Alterations in Intestinal Microbiota Correlate with Susceptibility to Type 1 Diabetes

Running title: Gut microbiota in subjects with islet autoimmunity

Aimon K. Alkanani[1], Naoko Hara[1], Peter A. Gottlieb[1], Diana Ir[2], Charles E. Robertson[2,3], Brandie D. Wagner[3,4], Daniel N. Frank[2,3], and Danny Zipris[1*]

[1]Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, Aurora, CO 80045

[2]Division of Infectious Diseases, University of Colorado Denver, Aurora, CO 80045

[3]University of Colorado Microbiome Research Consortium (MiRC), Aurora, CO 80045

[4]Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Denver, Aurora, CO 80045

*Address correspondence and reprint request: Dr. Danny Zipris, Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, 1775 Aurora Ct., Mail Stop B-140, Aurora, CO 80045; Tel. 303-724-0108; fax: 303-724-6839; Email address: danny.zipris@ucdenver.edu

Diabetes Publish Ahead of Print, published online June 11, 2015

ABSTRACT

We tested the hypothesis that alterations in the intestinal microbiota are linked with the progression of type 1 diabetes (T1D). Herein, we present results from a study performed in subjects with islet autoimmunity living in the United States. High-throughput sequencing of bacterial 16S rRNA genes and adjustment for gender, age, autoantibody presence, and HLA indicated that the gut microbiomes of seropositive subjects differed from those of autoantibody-free first-degree relatives (FDRs) in the abundance of four taxa. Furthermore, subjects with autoantibodies, seronegative FDRs, and new onset patients had different levels of the Firmicutes genera *Lactobacillus* and *Staphylococcus* compared with healthy controls with no family history of autoimmunity. Further analysis revealed trends towards increased and reduced abundances of the Bacteroidetes genera *Bacteroides* and *Prevotella*, respectively, in seropositive subjects with multiple versus one autoantibody. Canonical discriminant analysis suggested that the gut microbiomes of autoantibody positive individuals and seronegative FDRs clustered together but separate from those of new onset patients and unrelated healthy controls. Finally, no differences in biodiversity were evident in seropositive versus seronegative FDRs. These observations suggest that altered intestinal microbiota may be associated with disease susceptibility.

INTRODUCTION

Type 1 diabetes (T1D) is an autoimmune disorder that involves beta cell inflammation and destruction (1). Although the mechanisms that trigger the disease are not yet clear, human and animal studies implicate both genetic and environmental factors in the disease process. The intestinal microbiota plays a key role in the development and function of the immune system (2). Data from human and animal studies have led to the hypothesis that altered gut microbiota ("dysbiosis") could be associated with mechanisms of metabolic and immune mediated disorders, such as obesity, celiac disease, type 2 diabetes, and inflammatory bowel disease (IBD) (3,4).

Dysbiosis has been postulated to be associated with mechanisms of T1D (5-11). The development of T1D in animal models, such as the NOD (6) and RIP-B7.1 (12) mice and the diabetes-prone BioBreeding (13) and LEW1.WR1 (10) rats is linked with changes in the intestinal microbiome. Human studies performed in subjects at risk for T1D provided evidence for a decline and an increase in the abundance of the bacterial phyla Firmicutes and Bacteroidetes, respectively (14). Others reported the absence of *Bifidobacterium* species and increased levels of the genus *Bacteroides* in children with 2 or more islet autoantibodies (15). A study performed in subjects with T1D further demonstrated an increase in the abundances of Bacteroidetes and *Clostridium spp.* in addition to a reduction in the genera *Lactobacillus* and *Bifidobacterium* (16). Finally, significant alterations in microbial interaction networks with no differences in bacterial diversity, microbial composition, or the level of bacterial genera were observed in seropositive and seronegative children (17).

In this study, we tested the hypothesis that alterations in the human intestinal microbiota are linked with the progression of human T1D. Analysis of 16S bacterial rRNA sequencing data demonstrated that the abundances of four bacterial genera were altered in seropositive subjects compared to

seronegative FDRs. Furthermore, seropositive subjects and seronegative FDRs as well as new onset patients have a reduction in the abundance of the Firmicutes genera *Lactobacillus* or *Staphylococcus* compared with unrelated healthy controls. The data suggest that gastrointestinal tract dysbiosis may be associated with disease progression.

RESEARCH DESIGN AND METHODS

Subject characteristics

We analyzed the gut microbiome from four subject cohorts with or without evidence of islet autoimmunity residing in the Denver metro area. These studies included 35 subjects with newly diagnosed T1D (up to 6 months from disease diagnosis) recruited during their routine visits to the Barbara Davis Center for Diabetes. Twenty one individuals with 1-4 autoantibodies and 32 seronegative first-degree relatives (FDRs) of subjects with islet autoimmunity were recruited from the Type 1 Diabetes TrialNet Natural History Study (Table 1). Only two seronegative subjects had siblings in the seropositive group. Twenty three individuals without any family history of autoimmunity (unrelated healthy controls) were recruited from university employees and children of employees. The presence of antibodies against GAD65, insulin, ICA512, and ZnT8 was determined by the Autoantibody Core at the Barbara Davis Center for Diabetes. Typing for HLA DR3 and DR4 of study subjects was performed by PCR analysis (18). Thirty-two of the 35 new onset patients (91%) had DR3 and/or DR4 HLA diabetes risk alleles. All autoantibody seropositive subjects (21/21) and 29/32 seronegative FDRs carried DR3 and/or DR4 (91%). Twelve out of the 23 unrelated healthy subjects had DR3 and/or DR4 (52%). The median hemoglobin A1C level in the seropositive and new onset subject cohorts was 5.0% (range, 4.2-5.3) and 7.9% (range, 5.5-15). Subjects who received antibiotic therapy up to 4 weeks prior to sample collection, individuals with known infections, or gastrointestinal disorders were not included in the study. One subject in each of the seropositive and new onset groups and three from the seronegative cohort reported being vegetarians. The study was approved by the institutional review board at the University of Colorado Denver. Fecal samples were collected at the Barbara Davis Center or at the home of the study participants using stool collection tubes and stored in -80^{o}C until use. Samples collected at home were stored at -20^{o}C and delivered to the Center on ice by study subjects or lab personnel.

Microbiome analysis

Bacterial profiles were determined by broad-range amplification and sequence analysis of 16S rRNA genes. In brief, amplicons were generated using primers that target approximately 280 b.p. of the V4 variable region of the 16S rRNA gene (19). PCR products were normalized using a SequalPrep^{TM} kit (Invitrogen, Carlsbad, CA), pooled, lyophilized, purified and concentrated using a DNA Clean and Concentrator Kit (Zymo, Irvine, CA). The amplicon pool was quantified using Qubit Fluorometer 2.0 (Invitrogen, Carlsbad, CA), diluted to 4nM and denatured with 0.2 N NaOH at room temperature. The denatured DNA was diluted to 15pM and spiked with 25% of the Illumina PhiX control DNA prior to loading the sequencer. Illumina paired-end sequencing was performed on the Miseq platform with version v2.3.0.8 of the Miseq Control Software and version v2.3.32 of MiSeq Reporter, using a 600-cycle version 3 reagent kit.

As previously described (9), Illumina Miseq paired-end sequences were sorted by sample via barcodes in the paired reads with a python script. The sorted paired reads were assembled using phrap (20,21). Pairs that did not assemble were discarded. Assembled sequence ends were trimmed over a moving window of 5 nucleotides until average quality met or exceeded 20. Trimmed sequences with more than 1 ambiguity or shorter than 200 nucleotides were discarded. Potential chimeras identified with Uchime (usearch6.0.203_i86linux32) (22) using the Schloss (23) Silva reference sequences were removed from subsequent analyses. Assembled sequences were aligned and classified with SINA (1.2.11) using the 629,124 bacterial sequences in Silva 111 (24) as reference configured to yield the Silva taxonomy. Operational taxonomic units (OTUs) were produced by clustering sequences with identical taxonomic assignments. A median of 131,000 sequences (range: 8,100 to 313,000) were generated per sample, with a median Good's coverage score of 99.97%. The software package

Explicet (v2.9.4, www.explicet.org) (25) was used for display, analysis (rarefied values for median Good's coverage), and figure generation of results.

Statistical analysis

Patient demographics were described using percentages and medians and ranges across the 4 groups for the categorical and continuous variables, respectively. Shannon diversity was calculated using rarefaction in Explicet and compared across groups using an Analysis of Variance. For ease of interpretation, effective number of species (EFS) was calculated from Shannon's diversity (26,26). These values represent the expected number of taxa in an even community. Wilcoxon rank based tests were used to compare the relative abundance of each taxa across groups after adjusting for clinical covariates (age, gender, autoantibody presence, and HLA genotype). The relative abundance for each taxa was transformed to ranks and these transformed variables were used as dependent variables in a linear regression model. If the overall *p* value of the differences between the groups was significant, then the pairwise differences were calculated using linear contrasts and a false discovery rate (FDR) adjustment for multiple comparisons was used. An unadjusted *p* value for the overall test across groups and the subgroup analysis used a 0.05 cutoff and the adjusted *p* value used for the individual comparisons used a 0.1 cutoff to identify differentially expressed bacterial groups. Patterns in OTU prevalence and/or abundance in seropositive subjects with 1 (n=5) versus 2-4 autoantibodies (n=16) were analyzed using Wilcoxon rank based tests without the FDR correction. Canonical discriminant analysis was performed to multivariately assess differences in bacterial community across the four groups. A small constant (1/total) was added to the counts to eliminate zeros prior to the application of the centered log ratio transformation recommended for multivariate analysis of compositional data (27,28).

