The Gut Microbiome of Adults With Type 1 Diabetes and Its Association With the Host Glycemic Control

Diabetes Care 2022;45:1–9 | https://doi.org/10.2337/dc21-1656

OBJECTIVE

Previous studies have demonstrated an association between gut microbiota composition and type 1 diabetes (T1D) pathogenesis. However, little is known about the composition and function of the gut microbiome in adults with longstanding T1D or its association with host glycemic control.

RESEARCH DESIGN AND METHODS

We performed a metagenomic analysis of the gut microbiome obtained from fecal samples of 74 adults with T1D, 14.6 ± 9.6 years following diagnosis, and compared their microbial composition and function to 296 age-matched healthy control subjects (1:4 ratio). We further analyzed the association between microbial taxa and indices of glycemic control derived from continuous glucose monitoring measurements and blood tests and constructed a prediction model that solely takes microbiome features as input to evaluate the discriminative power of microbial composition for distinguishing individuals with T1D from control subjects.

RESULTS

Adults with T1D had a distinct microbial signature that separated them from control subjects when using prediction algorithms on held-out subjects (area under the receiver operating characteristic curve = 0.89 ± 0.03). Linear discriminant analysis showed several bacterial species with significantly higher scores in T1D, including Prevotella copri and Eubacterium siraeum, and species with higher scores in control subjects, including Firmicutes bacterium and Faecalibacterium prausnitzii (P < 0.05, false discovery rate corrected for all). On the functional level, several metabolic pathways were significantly lower in adults with T1D. Several bacterial taxa and metabolic pathways were associated with the host’s glycemic control.

CONCLUSIONS

We identified a distinct gut microbial signature in adults with longstanding T1D and associations between microbial taxa, metabolic pathways, and glycemic control indices. Additional mechanistic studies are needed to identify the role of these bacteria for potential therapeutic strategies.

Type 1 diabetes (T1D) is a common chronic disease in children and adolescents. The incidence of T1D has been rapidly rising in the past decade, especially in young
children (1). While a genetic predisposition for T1D exists, this rapid increase in the prevalence of the disease and the fact that <10% of genetically susceptible individuals will eventually develop T1D are suggestive of a large contribution of environmental factors to disease pathogenesis. These may include viral infections and nutritional factors (2). Gut microbiota composition has also been highlighted as a possible risk factor, with several studies in humans and animal models implicating its potential role in disease pathogenesis (3–5). These observations have further led to the “balanced signal” hypothesis, stating that microbiome composition may promote or inhibit T1D development (6). Several suggested mechanisms for the possible influence of the gut microbiome on T1D pathogenesis include immunological deregulation mediated by gut dysbiosis, as there is evidence that the microbiome plays an important role in the development and maturation of the immune system (7), and gut leakiness, as structural mucosal alterations and gut dysfunction was observed in both human and animal studies on T1D (8).

In recent years, a rapidly growing number of studies have investigated the role of the gut microbiome in T1D (9). However, most studies focused on disease pathogenesis, while only a few studies thus far have investigated the microbiome composition of individuals with a longstanding diagnosis, and those were mostly conducted on small cohorts and used a variety of computational analysis methods (10,11). In addition, while evidence on the regulating roles of the microbiome in normal and impaired glycemic response is accumulating in both animal models and humans (12), little is known on the role of the microbiome in glycemic control in individuals with longstanding T1D. Here, we analyzed microbial composition and function in a cohort of individuals with T1D who were at least 1 year following diagnosis and the associations between microbial taxa, functional pathways, and glycemic indices in individuals with T1D.

