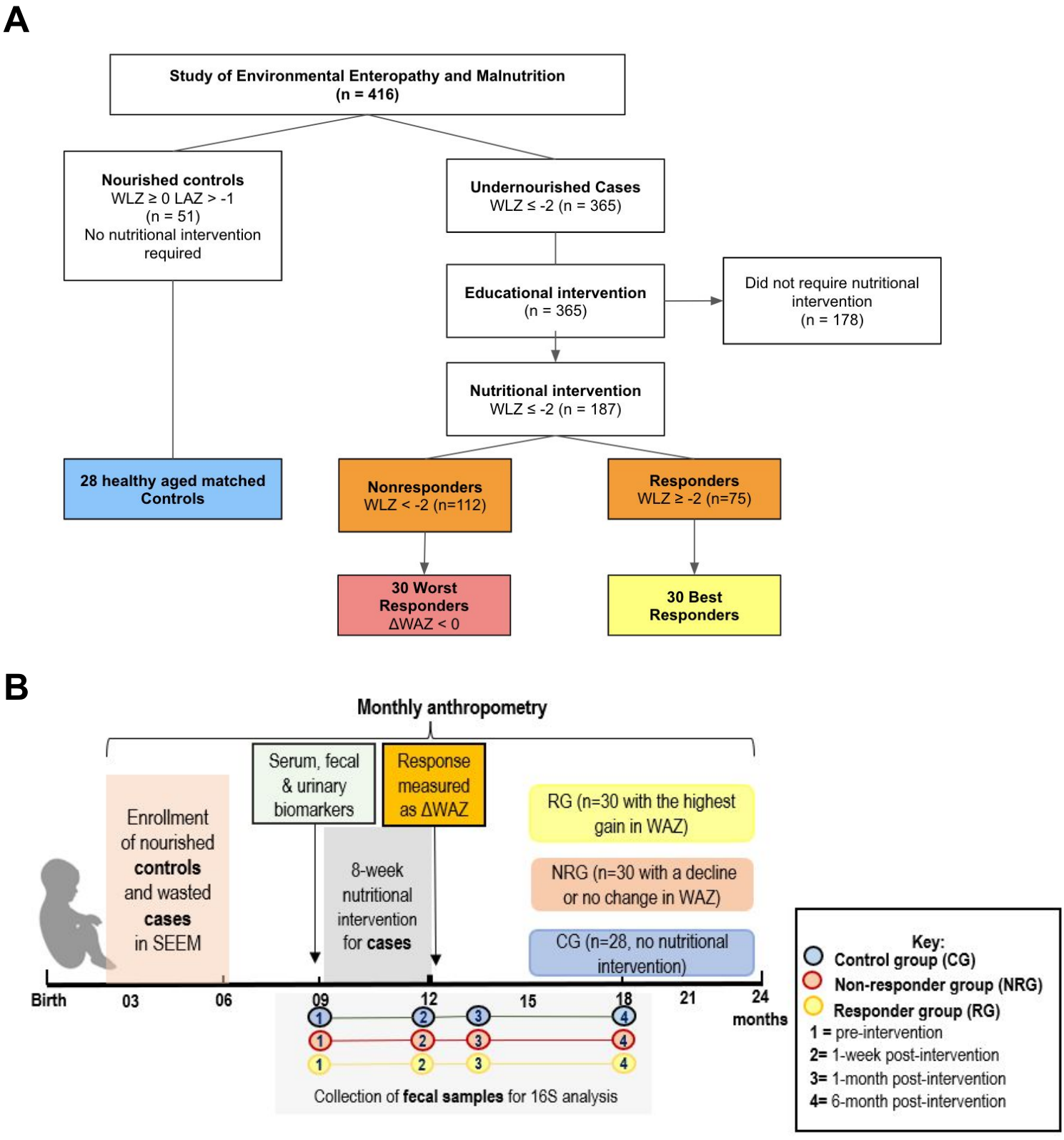

Figure 1: Overview of study design. A: decision tree/ description of study **B:** infogram of study enrollment and sample collection.

Figure 2

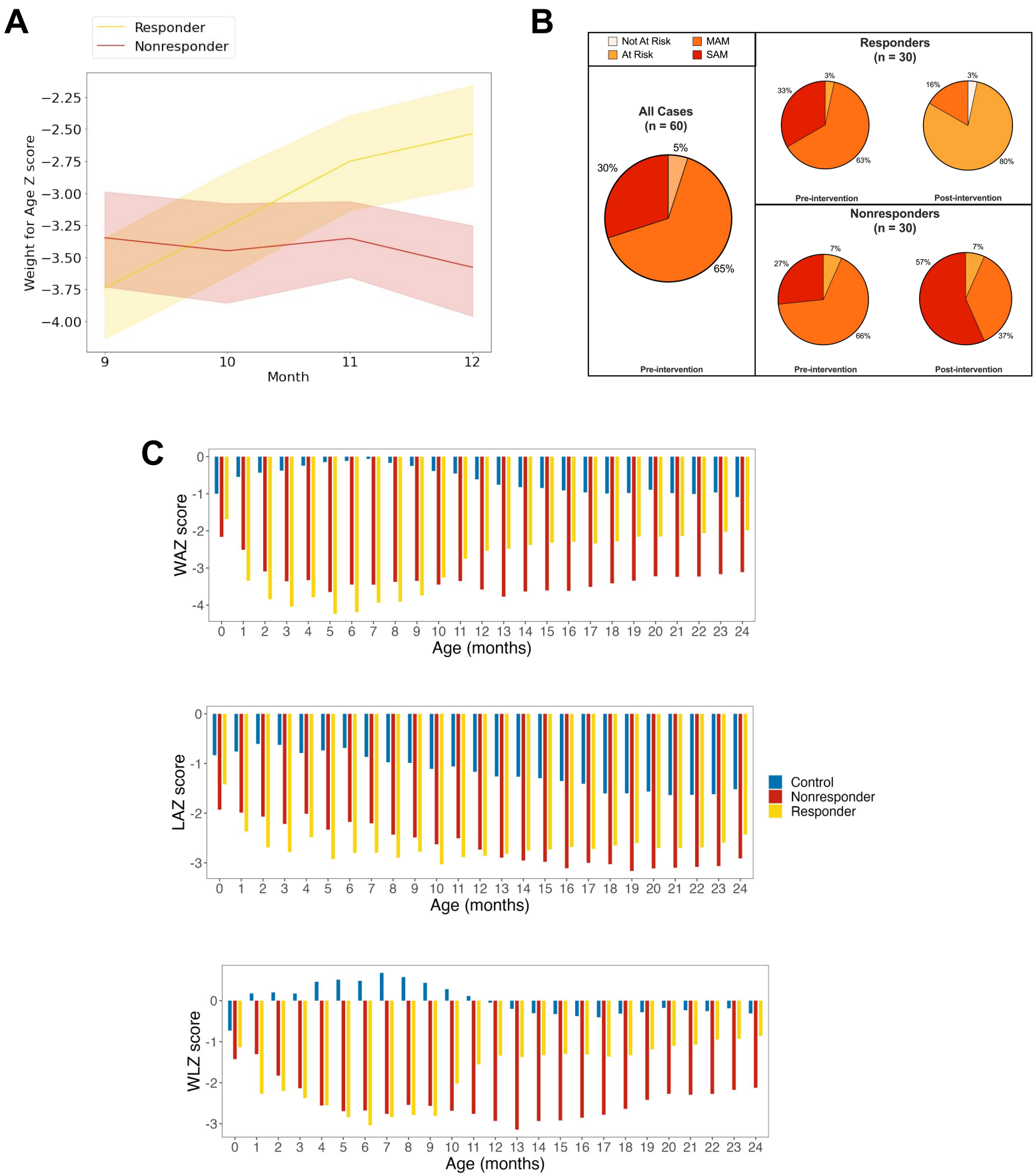

Figure 2: Children selected for microbiome analysis show diverging yet durable responses to NI. (A) Average change in WAZ score per month during NI. **(B)** Nourishment status of responders and nonresponders in 16s cohort before and after NI. **(C)** On average, responders had worse anthropometrics prior to NI, but showed persisting growth improvements in response to NI as measured by weight for age z scores, length for age z scores and weight for length z scores. Aged and region matched control cohort shown in blue.