RESULTS

Taxa in the stool from subjects with and without islet autoimmunity

We postulated that alterations in the intestinal microbiome would be involved in the progression of human T1D. To test this possibility, we sequenced the V4 region of 16S rRNA genes (primers 534F-805R (9,10)) from fecal samples of seropositive individuals with 1-4 autoantibodies (n=21) compared to seronegative FDRs (n=32). Comparisons were also made between the gut bacterial content of autoantibody positive individuals and seronegative FDRs to that of new onset patients (n=35) and autoantibody free healthy controls with no family history of autoimmunity (n=23). A median of 131,000 (range: 8,100 to 313,000) high-quality 16S rRNA gene sequences were obtained for each sample. The data presented in Fig. 1A demonstrate that similar overall gut bacterial profiles were observed among the subject cohorts. Furthermore, comparable Shannon bacterial diversity indices were measured in seropositive versus seronegative FDRs (Fig. 1B, $p = 0.44$). These observations suggest that islet autoimmunity is probably not associated with altered bacterial diversity or striking differences in the intestinal microbiota.

Gut bacterial communities in individuals with islet autoimmunity

We next compared the abundance of individual bacterial taxa from autoantibody positive versus autoantibody-free subjects. Similar numbers of bacterial phyla and genera (i.e., OTU richness) were observed in the subject cohorts (data not shown). Wilcoxon rank-based tests with adjustment for covariates, i.e. age, gender, autoantibody presence, and HLA and multiple comparisons across groups revealed significant differences in the abundance of nine bacterial taxa of which four were found to be different in seropositive versus seronegative cohorts (Figure 2, Table 2 and Supplemental Table 1). Parameter estimates further demonstrated that the level of the majority of the taxa with significant differences across groups was not associated with age, gender, or HLA3/4 (Supplemental Table 2).

Only the level of a group of bacteria that belong to the Bacteroidetes phylum and could not be classified to lower levels (termed by us Bacteroidetes other) was observed to be associated with the age (p = 0.02). The limited sample size of this study did not provide sufficient power to assess the effect of the autoantibody number on the gut bacterial composition. The covariate adjusted analysis indicated that a significant increase was detectable in the relative median abundance of the Firmicutes genera *Catenibacterium* and of the Bacteroidetes genera *Prevotellaceae* and *RC9-gut-group* in seropositive compared with seronegative FDRs (Figure 2, Table 2 and Supplemental Table 1, adjusted p = 0.02 for all). In contrast, a reduction in the abundance of the taxa classified only to Bacteroidetes (Bacteroidetes other) was evident in seropositive versus seronegative FDRs (adjusted p = 0.09).

We further analyzed the gut microbiomes of seropositive and seronegative subjects to those of new onset patients and healthy controls (Fig. 2, Table 2 and Supplemental Table 1). Differences in the bacterial level between these cohorts were noted in genera that belong to the phylum Firmicutes. A reduction and an increase were observed in the relative abundance of *Succiniclasticum* in new onset subjects compared with seropositive and seronegative subjects, respectively (adjusted p = 0.01 and p = 0.03, respectively). A decrease of *Catenibacterium* was observed in new onset subjects compared with autoantibody positive subjects (adjusted p = 0.04). An increase in the abundance of the Bacteroidetes genera *Alistipes* was detected in seronegative FDRs versus new onset patients (adjusted p = 0.02). A decrease in the abundance of the Bacteroidetes genera *Prevotellaceae* and the Firmicutes genera *Lactobacillus* and *Succiniclasticum* was seen in seronegative FDRs versus new onset patients (adjusted p = 0.07 for *Prevotellaceae* and *Lactobacillus* and p = 0.01 for *Succiniclasticum*). Spearman's rank correlation coefficients indicated that the relative abundance of all taxa found to be significantly different in new onset subjects versus seropositive or seronegative FDRs was not associated with the HbA1c level, age of disease onset, or disease duration (data not shown). The level

of the Firmicutes genera *Lactobacillus* and *Staphylococcus* were elevated in unrelated healthy control subjects compared to seropositive individuals (adjusted $p = 0.02$ and 0.03, respectively), whereas the level of the Bacteroidetes genera *Prevotellaceae* and *RC9-gut-group* were elevated in seropositive individuals compared to unrelated healthy control subjects (adjusted $p = 0.03$ for both). The abundances of *Lactobacillus* and *Staphylococcus* were also higher in the unrelated healthy cohort compared with the new onset and seropositive cohorts (adjusted $p = 0.07$ and 0.04, respectively).

Because of the wide age range of subjects in the seropositive and seronegative FDR groups (Table 1), we assessed whether the differences observed in the gut bacterial composition may be related to age. To this end, we re-analyzed the intestinal microbiome data after excluding those subjects older than 18 years of age (n=16 for unrelated healthy controls, 34 for new onset, 17 for seropositive and 21 for seronegative). We found that the abundances of 6 out of 9 bacterial taxa found to be different between cohorts (Fig. 2 and Table 2 and Supplemental Table 1) were also different when comparing cohorts of individuals younger than 18 years (data not shown). Similar to cohorts of subjects from all age groups (Fig. 2, Table 2 and Supplemental Table 1), the median abundances of *Staphylococcus* and *Lactobacillus* were diminished in genetically susceptible individuals under the age of 18 years with or without autoimmunity compared with unrelated healthy controls, but these differences did not reach statistical significance. Furthermore, unlike observed in groups comprising both children and adults, the abundance of the bacterial genus *Alistipes* was similar among individuals under the age of 18 years. Finally, the abundance of the Bacteroidetes genus *Barnesiella* was found to be significantly different among children and adolescents (data not shown). A similar analysis in subjects older than 18 years was not performed due to a limited sample number in each cohort.

Collectively, these findings may suggest that alterations in the abundances of specific bacterial groups can be observed in the intestine of genetically-susceptible individuals prior to and following disease onset and some of these changes may be age related.

The gut microbiome in subjects at low- versus high-risk for disease development

Next, we sought to identify trends in the gut microbiota of subjects at high-risk versus low-risk for T1D. To do so, we analyzed the intestinal microbiome of subjects with 2-4 autoantibodies (n=16) versus 1 autoantibody (n=5) using less stringent criteria, without adjustment for multiple comparisons, for this preliminary sub-analysis. The data shown in Table 3 demonstrate a 6- and 12-fold increase in the abundances of the genera *Bacteroides* and *Akkermansia*, respectively, in subjects with 2-4 versus 1 autoantibody (unadjusted $p = 0.01$ and $p = 0.04$, respectively). In contrast, a considerable 260-fold reduction was observed in the abundance of *Prevotella* in subjects with multiple versus 1 autoantibody (unadjusted $p = 0.01$). Finally, reduced abundances, albeit to a lesser degree than that seen for *Prevotella,* were observed in the level of the genera *Butyricimonas, Coprococcus,* and *Butyrivibrio* (unadjusted $p = 0.03$ for *Butyricimonas* and *Coprococcus* and $p < 0.01$ for *Butyrivibrio*). These findings suggest that seropositivity with more than one autoantibody may be linked with alterations in the abundance of bacteria that belong to the Bacteroidetes and Firmicutes phyla.

Canonical discriminant analysis

We used canonical discriminant analysis to identify constituents of the gut microbiota that multivariately differentiate across the groups. In this analysis, taxa with relative abundance less than 1% were combined into a single rare category. Healthy controls and new onset patients formed fairly-well separated clusters, whereas seronegative FDRs and seropositive subjects clustered together but separate from the other two groups (Figure 3, supplemental Table 3 and Supplemental Table 4). The majority of the separation was due to the first component (p=0.02 for the first component and p=0.15

for the first two). The first component separated the new onset from the remaining groups and the second component separated the healthy controls and new onset cluster from the seropositive and seronegative FDRs. These data raise the hypothesis that the gut microbiomes of seropositive subjects and seronegative FDRs are similar to each other, but distinct from those of new onset patients and unrelated healthy individuals. The data may also imply that the gut microbiota of new onset subjects is different from that of unrelated healthy individuals.

