Early Childhood Gut Microbiomes Show Strong Geographic Differences Among Subjects at High Risk for Type 1 Diabetes

DOI: 10.2337/dc14-0850

OBJECTIVE
Gut microbiome dysbiosis is associated with numerous diseases, including type 1 diabetes. This pilot study determines how geographical location affects the microbiome of infants at high risk for type 1 diabetes in a population of homogenous HLA class II genotypes.

RESEARCH DESIGN AND METHODS
High-throughput 16S rRNA sequencing was performed on stool samples collected from 90 high-risk, nonautoimmune infants participating in The Environmental Determinants of Diabetes in the Young (TEDDY) study in the U.S., Germany, Sweden, and Finland.

RESULTS
Study site-specific patterns of gut colonization share characteristics across continents. Finland and Colorado have a significantly lower bacterial diversity, while Sweden and Washington state are dominated by *Bifidobacterium* in early life. Bacterial community diversity over time is significantly different by geographical location.

CONCLUSIONS
The microbiome of high-risk infants is associated with geographical location. Future studies aiming to identify the microbiome disease phenotype need to carefully consider the geographical origin of subjects.

The Environmental Determinants of Diabetes in the Young (TEDDY) study was formed to investigate environmental factors that trigger type 1 diabetes in genetically at-risk children (1). The gut microbiome is of interest, as several studies (2–6) have shown that dysbiosis of the microbiome is associated with type 1 diabetes autoimmunity. The composition of the fecal microbiome is dependent on numerous external factors, including geographical location (7–11). This work presents the first geographical assessment of the gut microbiome in these genetically higher risk children.

RESEARCH DESIGN AND METHODS
The TEDDY study prospectively observes children at six clinical centers in Europe (Finland, Sweden, and Germany) and the U.S. (Colorado, Washington state, and Georgia/Florida) (1). A total of 1,129 stool samples from 90 children, 15 from Finland and Colorado, 30 from Sweden and Washington state, and 45 from Germany and Florida, were analyzed. The sequences were compared for bacterial diversity and ecological profiles. The findings indicate that geographical location significantly impacts the early gut microbiome.

**Kaisa M. Kemppainen,1 Alexandria N. Ardissone,2 Austin G. Davis-Richardson,1 Jennie R. Fagen,1 Kelsey A. Gano,1 Luis G. León-Novelo,7 Kendra Vehik,3 George Casella,4* Olli Simell,5 Anette G. Ziegler,7 Marian J. Rewers,7 Åke Lernmark,4 William Hagopian,9 Jin-Xiong She,10 Jeffrey P. Krischer,3 Beena Akolkar,11 Desmond A. Schatz,12 Mark A. Atkinson,13 and Eric W. Triplett,1 for the TEDDY Study Group**

1Department of Microbiology and Cell Science, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL
2Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA
3Health Informatics Institute, University of South Florida, Tampa, FL
4Department of Statistics, University of Florida, Gainesville, FL
5Department of Pediatrics, Turku University Hospital, University of Turku, Turku, Finland
6Institute of Diabetes Research, Helmholtz Zentrum München, and Klinikum Rechts der Isar, Technische Universität München, and Forschergruppe Diabetes e.V., Neuherberg, Germany
7Barbara Davis Center for Diabetes, School of Medicine, University of Colorado, Aurora, CO
8Department of Clinical Sciences, Lund University/Clinical Research Center, Skåne University Hospital, Malmö, Sweden
9Pacific Northwest Diabetes Research Institute, Seattle, WA
10Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Georgia Regents University, Augusta, GA
11National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD
12Department of Pediatrics, College of Medicine, University of Florida, Gainesville, FL
13Department of Pathology, Immunology, and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL

*Corresponding author: Eric W. Triplett, ewt@ufl.edu.

Received 4 April 2014 and accepted 22 October 2014.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc14-0850/-/DC1.

*A complete list of the TEDDY Study Group can be found in the Supplementary Data online.

†Deceased June 17, 2012

© 2014 by the American Diabetes Association.

Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered.
each study site, were collected monthly starting, on average, at 151.1 days after birth (SE 5.5 days after birth) until the average last sampling at 537 days after birth (SE 4.5 days after birth). Samples were collected at home and mailed to a TEDDY repository within 72 h, with ice packs during the summer months (1,12). Fecal sample storage at room temperature for up to 72 h does not affect bacterial composition by >10% (13). Subjects were determined to have the highest risk HLA class II genotype (DR4-DQA1*0301-DQB1*0302/DR3-DQA1*0501-DQB1*0201) by genotyping of cord blood (1), but neither autointi- bodies nor disease developed during the sample collection period. Clinical data were collected on gestational age, delivery mode, sex, and early feeding practices (age at first introduction to formula, and duration of exclusive and any breastfeeding), and later on diet (age at first introduction to oats, gluten, milk products, cow milk, and solid food) (1).

DNA was isolated from frozen stool samples as previously described (13). Extracted DNA was purified using the PowerClean DNA Kit (MO BIO Laboratories, Inc., Carlsbad, CA). 16S rRNA amplification, sequencing using a barcoded Illumina approach, sequence analysis, read trimming, and taxonomic classification were performed as previously described (14).

Samples with <10,000 reads and any operational taxonomic units with <50 reads in at least one sample were removed from the data set. This resulted in an average of 102,147 reads per sample (SE 1,151 reads per sample), of which on average 43.9% (SE 0.1%) were successfully classified at the genus level. The relative abundances of bacterial genera were calculated as the percentage of classified reads. Sequences that did not map to known genera were clustered to each other at 95% similarity. The original sequences were submitted to MG-RAST under project identification #3229. The bacterial diversity of each sample was determined by calculating the Shannon diversity index (SDI).

Data analysis was performed using R statistical software version 3.0.0 (15) or SAS version 9.3 (SAS Institute Inc., Cary, NC). Demographic, clinical, and dietary variables were assessed by site. Categorical variables were analyzed using Pearson $\chi^2$ test or Fisher exact test. Continuous variables were tested using the one-way ANOVA or Kruskal-Wallis test for differences in means. Generalized estimating equations for longitudinal correlated data were used to assess the association between geographical location and bacterial abundance and diversity adjusting for demographic, clinical, and dietary variables. Separate models were examined for each bacterial genus under study. A permutation test and the $F$ statistic were used to determine whether SDI differed among the six TEDDY study sites, as previously described (16). $P$ values <0.05 were considered significant.

RESULTS

Sex ($P = 0.0092$) and age at first introduction to oats ($P = 0.0042$), gluten ($P = 0.0001$), and milk products ($P < 0.0001$) were the only clinical characteristics significantly different by geographical location. To characterize the development of the gut microbiome over time, 16S rRNA sequencing read values of bacterial genera were grouped according to age of subject (in months) at the time of sample collection (Fig. 1A). Bacteroides was the predominant genera at all sites (average abundance 22.7%, SE 0.7%). The abundance of Bifidobacterium ($P = 0.0172$), Veillonella ($P = 0.0048$), Faecalibacterium ($P = 0.0122$), Streptococcus ($P = 0.0003$), and Akkermansia ($P = 0.0196$) was significantly different by geographical location after adjusting for significant clinical and dietary variables. Although Bacteroides abundance was not significantly different ($P = 0.0530$) by site, Colorado had a significantly higher abundance than all other sites ($P = 0.0126$), except Finland.

The permutation test of the SDI of bacterial genera identified at each site showed that all sites differ from each other across time (Fig. 1B). Furthermore, the difference remained significant after adjustment for delivery mode and age at first introduction to milk products ($P = 0.0045$). Colorado and Finland had a significantly lower SDI than all other sites ($P = 0.0258$). Georgia/Florida and Germany had a more diverse profile, characterized by a relative abundance of Clostridium, Bifidobacterium, and Veillonella of >8.0% each (Fig. 1A). In Sweden and Washington state, the fecal profile was dominated by Bifidobacterium until 8 and 10 months of age, respectively, and the overall abundance was significantly higher than in Colorado and Finland ($P = 0.0199$).