RESEARCH DESIGN AND METHODS

Study Design

We conducted a prospective clinical cohort originally designed to study the postprandial glycemic responses (PPGRs) of individuals with T1D. Full details on recruitment and the study protocols are specified in a companion paper by Shilo et al. (13), focused exclusively on modeling the PPGR in individuals with T1D. In brief, on the first day of the study, participants were invited to a study initiation meeting at the medical center. In this meeting, a physician-authorized participation and acquired informed consent, anthropometric measurements were obtained, and blood tests were drawn and analyzed in the hospital’s laboratories. Health and lifestyle questionnaires were completed by the participants. Throughout the 2 weeks of study participation, participants used a proprietary smartphone app (www.personalnutrition.org), to log, in real-time, food intake, sleep times, physical activity, and medication intake with the exception of insulin, which was recorded in the continuous subcutaneous insulin infusion devices. Participants were asked to follow their normal routine and dietary habits, with the exception of seven standardized meals. Participants were asked to provide one microbiome sample collected during the 2 weeks of study participation.

Participant Recruitment

Enrollment and recruitment were conducted in three medical centers in Israel between March 2017 and April 2019 (Fig. 1). The inclusion criteria for the study included age between 3 and 70 years old (13). However, as previous studies demonstrated that the interpersonal variation in the composition of the bacterial communities is significantly greater among children (14) and as a large variation exists in clinical phenotypes between children and adults (15), we choose to include only adults (18–70 years old) in the analyses presented here. Additional inclusion criteria were >1 year following T1D diagnosis, using continuous glucose monitoring (CGM) and continuous subcutaneous insulin infusion devices simultaneously, and a capability to work with a mobile phone app daily for the recording of the dietary intake. Exclusion criteria included an active inflammatory or neoplastic disease, pregnancy, and antibiotic use 3 months prior to participation in the study. Participants who reported a diagnosis of celiac disease were excluded from all microbiome analyses since several studies previously showed that celiac disease is correlated with a change in gut microbial composition (16).

Study Population

Overall, 142 individuals with T1D were recruited to the study, and 124 participants provided a stool sample. Seven participants reported a diagnosis of celiac disease and were therefore excluded from microbiome analyses, resulting in 117 individuals. From them, 74 were >18 years and were included in the analyses (Fig. 1). The average age was 32.3 ± 14.4 years (median 26 [interquartile range 21–43] years), and average disease duration was 14.6 ± 9.6 years (median 12 [interquartile range 7.8–18.3] years). Mean HbA1c level was 7.3 ± 1% (56.3 ± 10.9 mmol/mol) (see Supplementary Table 1 for mean values of all blood test results at study initiation). The mean BMI value was 25.1 ± 4 kg/m². Of the 74 participants, 33 (44.6%) had at least one additional comorbidity. The most common comorbidities were hypothyroidism (12 participants [16.22%]) and hyperlipidemia (10 participants [13.51%]). Thirty-nine participants (52.7%) consumed additional medications apart from insulin during the study. The most common medications were levothyroxine (12 participants [16.22%]), oral contraceptives (8 participants, 10.21%), and antilipemic drugs (8 participants, 10.21%) (see Supplementary Table 2 for a full list of medical conditions and medications consumed by the participants during the study). Cohort characteristics are presented in Table 1. Of the 74 individuals, 73 logged meals in real-time during the 2 weeks of study participation (see a companion paper by Shilo et al. [13]). Total energy intake was 1,666,610 kcal (22,830 per person). Average carbohydrate, fat, and protein consumption was 43 ± 1%, 38 ± 7%, and 17 ± 4% from the total energy, respectively.

Cohort Matching

To compare between the composition of the microbiome in adults with T1D and healthy adults, we used gut metagenomic profiles obtained from Israeli adults (17). Of 35,304 Israeli adults who submitted their sample between 13 January 2017 and 1 May 2021, 14,012 were excluded due to a different sequencing
method, and 13,295 were excluded due to the presence of one of the following metabolic, gastrointestinal, or systemic diseases: type 2 diabetes, T1D, gestational diabetes, prediabetes, impaired glucose tolerance or impaired fasting glucose, metabolic syndrome, fatty liver disease, morbid obesity, inflammatory bowel disease, Crohn disease, ulcerative colitis, undetermined colitis, pancreatic diseases, celiac disease, irritable bowel syndrome, diverticulosis, hepatitis or other liver diseases, cholangitis or other bile-related disease, HIV, autoimmune disease, and cancer.