DISCUSSION

The intestinal microbiota plays a pivotal role in maintaining immune homeostasis in the gut and periphery, protecting against microbial infections, and promoting the development and functionality of the immune system (29). A shift in the compositional structure of the gut microbiome could potentially lead to the disruption of the normal interplay between the microbiota and the host, resulting in adverse effects on health (30). Indeed, recent human and animal studies have linked alterations in the gut microbiota to both proinflammatory and metabolic disorders (29). The composition of the gut microbiota is influenced by environmental factors, such as diet, exposure to microbes, geography and cultural differences (reviewed in ref. 31). Herein, we addressed the potential link between compositional changes in the intestinal microbiota and disease development. We present for the first time microbiome data from at-risk and new onset subjects living in the United States. We made four key observations. First, no clear bacterial signature of predominant OTUs was detected in individuals with islet autoimmunity prior to and following disease onset (seropositive and newly diagnosed subjects). Second, the abundances of four bacterial genera were altered in seropositive subjects with ≥1 autoantibodies compared with seronegative FDRs. Third, the gut microbiota of healthy control subjects with no family history of autoimmunity had increased abundances of *Lactobacillus spp.* and *Staphylococcus spp.* versus new onset patients, seropositive and seronegative FDRs. Lastly, canonical discriminant analysis suggested that the intestinal microbiomes of seropositive subjects and seronegative FDR cohorts are similar to each other but distinct from those of new onset patients and unrelated healthy controls, and that the microbiota of new onset patients is different than that of unrelated healthy subjects. Based on these data, we hypothesize that alterations in the intestinal microbiome may be linked with diabetes susceptibility and T1D onset.

How alterations in the abundances of the bacterial phylum Bacteroidetes (Bacteroidetes other) and the Bacteroidetes genus and family *RC9-gut group* and *Prevotellaceae*, respectively, we observed in seropositive subjects compared with seronegative FDRs is not yet clear. Changes in the level of the *RC9-gut-group* was observed by us in diabetes-susceptible TLR9 deficient versus diabetes-resistant mice (11). The altered abundances of the bacterial genus and family *Catenibacterium* and Prevotellaceae*,* respectively, in seropositive versus negative individuals is reminiscent of observations made in subjects with HIV. In patients with HIV, the fecal microbiota contained higher levels of *Catenibacterium* and *Prevotellaceae* (32). It remains to be seen whether the increase in these microbes in subjects with islet autoimmunity may be caused at least in part by proinflammatory responses, as was recently suggested for individuals with HIV (32).

We observed a reduction in the relative abundances of *staphylococci* and *lactobacilli* in the new onset, seropositive, and seronegative FDR groups versus the unrelated healthy control group. The data could raise the hypothesis that subjects in the former groups have relatively limited ability to regulate proinflammatory responses. Indeed, potentially probiotic members of lactobacilli (33) have been associated with beneficial effects on pro-inflammatory disorders (34). Support for the possibility that lactobacilli could potentially downmodulate inflammation is provided by recent data that dendritic cells co-cultured with species of lactobacilli induce polarization of regulatory T cells (35,36). As to *Staphylococcus,* it is an important inhabitant of the human skin and the intestine (37). Its beneficial effect on the gut microbiome could potentially be associated with promoting the growth of anaerobic bacteria including species of the genera *Bifidobacterium*, *Clostridium* and *Bacteroides*, bacterial groups that promote maturation of the neonatal gut (33,38). Adding to the complexity of a possible role of *Staphylococcus* in T1D is the fact that these bacteria were also associated with the ability to

promote proinflammatory responses (39). Thus, future studies are required to elucidate whether and how changes in the abundance of *staphylococci* and *lactobacilli* are linked with disease progression.

The mechanisms leading to alterations in the intestinal microbiota or the timing of these changes in genetically susceptible individuals are not yet known. Our recent animal studies suggested that dysbiosis may develop as a result of microbial infections (10,11). Another potential mechanism could be associated, at least in part, with changes in immunity (32). We have recently shown that monocytes and dendritic cells from autoantibody positive and new onset subjects have enhanced TLR-induced IL-1β responses (12). That the immune system can reshape the gut microbiome with subsequent health changes is based on recent mouse studies. For example, mice with aberrant expression of TLR5, IL-22, or the inflammasome have altered gut bacteria associated with colitis (40-42) and metabolic syndrome (43). Furthermore, Wen and co-workers (6) and our studies (11) demonstrated that diabetes-susceptible mice deficient of various TLR pathways have altered intestinal microbiota. The possibility that the altered gut bacterial composition observed in seropositive subjects and autoantibody-free FDRs compared to new onset patients and unrelated healthy individuals is linked with differences in the expression of high-risk diabetes HLA alleles is not supported by our data, as we did not find significant association between bacterial abundances and the expression HLA3 and/or HLA4 versus other HLA types. Previous studies in mice (44) and rats (45) have demonstrated that MHC expression is one of the factors that governs the gut bacterial composition. It remains to be determined whether, like mice and rats, the gut bacterial composition in humans is influenced by the HLA profile.

Whether the altered gut microbiome observed in new onset patients is associated with disease onset is unknown, since it could be linked at least in part with inflammation observed in patients following disease onset (46). Our data do not support the possibility that a shift in the gut microbiome in this cohort is linked with levels of HbA1c, age at onset, or disease duration (data not shown). It could be

that alterations in the gut microbiota of new onset patients are a result of changes in diet that often ensue following disease onset (reviewed in ref. 47).

We found a pattern towards increased abundances of the Bacteroidetes genus *Bacteroides* and a reduction in the Bacteroidetes genus *Prevotella* and the phylum Firmicutes in subjects with multiple autoantibodies versus 1 autoantibody. These data are consistent with two Finnish studies showing that seropositive individuals with multiple autoantibodies have altered abundance of the *Bacteroides* genus (48). An increase in *Bacteroides* has been linked with the "Western diet" characterized by content high in protein and fat and low in plant fiber, whereas an increase in *Prevotella* has been linked to diet rich in plant fibers (49). *Prevotella* is highly prevalent in African children with diet rich in grains (50) and taxa from both *Prevotella* and Firmicutes can digest plant polysaccharides (49,51) and promote the production of short chain acids known for their anti-inflammatory properties (52). Through similar mechanisms, a Western diet may also be involved in the global rise in the incidence of T1D observed in recent decades (47,53). Our data further demonstrate a trend towards a reduction in the level of *Butyricimonas* and *Coprococcus* and an increase in *Akkermansia* in subjects with multiple autoantibodies compared with one autoantibody. How these changes may be involved in disease progression is unclear. The abundance of the Firmicutes genus *Coprococcus* was shown to be elevated in Crohn's disease and decreased in HIV individuals (54) whereas taxa that belong to *Butyricimonas Butyrivibrio*, and *Akkermansia* were linked with the synthesis of short chain fatty acids (55-57).

Our observations are somewhat different than data from previous reports that demonstrated increased microbial diversity in diabetic patients (58), or increased abundances of the Bacteroidaceae family and the *Bacteroides* genus in seropositive individuals compared with autoantibody-negative controls (48). There could be various reasons for these seemingly disparate observations as the gut microbiome is influenced by multiple environmental factors, such as diet, exposure to microbes,

geography, climate, and cultural differences that exist between different countries and communities (reviewed in ref. 31). Finally, the use of different data analysis approaches in previous studies compared with those used by us could also have led to a different outcome.

This study has a number of potential limitations. First, our results must be confirmed in larger, multi-center subject cohorts to assess their validity and generalizability. The relatively small sample size limited our ability to detect significant differences in the microbiota of individuals with multiple autoantibodies versus one autoantibody. Second, due to the cross-sectional design of the study, we were unable to determine whether alterations in the microbiome within an individual were associated with disease progression. Despite these limitations, our cohorts were very well characterized and the approach used enabled us to analyze for the first time the gut microbiome of genetically susceptible individuals with and without islet autoimmunity versus the general population.

In summary, our data raise the hypothesis that alterations in the structural composition of the intestinal microbiota are associated with T1D progression. Identifying bacteria and immune pathways associated with early diabetes may lead to a new class of immunotherapies to modulate the gut microbiota and prevent islet destruction.

ACKNOWLEDGMENTS

This study was supported by grant 17-2011-655 from JDRF (DZ), and NIH grant HG005964 (DNF).

A.K.A. recruited study participants. N.H. was involved in sample storage and performed HLA typing. P.A.G. contributed to subject recruitment. D.I. produced the DNA library and performed the sequencing. C.E.R. and B.D.W. performed data and bioinformatic analyses. D.N.F. designed and oversaw all aspects of the sequencing experiments and data analysis, researched the data and participated in writing the manuscript. D.Z. oversaw the project and wrote the manuscript and takes full responsibility for the work as a whole, including the study design, access to data, and the decision to submit and publish the manuscript.

The authors report no conflict of interest.

Figure Legends

Figure 1.

Panel A: Stacked bar chart of median percent counts of Operational Taxonomic Units (OTU) representing bacterial genera with a frequency of ≥1% of total counts in the stool from subjects with and without islet autoimmunity as indicated in the figure. The relative abundances are inferred from 16S rRNA sequence counts in datasets. The X and Y axes represent the sample name and percentages of bacterial taxa, respectively.

