Plasma Metabolome and Circulating Vitamins Stratified Onset Age of an Initial Islet Autoantibody and Progression to Type 1 Diabetes: The TEDDY Study

Qian Li,[1] Xiang Liu,[1] Jimin Yang,[1] Iris Erlund,[2] Åke Lernmark,[3] William Hagopian,[4] Marian Rewers,[5] Jin-Xiong She,[6] Jorma Toppari,[7,8] Anette-G. Ziegler,[9,10,11] Beena Akolkar,[12] Jeffrey P. Krischer,[1] and the TEDDY Study Group*

Diabetes 2021;70:282–292 | https://doi.org/10.2337/db20-0696

Children's plasma metabolome, especially lipidome, reflects gene regulation and dietary exposures, heralding the development of islet autoantibodies (IA) and type 1 diabetes (T1D). The Environmental Determinants of Diabetes in the Young (TEDDY) study enrolled 8,676 newborns by screening of HLA-DR-DQ genotypes at six clinical centers in four countries, profiled metabolome, and measured concentrations of ascorbic acid, 25-hydroxyvitamin D [25(OH)D], and erythrocyte membrane fatty acids following birth until IA seroconversion under a nested case-control design. We grouped children having an initial autoantibody only against insulin (IAA-first) or GAD (GADA-first) by unsupervised clustering of temporal lipidome, identifying a subgroup of children having early onset of each initial autoantibody, i.e., IAA-first by 12 months and GADA-first by 21 months, consistent with population-wide early seroconversion age. Differential analysis showed that infants having reduced plasma ascorbic acid and cholesterol experienced IAA-first earlier, while early onset of GADA-first was preceded by reduced sphingomyelins at infancy. Plasma 25(OH)D prior to either autoantibody was lower in T1D progressors compared with nonprogressors, with simultaneous lower diglycerides, lysophosphatidylcholines, triglycerides, and alanine before GADA-first. Plasma ascorbic acid and 25(OH)D at infancy were lower in HLA-DR3/DR4 children among IA case subjects but not in matched control subjects, implying gene expression dysregulation of circulating vitamins as latent signals for IA or T1D progression.

Type 1 diabetes (T1D) is an autoimmune disease developing in young children due to the interaction between genetic risks and environmental exposures, with appearance of β-cell autoantibodies or islet autoantibodies (IA) as the established precursors. The population-wide early onset of distinct initial autoantibodies had been reported at different ages throughout early childhood, i.e., an autoantibody against insulin (IAA-first) by age of 15 months and an autoantibody against GAD (GADA-first) before 24 months (1). Distinct type or onset age of an initial autoantibody can be a result of either genetic variation or immune response to early environmental exposures.

IAA-first was found to be related to HLA DR4-DQ8 and the *INS* gene polymorphism, while GADA-first was related

[1]Health Informatics Institute, University of South Florida, Tampa, FL
[2]Department of Government Services, Finnish Institute for Health and Welfare, Helsinki, Finland
[3]Department of Clinical Sciences, Clinical Research Centre, Skåne University Hospital, Lund University, Malmö, Sweden
[4]Pacific Northwest Research Institute, Seattle, WA
[5]Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, Aurora, CO
[6]Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta University, Augusta, GA
[7]Department of Pediatrics, Turku University Hospital, Turku, Finland
[8]Department of Physiology, University of Turku, Turku, Finland
[9]Institute of Diabetes Research, Helmholtz Zentrum München, Munich, Germany
[10]Forschergruppe Diabetes, Technical University of Munich, Klinikum Rechts der Isar, Munich, Germany
[11]Forschergruppe Diabetes e.V. at Helmholtz Zentrum München, Munich, Germany
[12]National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

Corresponding author: Qian Li, qian.li@epi.usf.edu

Received 1 July 2020 and accepted 20 October 2020

This article contains supplementary material online at https://doi.org/10.2337/figshare.13122524.

Q.L. and X.L. contributed equally.