**Stool Sample Collection and Genomic DNA Extraction**

Participants entering the study received a verbal explanation from the study coordinators and detailed printed instructions for stool collection. Microbiome sampling was done using a swab and an OMNIlgene-GUT (OMR-200; DNA Genotek) stool collection kit. Each participant was requested to collect stool via one swab and one separate OMNIlgene-GUT kit. However, only samples collected by OMNIlgene-GUT kits were sequenced and analyzed since it has the advantage of maintaining DNA integrity in typical ambient temperature fluctuations and since samples of the control group were collected only by the OMNIlgene-GUT kits. Collected samples were immediately stored in a home freezer (−18°C) and transferred in a provided cooler to our facilities where they were stored at −80°C (−20°C for OMNIlgene-GUT kits) until DNA extraction. Samples from adults with T1D were sequenced between April 2019 and August 2019, and samples from healthy control subjects were sequenced between April 2019 and May 2021. All samples analyzed in this study were sequenced using the same sequencing methods, including sequencing protocols of DNA extraction, library preparation, and sequencing machine. Control samples demonstrated that performing the process on different days had no effect on the results when the sequencing protocols were kept the same.

Metagenomic DNA was purified using MagAttract PowerSoil DNA extraction kit (QIAGEN) optimized for the Tecan automated platform. Next-generation sequencing libraries were prepared using Nextera DNA library prep (Illumina) and sequenced on a NovaSeq sequencing platform (Illumina). Sequencing was performed with 100 base pair single end reads with a depth of 10 million reads per sample. We filtered metagenomic reads containing Illumina adapters, filtered low-quality reads, and trimmed low-quality read edges. We detected host DNA by mapping with Bowtie 2 (19) to the human genome with inclusive parameters and removed those reads. Bacterial relative abundance (RA) estimation was performed by mapping bacterial reads to species-level genome bins (SGB) representative genomes (20). We selected all SGB representatives with at least five genomes in a group, and for these representative genomes kept unique regions as a reference data set. Mapping was performed using Bowtie 2 (19), and abundance was estimated by calculating the mean coverage of unique genomic regions across the 50% most densely covered areas, as previously described (21). Feature names include the lowest taxonomy level identified. In addition,
we also estimated the RA of bacterial groups, such as Akkermansia, Alistipes, Roseburia, Eubacterium, and Faecalibacterium prausnitzii as a summation of the abundances of SGBs belonging to the relevant species by National Center for Biotechnology Information classification.

Microbial Biodiversity Indices and Functional Analysis
Microbiome α-diversity was calculated by the Shannon diversity index. Richness was calculated as a number of species in the sample detected with an abundance of at least 1e-4. Comparison between microbial indices and RA of microbial taxa were performed using Mann-Whitney U test. HUMANv2 v2.8.2 (22) was used to integrate taxonomic information with functional profiles.

Associations With Clinical Phenotypes
We used several indices to analyze the association between clinical and microbial features and measures of glycemic control. These included fasting glucose, HbA1c level, and lipids measured by a blood test at study initiation and indices calculated based on CGM measurements during the 2 weeks of study participation, available for 73 participants. CGM-derived features included the percentage of the time spent in hypoglycemia and hyperglycemia defined as glucose values <70 mg/dL (3.9 mmol/L) and >180 mg/dL (10 mmol/L), respectively, time in range, defined as time spent in glucose values between 70 and 180 mg/dL (3.9–10 mmol/L) (23), and coefficient of variation (CV) as a measure of glucose variability (24). For the 73 participants who also logged meals throughout the study period, PPGRs were calculated (see in a companion paper by Shilo et al. [13]). Pearson correlations between the clinical phenotypes, RA converted to a log space of microbial taxa, and metabolic pathways were calculated.