Figure 3

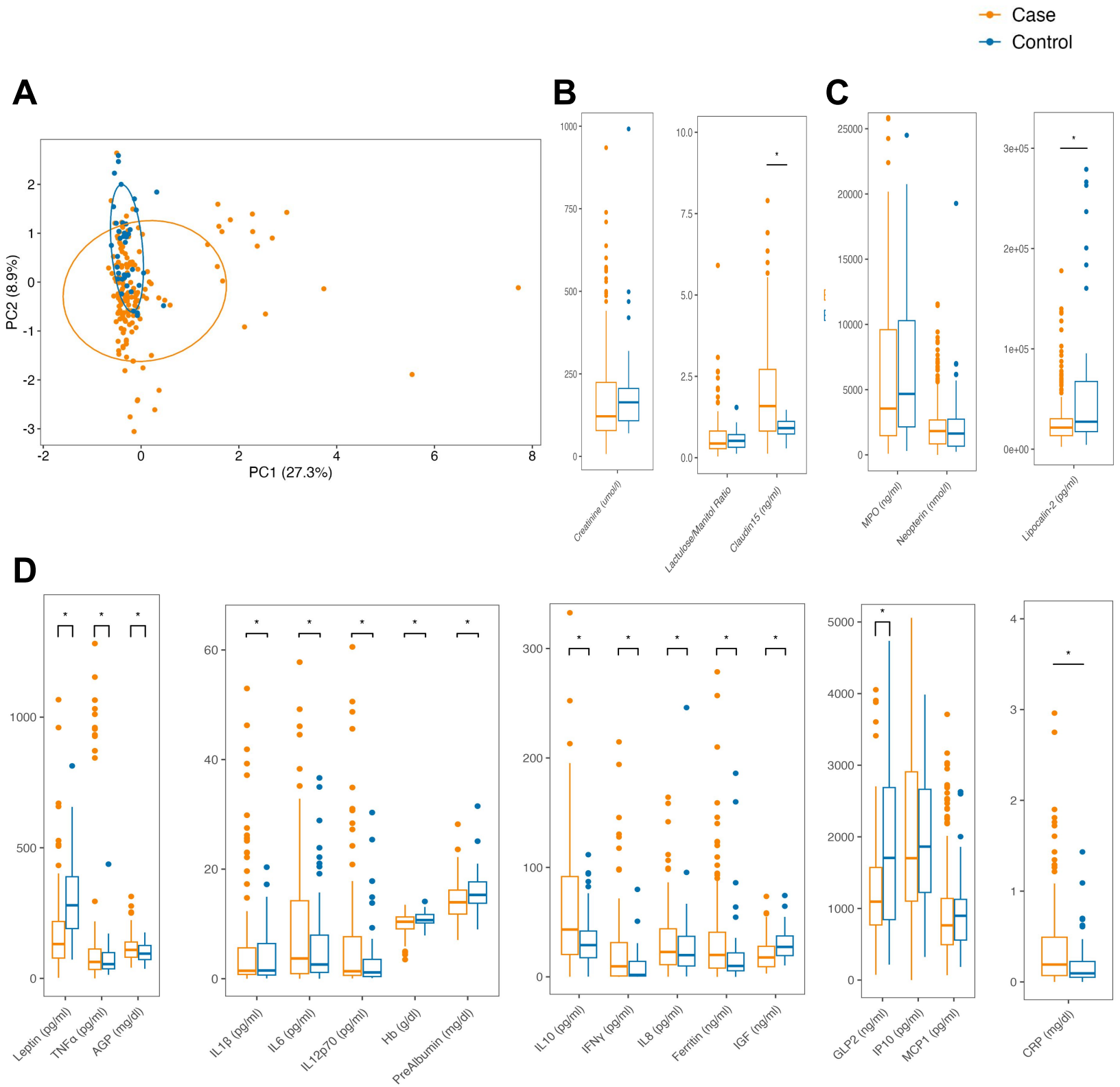

Figure 3: Stunted children exhibit increased systemic inflammation compared to healthy age matched controls. (A) Principal component analysis of biomarker profiles from stunted children (n=148) and healthy aged-matched children (n=39). **(B)** Urine, **(C)** fecal, and **(D)** serum biomarker/cytokine levels from children at 9 months of age prior to NI in SEEM cohort.

Figure 4

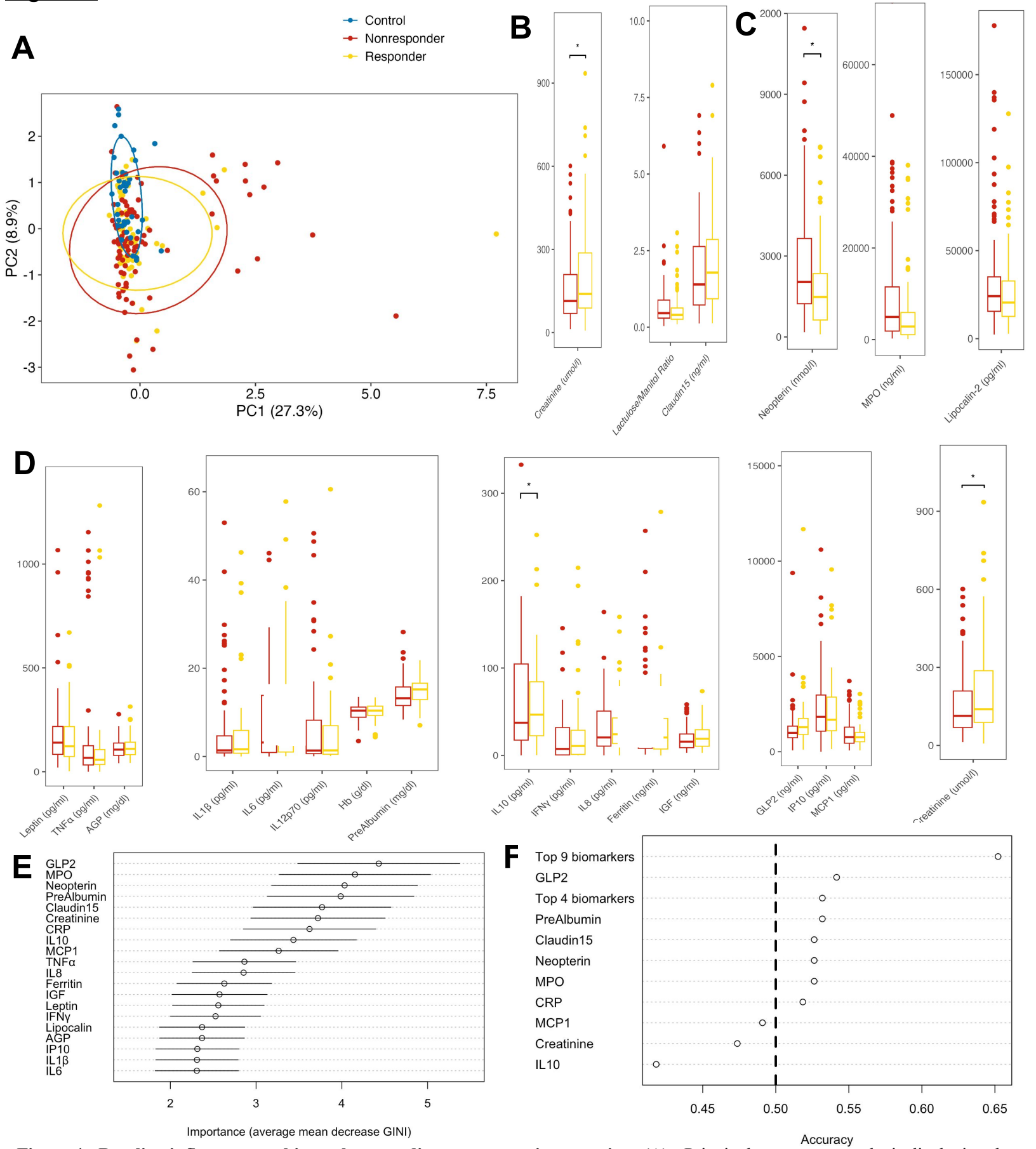

Figure 4: Baseline inflammatory biomarkers predict response to intervention: (A): Principal component analysis displaying the relationship between biomarkers and cytokines in the responders (n=66), nonresponders (n=82) and controls (n=38) at 9 months of age. **(B)** Urine, **(C)** fecal, and **(D)** serum biomarker levels from stunted children at 9 months of age prior to NI in SEEM cohort. Asterisk indicating $p < 0.05$. **(E)** Variable importance as shown by mean average decrease in GINI for each biomarker in Random Forest Model trained to predict whether a child would respond to the nutritional intervention. Circles indicate average importance, bars indicate standard deviation of 20-fold cross validation. **(F)** Accuracy of logistic regression models trained to predict response from individual biomarkers or combinations of biomarkers deemed important from Random Forest analysis.