Panel B: The distribution of Shannon indices across groups is displayed using box plots. The area inside the box represents the interquartile range (IQR: 25th to 75th percentiles), the median and mean are denoted by a line and a circle, respectively. The whiskers extend 1.5 IQR from the box, the observations outside of this range are displayed as points.

Figure 2. Median percent abundance of bacterial communities in subjects with and without islet autoimmunity. The plot displays those taxa that were significantly different across groups following adjustment for covariates (Statistical significance and numeric values are shown in Table 2 and Supplemental Table 1, respectively).

Figure 3. Plot from canonical discriminant analysis used to discriminate between seropositive and seronegative subjects. The weights attributed to the taxa which contributed the most to the components are displayed as vectors in the plot.

TABLE 1. Cohort characteristics

Variable	Unrelated healthy controls (N = 23)	New onset patients (N = 35)	Seropositive (N =21)	Seronegative FDRs (N = 32)
Females/Males	9/14	16/19	12/9	14/18
Age in years, median (range)	12 (4 – 24)	11 (2 – 20)	9 (4 – 49)	12 (3 – 45)
HLA				
3 alone	5 (22%)	10 (29%)	5 (24%)	11 (34%)
4 alone	3 (13%)	13 (37%)	12 (57%)	8 (25%)
3 plus 4	4 (17%)	9 (26%)	4 (19%)	10 (31%)
x/x	11 (48%)	3 (8%)	0	3 (10%)
Autoantibodies				
0	--	2 (6%)	0	32 (100%)
1	--	12 (34%)	5 (24%)	0
2-4	--	21 (60%)	16 (76%)	0
HbA1C, median (range)	--	7.9 (5.5 – 15.0)	5.0 (4.2 – 5.3)	--
Age at onset, years	--	10 (2 – 20)	--	--
Disease duration, weeks	--	5.1 (0.3 – 17.3)	--	--
Impaired glucose metabolism	--	--	5 (24%)	--

TABLE 2. Results from Wilcoxon tests after adjustment for covariates indicating the pairwise comparisons for the genera with a statistically significant difference across the groups

Taxa	Overall p-value	Comparison[a]	F-value	p-value[b]	FDR p-value
Bacteroidetes other[c]	0.006	NO vs Ab+ vs Ab-	2.42	0.094	0.154
		Ab+ vs Ab- vs Co	4.82	0.010	0.029
		Ab+ vs Co	9.48	0.003	0.019
		Ab- vs Co	1.69	0.197	0.276
		NO vs Ab-	2.86	0.094	0.154
		NO vs Ab+	0.26	0.614	0.682
		Ab+ vs Ab-	4.13	0.045	0.090
		Co vs NO	8.65	0.004	0.019
Alistipes	0.013	NO vs Ab+ vs Ab-	4.52	0.013	0.034
		Ab+ vs Ab- vs Co	1.10	0.338	0.422
		Ab+ vs Co	1.30	0.256	0.336
		Ab- vs Co	0.04	0.848	0.893
		NO vs Ab-	9.04	0.003	0.019
		NO vs Ab+	1.73	0.191	0.273
		Ab+ vs Ab-	2.01	0.159	0.231
		Co vs NO	7.02	0.009	0.028
RC9-gut-group	0.024	NO vs Ab+ vs Ab-	3.93	0.023	0.053
		Ab+ vs Ab- vs Co	4.91	0.009	0.028
		Ab+ vs Co	7.52	0.007	0.025
		Ab- vs Co	0.03	0.864	0.893
		NO vs Ab-	0.89	0.348	0.422
		NO vs Ab+	3.77	0.055	0.107
		Ab+ vs Ab-	7.72	0.006	0.024
		Co vs NO	1.15	0.287	0.369
Prevotellaceae	0.006	NO vs Ab+ vs Ab-	5.06	0.008	0.026
		Ab+ vs Ab- vs Co	5.53	0.005	0.022
		Ab+ vs Co	7.59	0.007	0.025
		Ab- vs Co	0.01	0.912	0.912
		NO vs Ab-	4.61	0.034	0.070
		NO vs Ab+	1.30	0.257	0.336
		Ab+ vs Ab-	9.44	0.003	0.019
		Co vs NO	3.69	0.058	0.108
Catenibacterium	0.015	NO vs Ab+ vs Ab-	5.37	0.006	0.024
		Ab+ vs Ab- vs Co	5.13	0.007	0.025
		Ab+ vs Co	2.59	0.111	0.172
		Ab- vs Co	2.22	0.139	0.209
		NO vs Ab-	0.71	0.402	0.469

		NO vs Ab+	6.03	0.016	0.039
		Ab+ vs Ab-	10.23	0.002	0.017
		Co vs NO	0.55	0.462	0.533
Lactobacillus	<.001	NO vs Ab+ vs Ab-	2.32	0.104	0.164
		Ab+ vs Ab- vs CO	8.82	<.001	0.008
		Ab+ vs Co	8.42	0.005	0.020
		Ab- vs Co	16.94	<.001	0.004
		NO vs b-	4.63	0.034	0.070
		NO vs Ab+	1.04	0.310	0.393
		Ab+ vs Ab-	0.87	0.352	0.422
		Co vs NO	4.87	0.030	0.066
Staphylococcus	0.001	NO vs Ab+ vs Ab-	1.88	0.158	0.231
		Ab+ vs Ab- vs Co	8.52	<.001	0.008
		Ab+ vs Co	6.79	0.011	0.029
		Ab- vs Co	16.79	<.001	0.004
		NO vs Ab-	3.59	0.061	0.109
		NO vs Ab+	0.23	0.630	0.691
		Ab+ vs Ab-	1.53	0.219	0.294
		Co vs NO	5.91	0.017	0.041
Succiniclasticum	0.001	NO vs Ab+ vs Ab-	6.58	0.002	0.017
		Ab+ vs Ab- vs Co	0.20	0.818	0.877
		Ab+ vs Co	0.35	0.556	0.634
		Ab- vs Co	0.01	0.907	0.912
		NO vs Ab-	11.73	<.001	0.011
		NO vs Ab+	6.77	0.011	0.029
		Ab+ vs Ab-	0.27	0.605	0.681
		Co vs NO	11.33	0.001	0.012
Thalassospira	0.008	NO vs Ab+ vs Ab-	1.61	0.204	0.283
		Ab+ vs Ab- vs Co	4.73	0.011	0.029
		Ab+ vs Co	1.54	0.217	0.294
		Ab- vs Co	9.26	0.003	0.019
		NO vs Ab-	0.03	0.873	0.893
		NO vs Ab+	2.34	0.129	0.197
		Ab+ vs Ab-	2.76	0.100	0.160
		Co vs NO	8.92	0.004	0.019

[a]Ab-, Seronegative; Ab+, Seropositive; CO, Unrelated control; NO, New onset
[b]Significant differences are highlighted in gray
[c]A group of bacteria that belong to the Bacteroidetes phylum and could not be classified to lower taxa levels

TABLE 3. Median percent abundances of bacterial taxa in subjects with 1 versus multiple autoantibodies

Taxa[a]	1 Autoantibody (n = 5)	2-4 Autoantibodi (n = 16)	Unadjusted *p*-value
Bacteria/Bacteroidetes/Bacteroidia/ Bacteroidales/Bacteroidaceae/ Bacteroides	6.37[b] (5.67 – 6.96)	35.62 (13.20 – 49.29)	0.01[c]
Bacteria/Bacteroidetes/Bacteroidia/ Bacteroidales/Prevotellaceae/ Prevotella	12.91 (9.62 – 26.77)	0.05 (0.01 – 1.82)	0.01
Bacteria/Bacteroidetes/Bacteroidia/ Bacteroidales/Porphyromonadaceae/ Butyricimonas	0.53 (0.28 – 0.54)	0.23 (0.00 – 0.33)	0.03
Bacteria/Firmicutes/Clostridia/ Clostridiales/Lachnospiraceae/ Coprococcus	1.37 (0.62 – 2.16)	0.26 (0.09 – 1.02)	0.03
Bacteria/Firmicutes/Clostridia/ Clostridiales/Lachnospiraceae/ Butyrivibrio	0.08 (0.04 – 0.44)	0.01 (0.00 – 0.00)	<0.01
Bacteria/Verrucomicrobia/ Verrucomicrobiae/Verrucomicrobiales/ Verrucomicrobiaceae/ Akkermansia	0.06 (0.02 – 0.08)	0.74 (0.15 – 4.46)	0.04

[a]Phylogeny of bacterial genera, assigned by 16S rRNA sequencing analysis
[b]Median relative abundance of 16S rRNA sequences for bacterial clade. Shown in brackets are the ranges of the bacterial abundance
[c]Results from Wilcoxon rank-based test

Reference List

1. Gianani,R, Eisenbarth,GS: The stages of type 1A diabetes: 2005. *Immunol Rev* 204:232-249, 2005

2. Chervonsky,A: Innate receptors and microbes in induction of autoimmunity. *Curr Opin Immunol* 21:641-647, 2009

3. Frank,DN, Zhu,W, Sartor,RB, Li,E: Investigating the biological and clinical significance of human dysbioses. *Trends Microbiol* 19:427-434, 2011

