Ross et al. Microbiome (2015) 3:7
DOI 10.1186/s40168-015-0072-y

Microbiome

RESEARCH Open Access

16S gut community of the Cameron County Hispanic Cohort

Matthew C Ross[1,2], Donna M Muzny[3], Joseph B McCormick[4], Richard A Gibbs[3], Susan P Fisher-Hoch[4] and Joseph F Petrosino[1,2,3*]

Abstract

Background: Obesity and type 2 diabetes (T2D) are major public health concerns worldwide, and their prevalence has only increased in recent years. Mexican Americans are disproportionately afflicted by obesity and T2D, and rates are even higher in the United States-Mexico border region. To determine the factors associated with the increased risk of T2D, obesity, and other diseases in this population, the Cameron County Hispanic Cohort was established in 2004.

Results: In this study, we characterized the 16S gut community of a subset of 63 subjects from this unique cohort. We found that these communities, when compared to Human Microbiome Project subjects, exhibit community shifts often observed in obese and T2D individuals in published studies. We also examined microbial network relationships between operational taxonomic units (OTUs) in the Cameron County Hispanic Cohort (CCHC) and three additional datasets. We identified a group of seven genera that form a tightly interconnected network present in all four tested datasets, dominated by butyrate producers, which are often increased in obese individuals while being depleted in T2D patients.

Conclusions: Through a combination of increased disease prevalence and relatively high gut microbial homogeneity in the subset of CCHC members we examined, we believe that the CCHC may represent an ideal community to dissect mechanisms underlying the role of the gut microbiome in human health and disease. The lack of CCHC subject gut community segregation based on all tested metadata suggests that the community structure we observe in the CCHC likely occurs early in life, and endures. This persistent 'disease'-related gut microbial community in CCHC subjects may enhance existing genetic or lifestyle predispositions to the prevalent diseases of the CCHC, leading to increased attack rates of obesity, T2D, non-alcoholic fatty liver disease, and others.

Background

Obesity is a major public health problem worldwide and the number one risk factor associated with multiple diseases, including type 2 diabetes (T2D). T2D is currently the most prevalent endocrine disease in the world and is estimated to afflict 430+ million people by 2030 [1]. T2D is a multifactorial disorder, with pathogenic contributions from genetics, the environment, and lifestyle [2,3]. Accumulating evidence shows metabolic diseases like T2D develop because of chronic, low grade, systemic inflammation that leads to disruption of the normal gut microbiota [4]. Recent studies have revealed gut microbiome signatures associated with obesity and T2D patients [5,6]. Two independent studies in European and Chinese populations revealed increased abundances of opportunistically pathogenic *Clostridium* species and decreased abundances of butyrate-producing *Roseburia*, *Faecalibacterium*, and *Eubacterium* species associated with T2D patients [7,8]. Karlsson *et al.* [7] also found increased abundances of *Lactobacillus gasseri* and *Streptococcus mutans* predictive of insulin resistance while Qin *et al.* [8] found enrichment in *Escherichia coli* associated with current T2D patients. There is mounting evidence that the structure of the gut microbial community has significant implications for health and disease, and therapeutic manipulations of these communities can have immediate effects [9-14]. For example, in mice, deficiency of Toll-like receptor 5

* Correspondence: jpetrosi@bcm.edu
[1]Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine, Houston, TX, USA
[2]Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX, USA
Full list of author information is available at the end of the article

(TLR-5), an innate immune sensor of flagellin, results in mice that mimic the symptoms of metabolic syndrome including hyperlipidemia, hypertension, and insulin resistance [15]. These symptoms were induced in wild-type germ-free mice upon fecal transfer from the TLR-5 deficient donors. Likewise, there is evidence that direct alteration of the gut microbiome through fecal transplantation can temporarily reverse many of the symptoms associated with metabolic syndrome and other diseases including ulcerative colitis, irritable bowel syndrome, and chronic fatigue syndrome [12,14,16]. Diet can have a major impact on the composition of the gut microbial community and has been implicated in the establishment of gut enterotypes, which are distinct microbial community signatures driven by the differential abundance of certain key taxa [17]. However, once established, enterotype-defining taxa appear to be resistant to modification through dietary intervention [18], suggesting that events early in life establish enduring signatures in the gut microbial composition [19]. However, studies in animals and humans implementing dietary interventions aimed at improving metabolic markers have noted non-enterotype-defining shifts in the gut microbiome associated with improved health [13,20-23]. For example, *Akkermansia muciniphila*, a closely adherent mucin-degrading bacterium, is depleted after administration of a high-fat diet and is strongly negatively associated with obesity and T2D [24]. Dietary supplementation of mice with *A. muciniphila* reversed the high-fat-diet effects, such as inflammation and insulin resistance, presumably through improved gut barrier function. As a result, there is mounting enthusiasm for altering the microbiome through changes in lifestyle, diet and/or probiotics to help prevent and alleviate many of the disease risks associated with obesity and T2D.

Americans of Mexican descent (MAm) are at an increased risk of obesity and T2D compared to all Americans (Am) nationally (MAm 39.1% BMI ≥30, T2D 12.8%) (Am 35.7% BMI ≥30, T2D 8.3%) [25-27]. Diabetes in Hispanics occurs earlier (mean age of diagnosis 49.4 years for Hispanics, 53.8 years for all Americans), manifests with higher complication rates, and attacks at nearly twice the rate of non-Hispanic whites (WAm) (T2D 12.8% MAm, T2D 7.6% WAm [28]). Mexican Americans living along the United States (US)-Mexico border are at even greater risk of developing T2D compared to Mexican Americans nationally (15.7% along the border vs. 12.8% nationally) [29-33]. This likely has roots in the genetic makeup of this population, lifestyle, diet, and socioeconomic status among other factors.

The US-Mexico border region comprises a diverse mixture of economies and disease burdens owing to the very unique countries that lie on either side. Driving this dichotomy is the greater than fivefold disparity between the GDP-per-capita of these two countries [34]. To identify the important risk factors for obesity and T2D of Mexican Americans living in the lower Rio Grande Valley, the Cameron County Hispanic Cohort (CCHC) was established in 2004 [35] (Figure 1). Overall, members of this community have much higher obesity (50.9% BMI ≥30, 9.0% BMI ≥40) and T2D rates (28.0%) versus the average American population (35.7% BMI ≥30, 6.3% BMI ≥40, T2D 8.3%) [36-38]. These statistics reflect the general trend of Mexican Americans living along the entire US-Mexico border [27]. Participants in this study were slightly heavier but had lower T2D rates than the CCHC as a whole (60% BMI ≥30, 12.9% T2D). The CCHC represents the first exclusively Mexican American group from a border city with poor overall

Figure 1 Location of Cameron County, Texas. MEX, Mexico; TX, Texas.

health. In light of the previously described recent studies, and the disproportionately increased prevalence of T2D and obesity in this population, we sought to characterize the gut microbiome in 63 subjects belonging to the CCHC. For this study, we chose to utilize Human Microbiome Project (HMP) stool data for comparative analysis [39-41] (Figure 2) (Figure 3). The 300 participants of the HMP were subjected to a lengthy list of exclusion criteria and here represent a healthy Western microbiome [42]. Using the HMP stool data as a reference will help determine associations with the gut microbial structure and increase our understanding about the observed predisposition to obesity and T2D in the CCHC.