Figure 5

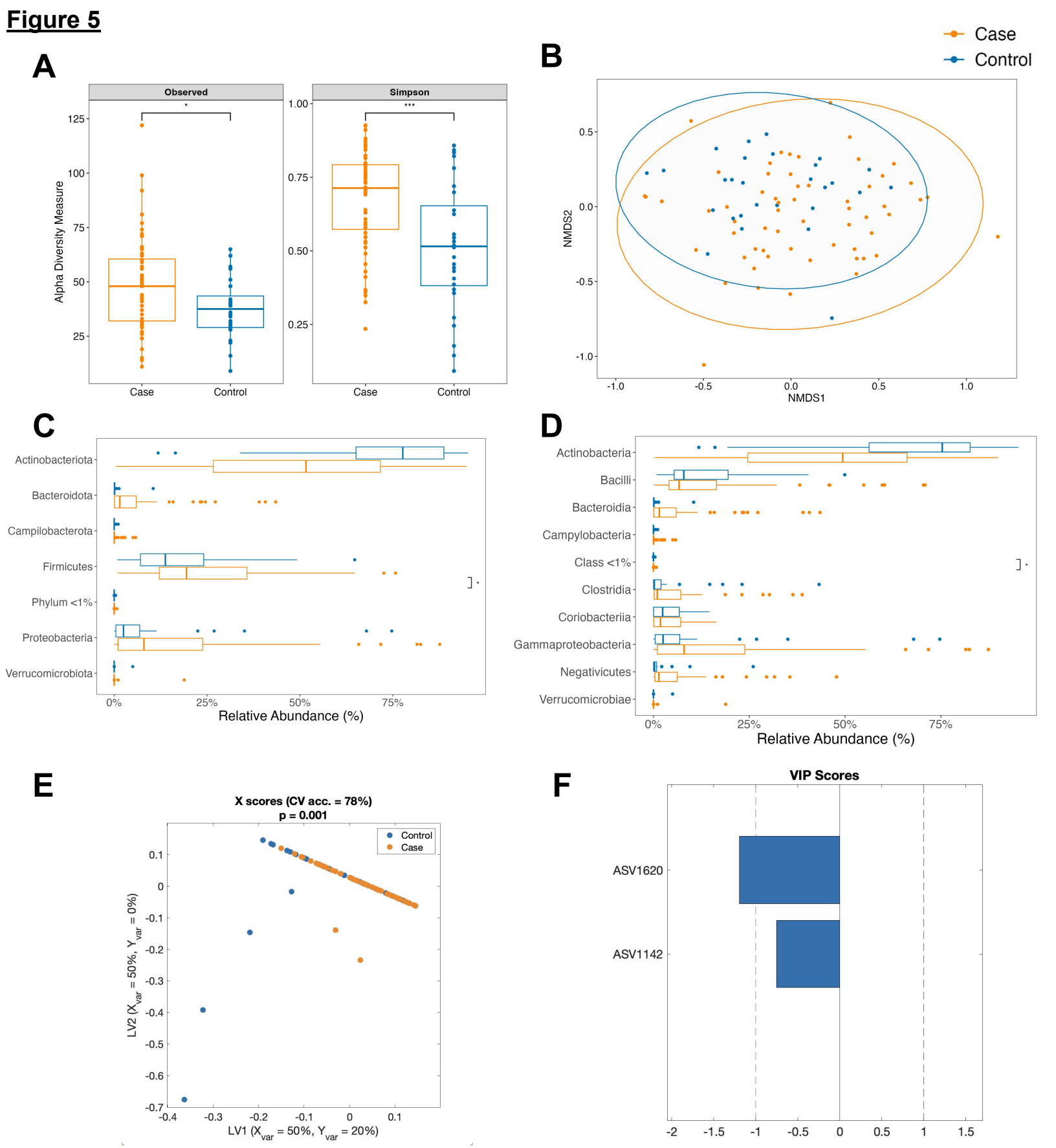

Figure 5: Fecal microbiomes of stunted children are more diverse than healthy controls : (A) Alpha diversity measurements (Number of observed species, Simpson index) of stunted cases and healthy controls at 9 months of age prior to NI. **(B):** Non-metric multidimensional scaling plot on Bray-Curtis matrices comparing the community composition of the cases and controls prior to the intervention. Relative abundance of **(C)** Phylum and **(D)** Class level taxonomy between cases and controls prior to the nutritional intervention. An OPLSDA model was constructed to discriminate between cases and controls using the relative abundance of ASVS from their fecal microbiome. The model outperformed all of 1000 randomly permuted models ($p < 0.001$). **(E)** Scatter plot of the X scores on latent variables 1 and 2 (LV1 & LV2), where each point represents one sample. **(F)** Bar plot shows the Variable Importance in Projection (VIP) scores, artificially oriented in the direction of loadings on LV1 and colored according to their association with case or control samples. VIP scores > 1 indicates a variable with greater than average influence on the projection.

Figure 6

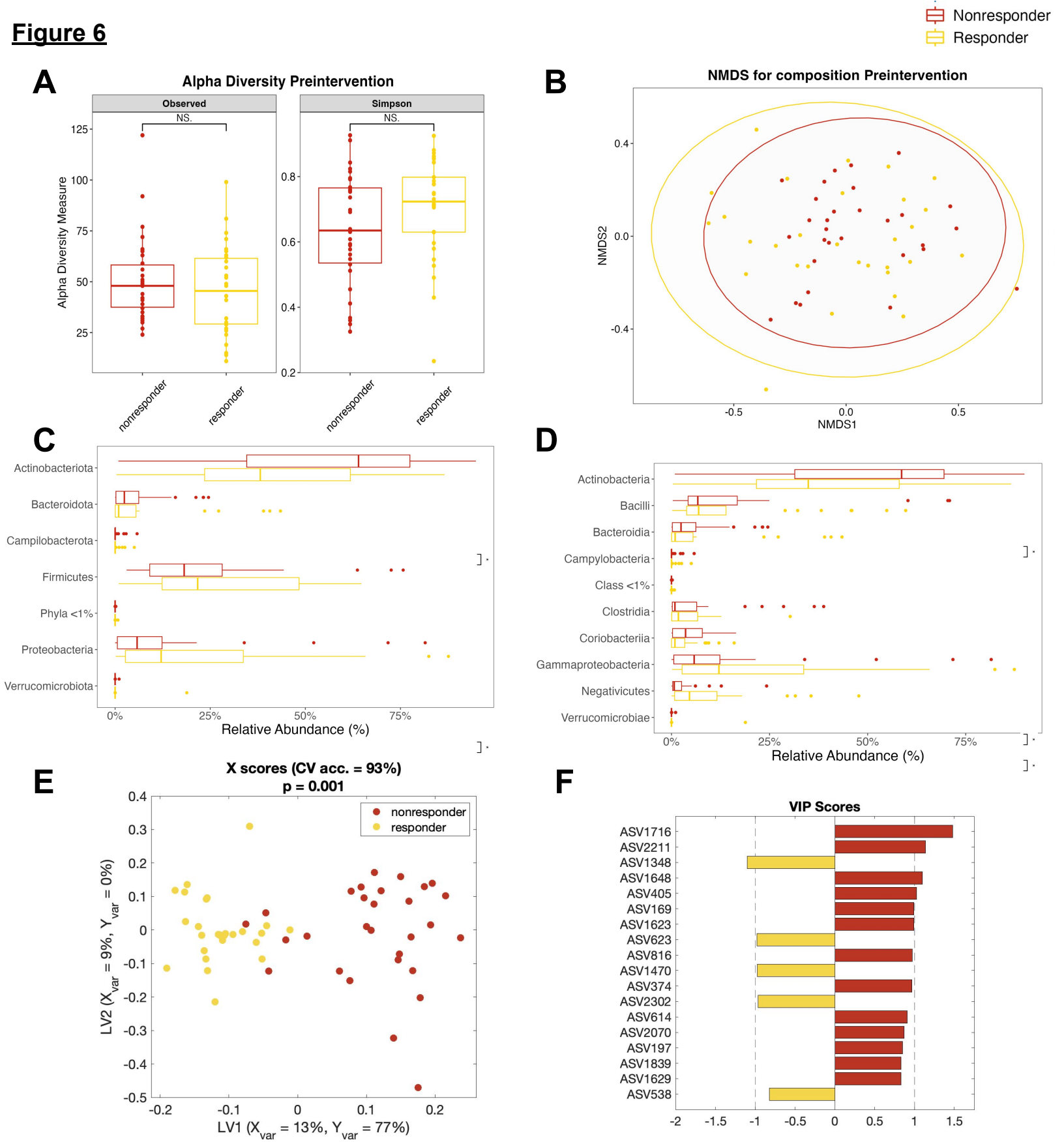

Figure 6: Responders and nonresponders differ in their fecal microbiome composition prior to the nutritional intervention. **(A)** Alpha diversity measurements (Number of observed species, Simpson index) of responders and nonresponders at 9 months of age prior to NI. **(B):** Non-metric multidimensional scaling plot on Bray-Curtis matrices comparing the community composition of the responders and nonresponders prior to the intervention. Relative abundance of **(C)** Phylum and **(D)** Class level taxonomy between cases and controls prior to the nutritional intervention. An OPLSDA model was constructed to discriminate between responders and nonresponders using the relative abundance of ASVS from their fecal microbiome. The model outperformed all of 1000 randomly permuted models ($p < 0.001$). **(E)** Scatter plot of the X scores on latent variables 1 and 2 (LV1 & LV2), where each point represents one sample. **(F)** Bar plot shows the Variable Importance in Projection (VIP) scores, artificially oriented in the direction of loadings on LV1 and colored according to their association with responder or nonresponder samples. VIP scores > 1 indicates a variable with greater than average influence on the projection.