*A complete list of the TEDDY Study Group can be found in supplementary material online.

to HLA DR3-DQ2 polymorphisms in the *ERBB3*, *SH2B3*, and *BACH2* gene polymorphisms (1). Age of IA seroconversion was recently found to be associated with progression afterward to T1D (2), regardless of the type of initial autoantibodies. Furthermore, nutrients or dietary pattern had been found to be one of the critical exposure-related factors predictive of IA or T1D in early childhood (3,4). Recent studies in T1D have reported the impact of diet or blood concentrations of nutrient biomarkers on disease outcomes, such as infant formula type (5), gluten intake amount (6), vitamin C (ascorbic acid) (7), vitamin D (8), probiotics (9) and n-3 fatty acids (10–12). The aforementioned results had not reported potential etiology, causes, or molecular signatures for the population-wide skewed seroconversion age of either autoantibody, which might also be predictive of further progression to T1D.

The exposures to ascorbic acid, vitamin D, and fatty acids can be assessed either at the intake level using dietary assessment tools or at biomarker level quantified by immunoassay platforms (13) and chromatographic methods (14,15). These micronutrients and fatty acids are usually considered positively associated with the actual intake amount, as well as regulated by protein-coding genes and metabolic pathways, such as vitamin C transport genes *SLC23A1*, *SLC2A1*, and *SLC2A2* and 25-hydroxyvitamin D [25(OH)D] receptor genes *VDR* and *RXRA*. A secondary set of biomarkers for nutrient sources is the lipidome profiled from plasma or serum, which is related to both nutrients and energy sources and metabolic activity. As structural components of cell membranes, complex lipids are involved in a variety of metabolic pathways (16) and play an important role in children's growth and disease development. Previous studies had found that heterogeneity of lipidome profiles signaled the underlying development of or recovery from diseases (17–19). Therefore, children's lipidome reflects metabolism of nutrients as a link between gene regulation of nutrient biomarkers and characteristics related to primary end points, i.e., IA or T1D.

Previous studies in T1D rarely integrated dietary biomarkers and lipidome profiles due to either limited blood samples or cohort size (4,10,18). The Environmental Determinants of Diabetes in the Young (TEDDY) study measured concentrations of dietary biomarkers in plasma or on red blood cell membrane for the enrolled participants by immunoassay and chromatographic methods and profiled untargeted plasma metabolome by mass spectrometry platforms (20). In the current study, we hypothesized that plasma ascorbic acid, 25(OH)D, and erythrocyte n-3 fatty acids, i.e., α-linolenic acid (ALA), eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA), and docosapentaenoic acid (DPA), are nutrients (or biomarkers) interacting with lipidome followed by the development of IA and progression to T1D.

Our present analyses were performed on GADA-first and IAA-first children and their matched control subjects. We first adopted an unsupervised clustering approach to group GADA-first and IAA-first children based on their temporal lipidomic profiles at and prior to seroconversion and further identified nutritional and metabolic signatures for IA-related outcomes, as well as assessed the association of HLA-DR-DQ haplotypes or genotypes with the identified signatures.

RESEARCH DESIGN AND METHODS

TEDDY Nested Case-Control Study

TEDDY is an observational prospective study of children at increased genetic risk of T1D conducted in six clinical research centers: three in the U.S. (CO, GA/FL, and WA) and three in Europe (Finland, Germany, and Sweden). A total of 8,676 children were enrolled and followed prospectively from 3 months to 15 years of age for blood samples collection and IA measurement at 3-month intervals before 4 years of age and at 6-month intervals thereafter. All children who were persistently positive for any autoantibody were followed every 3 months until the age of 15 years. Detailed study design and methods have previously been published (21). Written informed consents were obtained for all study participants from a parent or a primary caretaker, separately, for genetic screening and participation in the prospective follow-up, beginning at birth. Autoantibodies were measured in two laboratories by radiobinding assays (21,22). Persistent IA was defined as confirmed positive autoantibodies (IAA, GADA, or IA-2A) in at least two consecutive samples by both laboratories, with the date of seroconversion defined as that of the first positive sample. T1D was diagnosed with use of the American Diabetes Association criteria (23).