Results and discussion

16S rRNA gene sequencing

16S rRNA gene profiling of 63 CCHC subjects revealed that no taxa is significantly associated with any clinical measure evaluated in this study (after correcting for multiple comparisons). These included BMI, age, cholesterol, waist-to-hip ratio, diabetes status, triglycerides, sex, fasting glucose, and others (complete metadata included in Additional file 1). We found this apparent high level of homogeneity across subjects surprising, as the variance between subjects of the CCHC was about one third of that observed between HMP subjects.

In the CCHC subjects, we found significantly elevated levels of organisms belonging to the *Firmicutes* ($P < 0.001$, FDR <2%) and *Actinobacteria* ($P < 0.001$, FDR <2%) phyla, while there were significantly fewer *Bacteroidetes* ($P < 0.001$, FDR <2%) compared to subjects in the HMP (Table 1) (Figure 4). At the family level, the CCHC showed significantly increased *Lachnospiraceae* ($P < 0.001$, FDR <1%), *Veillonellaceae* ($P < 0.001$, FDR <1%), *Coriobacteriaceae* ($P < 0.001$, FDR <1%), *Ruminococcaceae* ($P < 0.002$, FDR <1%), and *Prevotellaceae* ($P < 0.005$, FDR <2%) compared to HMP subjects. Significantly decreased families include *Bacteroidaceae* ($P < 0.001$, FDR <1%) and *Rikenellaceae* ($P < 0.001$, FDR <1%). For those operational taxonomic units (OTUs) assignable to a genus, many of the significant differences occur in genera previously reported as associated with obesity and/or T2D. The groups of significantly increased genera in the CCHC include *Prevotella* ($P < 0.003$, FDR <4%), *Collinsella* ($P < 0.001$, FDR <2%), *Roseburia* ($P < 0.001$, FDR <2%), *Streptococcus* ($P < 0.001$, FDR <2%), and *Dialister* ($P < 0.003$, FDR <4%). Those significantly decreased in the CCHC include *Bacteroides* ($P < 0.001$, FDR <2%), *Alistipes* ($P < 0.001$, FDR <2%), and *Parabacteroides* ($P < 0.001$, FDR <2%). Many differences between the CCHC gut microbial community and HMP subjects noted above tend to mirror shifts often observed in studies comparing subjects who are obese or have T2D versus healthy controls. However, while T2D and obesity are pervasive in the CCHC, many members who were examined in this study are not obese or diabetic (40% BMI <30, 87.3% HbA1c <6). This suggests that the gut microbiome may be a contributing factor to the development of metabolic disease or that the microbial community structure may serve as a predictive biomarker of metabolic disease onset.

In this subset of CCHC members compared to the HMP, we find significantly increased *Coriobacteriaceae*, and specifically the genus *Collinsella* (Figure 5). These shifts are strongly correlated with high low-density lipoprotein levels and high total cholesterol in both human and animal studies [43,44]. For example, the observation was made decades ago that germ-free animals have higher serum cholesterol [45]. Individuals with borderline-high total cholesterol (≥200 mg/dL) and high total cholesterol (≥240 mg/dL) are more prevalent in these CCHC study subjects compared to Mexican Americans nationwide (CCHC 50% and 22.2% versus all Mexican Americans 46.4% and 14.3%, respectively). Additionally, a higher proportion of CCHC subjects in this study have elevated LDL cholesterol (≥130 mg/dL) versus Mexican Americans

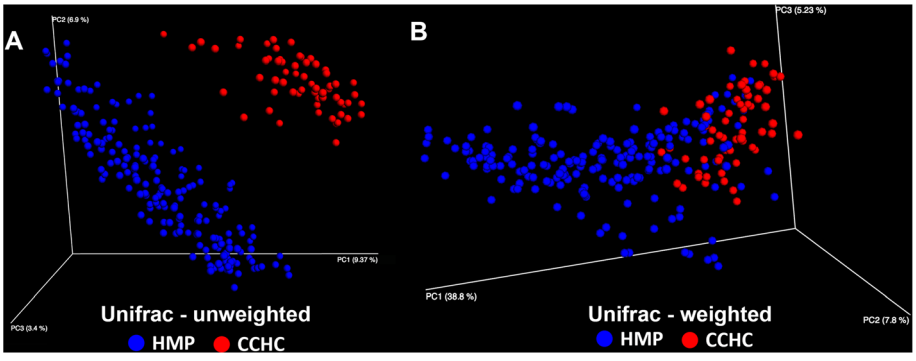

Figure 2 16S rRNA sequence principle coordinates analysis (PCoA) plots of CCHC and HMP stool samples. (A) Comparison of stool samples from the CCHC and HMP analyzed using the unweighted UniFrac metric. **(B)** Same stool samples as in 1A analyzed using the weighted UniFrac metric; HMP $n = 213$, CCHC $n = 63$. CCHC, Cameron County Hispanic Cohort; HMP, Human Microbiome Project; PC, principal coordinate.

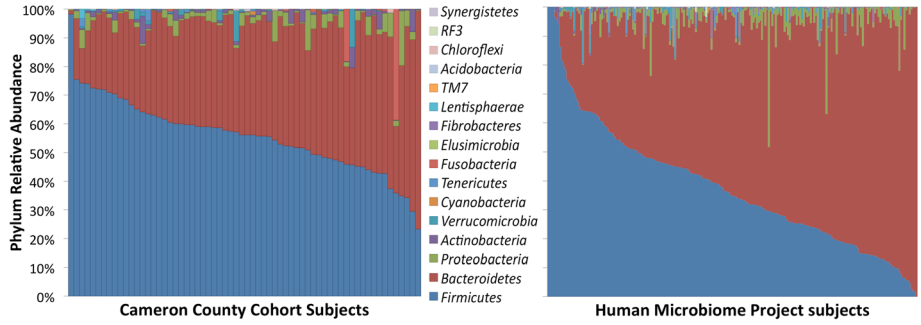

Figure 3 Phylum level analysis of CCHC and HMP subject stool samples. HMP $n = 213$, CCHC $n = 63$.

nationally (CCHC 30.2% versus all Mexican Americans 27.7%). Members of the *Coriobacteriaceae* family respond to dietary interventions involving grain sorghum lipids and dietary whole grains that decrease cholesterol absorption by the host [21,46]. *Coriobacteriaceae* levels were negatively associated with improved metabolic and immunological markers after dietary intervention [21]. Similarly, studies also noted significantly higher *Collinsella* in omnivores compared to vegetarians [47] and significant reductions of *Collinsella* on a low-carb weight loss diet [48]. Together these observations show that *Collinsella* and perhaps other members of *Coriobacteriaceae* are often positively correlated with disease, particularly elevated cholesterol, and may be a worthwhile target for behavioral, probiotic, and/or prebiotic manipulation or as a diagnostic biomarker.