Figure 7

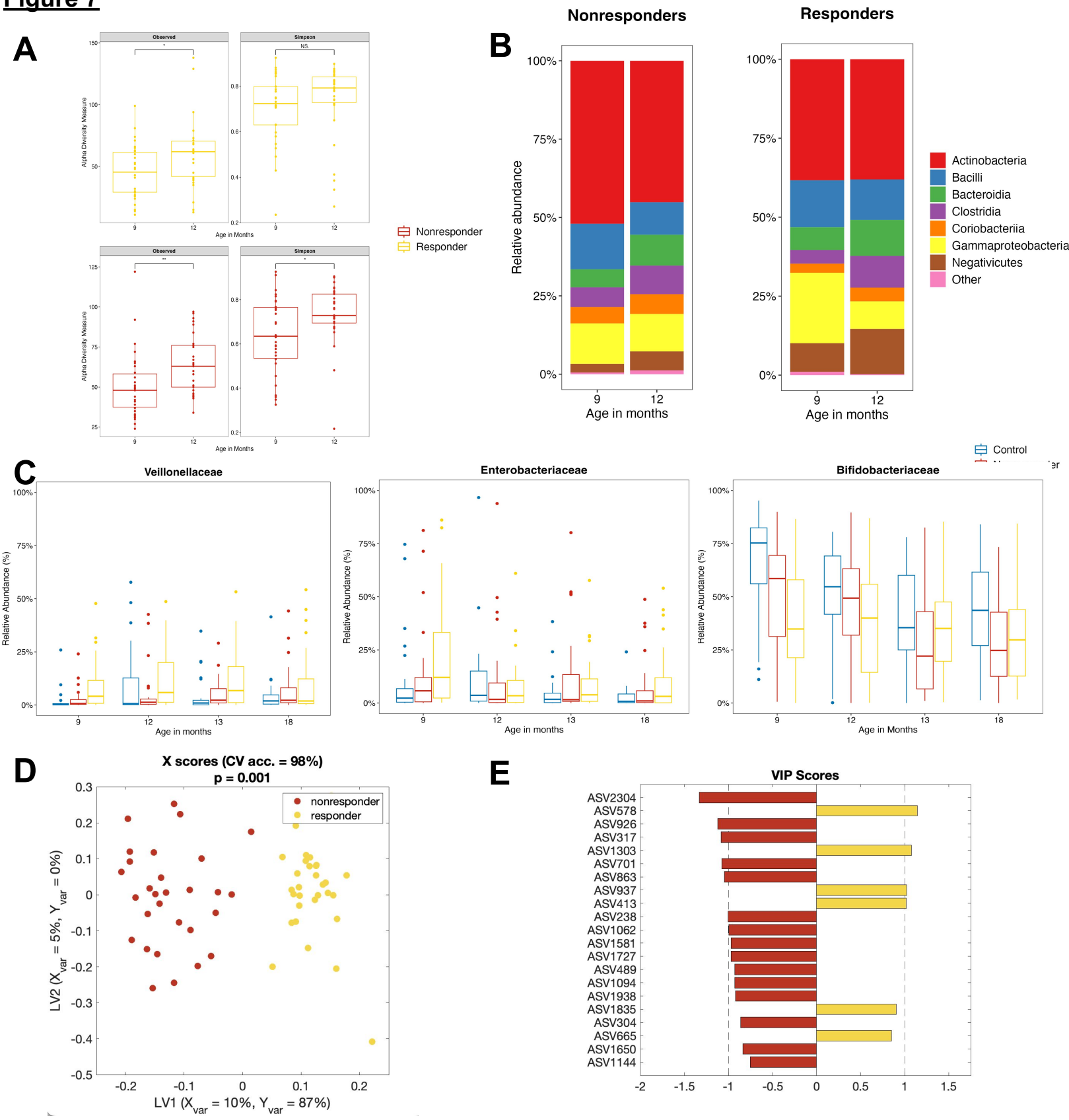

Figure 7: Responders exhibit a distinct shift in microbiome composition over the course of the nutritional intervention. (A) Alpha diversity measurements (Number of observed species, Simpson index) of responders and nonresponders pre and post intervention. **(B)** Comparison of relative abundance of microbial taxa at the phylum level in responder and nonresponder stool samples over the course of the study. **(C)** Comparison of relative abundance of select microbial taxa at the family level in responder, nonresponder, and control stool samples throughout the course of the study. An OPLSDA model was constructed to discriminate between responders and nonresponders using the relative abundance of ASVS from their fecal microbiome. The model outperformed all of 1000 randomly permuted models ($p < 0.001$). **(D)** Scatter plot of the X scores on latent variables 1 and 2 (LV1 & LV2), where each point represents one sample. **(E)** Bar plot shows the Variable Importance in Projection (VIP) scores, artificially oriented in the direction of loadings on LV1 and colored according to their association with responder or nonresponder samples. VIP scores > 1 indicates a variable with greater than average influence on the projection

Figure 8

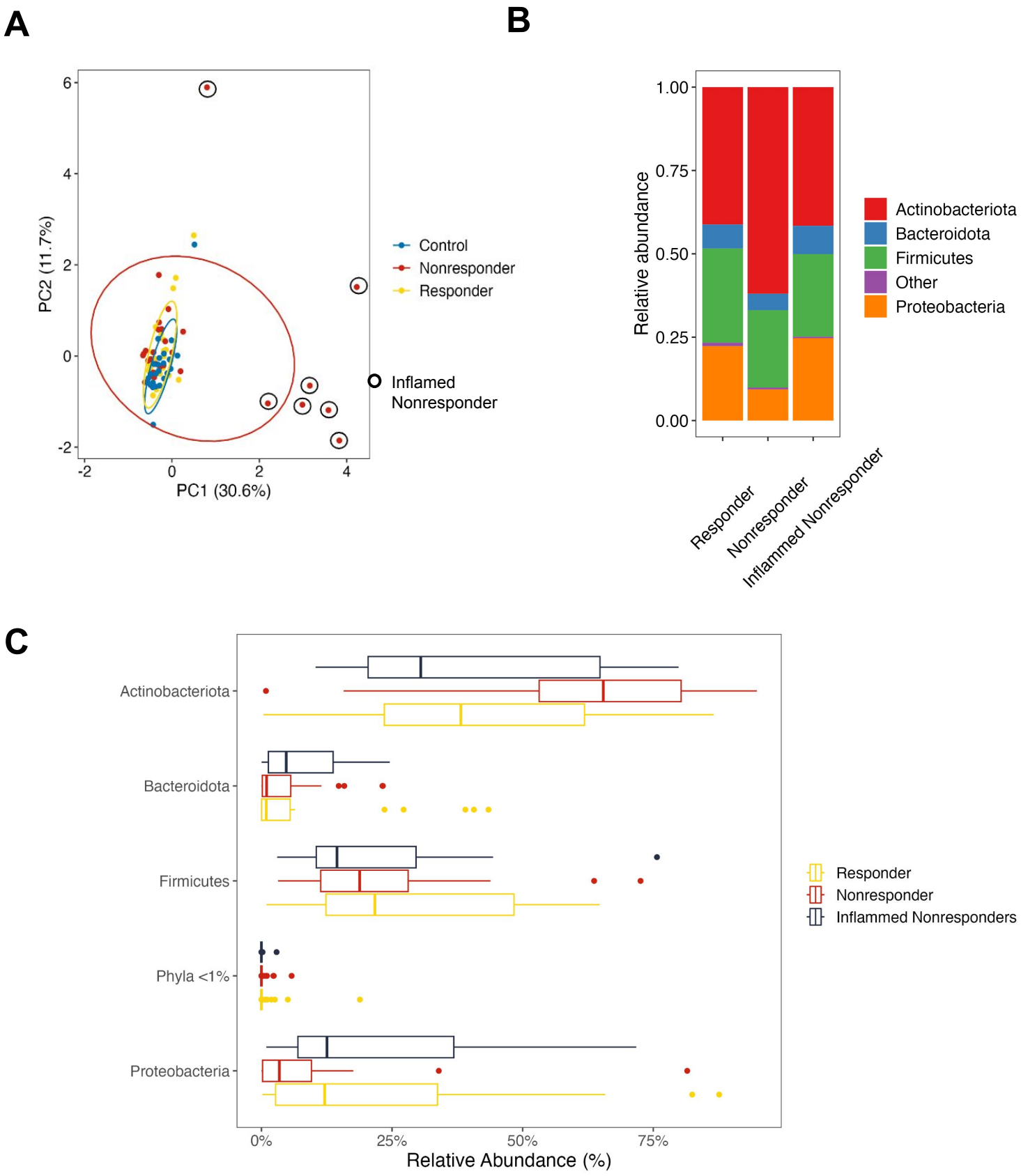

Figure 8: Nonresponders with high levels of inflammation have fecal microbiomes that resemble those of responders. (A) Principal component analysis displaying the relationship between biomarkers and cytokines measured pre-intervention from patients with matching microbiome data in the responders (n=30), nonresponders (n=30) and control groups (n=28). Black circles denote nonresponder samples with high levels of inflammation selected for downstream analysis. **(B)** Comparison of the average relative abundance of microbial taxa at the phylum level in responder, nonresponder, and inflamed nonresponder stool samples prior to the nutritional intervention. **(C)** Box plots showing distributions of fecal microbiome composition in responders, nonresponders, and inflamed nonresponders.