T1D Prediction Model Based on Microbial Features
To evaluate the discriminative power of microbial composition for T1D, we constructed a prediction model based on XGBoost (25), which solely takes microbiome features as inputs. This model can capture nonlinear interactions between bacteria and was previously shown to outperform other methods for the classification of human microbiome data (26). The mean and SD of the receiver operating characteristic curve were computed by using the curves that were generated in fivefold cross-validation. In addition, we verified that when randomly swapping the target labels, the performances reflected a random prediction, hence an area under the curve (AUC) very close to 0.5, as an additional control. We analyzed feature attributes using SHAP (Shapley Additive exPlanation) to explore model interpretability. SHAP values represent the average change in the model’s output upon conditioning on a specific feature (27).
associations were found between nutritional parameters and bacterial taxa.

**Microbiome Composition in Individuals With T1D**

To improve our understanding of the composition of the gut microbiome in T1D, individuals with T1D were compared with healthy controls (1:4 matching by age, see Research Design and Methods). Overall, 74 adults with T1D were compared with 296 healthy adults. There were no statistically significant differences in sex, weight, or BMI between groups. As expected, healthy adults had significantly lower levels of HbA1c (Table 1). Microbial α-diversity was not significantly different between the groups (Fig. 3C), aligned with previous studies (11,28) but in contrast with others, reporting a lower diversity in individuals with T1D (29,30). In addition, species richness and the ratio of *Firmicutes-to-Bacteroidetes* of taxonomic profiles were not significantly different between the groups (Supplementary Table 3). Linear discriminant analysis showed a total of 17 bacterial taxa with significantly higher LDA scores in individuals with T1D and 15 bacterial taxa with significantly higher LDA scores in healthy adults (Fig. 3A and D and Supplementary Table 4). Bacterial species with significantly higher scores in individuals with T1D included *Prevotella copri*, *Eubacterium siraeum*, and *Alistipes inops*, and several species with a higher score in healthy adults, including *Firmicutes bacterium*, *Alistipes putredinis*, *Faecalibacterium prausnitzii*, and *Ruminococcus gnavus* ($P < 0.05$, FDR corrected). Dimensionality reduction techniques, including principal component analysis (PCA), in which the principal coordinate combination with the greatest contribution rate was PC1 = 7.7%, PC2 = 4.1%, and t-distributed stochastic neighbor embedding...
(t-SNE), did not reveal visually distinctive differences between individuals with T1D and control subjects (Supplementary Fig. 2). On the functional level, when comparing metabolic pathways, several metabolic pathways, including \( \gamma \)-glutamate and \( \gamma \)-glutamine biosynthesis, \( \gamma \)-ornithine de novo biosynthesis, and superpathway of hexuronide and hexuronate degradation,
were significantly lower in adults with T1D (P < 0.05, FDR corrected).

Classification of Individuals With T1D by Microbial Features
We next analyzed our ability to distinguish individuals with T1D from control subjects based solely on microbiome features. We constructed a prediction model based solely on microbial features and used cross-validation schemes for validation of the model (see Research Design and Methods). The discrimination performance of the model had an AUC of 0.89 ± 0.03 and permutations P < 0.001 (Fig. 3B). The most impactful microbial taxa for the prediction were *Prevotella copri*, which impacted the model toward the prediction of T1D, and *Ruminococcus*, which impacted the model toward the prediction of a healthy state (Supplementary Fig. 3).

CONCLUSIONS
In this study, we profiled the gut microbiome composition in adults with longstanding T1D and identified several associations between bacterial taxa, metabolic pathways, and the glycemic control of the host. While a growing body of evidence, mainly originating from studies on animal models, suggests that gut microbiota has a causal impact on host glycemic control (31), that may be mediated by mechanisms such as modulation of incretin secretion, short-chain fatty acid production, metabolism of bile acid, and regulation of adipose tissue (32), data regarding the role of the microbiome in the glycemic control of individuals with T1D are still sparse.