Table 1 16S rRNA relative abundance comparison between CCHC and HMP subject stool samples

	CCHC	HMP	*P* value	*q* value
Phylum				
Firmicutes	56.7%	38.1%	<0.001	<0.02
Bacteroides	37.2%	57.0%	<0.001	<0.02
Actinobacteria	1.28%	0.28%	<0.001	<0.02
Family				
Lachnospiraceae	26.8%	17.6%	<0.001	<0.01
Veillonellaceae	3.10%	1.46%	<0.001	<0.01
Coriobacteriaceae	1.22%	0.23%	<0.001	<0.01
Ruminococcaceae	20.6%	14.3%	<0.01	<0.01
Prevotellaceae	12.1%	4.4%	<0.01	<0.02
Bacteroideaceae	17.8%	40.2%	<0.001	<0.01
Rikenellaceae	1.64%	5.35%	<0.001	<0.01
Genus				
Prevotella	11.9%	4.10%	<0.01	<0.05
Collinsella	1.09%	0.16%	<0.001	<0.02
Roseburia	1.37%	0.95%	<0.001	<0.02
Streptococcus	0.68%	0.05%	<0.001	<0.02
Dialister	0.55%	0.13%	<0.01	<0.05
Bacteroides	17.7%	39.8%	<0.001	<0.02
Alistipes	1.51%	5.10%	<0.001	<0.02
Parabacteroides	1.92%	4.58%	<0.001	<0.02

P and *q* values were calculated via metastats (http://metastats.cbcb.umd.edu/detection.html).

In addition, among our study subjects, we observe a significantly higher *Firmicutes*:*Bacteroidetes* ratio compared to the HMP (Figure 4). Higher fecal concentrations of short-chain fatty acids (SCFA) in obese compared with lean individuals have been attributed to a higher *Firmicutes*:*Bacteroidetes* ratio [49]. In this study, significantly more reads were identified as *Lachnospiraceae* and *Roseburia* (*Firmicutes*) in the CCHC subjects compared to the HMP (Figure 6). Increased *Lachnospiraceae* in particular has been associated with obesity, non-alcoholic fatty liver disease (NAFLD), and protection from colorectal cancer [50,51]. The protection from colorectal cancer is attributed to higher butyrate production, an ability harbored by many species in the *Lachnospiraceae* family including species of the genus *Roseburia*. Interestingly, these observations correlate with the observation that Hispanics living along the Texas-Mexico border have significantly lower rates of colorectal cancer than those living in non-border counties [52]. In the CCHC, colorectal cancer ranks seventh amongst women and sixth amongst men in incidence, whereas colorectal cancer ranks second nationally for Hispanics of both sexes [53]. None of the subjects included in this study had colorectal cancer. Additionally, NAFLD is found at much higher rates amongst the CCHC members (46% of study subjects had elevated alanine aminotransferase), and Hispanic individuals in general [54]. These observations demonstrate the importance of understanding host/microbial relationships in order to prevent unforeseen effects of targeted gut microbial community manipulation, because gut microbial composition may be concurrently protective of and predisposing to certain diseases.

The *Prevotellaceae* family has exhibited strong positive associations with obesity [55] and impaired glucose

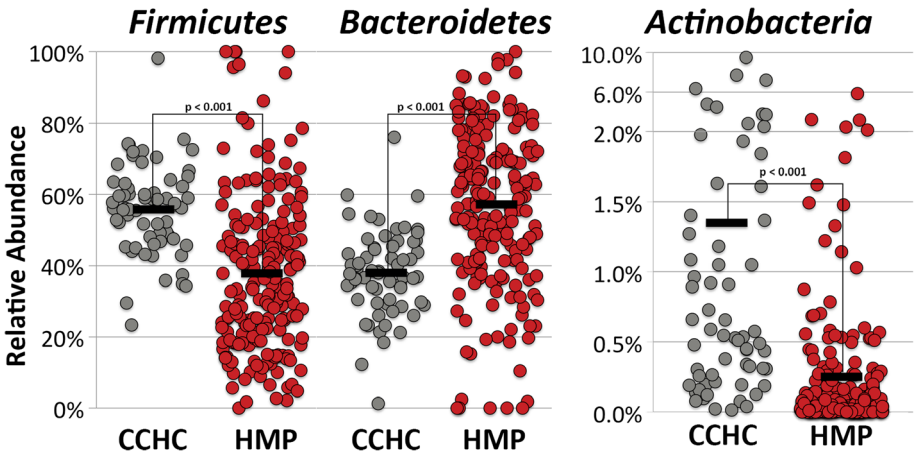

Figure 4 Breakdown of phylum level analysis showing the three most abundant phyla. Scatter plots representing 16S rRNA relative abundance of the three most prevalent phyla: *Firmicutes*, *Bacteroidetes*, and *Actinobacteria*. Black bars represent mean relative abundance. All *P* values were calculated by Mann-Whitney *U* test. HMP $n = 213$, CCHC $n = 63$. CCHC, Cameron County Hispanic Cohort; HMP, Human Microbiome Project.

tolerance [56], while having a negative association with type 1 diabetes [57,58]. In the CCHC, subjects clustered distinctly by their relative abundance of *Prevotella* into two groups (Figure 7). Of note, Bergstrom *et al.* [19] found a bimodal distribution of *Prevotella* apparent by the third year of life, and Roager *et al.* [18] showed that, in adults, this high/low *Prevotella* grouping remained stable during a 6-month dietary intervention, even after courses of antibiotics. Numerous recent studies have linked high prevalence of this family to a predominately plant-based diet [47,59-61]. However, it seems unlikely that differences in diet are the sole explanation for the bimodal distribution of *Prevotella* within the CCHC. All of the CCHC subjects are from the same general neighborhood, are of the same ethnicity, share similar socioeconomic status, and in some cases are family members. Interestingly, family members were as often discordant for high/low *Prevotella* grouping as concordant. This suggests factors additional to diet may determine the prevalence of *Prevotella* in the gut microbial community. In the CCHC, the bimodal distribution was not explained by age, sex, BMI, waist-to-hip ratio, income quartile, birth country, total cholesterol, triglycerides, high blood pressure, diabetes status, fasting glucose, or weight change over 5 years (as percentage of body weight). This recurring observation of sample division by high/low *Prevotella* abundance remains intriguing; however, none of the metadata we examined suggests a reason for the distinct grouping.