Supplemental Table 1: Baseline characteristics and comparison of SEEM and 16S cohorts

	All Wasted children receiving NI						Controls		
	Responders			Non-responder					
Variables	all responders in SEEM study units	only those with fecal 16S data (n=30)	p	all non-responders in SEEM study	only those with fecal 16S data (n=30)	p	all controls in SEEM study	only those with fecal 16S data (n=28)	p
Total Number	87	30		99	30		48	30	
Demographics									
Gender	59	19	0.660	58	20	0.426	27	18	0.496
Birthplace	66	22 vs 8	0.790	75	21 vs 9	0.550	40	24 vs 3	0.784
Gestational age	39 (39 39)	39 (38 39)	0.630	39 (39 39)	39 (39 39)	0.170	39 (39 39)	39 (39 39)	0.442
Birth anthropometrics									
WAZ	-2.2 -3.17 -1.33	-1.645 -2.34 -0.96	0.062	-2.21 -2.795 -1.655	-2.09 -2.69 -1.74	0.987	-1.03 -1.585 -0.39	-1.03 -1.385 -0.305	0.755
LAZ	-2.13 -2.94 -1.15	-1.51 -2.14 -0.93	**0.032**	-1.83 -2.63 -1.195	-1.765 -2.83 -1.07	0.967	-0.895 -1.56 -0.065	-0.88 -1.54 -0.16	0.667
WLZ	-1.26 -1.96 -0.48	-1.02 -1.81 -0.32	0.521	-1.59 -2.15 -0.89	-1.855 -2.12 -1.17	0.597	-0.48 -1.13 -0.095	-0.445 -1.155 -0.095	1.000
Maternal factors									
age (at the time of delivery)	29 (24, 35)	30 (24, 38)	0.443	26 (23, 33.5)	26 (23, 35)	0.659	26 (23, 30)	27 (25, 30)	0.653
weight (at the time /after	46 (41.75, 51.5)	47 (42.5, 54)	0.426	44.5 (40, 48.79)	45 40 48.5	0.628	49.55 (44.25, 53.85)	49.25 (44.5, 53.60)	0.693
height	152.4 (149.75, 155.35)	154.09 (151, 156.59)	0.058	151.801(47.95, 155)	152.19 (149.5, 156)	0.148	152.05 (148.35, 157)	151.95 (148.25, 155.55)	0.473
BMI	19.73 (18.58, 21.78)	19.98 (17.97, 22.04)	0.734	19.35 (17.53, 21.30)	18.69 (17.15, 21.03)	0.433	20.99 (19.59, 24.28)	20.92 (19.76, 24.46)	0.886
Paternal BMI	19.90 (18.71, 22.37)	21.22 (18.71, 22.88)	0.903	19.34 (18.03, 21.76)	18.91 (18.07, 20.05)	0.452	21.86 (19.85, 25.17)	22.45 (20.27, 24.32)	0.731
Breast feeding status		30			30			28	
exclusive	38	12		34	10		16	10	
partial	45	16		54	17		28	15	
none	4	2		11	3		3	2	
		0			0		1	1	
Pre-intervention anthropometrics									
WAZ 9mo	-3.650	-3.750	0.896	-3.280	-3.155	0.770	-0.005	-0.110	0.417
LAZ 9mo	-2.890	-2.965	0.963	-2.490	-2.505	0.680	-0.900	-0.880	0.797
WLZ 9mo	-2.620	-2.655	0.724	-2.370	-2.330	0.763	0.595	0.530	0.595
Pre-intervention status n(%)									
Wasted (≤-2)	84	29		94	28		2	0	
Underweight (≤-2)	61	19		63	20		8	5	
Stunted (≤-2)	81	30		82	23		0	0	
Pre-intervention markers									
Hemoglobin	10.4 (9.3, 11.4)	10.5 (9.8, 11.1)	0.954	10.4 (8.9, 11.2)	10.3 (8.8, 11.6)	0.915	10.7 (10.15, 11.75)	10.45 (9.5, 11.3)	0.461
Pre-albumin	15.2 (12.9, 16.6)	15.2 (11.7, 16.2)	0.503	13.2 (11.55, 15.75)	13.5 (11.55, 15.75)	0.661	15.3 (13.7, 17.7)	15.2 (13.6, 16.35)	0.173
Serum AGP mg/dl	110.5 (82, 142.49)	116 (93, 137)	0.580	106.5 (78.5, 138.44)	123 (85, 138.05)	0.656	94.77 (72, 126)	115 (85, 134)	0.166
Serum IGF-1 ng/ml	19.06 (10.48, 29.44)	20.13 (10.41, 35.39)	0.567	15.81 (8.87, 24.52)	15.81 (11.28, 26.34)	0.946	27.35 (19.26, 37.65)	24.46 (17.57, 36.48)	0.528
Serum CRP	0.136 (0.05, 0.39)	0.118 (0.05, 0.44)	0.628	0.24 (0.11, 0.64)	0.24 (0.14, 0.58)	0.941	0.095 (0.05, 0.23)	0.13 (0.07, 0.37)	0.888
Serum Ferritin	20.5 (7.2, 42.5)	21.6 (16, 50)	0.387	19 (8.05, 39.5)	22 (8, 55)	0.834	9.95 (5.5, 22)	16.95 (5.8, 34.8)	0.360
Serum Leptin	122.5 (73.22, 217.48)	129.76 (69.81, 188.56)	0.575	140.03 (83.68, 217.91)	180.81 (115.28, 264.97)	0.215	279.85 (190.71, 389.19)	258.17 (179.39, 345.51)	0.381
Serum GLP-2	1288.34 (908.43, 1747.31)	1169.88 (976.08, 1607.18)	0.583	993.43 (723.62, 1350.44)	1101.08 (699.99, 1396.07)	0.955	1705.89 (843.92, 2688.24)	1260.87 (649.51, 2286.14)	0.426
Fecal MPO ng/ml	2858.25 (1050, 6000)	3034.25 (1250, 7100)	0.594	4933 (1850, 11600)	4775 (2250, 10250)	0.197	4672.25 (2079.5, 10575)	4672.25 (2640.25, 9275)	0.752
Fecal NEO nmol/L	1482.75 (627.5, 2350)	1675 (850, 2150)	0.962	2037.5 (1228.5, 3651.25)	2375 (1525, 5225)	0.204	1633.5 (630, 2750)	1341.5 (485.95, 2600)	0.966
Fecal Lipocalin	20526.15 (12647.85, 32778.5)	25292.58 (16169.55, 33204.25)	0.739	24134.53 (15475.5, 35537)	28370.73 (17331.3, 87550)	**0.041**	27370 (17347.08, 67500)	40097.93 (20119.5, 94775)	0.293
Urine claudin-15	1.783 (0.93, 2.866)	2.01 (0.92, 2.89)	0.579	1.4 (0.72, 2.66)	2.11 (0.82, 2.87)	0.462	0.909 (0.73, 1.13)	0.92 (0.69, 1.15)	0.786
Urine creatinine	139.36 (88.38, 287.24)	174.32 (92.70, 469.29)	0.329	113.80 (68.95, 212.04)	121.52 (91.19, 220.25)	0.614	163.79 (107.98, 205.83)	176.57 (129.65, 212.41)	0.540

Supplemental Table 2: Baseline Characteristics of 16S substudy

Variables	Wasted children (16S)		P-values
	Responder (n=30)	nonresponder (n=30)	
	Median (IQR)	Median (IQR)	
Demographics			
Gender (# male)	19	20	0.791
Birthplace (hospital vs home)	22 vs 8	21 vs 9	0.779
Gestational age	39 (38 39)	39 (39 39)	0.683
Birth anthropometrics			
WAZ	-1.65 (-2.34, -0.96)	-2.09 (-2.69, -1.74)	0.980
LAZ	-1.51 (-2.14, -0.93)	-1.77 (-2.83, -1.07)	0.447
WLZ	-1.02 (-1.81, -0.32)	-1.86 (-2.12, -1.17)	**0.041**
Maternal factors			
age (at the time of delivery)	30 (24, 38)	26 (23, 35)	0.582
weight (at the time /after delivery)	47 (42.5, 54)	45 (40, 48.5)	0.448
height	154.09 (151, 156.59)	152.19 (149.5, 156)	0.171
BMI	19.98 (17.97, 22.04)	18.69 (17.15, 21.03)	0.139
Paternal BMI	21.22 (18.71, 22.88)	18.91 (18.07, 20.05)	0.143
Breast feeding status			
Exclusive (n)	12	10	
partial(n)	16	17	
none(n)	2	3	
N/A	0	0	
Pre-intervention anthropometrics			
WAZ 9mo	-3.75	-3.155	0.060
LAZ 9mo	-2.965	-2.505	0.190
WLZ 9mo	-2.655	-2.33	**0.047**
Pre-intervention status n(%)			
Underweight (≤-2)	29	28	
Stunted (≤-2)	19	20	
Wasted (≤-2)	30	23	
Pre-intervention markers			
Hemoglobin	10.5 (9.8, 11.1)	10.3 (8.8, 11.6)	0.986
Pre-albumin	15.2 (11.7, 16.2)	13.5 (11.55, 15.75)	0.103
Serum AGP mg/dl	116 (93, 137)	123 (85, 138.05)	0.895
Serum IGF-1 ng/ml	20.13 (10.41, 35.39)	15.81 (11.28, 26.34)	0.360
Serum CRP	0.118 (0.05, 0.44)	0.24 (0.14, 0.58)	0.166
Serum Ferritin	21.6 (16, 50)	22 (8, 55)	0.134
Serum Leptin	129.76 (69.81, 188.56)	180.81 (115.28, 264.97)	0.093
Serum GLP-2	1169.89 (976.08, 1607.18)	1101.08 (699.99, 1396.07)	0.175
Fecal MPO ng/ml	3034.25 (1250, 7100)	4775 (2250, 10250)	0.551
Fecal NEO nmol/L	1675 (850, 2150)	2375 (1525, 5225)	**0.004**
Fecal Lipocalin	25292.56 (16169.55, 33204.25)	28370.73 (17331.3, 87550)	**0.008**
Urine claudin-15	2.01 (0.92, 2.88)	2.11 (0.82, 2.87)	0.655
Urine creatinine	174.32 (92.70, 469.3)	121.52 (91.19, 220.25)	0.168