4. Peterson,DA, Frank,DN, Pace,NR, Gordon,JI: Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host Microbe* 3:417-427, 2008

5. Dunne,JL, Triplett,EW, Gevers,D, Xavier,R, Insel,R, Danska,J, Atkinson,MA: The intestinal microbiome in type 1 diabetes. *Clin Exp Immunol* 177:30-37, 2014

6. Wen,L, Ley,RE, Volchkov,PY, Stranges,PB, Avanesyan,L, Stonebraker,AC, Hu,C, Wong,FS, Szot,GL, Bluestone,JA, Gordon,JI, Chervonsky,AV: Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* 455:1109-1113, 2008

7. Brugman,S, Klatter,FA, Visser,JT, Wildeboer-Veloo,AC, Harmsen,HJ, Rozing,J, Bos,NA: Antibiotic treatment partially protects against type 1 diabetes in the Bio-Breeding diabetes-prone rat. Is the gut flora involved in the development of type 1 diabetes? *Diabetologia* 49:2105-2108, 2006

8. Roesch,LF, Lorca,GL, Casella,G, Giongo,A, Naranjo,A, Pionzio,AM, Li,N, Mai,V, Wasserfall,CH, Schatz D, Atkinson,MA, Neu,J, Triplett,EW: Culture-independent identification of gut bacteria correlated with the onset of diabetes in a rat model. *ISME J* 3:536-548, 2009

9. Markle,JGM, Frank,DN, Mortin-Toth,S, Robertson,CE, Feazel,LM, Rolle-Kampczyk,U, von Bergen,M, McCoy,KD, Macpherson,AJ, Danska,JS: Sex Differences in the Gut Microbiome Drive Hormone-Dependent Regulation of Autoimmunity. *Science* 339:1088, 2013

10. Hara,N, Alkanani,AK, Ir,D, Robertson,CE, Wagner,BD, Frank,DN, Zipris,D: Prevention of Virus-Induced Type 1 Diabetes with Antibiotic Therapy. *J Immunol* 189:3805-3814, 2012

11. Alkanani,AK, Hara,N, Lien,E, Ir,D, Kotter,CV, Robertson,CE, Wagner,BD, Frank,DN, Zipris,D: Induction of Diabetes in the RIP-B7.1 Mouse Model Is Critically Dependent on TLR3 and MyD88 Pathways and Is Associated With Alterations in the Intestinal Microbiome. *Diabetes* 63:619-631, 2014

12. Alkanani,AK, Rewers,M, Dong,F, Waugh,K, Gottlieb,PA, Zipris,D: Dysregulated Toll-Like Receptor-Induced Interleukin-1ß and Interleukin-6 Responses in Subjects at Risk for the Development of Type 1 Diabetes . *Diabetes* 61:2525-2533, 2012

13. Lau,K, Benitez,P, Ardissone,A, Wilson,TD, Collins,EL, Lorca,G, Li,N, Sankar,D, Wasserfall,C, Neu,J, Atkinson,MA, Shatz,D, Triplett,EW, Larkin,J: Inhibition of Type 1

Diabetes Correlated to a Lactobacillus johnsonii N6.2-Mediated Th17 Bias. *J Immunol* 186:3538-3546, 2011

14. Giongo,A, Gano,KA, Crabb,DB, Mukherjee,N, Novelo,LL, Casella,G, Drew,JC, Ilonen,J, Knip,M, Hyoty,H, Veijola,R, Simell,T, Simell,O, Neu,J, Wasserfall,CH, Schatz D, Atkinson,MA, Triplett,EW: Toward defining the autoimmune microbiome for type 1 diabetes. *ISME J* 5:82-91, 2011

15. de Goffau,M, Fuentes,S, van den Bogert,B, Honkanen,H, de Vos,W, Welling,G, Hy+¦ty,H, Harmsen,H: Aberrant gut microbiota composition at the onset of type 1 diabetes in young children. *Diabetologia* 57:1569-1577, 2014

16. Murri,M, Leiva,I, Gomez-Zumaquero,JM, Tinahones,F, Cardona,F, Soriguer,F, Queipo-Ortuno,MI: Gut microbiota in children with type 1 diabetes differs from that in healthy children: a case-control study. *BMC Medicine* 11:46, 2013

17. Endesfelder,D, Castell,Wz, Ardissone,A, Davis-Richardson,AG, Achenbach,P, Hagen,M, Pflueger,M, Gano,KA, Fagen,JR, Drew,JC, Brown,CT, Kolaczkowski,B, Atkinson,M, Schatz,D, Bonifacio,E, Triplett,EW, Ziegler,AG: Compromised gut microbiota networks in children with anti-islet cell autoimmunity. *Diabetes* 63:2006-2014, 2014

18. Steck,AK, Armstrong,TK, Babu,SR, Eisenbarth,GS: Stepwise or linear decrease in penetrance of type 1 diabetes with lower-risk HLA genotypes over the past 40 years. *Diabetes* 60:1045-1049, 2011

19. Caporaso,JG, Lauber,CL, Walters,WA, Berg-Lyons,D, Lozupone,CA, Turnbaugh,PJ, Fierer,N, Knight,R: Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 108:4516-4522, 2011

20. Ewing,B, Green,P: Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8:186-194, 1998

21. Ewing,B, Hillier,L, Wendl,MC, Green,P: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175-185, 1998

22. Edgar,RC, Haas,BJ, Clemente,JC, Quince,C, Knight,R: UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194-2200, 2011

23. Schloss,PD, Westcott,SL: Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl Environ Microbiol* 77:3219-3226, 2011

24. Quast,C, Pruesse,E, Yilmaz,P, Gerken,J, Schweer,T, Yarza,P, Peplies,J, Glockner,FO: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590-D596, 2013

25. Robertson,CE, Harris,JK, Wagner,BD, Granger,D, Browne,K, Tatem,B, Feazel,LM, Park,K, Pace,NR, Frank,DN: Explicet: graphical user interface software for metadata-driven

management, analysis and visualization of microbiome data. *Bioinformatics* 29:3100-3101, 2013

26. Jost,L: Entropy and diversity. *Oikos* 113:363-375, 2006

27. Aitchinson J: *The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability.* Lomdon, Chapman & Hall, 1986,

28. Filzmoser,P, Hron,K, Reimann,C: Univariate statistical analysis of environmental (compositional) data: problems and possibilities. *Sci Total Environ* 407:6100-6108, 2009

29. Brenchley,JM, Douek,DC: Microbial Translocation Across the GI Tract*. *Annu Rev Immunol* 30:149-173, 2012

30. Kang,DW, Park,JG, Ilhan,ZE, Wallstrom,G, Labaer,J, Adams,JB, Krajmalnik-Brown,R: Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children. *Plos One* 8:e68322, 2013

31. Yatsunenko,T, Rey,FE, Manary,MJ, Trehan,I, Dominguez-Bello,MG, Contreras,M, Magris,M, Hidalgo,G, Baldassano,RN, Anokhin,AP, Heath,AC, Warner,B, Reeder,J, Kuczynski,J, Caporaso,JG, Lozupone,CA, Lauber,C, Clemente,JC, Knights,D, Knight,R, Gordon,JI: Human gut microbiome viewed across age and geography. *Nature* 486:222-227, 2012

32. Lozupone,C, Li,M, Campbell,T, Flores,S, Linderman,D, Gebert,M, Knight,R, Fontenot,A, Palmer,B: Alterations in the Gut Microbiota Associated with HIV-1 Infection. *Cell Host & Microbe* 14:329-339, 2013

33. Penders,J, Thijs,C, Vink,C, Stelma,FF, Snijders,B, Kummeling,I, van den Brandt,PA, Stobberingh,EE: Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* 118:511-521, 2006

34. Hammer,GE, Ma,A: Molecular control of steady-state dendritic cell maturation and immune homeostasis. *Annu Rev Immunol* 31:743-791, 2013

35. Bron,PA, van Baarlen,P, Kleerebezem,M: Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Micro* 10:66-78, 2012

36. Smits,HH, Engering,A, van der Kleij,D, De Jong,EC, Schipper,K, van Capel,TMM, Zaat,BAJ, Yazdanbakhsh,M, Wierenga,EA, van Kooyk,Y, Kapsenberg,ML: Selective probiotic bacteria induce IL-10–producing regulatory T cells in vitro by modulating dendritic cell function through dendritic cell–specific intercellular adhesion molecule 3–grabbing nonintegrin. *J Allergy Clin Immunol* 115:1260-1267, 2005

37. Sanchez,E, Ribes-Koninckx,C, Calabuig,M, Sanz,Y: Intestinal Staphylococcus spp. and virulent features associated with coeliac disease. *J Clin Pathol* 2012

38. Marques,TM, Wall,R, Ross,RP, Fitzgerald,GF, Ryan,CA, Stanton,C: Programming infant gut microbiota: influence of dietary and environmental factors. *Curr Opin Biotechnol* 21:149-156, 2010