OTU network analysis

The gut microbial environment is influenced by host genetics, host diet, and other factors and may influence the way the gut microbiota interact with each other and the host [20,62]. We sought to determine whether OTUs exhibited correlative associations across individuals in the CCHC, possibly revealing biologically relevant relationships between taxa not discernable from other types

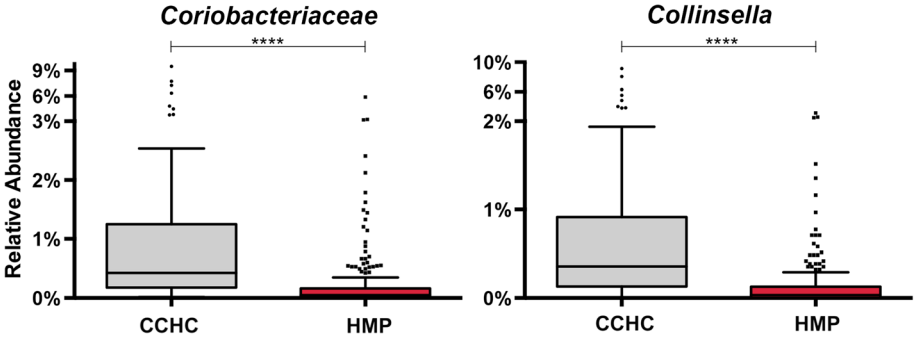

Figure 5 *Coriobacteriaceae* and *Collinsella* 16S rRNA relative abundance differences between CCHC and HMP stool samples. Tukey style box plots of the differences in 16S rRNA relative abundance of the family *Coriobacteriaceae* and the genus *Collinsella* between CCHC and HMP subject stool samples. Statistical significance was evaluated by Mann-Whitney *U* test where **** = $P < 0.0001$; HMP $n = 213$, CCHC $n = 63$. CCHC, Cameron County Hispanic Cohort; HMP, Human Microbiome Project.

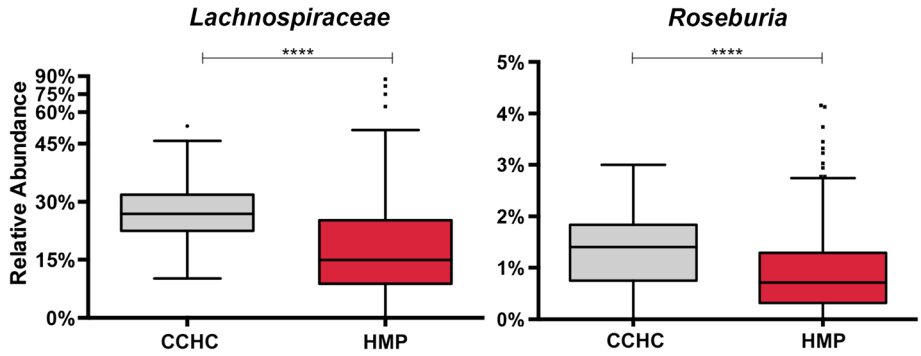

Figure 6 *Lachnospiraceae* and *Roseburia* 16S rRNA relative abundance differences between CCHC and HMP stool samples. Tukey style box plots of the differences in 16S rRNA relative abundance of the family *Lachnospiraceae* and the genus *Roseburia* between CCHC and HMP subject stool samples. Statistical significance was evaluated by Mann-Whitney *U* test where **** = $P < 0.0001$; HMP $n = 213$, CCHC $n = 63$. CCHC, Cameron County Hispanic Cohort; HMP, Human Microbiome Project.

of analysis. Among other applications, this information can prove useful for *in vitro* manipulation of organisms currently recalcitrant to culture [63,64] and will likely be indispensible as probiotics and personalized medical treatments are developed.

We incorporated two additional, unrelated datasets to diversify the population for the OTU correlation analysis to determine whether correlated OTUs could either remain so across diverse populations, age groups, and socioeconomic backgrounds or whether these correlations may only hold true for certain population strata. These additional datasets include a type 1 diabetes cohort from Mexico [57] and an elderly cohort from Ireland [65] (Figure 8). We identified 22 OTU pairs that were positively correlated in at least three of the four datasets, suggesting a mutualistic relationship or possibly codependency (Figure 9). Additionally, these 22 OTU pairs only comprise 11 distinct genera, with only 7 genera contained in 18 of the pairings. This highlights a small network of highly correlated OTUs present in the human gut. Of these seven highly correlated OTUs, five belong to the butyrate-producing family *Lachnospiraceae*, while the remaining two belong to *Ruminococcaceae.* Of these seven genera, all but one, *Lachnospira*, were significantly increased in the CCHC compared to the HMP. Many of these genera are noted for containing species that produce butyrate through fermentation of hydrolyzed polysaccharides and may explain the low rates of colon cancer seen in this population. Additional studies have observed depletion of these genera in T2D, suggesting that diminished butyrate production in the gut may play a role in the pathogenesis of T2D [66].

Negative correlations were less abundant and no OTU pairs were negatively correlated across all four datasets; however, there were nine pairs that were negatively correlated in two or three of the datasets. Many studies have noted a strong co-exclusive relationship between *Prevotella* and *Bacteroides*; however, these were only strongly negatively correlated in two of the four datasets (CCHC and T1D). In the two datasets in which there was little correlation between *Prevotella* and *Bacteroides* (HMP and elderly), *Prevotella* was present at very low

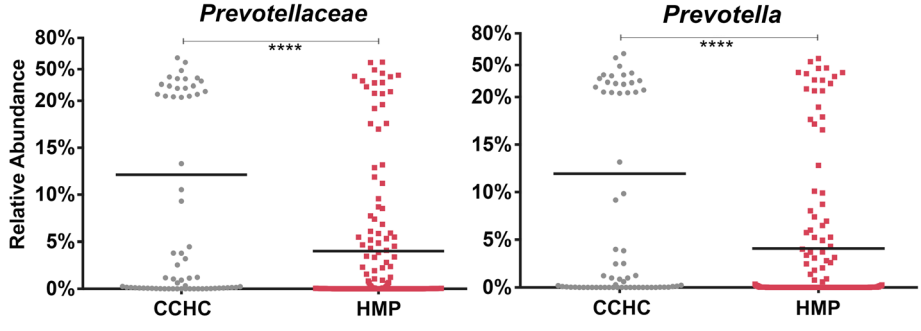

Figure 7 *Prevotellaceae* and *Prevotella* 16S rRNA relative abundance differences between CCHC and HMP stool samples. Scatter plots representing 16S rRNA relative abundance of the family *Prevotellaceae* and the genus *Prevotella*. Black bars represent mean relative abundance. All *P* values were calculated by Mann-Whitney *U* test where **** = $P < 0.0001$; HMP $n = 213$, CCHC $n = 63$. CCHC, Cameron County Hispanic Cohort; HMP, Human Microbiome Project.