For biomarker discovery, a nested case-control (NCC) design with risk set sampling was used based on data collected as of 31 May 2012 (24). IA case subjects (418) were selected, and each was matched with 3 control subjects who had not developed IA by the age when the case subject experienced the first-appearing autoantibodies. Matching factors were clinical centers, sex, and first-degree relatives' status of T1D. For case subjects selected in the NCC design, plasma ascorbic acid, 25(OH)D, and erythrocyte fatty acids were measured from the children's blood samples annually from the end of infancy (i.e., at 12, 24, 36, 48, and 60 months), while the children's metabolomes including lipidome were profiled at each 3-month visit following birth up to seroconversion. Nutrient biomarkers and metabolome were also measured at the same age (or visits) for matched control subjects. The concentrations of nutrient biomarkers at infant ages (i.e., at 3, 6, and 9 months) were also available at different frequencies (8,25). Bar plots in Fig. 1 present the numbers of subjects having biomarkers and lipidome measured, individually, per time point within 1 year prior to seroconversion. Each matched case-control pair had blood samples collected at the same ages, although random missing visits occurred.

Nutrient Biomarkers and Metabolome Profiling

Nutrient biomarkers were profiled at the Biochemistry Laboratory, Genomics and Biomarkers Unit, National

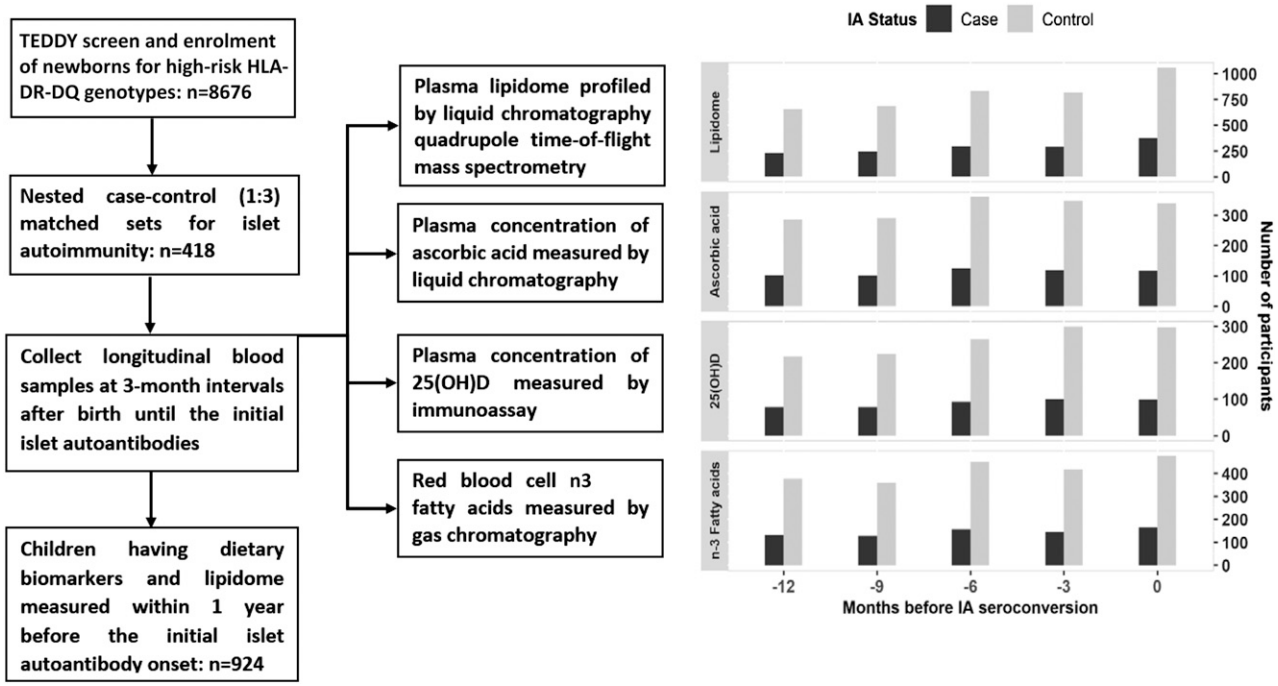

Figure 1—TEDDY study design and distribution of participants having nutrient biomarkers and lipidome measured.

Institute for Health and Welfare, Helsinki, Finland. Plasma ascorbic acid was determined by an ion-paired, reversed-phase, high-performance liquid chromatographic method using electrochemical detection, while plasma 25(OH)D was measured with the ARCHITECT 25-OH Vitamin D chemiluminescent microparticle immunoassay. n-3 fatty acids were analyzed from erythrocytes stabilized in 2-propanol and butylated hydroxytoluene. Erythrocyte fatty acids composition was analyzed with an Agilent 6890 gas chromatograph with a split injector and hydrogen as the carrier gas, using a capillary column Omegawax 320 (length 30 m, inner diameter 0.32 mm, phase layer 0.25 μm) (Supelco, Bellefonte, PA). Data were analyzed with OpenLab CDS ChemStation software (Agilent, Palo Alto, CA). The percentage composition of fatty acid methyl esters was normalized to 100% in each sample. Samples of the case subjects and their control subjects at each age point were processed in the same batch for all analytes.