39. Edwards,LA, O'Neill,C, Furman,MA, Hicks,S, Torrente,F, P+¬rez-Machado,M, Wellington,EM, Phillips,AD, Murch,SH: Enterotoxin-producing staphylococci cause intestinal inflammation by a combination of direct epithelial cytopathy and superantigen-mediated T-cell activation. *Inflamm Bowel Dis* 18:624-640, 2012

40. Carvalho,F, Koren,O, Goodrich,J, Johansson,M, Nalbantoglu,I, Aitken,J, Su,Y, Chassaing,B, Walters,W, Gonz+ílez,A, Clemente,J, Cullender,T, Barnich,N, Darfeuille-Michaud,A, Vijay-Kumar,M, Knight,R, Ley,R, Gewirtz,A: Transient Inability to Manage Proteobacteria Promotes Chronic Gut Inflammation in TLR5-Deficient Mice. *Cell Host & Microbe* 12:139-152, 2012

41. Elinav E, Strowig T, Kau A, Henao-Mejia J, Thaiss C, Booth C, Peaper D, Bertin J, Eisenbarth S, Gordon J, Flavell R: NLRP6 Inflammasome Regulates Colonic Microbial Ecology and Risk for Colitis (Abstract). *Cell* 145: 5, 2011

42. Zenewicz,LA, Yin,X, Wang,G, Elinav,E, Hao,L, Zhao,L, Flavell,RA: IL-22 Deficiency Alters Colonic Microbiota To Be Transmissible and Colitogenic. *J Immunol* 190:5306-5312, 2013

43. Vijay-Kumar,M, Aitken,JD, Carvalho,FA, Cullender,TC, Mwangi,S, Srinivasan,S, Sitaraman,SV, Knight,R, Ley,RE, Gewirtz,AT: Metabolic Syndrome and Altered Gut Microbiota in Mice Lacking Toll-Like Receptor 5. *Science* 328:228-231, 2010

44. Toivanen,P, Vaahtovuo,J, Eerola,E: Influence of Major Histocompatibility Complex on Bacterial Composition of Fecal Flora. *Infect Immun* 69:2372-2377, 2001

45. Lin,P, Bach,M, Asquith,M, Lee,AY, Akileswaran,L, Stauffer,P, Davin,S, Pan,Y, Cambronne,ED, Dorris,M, Debelius,JW, Lauber,CL, Ackermann,G, Baeza,YV, Gill,T, Knight,R, Colbert,RA, Taurog,JD, Van Gelder,RN, Rosenbaum,JT: HLA-B27 and human beta2-microglobulin affect the gut microbiota of transgenic rats. *Plos One* 9:e105684, 2014

46. Devaraj,S, Dasu,MR, Rockwood,J, Winter,W, Griffen,SC, Jialal,I: Increased Toll-like receptor (TLR) 2 and TLR4 expression in monocytes from patients with Type 1 diabetes: further evidence of a proinflammatory state. *J Clin Endocrinol Metab* 93:578-583, 2008

47. Maslowski,KM, Mackay,CR: Diet, gut microbiota and immune responses. *Nat Immunol* 12:5-9, 2011

48. de Goffau,MC, Luopajarvi,K, Knip,M, Ilonen,J, Ruohtula,T, Harkonen,T, Orivuori,L, Hakala,S, Welling,GW, Harmsen,HJ, Vaarala,O: Fecal Microbiota Composition Differs Between Children With β-Cell Autoimmunity and Those Without. *Diabetes* 2013

49. Wu,GD, Chen,J, Hoffmann,C, Bittinger,K, Chen,YY, Keilbaugh,SA, Bewtra,M, Knights,D, Walters,WA, Knight,R, Sinha,R, Gilroy,E, Gupta,K, Baldassano,R, Nessel,L, Li,H, Bushman,FD, Lewis,JD: Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science* 334:105-108, 2011

50. De Filippo,C, Cavalieri,D, Di Paola,M, Ramazzotti,M, Poullet,JB, Massart,S, Collini,S, Pieraccini,G, Lionetti,P: Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *PNAS* 107:14691-14696, 2010

51. Holmes,E, Li,J, Marchesi,J, Nicholson,J: Gut Microbiota Composition and Activity in Relation to Host Metabolic Phenotype and Disease Risk. *Cell Metabolism* 16:559-564, 2012

52. Smith,PM, Howitt,MR, Panikov,N, Michaud,M, Gallini,CA, Bohlooly,Y, Glickman,JN, Garrett,WS: The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science* 341:569-573, 2013

53. Brown,CT, Davis-Richardson,AG, Giongo,A, Gano,KA, Crabb,DB, Mukherjee,N, Casella,G, Drew,JC, Ilonen,J, Knip,M, Hyoty,H, Veijola,R, Simell,T, Simell,O, Neu,J, Wasserfall,CH, Schatz,D, Atkinson,MA, Triplett,EW: Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. *Plos One* 6:e25792, 2011

54. Sands,S, Tsau,S, Yankee,T, Parker,B, Ericsson,A, LeVine,S: The effect of omeprazole on the development of experimental autoimmune encephalomyelitis in C57BL/6J and SJL/J mice. *BMC Research Notes* 7:605, 2014

55. Belzer,C, de Vos,WM: Microbes inside[mdash]from diversity to function: the case of Akkermansia. *ISME J* 6:1449-1458, 2012

56. Sakamoto,M, Takagaki,A, Matsumoto,K, Kato,Y, Goto,K, Benno,Y: Butyricimonas synergistica gen. nov., sp. nov. and Butyricimonas virosa sp. nov., butyric acid-producing bacteria in the family 'Porphyromonadaceae' isolated from rat faeces. *International Journal of Systematic and Evolutionary Microbiology* 59:1748-1753, 2009

57. Lee,WJ, Hase,K: Gut microbiota-generated metabolites in animal health and disease. *Nat Chem Biol* 10:416-424, 2014

58. de Goffau,MC, Fuentes,S, van den Bogert,B, Honkanen,H, de Vos,WM, Welling,GW, Hyoty,H, Harmsen,HJ: Aberrant gut microbiota composition at the onset of type 1 diabetes in young children. *Diabetologia* 57:1569-1577, 2014