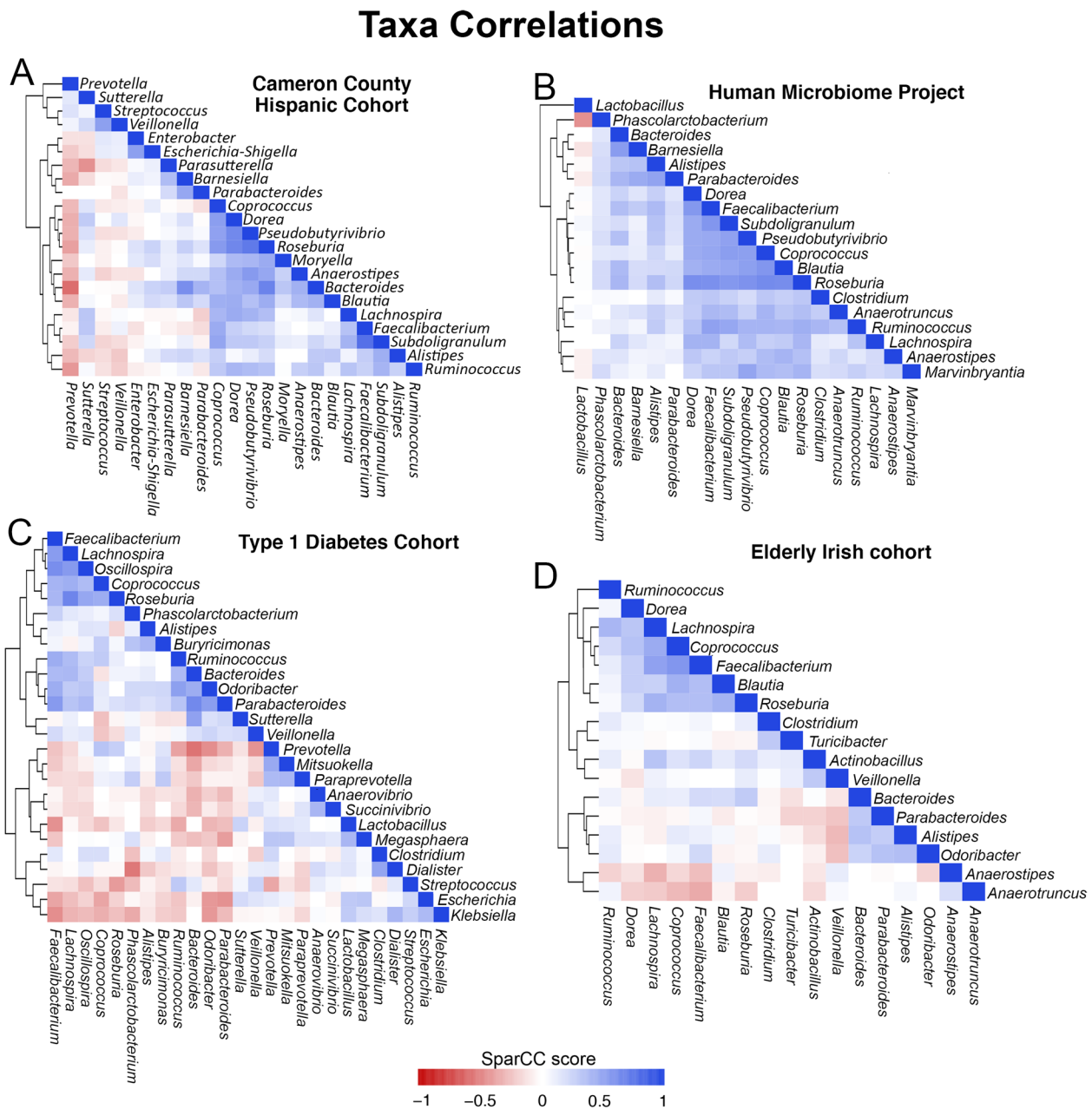

Figure 8 Heat maps showing the strength of OTU correlations calculated by the SparCC algorithm. These heat maps represent the strength of pairwise OTU-OTU co-occurrence/co-exclusion calculations made by the SparCC algorithm. Scores range from −1, a perfect negative correlation, to 1, a perfect positive correlation. **(A)** Cameron County Hispanic Cohort subjects. **(B)** Human Microbiome Project subjects. **(C)** Type 1 diabetes cohort subjects. **(D)** Elderly Irish cohort subjects. For **(A)**, **(B)**, and **(C)** only OTUs with a correlation above 0.4 were plotted for clarity. For **(D)**, OTUs with a correlation above 0.3 were plotted.

abundance across subjects. This suggests that the host environment, such as differences imparted by genetic factors or by diet, may tip the balance in a competition between species as has been suggested for *Prevotella*. In both the T1D (children) and elderly cohorts, we found that strong OTU correlations, both positive and negative, were much less abundant. This may be explained by the increased flux in the microbiota in the very young and the very old as well as extreme interpersonal variability in these age groups [65,67].