Here, bacterial taxa and metabolic pathways that were significantly associated with glycemic indices of the host included *Enterobacterales* species and pathways relating to aromatic acid and chorismate biosynthesis, which were correlated with glucose average, *Prevotellaceae* species that were inversely correlated with HbA1c level, and the pyrimidine nucleobases salvage pathways that were inversely correlated with glucose average (P < 0.05, FDR corrected for all). Several small-scale studies previously showed different associations, including a study of 12 Chinese subjects with T1D (33) that demonstrated an inverse correlation of the abundance of *Faecalibacterium* and HbA1c levels, and a study conducted in Brazil that included 20 individuals with T1D and demonstrated a correlation between the relative abundances of *Bacteroidetes*, *Lactobacillales*, and *Bacteroides dorei* and HbA1c levels (34). Importantly, the correlations observed in this study were not strong, and further studies integrating multomic data, including metagenomic, metatranscriptomic, and metaproteomic, along with high-quality clinical and nutritional data, are needed in order to identify the potential role of these bacteria and metabolic pathways and their influence on the host’s glycemic control.

We identified a distinct gut microbial signature in adults with longstanding T1D compared with healthy adults. By using an expanded reference set (20) for the first time in individuals with T1D, as well as a relatively large control group, we show a total of 17 bacterial taxa with significantly higher LDA scores in T1D and 15 bacterial taxa with significantly higher LDA scores in control subjects (Fig. 3). Although dimensionality reduction analyses did not reveal visually distinctive differences (Supplementary Fig. 2) and the diversity and richness were not statistically different between groups, we were able to devise a model that accurately distinguishes between adults with T1D and healthy control subjects using only microbiome features (AUC = 0.89 ± 0.03) (Fig. 3B). Interestingly, the most impactful microbial taxa for the prediction were *Prevotella copri*, which impacted the model toward the prediction of T1D, and *Ruminococcus*, which impacted the model toward the prediction of a healthy state (Supplementary Fig. 2), aligned with the results of the LDA analysis, showing higher scores for *Prevotella copri* in T1D and *Ruminococcus gravis* in healthy adults (Fig. 3).

Previous studies (9) reported various results regarding the taxonomic composition of the gut microbiome in individuals with T1D compared with healthy control subjects and their interpretation is challenging due to a large heterogeneity in both study population and analytic approaches. It is also worthy to note that gut microbiota in T1D was previously shown to differ at taxonomic and functional levels compared not only with healthy subjects but also with nonautoimmune diabetes (35). The most common findings in individuals with T1D included alterations in the following bacterial species: *Bacteroides*, *Streptococcus*, *Clostridium*, *Bifidobacterium*, *Prevotella*, *Staphylococcus*, *Blautia*, *Faecalibacterium*, *Roseburia*, and *Lactobacillus* (36). In the largest human cohort to date, no particular taxon was associated with the T1D development, but the microbiome of control children was found to contain more genes related to fermentation and biosynthesis of short-chain fatty acid (SCFA) compared with children who eventually developed T1D (4). An additional study also reported a decrease in SCFA producers in individuals with longstanding T1D. Moreover, it was previously shown that feeding NOD mice with SCFA-rich (butyrate and acetate) diets had substantial effects on their immune system and a protective effect from the development of diabetes (37). In this cohort, when comparing metabolic pathways, we found several metabolic pathways, including L-glutamate and L-glutamine biosynthesis, l-ornithine de novo biosynthesis, and superpathway of hexuronide and hexuronate degradation, that were significantly lower in adults with T1D (P < 0.05, FDR corrected). To the best of our knowledge, these findings have not been previously described in individuals with T1D, and their role should be further explored in future work.