Supplemental Table 3: Epidemiology of Inflamed Nonresponders from figure 8

Non-responder outliers	Prevalence of diarrheal days per year	Prevalence of ARI days per year	Diarrhea episodes per year	ARI episodes per year
01-027-0171	16	6	6.62	1
01-027-0823	46	19	11.03	3
01-029-0976	37	37	7.95	3
02-055-0529	22	9	7.29	1
02-073-0644	73	4	12.10	1
03-046-0691	113	98	19.93	4
03-050-0575	73	114	13.81	4

Fig. S1

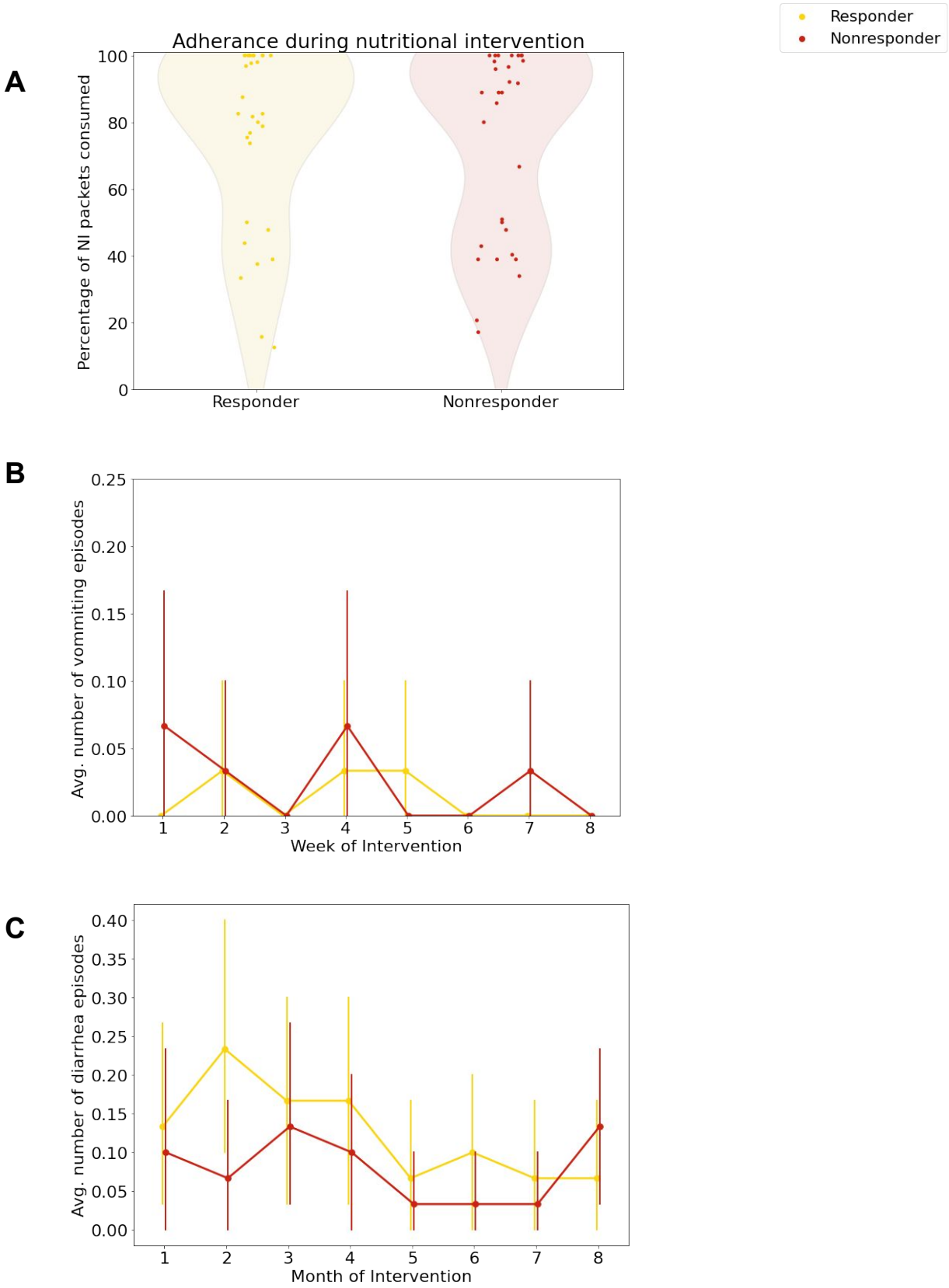

Fig S1: Responders and Nonresponders show similar patterns of compliance during nutritional intervention: (A) Adherence to nutritional intervention by group as shown by percentage of Acha-mum packets consumed out of total packets distributed for each child during the intervention. **(B)** Mean number of vommitting episodes per week of the intervention in each group, error bars show standard deviation. **(C)** Mean number of diahreal episodes per week of the intervention in each group, error bars show standard deviation.

Figure S2

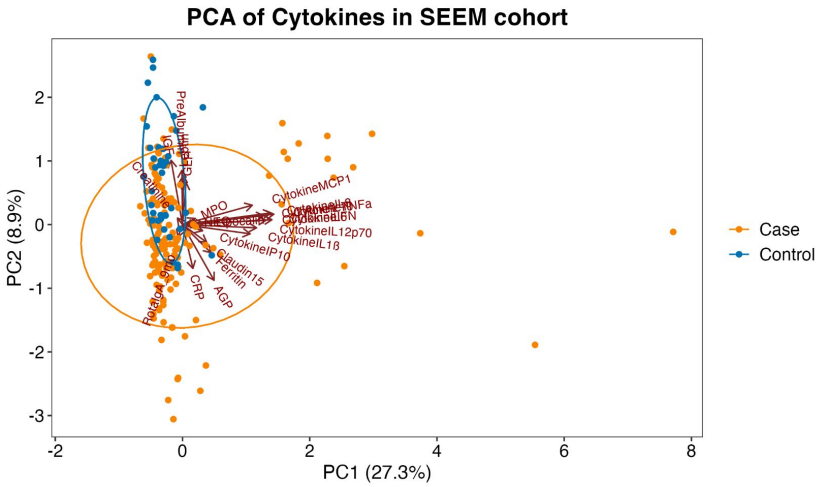

Fig S2: PCA of cytokines in SEEM cohort measured at 9 months colored by nutritional status. Arrows show loadings for each variable.

Figure S3

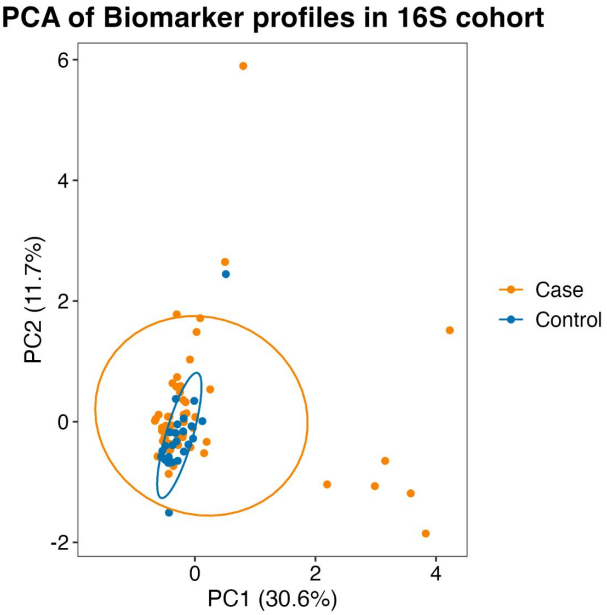

Fig S3:PCA of cytokines in patients for which 16s sequencing was performed colored by nutritional status.