Conclusions

Within the CCHC, there are no significant gut microbiome shifts associated with age, sex, disease status, or

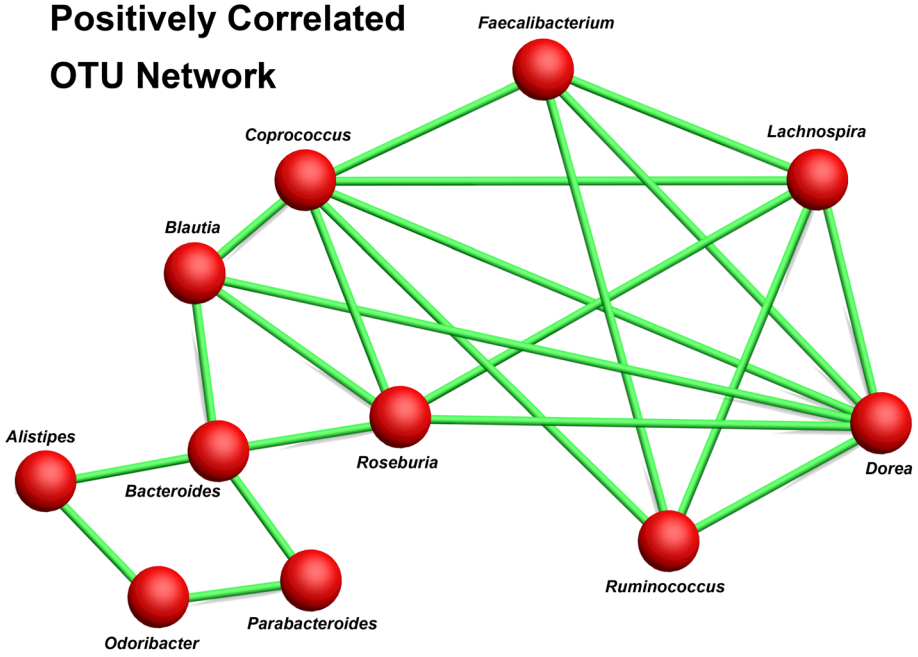

Figure 9 Network plot of commonly positively correlated taxa. This edge/node style plot shows OTU-OTU co-occurrence correlations that appeared in at least three of the four datasets tested. Nodes (red spheres) represent genera and edges (green lines) represent a correlation across multiple datasets. OTU, operational taxonomic unit.

any other available measure. This suggests that the often-reported 'disease' associated gut community shifts we found in the CCHC manifest early and likely persist for life. The youngest CCHC member of this analysis set is 28 years old, thus sampling this population at earlier time points would provide insight into the age at which these observed signatures become apparent. Whether causal or an effect of outside influences, this persistent community structure might compound other predisposing genetic or lifestyle factors, leading to the higher rates of obesity, T2D, NAFLD, and other diseases observed in the CCHC. However, this gut community structure may contribute to the low rates of colorectal cancer observed in the CCHC through increased production of butyrate, suggesting that a particular gut microbiome composition can be both predisposing and protective of different diseases simultaneously. This highlights the need for a more complete understanding of host/microbe relationships when implementing targeted manipulation of the human gut community.

Among other observations, we identified a core group of taxa that appear to be tightly correlated across many populations. A large portion of these taxa belong to the butyrate-producing family *Lachnospiraceae* and were found to be increased in relative abundance in the CCHC compared to the HMP. This correlated group of taxa is also often depleted in T2D patients, suggesting a possible link between the metabolic functions of these taxa and progression of T2D. Information about microbial codependence will be indispensable in the creation of synthetic communities for probiotics or other purposes.

Manipulation of the intestinal microbiome shows promise as a therapy for many diseases; nonetheless, a comprehensive understanding of the mechanisms of action remains to be deciphered. Members of the CCHC exhibit increased prevalence of obesity and T2D, which make them an interesting population to utilize for unraveling related host/microbiome relationships. However, this was a cross-sectional study and thus lacks the resolution necessary to elucidate how the microbiome signatures in the CCHC specifically impact these diseases.

Additionally, we found a higher level of gut microbiome homogeneity within the CCHC compared with other sampled populations such as the HMP (average variance between CCHC subjects across phyla was more than threefold lower than HMP subjects), perhaps making the CCHC an ideal community to test prebiotic or probiotic interventions. Future directions will include longitudinal sampling of CCHC members so that we may assess the low variability and overall stability of the gut microbiome signatures observed in this cohort and its involvement in disease states.

Methods

A subset of participants from the original CCHC study [35] were re-solicited for a follow-up visit and asked to provide a stool sample. This follow-up visit occurred approximately 5 years after the original study visit. Stools

were self-collected at home, delivered to the clinic, aliquoted into 50 mL tubes, mixed to a 50% suspension in RNAlater (Invitrogen, Carlsbad, CA, USA) and stored at −20°C within 24 h of collection. A total of 69 stool samples were delivered to Baylor College of Medicine for extraction and sequencing, and sequence data was generated for 63 of them. Samples for both the HMP and CCHC studies were processed via the same protocols [39] and, in most cases, by the same personnel. The samples were thawed, vortexed, and a wide-bore tip was used to transfer approximately 750 μL of the slurry to a MoBio PowerSoil garnet bead tube (Mo Bio Laboratories, Carlsbad, CA, USA). The stool was processed according to the kit instructions and eluted into 50 μL. 16S rRNA sequencing was performed on the V1 to V3 region (primers 27f and 534r) and sequenced via Roche 454 pyrosequencing (454 Life Sciences, Branford, CT, USA) (average 12,900 reads/sample). Raw fastq files have been deposited into SRA (project accession SRP053023). Sequences were processed using QIIME software [68]. Reads were quality trimmed using default settings and normalized to 3,100 reads/sample. All CCHC clinical metadata are included in Additional file 1. OTUs were picked closed-reference using uclust-ref against the Silva database release 111 at 97%. OTU network correlation analysis was performed using 16S rRNA relative abundance data to identify clusters of highly correlated OTUs. Iterations of the sparse correlations for compositional data (SparCC) algorithm [69] were used to generate correlation matrices in this analysis. This algorithm is specifically designed to account for compositional effects often present in genomic survey data, such as 16S rRNA sequence data. We chose the HMP dataset to compare against, as this represents a standard gut microbiome of a healthy population [39-41]. For our analysis, we used all HMP stool samples that had 16S rRNA V1 to V3 region data generated and also contained at least 3,100 reads, which amounted to 213 samples.

Additional file

Additional file 1: CCHC clinical metadata. This file contains relevant clinical metadata on the subjects included in this study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MCR carried out all sample manipulations including nucleic acid extraction, amplification, and preparation for sequencing; performed all data analysis; generated all figures; and drafted the manuscript. DMM oversaw all sequence data generation and quality control. JBM designed and developed the CCHC. RAG oversaw all sequence data generation and quality control. SPF-H designed and developed the CCHC and participated in design of study. She is also responsible for collection of specimens and data and reviewed the manuscript. JFP helped conceive the study, participated in the design of the study, and helped coordinate and draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank our cohort recruitment team, particularly Rocio Uribe, Jodi Rodriguez, and Julie Ramirez. We also thank Marcela Montemayor and laboratory staff for their contribution, Pablo Sanchez for our database management, and Christina Villarreal for administrative support. We thank Valley Baptist Medical Center, Brownsville for providing us space for our Center for Clinical and Translational Science Clinical Research Unit. We finally thank the community of Brownsville and the participants who so willingly participated in this study in their city. This work was supported by MD000170 P20 funded from the National Institute on Minority Health and Health disparities (NIMHD), and the Centers for Clinical and Translational Science Award 1U54RR023417-01 from the National Center for Research Resources (NCRR).