TEDDY subjects' plasma samples selected under an NCC design were shipped to National Institutes of Health West Coast Metabolomics Center at University of California, Davis (UCD), and processed with standard laboratory protocols followed. Prior to sample shipment, the TEDDY Data Coordinating Center designed laboratory running batches by assigning plasma samples of matched subjects to the same batch. Mass spectrometry processing at University of California, Davis, followed the order of designed batches to ensure that disease outcomes did not confound with run order.

The primary metabolites assay was profiled by gas chromatography time-of-flight mass spectrometry (GC-TOF MS) on a LECO Pegasus III instrument annotated by BinBase (26). The complex lipids (lipidome) assay was analyzed by charged surface hybrid column with electrospray ionization (CSH-ESI) on Agilent ultra-high-pressure liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-QTOF MS) instruments annotated with LipidBlast (27). Peaks from GC-TOF MS platform were automatically detected and deconvoluted from coeluting peaks by the LECO ChromaTOF software (version 3.0). Raw data output from CSH-ESI quadrupole time-of-flight instruments were processed in an untargeted manner by Agilent's software MassHunter Qual (version B.05.00) to find peaks. Peak features were then imported into Mass Profiler Professional for peak alignments to seek which peaks are present in multiple chromatograms. These peaks are then collated and constrained within the MassHunter Quantitative Analysis software (version B.05.01) on the accurate mass precursor ion level, using the MS information and the LipidBlast library. We applied a comprehensive normalization pipeline, Systematic Error Removal using Random Forest (SERRF) (28), to the raw quantified intensity of metabolites, successfully removing laboratory run-order effect in terms of a quality control metric, i.e., reduced coefficient of variation of internal standard compounds. The missing compounds in the normalized data were filled by the values generated from a label-free missing value imputation tool, GMSimpute (29).

Unsupervised Clustering of Temporal Lipidomic Profiles

For screening potential linkage among lipidome, nutrients, and IA-related outcomes, an unsupervised clustering

approach was first applied to the GADA-first and IAA-first children ($n = 165$) having plasma lipidome profiled at three consecutive visits within 6 months before or at seroconversion, labeled as visits 0, −3, and −6. Previous results had shown a strong infant age pattern in TEDDY children's lipidome (20), which may be a result of growth and dietary exposure changes similar to in gut microbiome (30). The aim of this analysis was to detect latent subgroups of children having stratified blood concentrations of nutrients or distinct end point outcomes instead of grouping lipidome profiles only representative of age or growth phases (30). Hence, we merged every child's multivisit lipidomes into one temporal lipidomic profile, consisting of three consecutive measurements of 208 lipid species with dimension of $m = 3 \times 208 = 624$.