The strength of our study includes a relatively large sample size compared with previous studies, the integration of data on glucose measurements obtained from CGM devices, and the expanded reference set we used. The greatest limitation of our study is its observational nature. Further studies are needed in order to attribute causality to the gut microbiome alterations we describe, as currently, whether these taxa are a cause or an effect of the disease remains unclear. In addition, although the sample size of the cohort is relatively large, it may still be insufficient to reach robust associations with clinical phenotyping. Finally, several additional factors may influence the composition of the gut microbiome. For example, nutritional habits may differ between individuals with T1D compared with healthy individuals. While we did not have detailed nutritional data on our control group, macro-nutrient distribution in the T1D cohort...
was very similar to healthy adults in Israel as measured in a previous study performed by our group (38). In this study, healthy individuals logged meals during 1 week and consumed an average of 46 ± 8% carbohydrate, 36 ± 7% fat, and 15 ± 3% protein from the total energy, compared with an average of 43 ± 1% carbohydrate, 38 ± 7% fat, and 17 ± 4% protein consumed by the T1D cohort. Moreover, in the group of individuals with T1D, no associations were found between nutritional parameters and bacterial taxa. Medication consumption may also influence microbial composition, and we therefore excluded individuals with antibiotic use 3 months prior to participation. While other types of medications, such as proton pump inhibitors, may also have an effect (39), they were only consumed by a very small percentage of our cohort (Supplementary Table 2). Family kindred may also have a pronounced effect on the structural and functional composition of the gut microbiome (40). However, none of the adults with T1D included in this study were family members sharing the same household. Microbiome composition is also heavily influenced by geographic location (14), and therefore, additional studies are needed in order to determine whether our findings can be generalized to non-Israeli populations.

In conclusion, our study highlights a distinct gut microbial composition in individuals with longstanding T1D compared with healthy individuals. We identified unknown associations between microbial taxa, metabolic pathways, and clinical phenotypes and note the importance of expanding the gut microbiome reference set, as it allows us to also identify associations with unclassified bacterial strains that may play a part in disease pathogenesis. Our findings provide a foundation for additional large-scale analyses of the gut microbiome in individuals with T1D in order to identify host–microbe interactions and to identify the causal role of these bacterial taxa for the development of novel therapeutic strategies in T1D.

Acknowledgments. The authors thank the Segal group members for fruitful discussions.

Funding. This work is supported by the Israel Science Foundation (ISF) (grant no. 3-14762). E.S. is supported by the Crown Human Genome Center, Larson Charitable Foundation New Scientist Fund, Else Kroener Fresenius Foundation, White Rose International Foundation, Ben B. and Joyce E. Eisenberg Foundation, Nissenbaum Family, Marcos Pinheiro de Andrade and Vanessa Buchheim, Lady Michelle Michels, Aliza Moussaiif, and grants funded by the Minerva foundation with funding from the Federal German Ministry for Education and Research and by the European Research Council and the Israel Science Foundation.

These funding sources had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Duality of Interest. E.S. is a paid consultant for DayTwo. No other potential conflicts of interest relevant to this article were reported. No pharmaceutical manufacturers or companies from the industry contributed to the planning, design, or conduct of the trial.

Author Contributions. S.S. and A.G. conceived the project, designed and conducted the analyses, interpreted the results, and wrote the manuscript. M.R. provided data and interpreted the results. T.K. conceived the project, designed the analysis, and interpreted the results. Y.B. designed the analysis and interpreted the results. D.K. and T.K. designed and conducted the analyses. N.B. designed the analysis and interpreted the results. B.C.W. and Y.G.-G. coordinated and designed data collection. M.C., N.Z.L., N.S., N.G., N.L., and S.K. provided data and interpreted the results. A.W. conceived the project and directed sample sequencing. O.P.H. and E.S. conceived the project, designed and conducted the analyses, interpreted the results, and supervised the project and analyses. All authors reviewed and approved the manuscript and vouch for the accuracy and completeness of the data. E.S. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References