Figure S4

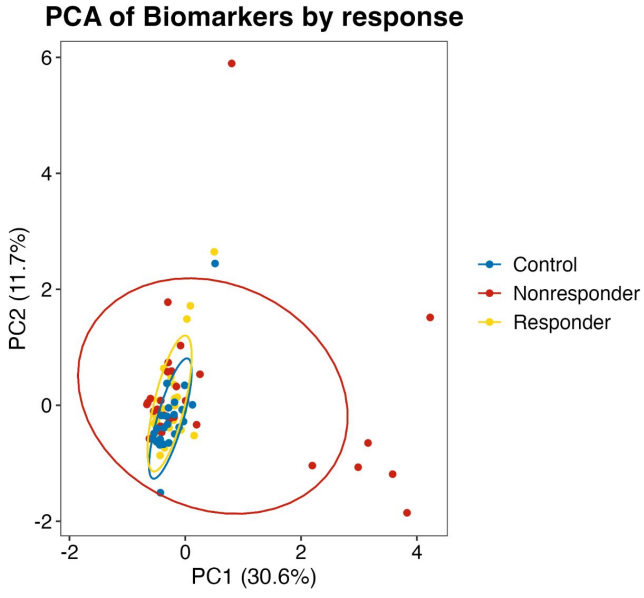

Fig S4: PCA of 16S cohort biomarkers by response

Fig. S5

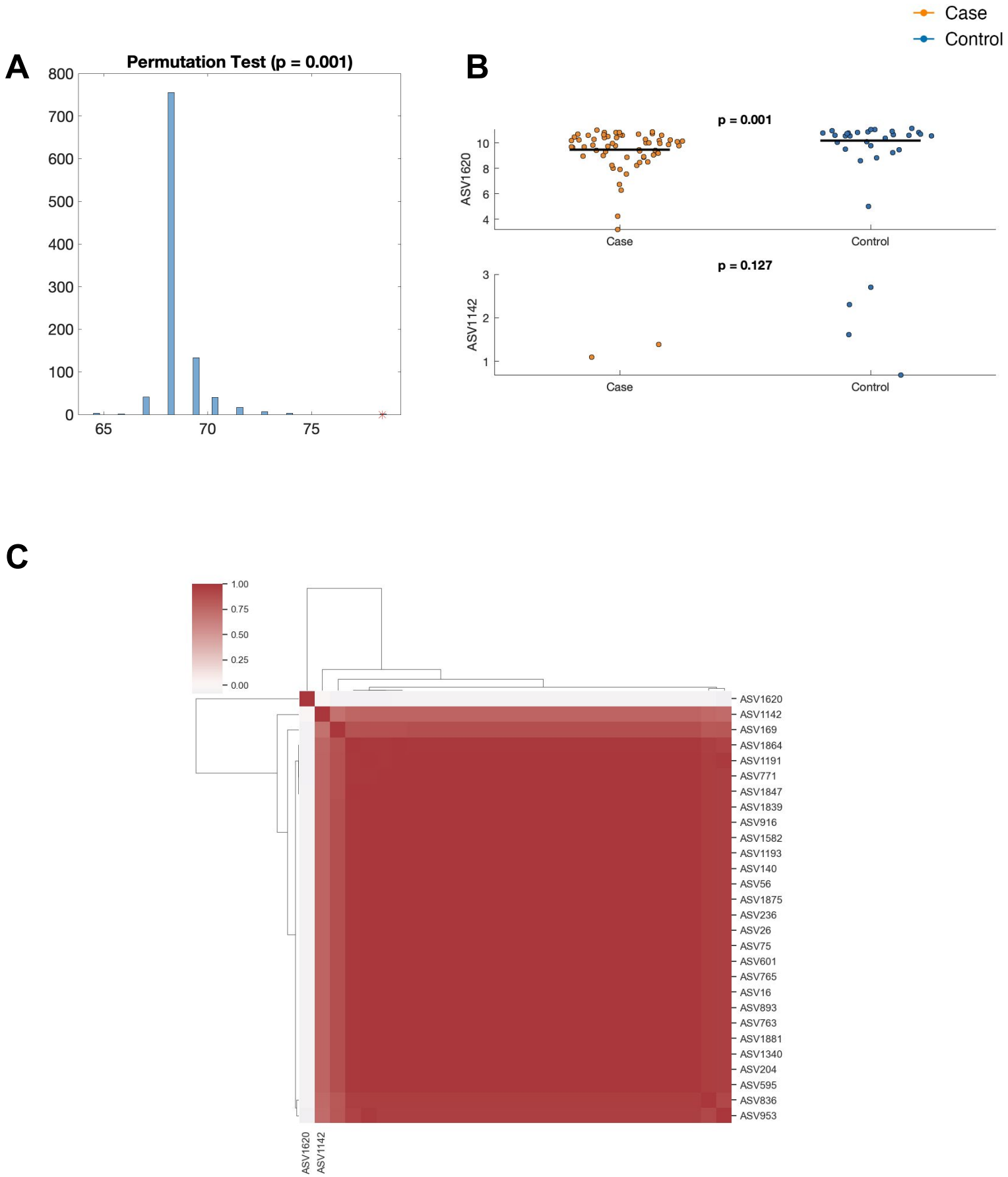

Fig S5: Additional OPLS-DA figures for Fig 5 (A) Results of permutation testing for OPLSDA model in figure 5. Plot shows a histogram of cross validation accuracy for 1000 randomly permuted models. Correctly labeled model is shown as a red star. **(B)** Jitter plots showing ASV counts per sample for VIPs identified in OPLSDA analysis. **(C)** Heatmap showing correlation between important ASVs identified in OPLS-DA and other ASVs with >70% correlation that were removed as being linearly correlated. X axis has only important features labeled, while y-axis has correlated and important features labeled.

Fig S6

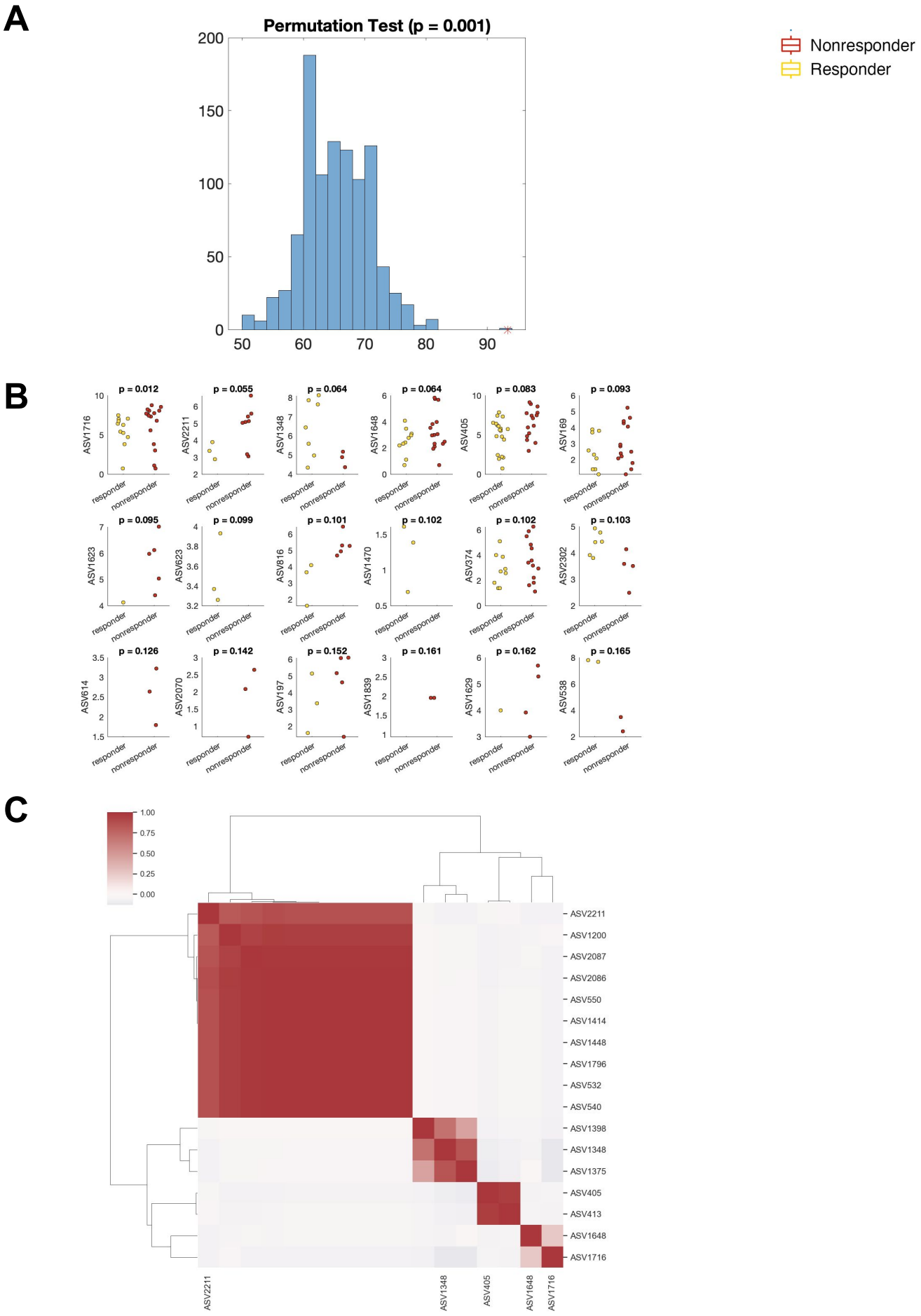

Fig S6: Additional OPLS-DA figures for Fig. 6 (A) Results of permutation testing for OPLSDA model in figure 5. Plot shows a histogram of cross validation accuracy for 1000 randomly permuted models. Correctly labeled model is shown as a red star. **(B)** Jitter plots showing ASV counts per sample for VIPs identified in OPLSDA analysis. **(C)** Heatmap showing correlation between important ASVs identified in OPLS-DA and other ASVs with >70% correlation that were removed as being linearly correlated. X axis has only important features labeled, while y-axis has correlated and important features labeled.