Temporal lipidomic profiles of the children who experienced seroconversion by the end of infancy might be grouped into a cluster due to infant age patterns. On the other hand, a cluster with both children with preinfancy seroconversion and children with postinfancy seroconversion may be representative of young age and other characteristics. Existing studies (31,32) have reported strong association among lipid species profiled from each sample. Thus, to account for the strong associations among lipid species and temporal measurements, we used Gaussian model–based clustering to group ($n = 165$) temporal lipidomic profiles. The clustering was performed by R package mclust (33), and the optimal number of clusters was selected between 1 and 10 according to Bayesian information criterion.

Cluster-Wise Lipidomic and Nutritional Signatures and Correlation Analysis

We first used the Bioconductor package limma (34) to perform lipidome-wide differential analysis for identification of lipidomic signatures that defined each cluster of children. Classes of lipidomic signatures were further confirmed by enrichment analysis from ChemRICH (35) with filters: altered lipid species ratio >0.3, increased or decreased species ratio >0.3, and Benjamini-Hochberg (BH) procedure–adjusted (36) P value <0.05. Next, the concentration of each nutrient biomarker measured within 12 months before seroconversion was compared between clusters of children by linear regression, adjusted by the age of biomarker measurement with statistical significance at $P < 0.05$. A nutrient biomarker with concentration stratified by the clusters is a nutritional signature for the corresponding subgroup(s) of children. The cluster-wise nutritional and lipidomic signatures revealed the latent relation between nutrient biomarkers and lipid species.

We further confirmed the lipidome-nutrients connectivity by Spearman correlation coefficients (ρ) of pairwise lipid species and nutrient biomarkers based on IAA-first and GADA-first case subjects ($n = 191$) and matched control subjects ($n = 517$), who had both types of analytes measured within 12 months before seroconversion. Connectivity (or correlation) threshold was determined by $\rho > 0.1$ or < -0.1. The concentration of each nutrient biomarker was the log-scaled measurement (or the average of more than one measurement) within 12 months prior to seroconversion for case subjects and at the same age for matched control subjects, while the lipidome was selected at the last visit (3 months) prior to seroconversion and at the same age of matched control subjects.

Differential Abundance Analysis for IA-Related Outcomes

Demographic characteristics and IA-related outcomes per cluster, such as the type of an initial autoantibody, seroconversion age, and further progression to T1D, were first identified by Kruskal-Wallis test, Fisher exact test, and Cox regression with statistical significance at $P < 0.05$. We further performed differential analysis on metabolome (i.e., primary metabolites and lipidome) and nutrient biomarkers against the identified end point outcomes among GADA-first and IAA-first case subjects, using limma or linear regression. Metabolites associated with end point outcomes were selected by adjusted $P < 0.2$, while associations between nutrient biomarkers and end point outcomes were analyzed by linear regression and selected by $P < 0.05$. Prognostic effect was further confirmed by Cox regression. Age at analytes measurements, country, family history of T1D, sex, and HLA DR-DQ haplotypes were included in differential or Cox analysis as covariates.

Data and Resource Availability

The data sets generated and analyzed during the current study will be made available in the National Institute of Diabetes and Digestive and Kidney Diseases Central Repository (https://www.niddkrepository.org/studies/teddy).