Author details

[1]Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine, Houston, TX, USA. [2]Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX, USA. [3]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA. [4]University of Texas School of Public Health, Brownsville, TX, USA.

Received: 21 October 2014 Accepted: 3 February 2015
Published online: 06 March 2015

References

1. Cho NG, Whiting D, Guariguata L, Montoya PA, Forouhi N, Hambleton I, et al. The Global Burden. In: Guariguata L, Nolan T, Beagley J, Linnenkamp U, Jacqmain, editors. IDF Diabetes Atlas, Sixth ed. International Diabetes Foundation; 2013. p. 29-49.
2. Mengual L, Roura P, Serra M, Montasell M, Prieto G, Bonet S. Multifactorial control and treatment intensity of type-2 diabetes in primary care settings in Catalonia. Cardiovasc Diabetol. 2010;9:14.
3. Saxena R, Elbers CC, Guo Y, Peter I, Gaunt TR, Mega JL, et al. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. Am J Hum Genet. 2012;90(3):410–25.
4. Chassaing B, Gewirtz AT. Gut microbiota, low-grade inflammation, and metabolic syndrome. Toxicol Pathol. 2014;42(1):49–53.
5. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027–31.
6. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. Nature. 2009;457(7228):480–4.
7. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature. 2013;498(7452):99–103.
8. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature. 2012;490(7418):55–60.
9. Zak-Golab A, Olszanecka-Glinianowicz M, Kocelak P, Chudek J. The role of gut microbiota in the pathogenesis of obesity. Postepy Hig Med Dosw. 2014;68:84–90.
10. Udayappan SD, Hartstra AV, Dallinga-Thie GM, Nieuwdorp M. Intestinal microbiota and fecal transplantation as treatment modality for insulin resistance and type 2 diabetes mellitus. Clin Exp Immunol. 2014;177(1):24–9. doi:10.1111/cei.12293.
11. Nieuwdorp M, Gilijamse PW, Pai N, Kaplan LM. Role of the microbiome in energy regulation and metabolism. Gastroenterology. 2014;146(6):1525–33. doi:10.1053/j.gastro.2014.02.008.
12. Khoruts A, Weingarden AR. Emergence of fecal microbiota transplantation as an approach to repair disrupted microbial gut ecology. Immunology letters. 2014;162(2 Pt A):77–81. doi:10.1016/j.imlet.2014.07.016.
13. Joyce SA, Gahan CG. The gut microbiota and the metabolic health of the host. Curr Opin Gastroenterol. 2014;30(2):120–7.
14. Smits LP, Bouter KE, de Vos WM, Borody TJ, Nieuwdorp M. Therapeutic potential of fecal microbiota transplantation. Gastroenterology. 2013;145(5):946–53.
15. Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, Srinivasan S, et al. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science. 2010;328(5975):228–31.
16. Vrieze A, Van Nood E, Holleman F, Salojarvi J, Kootte RS, Bartelsman JF, et al. Transfer of intestinal microbiota from lean donors increases insulin