CONCLUSIONS

This study highlights the great variability in the composition and diversity of gut microbiomes among the six TEDDY study sites. So far, studies examining the association between the gut microbiome and type 1 diabetes autoimmunity have focused on small geographical regions (2–6). Country- and lifestyle-specific factors are a major player in shaping the composition of the gut community (7–11), but the extent to which they affect the microbiome of infants at high risk for type 1 diabetes was previously unknown. Our study is the first to investigate the fecal microbial profile of high-risk children across two continents and multiple countries.

Our data suggest that children at high risk for type 1 diabetes have study site-specific patterns of gut colonization showing intercontinental similarities but intracontinental differences. Geographical origin significantly associated with the diversity of bacterial communities and the relative abundance of numerous bacterial genera (Fig. 1A). Low bacterial diversity was characteristic of subjects from Finland and Colorado. Previous studies have reported similar reductions in microbiome diversity in northern European infants compared with infants from southern European countries (Sweden vs. Spain and Finland vs. Germany, respectively) (9,10). While Finland has the highest incidence of type 1 diabetes in children (17), it remains to be seen whether SDI is associated with disease incidence. Perplexing differences, not explained by clinical characteristics, exist between the neighboring countries Sweden and Finland. Compared with their Swedish neighbors, Finnish subjects had a significantly higher abundance of Bacteroides ($P = 0.0508$) and Veillonella ($P = 0.0160$), and a lower abundance of Bifidobacteria ($P = 0.0199$), Akkermansia ($P = 0.0014$), and Ruminococcus ($P = 0.0248$). The fecal profile of Swedish subjects was more similar to those from Washington state (both groups of subjects were dominated by Bifidobacteria at early time points), suggesting a
are different and cultural factors that, on their own, determine the underlying environmental genotypes. Geography represents a cultural effect that could result in a population of homogenous HLA class II genotypes. With geographical location, even in a small infant gut (18).

more typical colonization pattern of the infant gut (18).

Clearly, microbiome diversity varies with geographical location, even in a population of homogenous HLA class II genotypes. Geography represents a cul-

mains unknown because significant differences in microbiome composition exist even after adjusting for numerous early-life and dietary variables. Future studies must carefully consider the impact of geographical location on the microbiome of children who are genetically at higher risk for type 1 diabetes as location may confound analyses of disease-associated microbiome states. Furthermore, whether the microbiome is causative or merely an indicator of underlying type 1 diabetes etiology, these microbiome differences suggest a geographically tailored approach to diagnostics or preventative therapies.

Funding. This study was funded by grants U01-DK-63829, U01-DK-63861, U01-DK-63821, U01-DK-63865, U01-DK-63836, U01-DK-63790, UC4-DK-63829, U4-DK-63861, U4-DK-63821, UC4-DK-63865, U4-DK-63863, UC4-DK-63864, and UC4-DK-95300, and Contract No. HHSN27220070001C from the National Institute of Diabetes and Digestive and Kidney Disease; the National Institute of Allergy and Infectious Diseases; the National Institute of Child Health and Human Development; the National Institute of Environmental Health Sciences; JDRF; and the Centers for Disease Control and Prevention. This work was supported in part by National Institutes of Health/National Center for Advancing Translational Sciences Clinical and Translational Science Awards to the University of Florida (UL1-TR-000064) and the University of Colorado (UL1-TR-001082).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. K.M.K. participated in the study design, carried out data analysis, drafted the manuscript, and read and approved the final manuscript. A.N.A. participated in the study conception and design, carried out data analysis, helped draft the manuscript, and read and approved the final manuscript. A.G.D.-R. and J.R.F. participated in the study design, carried out data analysis, and read and approved the final manuscript. K.A.G. participated in the study design, helped to draft the manuscript, and read and approved the final manuscript. K.V. participated in the study design and data analysis, and read and approved the final manuscript. G.C. participated in the study design and data analysis. O.S., A.G.Z., M.J.R., A.L., and M.A.A. conceived of the study, participated in the study design and acquisition of data, and read and approved the final manuscript. E.W.T. conceived of the study; participated in its design, data acquisition, and data analysis; helped to draft the manuscript; and read and approved the final manuscript. E.W.T. is the guarantor of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References