RESULTS

Clusters of Children Defined by Temporal Lipidomic Profiles

The unsupervised clustering algorithm grouped 165 TEDDY children's temporal lipidomic profiles before GADA-first or IAA-first as Clusters I–IV, visualized by t-distributed stochastic neighbor embedding (tSNE) plot in Fig. 2*A* and the heatmap in Fig. 2*B*. Children in Clusters I and II were not clearly distinguished from each other on the tSNE plot, but they were well separated from Clusters III to IV. In addition, a clear border between Cluster IV and the others was observed (Fig. 2*A*). The distribution of clusters in the tSNE plot agreed to the lipidomic patterns in Fig. 2*B*; i.e., Clusters I and II shared similar patterns in triacylglycerols (TG), but a remarkable difference between Clusters I–II and Clusters III–IV was observed in TG. A heatmap of visits −3 and −6 presented contrast of Clusters II versus I and Clusters IV versus III in TG.

Cluster-Wise Lipidomic and Nutrient Signatures and Connectivity

We further performed lipidome-wide differential and enrichment analyses at visits −3 and −6 to identify the most

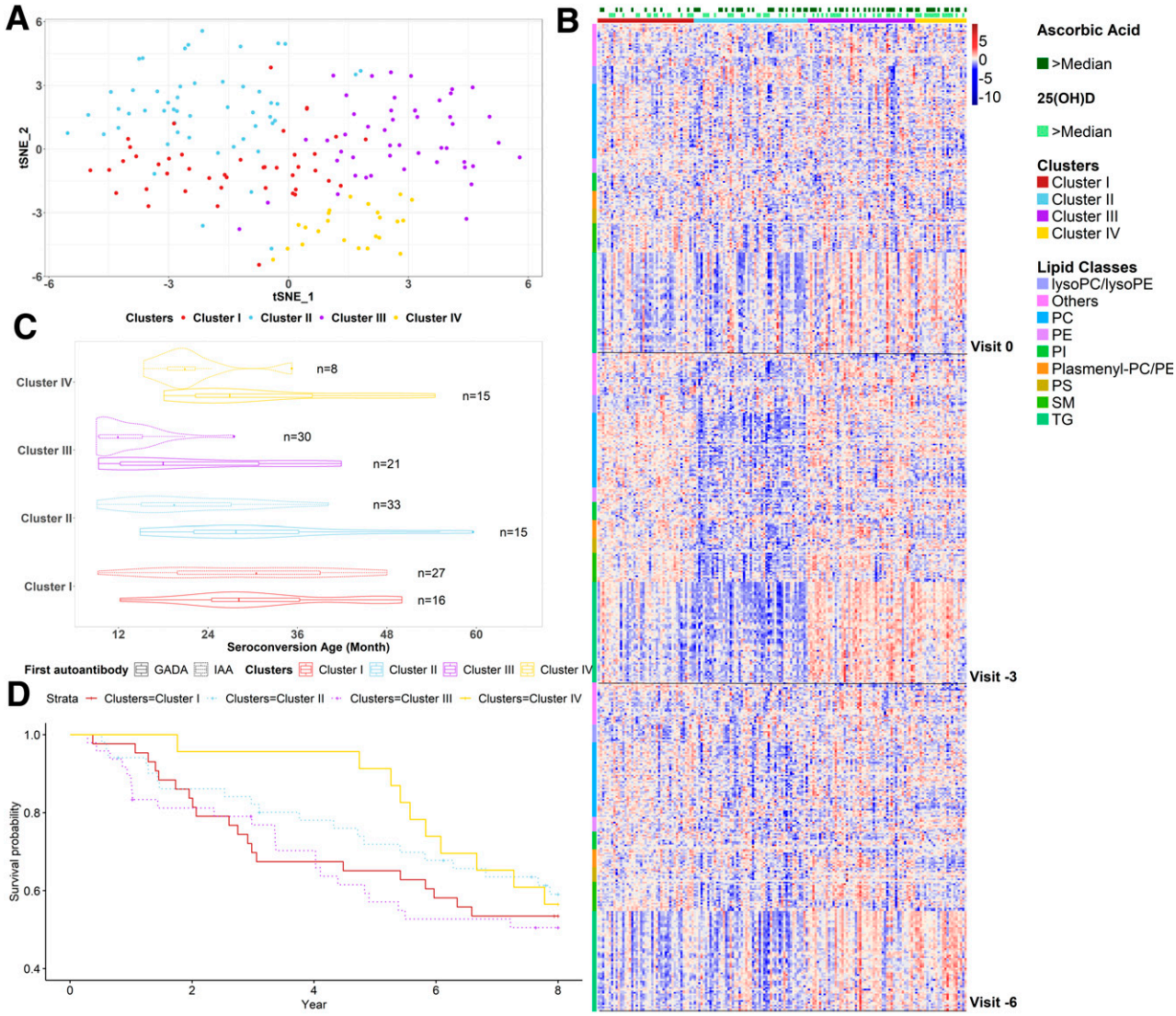

Figure 2—Clusters identified by temporal lipidomic profiles of 165 TEDDY children. *A*: tSNE plot based on longitudinal lipidome, with axes representing two dimensions of distances. *B*: Heatmap for longitudinal lipidome with columns as subjects and rows as lipids per visit. *C*: Distribution of seroconversion age for each type of initial autoantibody per cluster. *D*: Kaplan-Meier curve for progression from the initial autoantibody to T1D.

differentiated lipid classes as top signatures for Clusters III and IV versus Clusters I and II, Cluster II versus Cluster I, and Cluster IV versus Cluster III, Cluster III versus Cluster I, and Cluster IV versus Cluster I. Differential analysis results can be found in Supplementary Tables 1–10 and the differentially abundant lipid classes at both visits −3 