sensitivity in individuals with metabolic syndrome. Gastroenterology. 2012;143(4):913–6. e917.
17. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature. 2011;473(7346):174–80.
18. Roager HM, Licht TR, Poulsen SK, Larsen TM, Bahl MI. Microbial enterotypes, inferred by the prevotella-to-bacteroides ratio, remained stable during a 6-month randomized controlled diet intervention with the new nordic diet. Appl Environ Microbiol. 2014;80(3):1142–9.
19. Bergstrom A, Skov TH, Bahl MI, Roager HM, Christensen LB, Ejlerskov KT, et al. Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. Appl Environ Microbiol. 2014;80(9):2889–900. doi:10.1128/AEM. 00342-14.
20. Shen W, Gaskins HR, McIntosh MK. Influence of dietary fat on intestinal microbes, inflammation, barrier function and metabolic outcomes. J Nutr Biochem. 2014;25(3):270–80.
21. Martinez I, Lattimer JM, Hubach KL, Case JA, Yang J, Weber CG, et al. Gut microbiome composition is linked to whole grain-induced immunological improvements. ISME J. 2013;7(2):269–80.
22. Daniel H, Moghaddas Gholami A, Berry D, Desmarchelier C, Hahne H, Loh G, et al. High-fat diet alters gut microbiota physiology in mice. ISME J. 2014;8 (2):295–308.
23. Hooda S, Vester Boler BM, Kerr KR, Dowd SE, Swanson KS. The gut microbiome of kittens is affected by dietary protein:carbohydrate ratio and associated with blood metabolite and hormone concentrations. Brit J Nutr. 2013;109(9):1637–46.
24. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, et al. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. Proc Natl Acad Sci U S A. 2013;110(22):9066–71.
25. Fisher-Hoch SP, Vatcheva KP, Laing ST, Hossain MM, Rahbar MH, Hanis CL, et al. Missed opportunities for diagnosis and treatment of diabetes, hypertension, and hypercholesterolemia in a Mexican American population, Cameron County Hispanic Cohort, 2003-2008. Prev Chronic Disease. 2012;9:110298.
26. Stoddard P, He G, Vijayaraghavan M, Schillinger D. Disparities in undiagnosed diabetes among United States-Mexico border populations. Rev Panam Salud Publica. 2010;28(3):198–206.
27. Diaz-Apodaca BA, Ebrahim S, McCormack V, de Cosio FG, Ruiz-Holguin R. Prevalence of type 2 diabetes and impaired fasting glucose: cross-sectional study of multiethnic adult population at the United States-Mexico border. Rev Panam Salud Publica. 2010;28(3):174–81.
28. Umpierrez GE, Gonzalez A, Umpierrez D, Pimentel D. Diabetes mellitus in the Hispanic/Latino population: an increasing health care challenge in the United States. Am J Med Sci. 2007;334(4):274–82.
29. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ. 2011;343:d7163.
30. Shai I, Jiang R, Manson JE, Stampfer MJ, Willett WC, Colditz GA, et al. Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. Diabetes Care. 2006;29(7):1585–90.
31. van Dam RM. The epidemiology of lifestyle and risk for type 2 diabetes. Eur J Epidemiol. 2003;18(12):1115–25.
32. van Tilburg J, van Haeften TW, Pearson P, Wijmenga C. Defining the genetic contribution of type 2 diabetes mellitus. J Med Genet. 2001;38(9):569–78.
33. Permutt MA, Wasson J, Cox N. Genetic epidemiology of diabetes. J Clin Invest. 2005;115(6):1431–9.
34. World Bank: GDP per capita.(2014) http://data.worldbank.org/indicator/NY.GDP.PCAP.CD accessed 06 July 2014.
35. Fisher-Hoch SP, Rentfro AR, Salinas JJ, Perez A, Brown HS, Reininger BM, et al. Socioeconomic status and prevalence of obesity and diabetes in a Mexican American community, Cameron County, Texas, 2004-2007. Prev Chronic Dis. 2010;7(3):A53.
36. Cowie CC, Rust KF, Byrd-Holt DD, Eberhardt MS, Flegal KM, Engelgau MM, et al. Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: National Health And Nutrition Examination Survey 1999-2002. Diabetes Care. 2006;29(6):1263–8.
37. Choi JY. Prevalence of overweight and obesity among US immigrants: results of the 2003 New Immigrant Survey. J Immigr Minor Health. 2012;14(6):1112–8.
38. NIDDK: Overweight and obesity statistics. (2014) http://win.niddk.nih.gov/statistics/ accessed 25 July 2014
39. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207–14.
40. Human Microbiome Project C. A framework for human microbiome research. Nature. 2012;486(7402):215–21.
41. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449(7164):804–10.
42. Aagaard K, Petrosino J, Keitel W, Watson M, Katancik J, Garcia N, et al. The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. FASEB J. 2013;27(3):1012–22.
43. Martinez I, Wallace G, Zhang C, Legge R, Benson AK, Carr TP, et al. Diet-induced metabolic improvements in a hamster model of hypercholesterolemia are strongly linked to alterations of the gut microbiota. Appl Environ Microbiol. 2009;75(12):4175–84.
44. Lahti L, Salonen A, Kekkonen RA, Salojarvi J, Jalanka-Tuovinen J, Palva A, et al. Associations between the human intestinal microbiota, Lactobacillus rhamnosus GG and serum lipids indicated by integrated analysis of high-throughput profiling data. Peer J. 2013;1:e32.
45. Danielsson H, Gustafsson B. On serum-cholesterol levels and neutral fecal sterols in germ-free rats; bile acids and steroids 59. Arch Biochem Biophys. 1959;83:482–5.
46. Carr TP, Weller CL, Schlegel VL, Cuppett SL, Guderian Jr DM, Johnson KR. Grain sorghum lipid extract reduces cholesterol absorption and plasma non-HDL cholesterol concentration in hamsters. J Nutr. 2005;135(9):2236–40.
47. Matijasic BB, Obermajer T, Lipoglavsek L, Grabnar I, Avgustin G, Rogelj I. Association of dietary type with fecal microbiota in vegetarians and omnivores in Slovenia. European journal of nutrition. 2013;53(4):1051–64. doi:10.1007/s00394-013-0607-6.
48. Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze X, et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 2011;5(2):220–30.
49. Rahat-Rozenbloom S, Fernandes J, Gloor GB, Wolever TM. Evidence for greater production of colonic short chain fatty acids in overweight than lean humans. International journal of obesity. 2014. doi:10.1038/ijo.2014.46.
50. Meehan CJ, Beiko RG. A phylogenomic view of ecological specialization in the lachnospiraceae, a family of digestive tract-associated bacteria. Genome Biol Evol. 2014;6(3):703–13.
51. Raman M, Ahmed I, Gillevet PM, Probert CS, Ratcliffe NM, Smith S, et al. Fecal microbiome and volatile organic compound metabolome in obese humans with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2013;11(7):868–75. e861-863.
52. Colorectal Cancer in Texas: A closer look. (2010) http://www.cprit.state.tx.us/images/uploads/report_colorectal_a_closer_look.pdf accessed 06 July 2014
53. Centers for Disease Control and Prevention: United States Cancer Statistics; Public Information Data; CDC, (2011) http://apps.nccd.cdc.gov/uscs/ accessed 14 January 2015
54. Carrion AF, Ghanta R, Carrasquillo O, Martin P. Chronic liver disease in the Hispanic population of the United States. Clin Gastroenterol Hepatol. 2011;9(10):834–41. quiz e109-810.
55. Zhang H, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, Yu Y, et al. Human gut microbiota in obesity and after gastric bypass. Proc Natl Acad Sci U S A. 2009;106(7):2365–70.
56. Ellekilde M, Krych L, Hansen CH, Hufeldt MR, Dahl K, Hansen LH, et al. Characterization of the gut microbiota in leptin deficient obese mice - correlation to inflammatory and diabetic parameters. Res Vet Sci. 2014;96(2):241–50.
57. Mejia-Leon ME, Petrosino JF, Ajami NJ, Dominguez-Bello MG, de la Barca AM. Fecal microbiota imbalance in Mexican children with type 1 diabetes. Sci Rep. 2014;4:3814.
58. Murri M, Leiva I, Gomez-Zumaquero JM, Tinahones FJ, Cardona F, Soriguer F, et al. Gut microbiota in children with type 1 diabetes differs from that in healthy children: a case-control study. BMC Med. 2013;11:46.
59. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science. 2011;334(6052):105–8.
60. Grzeskowiak L, Collado MC, Mangani C, Maleta K, Laitinen K, Ashorn P, et al. Distinct gut microbiota in southeastern African and northern European infants. J Pediatr Gastroenterol Nutr. 2012;54(6):812–6.
61. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A. 2010;107(33):14691–6.

62. Ferreyra JA, Ng KM, Sonnenburg JL. The enteric two-step: nutritional strategies of bacterial pathogens within the gut. Cell Microbiol. 2014;16(7):993–1003.
63. Vartoukian SR, Palmer RM, Wade WG. Cultivation of a Synergistetes strain representing a previously uncultivated lineage. Environ Microbiol. 2010;12(4):916–28.
64. Morris JJ, Kirkegaard R, Szul MJ, Johnson ZI, Zinser ER. Facilitation of robust growth of Prochlorococcus colonies and dilute liquid cultures by 'helper' heterotrophic bacteria. Appl Environ Microbiol. 2008;74(14):4530–4.
65. Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, et al. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. Proc Natl Acad Sci U S A. 2011;108 Suppl 1:4586–91.
66. Fernandes J, Su W, Rahat-Rozenbloom S, Wolever TM, Comelli EM. Adiposity, gut microbiota and faecal short chain fatty acids are linked in adult humans. Nutr Diab. 2014;4:e121.
67. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222–7.
68. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335–6.
69. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. PLoS Comput Biol. 2012;8(9):e1002687.