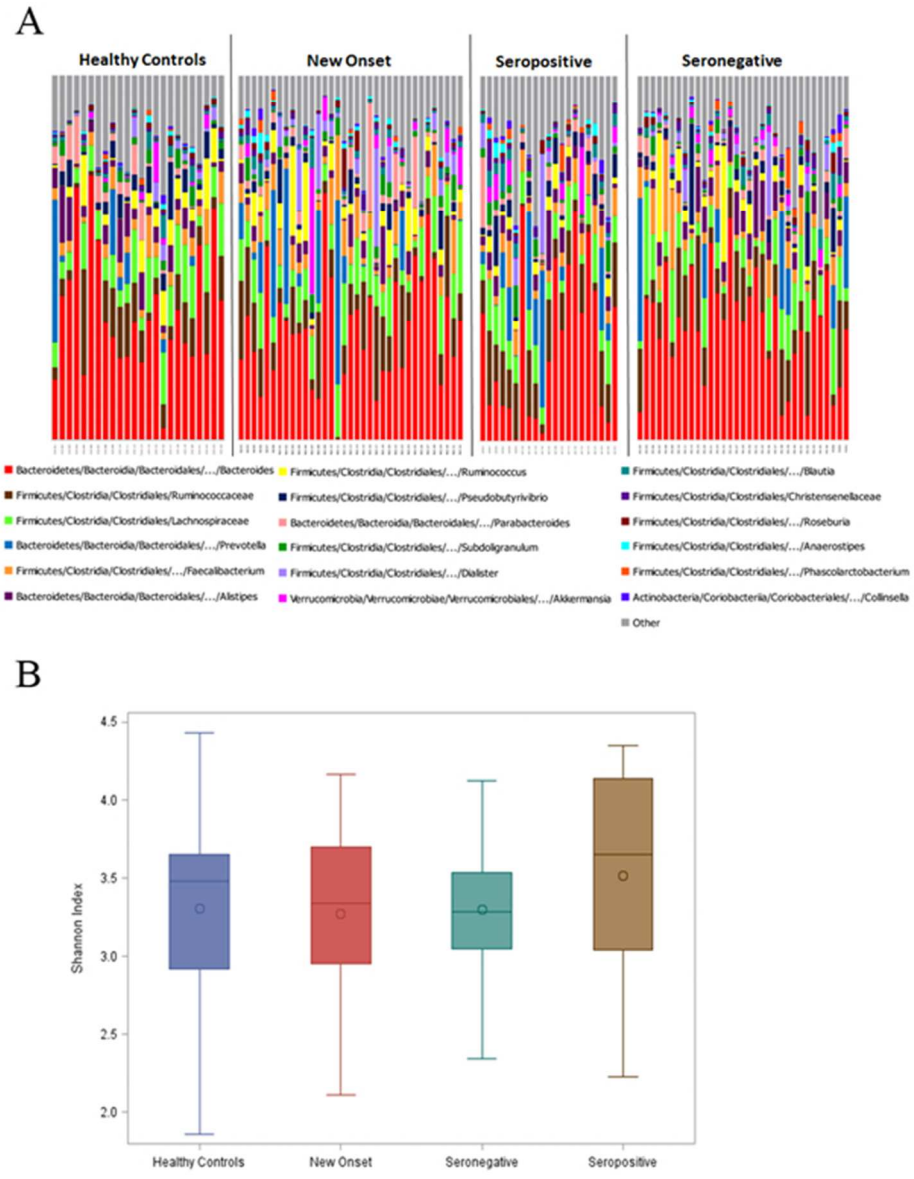

Figure 1. Panel A: Stacked bar chart of median percent counts of Operational Taxonomic Units (OTU) representing bacterial genera with a frequency of ≥1% of total counts in the stool from subjects with and without islet autoimmunity as indicated in the figure. The relative abundances are inferred from 16S rRNA sequence counts in datasets. The X and Y axes represent the sample name and percentages of bacterial taxa, respectively.
Panel B: The distribution of Shannon indices across groups is displayed using box plots. The area inside the box represents the interquartile range (IQR: 25th to 75th percentiles), the median and mean are denoted by a line and a circle, respectively. The whiskers extend 1.5 IQR from the box.

52x66mm (300 x 300 DPI)

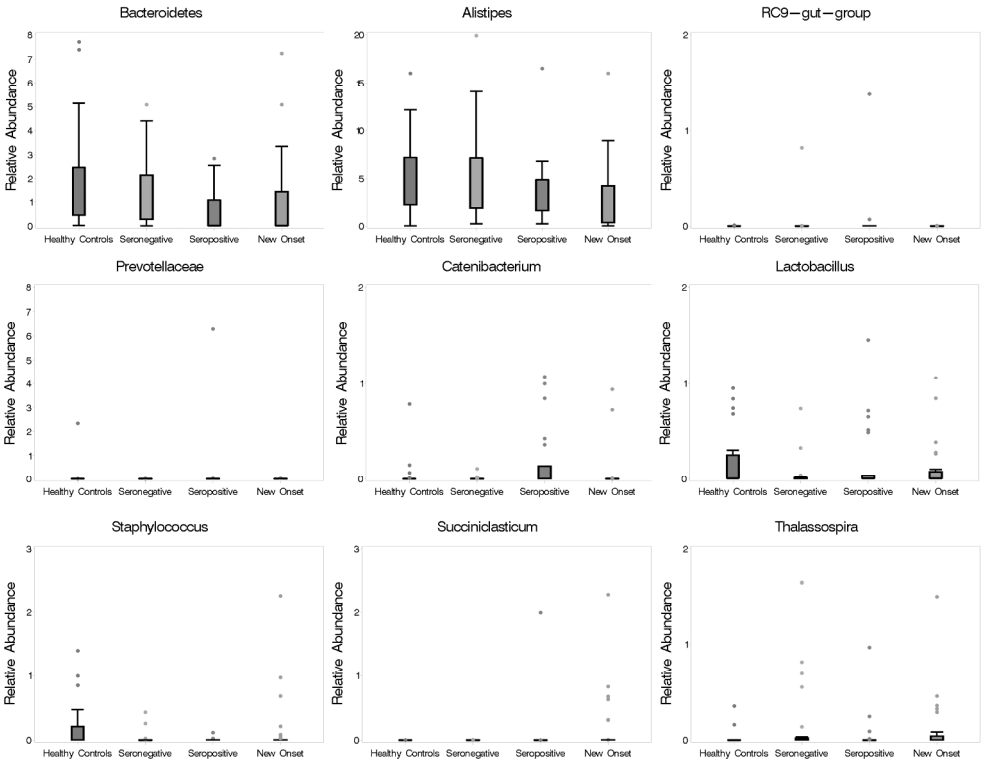

Figure 2. Median percent abundance of bacterial communities in subjects with and without islet autoimmunity. The plot displays those taxa that were significantly different across groups following adjustment for covariates (Significance and numeric values are shown in Table 2 and Supplementary Table 1, respectively).
207x161mm (300 x 300 DPI)

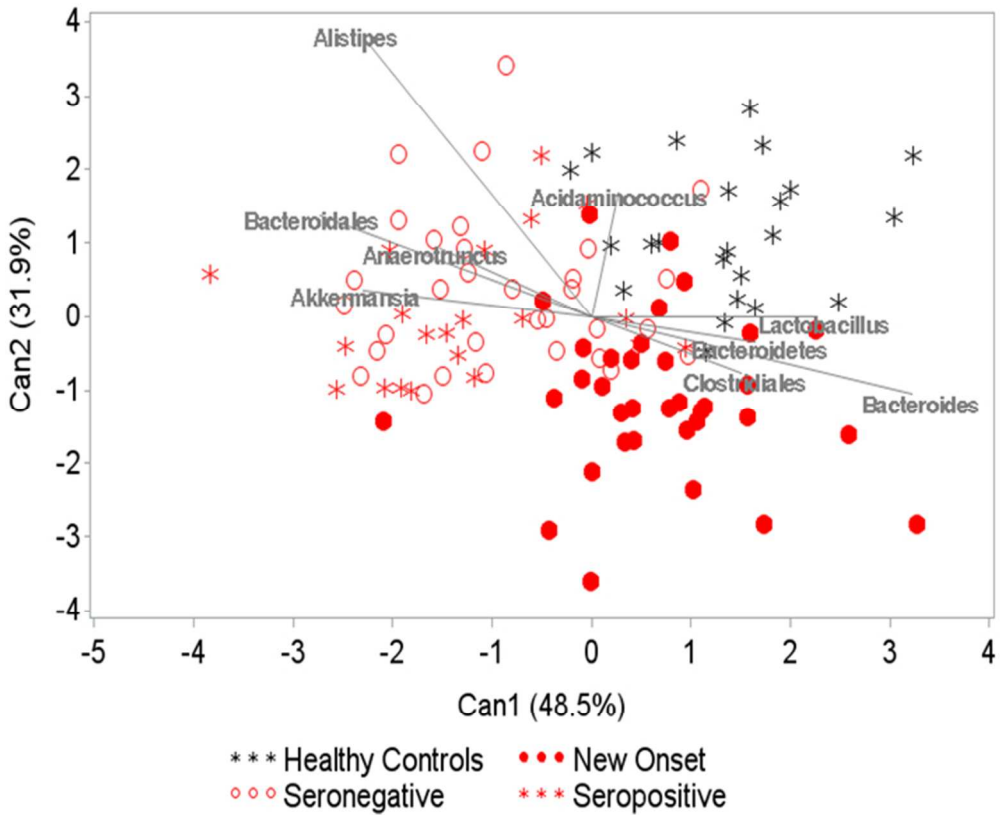

Figure 3. Plot from canonical discriminant analysis used to discriminate between seropositive and seronegative subjects. The weights attributed to the taxa which contributed the most to the components are displayed as vectors in the plot.
60x49mm (300 x 300 DPI)

Online Appendix

Supplemental Table 1. Median percent abundance and interquartile range of bacterial communities in subjects with and without Islet Autoimmunity

	Subject Cohort[a]			
Taxa[b]	Healthy Controls (n = 23)	New Onset (n = 35)	Seropositive (n = 21)	Seronegative FDRs (n = 32)
Bacteroidetes Other	1.65 (0.43 – 4.39) Max=16.71	0.25 (0.04 – 2.13) Max=10.31	0.43 (0.01 – 1.06) Max=2.84	0.74 (0.28 – 2.23) Max=13.49
Bacteroidetes/Bacteroidia/ Bacteroidales/Rikenellaceae/ Alistipes	3.23 (2.15 – 7.17) Max=16.00	1.95 (0.33 – 4.11) Max=16.02	3.65 (1.57 – 4.85) Max=16.53	4.82 (1.82 – 7.12) Max=19.89
Bacteroidetes/Bacteroidia/Bacteroid ales/Rikenellaceae/RC9-gut-group	0 (0 – 0) Max=0.01	0 (0 – 0) Max=0	0 (0 – 0) Max=1.38	0 (0 – 0) Max=0.82
Bacteroidetes/Bacteroidia/ Bacteroidales/Prevotellaceae	0 (0 – 0) Max=2.34	0 (0 – 0) Max=0.01	0 (0 – 0) Max=6.28	0 (0 – 0) Max=0.01
Firmicutes/Erysipelotrichi/Erysipelotr ichales/Erysipelotrichaceae/ Catenibacterium	0 (0 – 0) Max=0.78	0 (0 – 0) Max=0.94	0 (0 – 0.14) Max=1.08	0 (0 – 0) Max=0.11
Firmicutes/Bacilli/Lactobacillales/ Lactobacillaceae/Lactobacillus	0.05 (0.01 – 0.25) Max=0.95	0.01 (0 – 0.08) Max=1.06	0 (0 – 0.04) Max=1.46	0 (0 – 0.01) Max=0.74
Firmicutes/Bacilli/Bacillales/ Staphylococcaceae/Staphylococcus	0.01 (0 – 0.21) Max=1.40	0 (0 – 0) Max=2.25	0 (0 – 0) Max=0.12	0 (0 – 0) Max=0.43
Firmicutes/Clostridia/Clostridiales/ Veillonellaceae/Succiniclasticum	0 (0 – 0) Max=0	0 (0 – 0) Max=2.27	0 (0 – 0) Max=1.99	0 (0 – 0) Max=0
Proteobacteria/Alphaproteobacteria /Rhodospirillales/Rhodospirillaceae/ Thalassospira	0 (0 – 0) Max=0.35	0 (0 – 0.04) Max=1.49	0 (0 – 0) Max=0.97	0 (0 – 0.02) Max=1.64

[a]Median relative abundance of 16S rRNA sequences for bacterial clade and subject cohorts

[b]Phylogeny of bacterial phyla and genera, assigned by 16S rRNA equencing analysis.