Fig S7

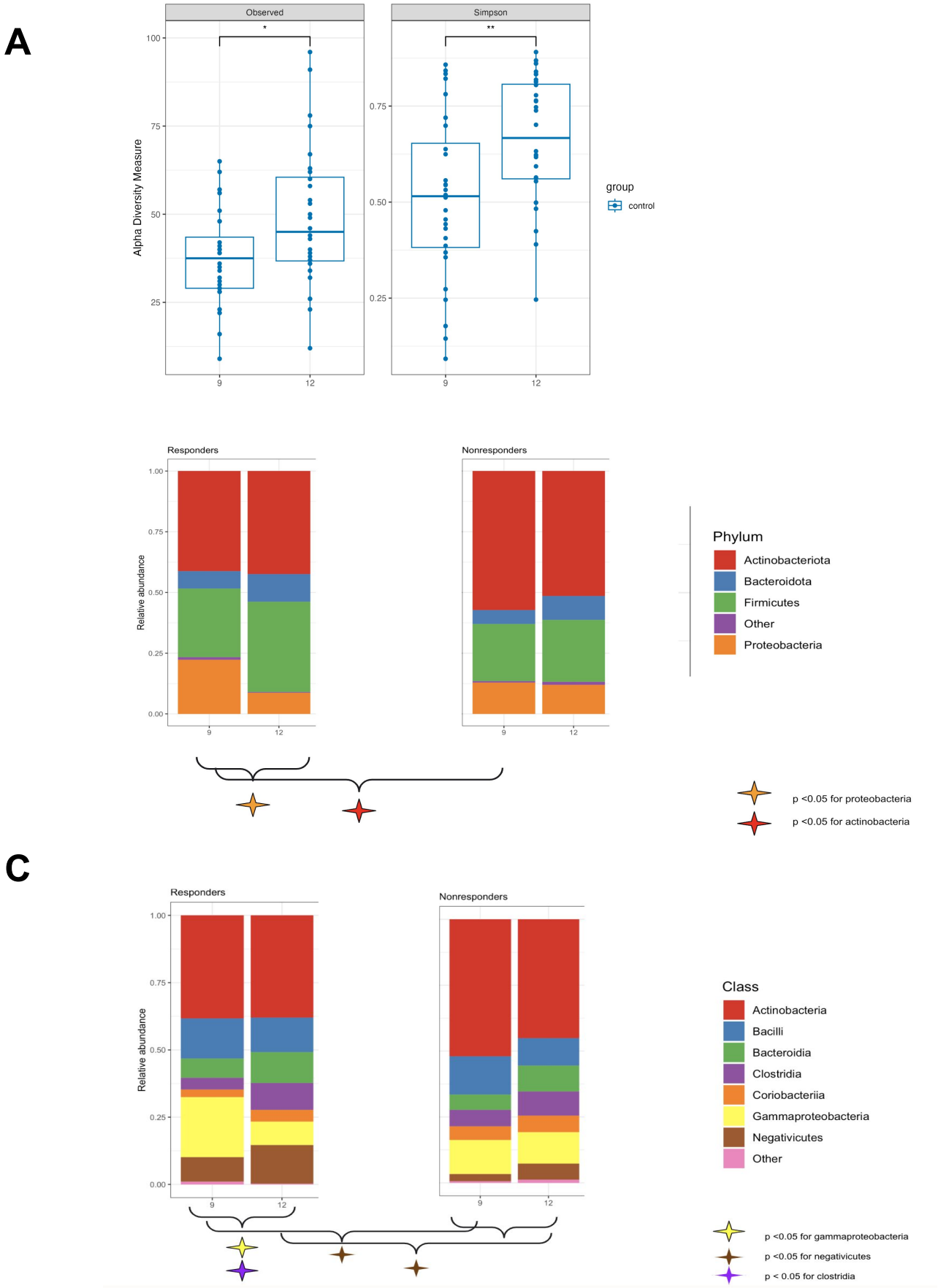

Fig S7: Microbiome shifts during intervention for Fig. 7 (A) Alpha diversity of control samples at 9 months and 12 months of life. **(B)** Bar plot showing average relative abundance at Phylum level in responders and nonresponders pre and post intervention. **(C)** Bar plot showing average relative abundance at Class level in responders and nonresponders at pre and post intervention.

Fig S8

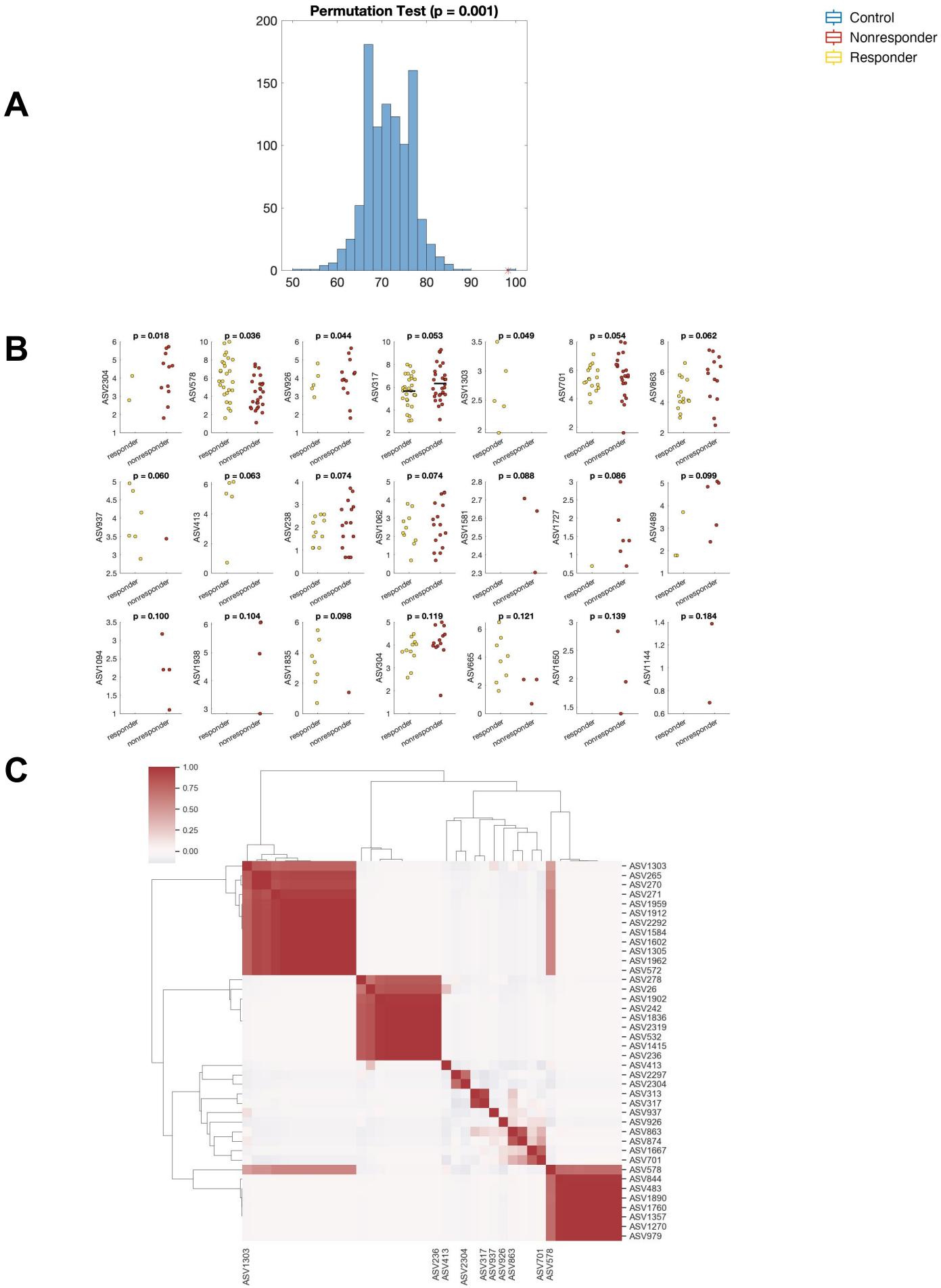

Fig S8: Additional OPLS-DA figures for Fig. 7 (A) Results of permutation testing for OPLSDA model in figure 5. Plot shows a histogram of cross validation accuracy for 1000 randomly permuted models. Correctly labeled model is shown as a red star. **(B)** Jitter plots showing ASV counts per sample for VIPs identified in OPLSDA analysis. **(C)** Correlation network of ASVs with >70% correlation of ASV identified in OPLS-DA. X axis has only important features labeled, while y-axis has correlated and important features labeled.