Supplemental Table 2. Parameter estimates and corresponding p-values for association between the taxa that were significantly different across groups and the variables included as covariates.

OTU_Name	Parameter	Estimate	StdErr	tValue	Probt
Bacteria/Bacteroidetes other	Gender F	14.23	5.84	2.44	0.02
	Age	-0.03	0.29	-0.11	0.91
	HLA34_geno	-9.05	6.89	-1.31	0.19
Bacteria/Bacteroidetes/Bacteroidia/ Bacteroidales/Prevotellaceae	Gender F	1.01	5.31	0.19	0.85
	Age	0.21	0.26	0.79	0.43
	HLA34_geno	-9.27	6.26	-1.48	0.14
Bacteria/Bacteroidetes/Bacteroidia/Bacteroi dales/Rikenellaceae/Alistipes	Gender F	7.17	5.84	1.23	0.22
	Age	0.50	0.29	1.73	0.09
	HLA34_geno	-9.39	6.88	-1.36	0.18
Bacteria/Bacteroidetes/Bacteroidia/Bacteroi dales/Rikenellaceae/RC9-gut-group	Gender F	1.86	4.48	0.42	0.68
	Age	0.31	0.22	1.40	0.16
	HLA34_geno	-5.38	5.28	-1.02	0.31
Bacteria/Firmicutes/Bacilli/Bacillales/Staph ylococcaceae/Staphylococcus	Gender F	0.50	5.43	0.09	0.93
	Age	-0.13	0.27	-0.47	0.64
	HLA34_geno	4.92	6.41	0.77	0.44
Bacteria/Firmicutes/Bacilli/Lactobacillales/ Lactobacillaceae/Lactobacillus	Gender F	8.07	5.83	1.38	0.17
	Age	-0.13	0.29	-0.46	0.65
	HLA34_geno	6.42	6.87	0.93	0.35
Bacteria/Firmicutes/Clostridia/Clostridiales/ Veillonellaceae/Succiniclasticum	Gender F	-3.59	3.69	-0.97	0.33
	Age	0.06	0.18	0.31	0.76
	HLA34_geno	2.45	4.36	0.56	0.57
Bacteria/Firmicutes/Erysipelotrichi/ Erysipelotrichales/Erysipelotrichaceae/ Catenibacterium	Gender F	5.52	4.18	1.32	0.19
	Age	0.38	0.21	1.81	0.07
	HLA34_geno	5.60	4.93	1.14	0.26
Bacteria/Proteobacteria/Alphaproteobacteria /Rhodospirillales/Rhodospirillaceae/ Thalassospira	Gender F	1.44	5.19	0.28	0.78
	Age	0.13	0.26	0.52	0.61
	HLA34_geno	-9.43	6.12	-1.54	0.13

Supplemental Table 3. The contribution of taxa to the components of the CDA. Taxa with the larger weights are displayed as vectors in Figure 3

		Total-Sample Standardized Canonical Coefficients	
Taxa	**Variable**	**Can1**	**Can2**
4C0d-2	**ccomp1**	0.299804254	0.633716389
Acidaminococcus	**ccomp2**	1.012205628	0.545035741
Akkermansia	**ccomp3**	0.477481505	-0.635678142
Alistipes	**ccomp4**	1.136123143	-0.325754334
Anaerostipes	**ccomp5**	0.364508810	-0.212586495
Anaerotruncus	**ccomp6**	0.309109720	0.155541113
Bacteroidales	**ccomp7**	0.707640095	-0.393053196
Bacteroides	**ccomp8**	-0.767625001	0.784472494
Bacteroidetes	**ccomp9**	-0.082986931	0.742527070
Barnesiella	**ccomp10**	0.614884971	-0.016553197
Bifidobacterium	**ccomp11**	0.279239684	0.684183275
Bilophila	**ccomp12**	-0.070005588	-0.336180224
Blautia	**ccomp13**	-0.474302716	0.154886636
Butyricimonas	**ccomp14**	0.277708610	0.808761348
Butyrivibrio	**ccomp15**	0.206668403	0.374353482
Catenibacterium	**ccomp16**	0.357204764	0.112580537
Christensenellaceae	**ccomp17**	0.423233579	0.388158117
Clostridiales	**ccomp18**	-0.102713275	0.874615808
Clostridium	**ccomp19**	0.390287113	-0.196114698
Collinsella	**ccomp20**	0.208308186	-0.200672444
Coprococcus	**ccomp21**	0.177414091	-0.023285721
Desulfovi	**ccomp22**	0.215282408	0.844102792
Dialister	**ccomp23**	-0.104782334	-0.112840009
Dorea	**ccomp24**	-0.031404346	0.281651560
Enterobacteriaceae	**ccomp25**	0.939062473	-0.310883551
Erysipelotrichaceae	**ccomp26**	0.555484256	-0.078314963
Escherichia	**ccomp27**	0.374316349	0.482396168
Faecalibacterium	**ccomp28**	-0.244162693	-0.295021491

		Total-Sample Standardized Canonical Coefficients	
Taxa	**Variable**	**Can1**	**Can2**
Family-XIII-Incertae-Sedis	**ccomp29**	0.532715785	0.503612389
Firmicutes	**ccomp30**	0.345802327	0.176107321
Fusobacteriales	**ccomp31**	0.868986203	0.655486318
Haemophilus	**ccomp32**	0.346381876	0.289265322
Lachnospira	**ccomp33**	0.034732233	-0.241607988
Lachnospiraceae	**ccomp34**	0.523346924	-0.275436813
Lactobacillus	**ccomp35**	-0.036789106	0.305454475
Megamonas	**ccomp36**	0.338486080	-0.481391905
Megasphaera	**ccomp37**	0.888426756	0.405048090
Odoribacter	**ccomp38**	0.222338119	-0.023734775
Parabacteroides	**ccomp39**	0.471641948	0.375825003
Parasutterella	**ccomp40**	0.026590202	-0.175482204
Peptostreptococcaceae	**ccomp41**	-0.049330887	0.315226803
Phascolarctobacterium	**ccomp42**	0.635199148	0.058990052
Prevotella	**ccomp43**	-0.172670031	0.358867138
Prevotellaceae	**ccomp44**	0.657601617	0.677164329
Pseudobutyrivibrio	**ccomp45**	1.093872519	0.293682695
RC9-gut-group	**ccomp46**	0.979880261	0.413101967
RF3	**ccomp47**	0.203984848	-0.573056456
RF9	**ccomp48**	0.941542213	0.241904500
Rare	**ccomp49**	-0.132724123	0.420280668
Roseburia	**ccomp50**	0.574926873	0.156891393
Ruminococcaceae	**ccomp51**	0.182121120	-0.198241990
Ruminococcus	**ccomp52**	-0.054171724	0.243959367
S24-7	**ccomp53**	0.070465419	-0.474602458
Sarcina	**ccomp54**	-0.001716459	0.485477916
Staphylococcus	**ccomp55**	0.375550283	0.847643372
Streptococcus	**ccomp56**	-0.560229826	-0.331514865
Subdoligranulum	**ccomp57**	0.511932369	1.075266721
Succiniclasticum	**ccomp58**	-0.032840072	0.095040965

		Total-Sample Standardized Canonical Coefficients	
Taxa	**Variable**	**Can1**	**Can2**
Succinivibrio	**ccomp59**	-0.262243313	0.318767805
Sutterella	**ccomp60**	0.472907952	-0.132417982
Thalassospira	**ccomp61**	-0.118320856	0.339149660
Turicibacter	**ccomp62**	0.606779304	0.102158881
Veillonella	**ccomp63**	0.000000000	0.000000000

Supplemental Table 4. Class means on canonical variables

Subject Cohort	Canonical Discriminant 1	Canonical Discriminant 2
Unrelated healthy con	1.37	1.17
New Onset	0.68	-1.10
Seropositive	-1.27	0.02
Seronegative	-0.89	0.35