and −6 are summarized in Supplementary Table 11, verifying lipidomic signatures illustrated in Fig. 2*B*. In comparisons with the reference Cluster I, lipidomic signatures for Cluster II were increased phosphatidylcholines (PC) and sphingomyelins (SM), while higher triacylglycerols (TG) and diglycerides (DG) and lower PC and cholesterol esters (CE) were signatures for Cluster IV. Cluster III compared with the other clusters had increased TG, DG, lysophosphatidylcholines (lysoPC), and plasmalogens, as a result of young age (20) lipidome profiles (at seroconversion) as shown in Table 1. A higher level of ascorbic acid was observed in Cluster III ($P = 0.0024$) and Cluster IV ($P = 0.0031$), while 25(OH)D concentration was elevated only in Cluster IV ($P = 0.0093$). Ascorbic acid was also found decreasing along with growth, being consistent with higher ascorbic acid and young age at measurement in Cluster III. Cluster IV had lower composition of n-3 fatty acids ALA, EPA, DHA, and DPA ($P = 0.0006$, <0.001, 0.0156, 0.0007), with mean levels shown in Table 1, and were positively associated with measurement age ($P < 0.001$).

Spearman correlation coefficients among nutrient biomarkers and lipidome are listed in Supplementary Table 12 and visualized in Supplementary Fig. 1, including the pairwise correlation between dietary biomarkers. Higher levels of TG, DG, ascorbic acid, and 25(OH)D in Clusters III and IV were consistent with TG and DG positive

Table 1—Characteristics of Clusters I–IV

	Cluster I	Cluster II	Cluster III	Cluster IV	*P**
N (%)	43 (26.1)	51 (30.9)	48 (29.1)	23 (13.9)	
ALA (composition %)	0.15 (0.04)	0.15 (0.06)	0.11 (0.04)	0.11 (0.04)	
EPA (composition %)	0.60 (0.29)	0.48 (0.37)	0.50 (0.25)	0.27 (0.14)	
DHA (composition %)	3.98 (1.04)	3.92 (1.18)	4.96 (1.04)	3.39 (0.83)	
DPA (composition %)	2.11 (0.40)	1.85 (0.54)	1.77 (0.44)	1.65 (0.56)	
Seroconversion age (month)	30.1 (10.9)	25.0 (11.3)	14.8 (8.4)	27.9 (10.8)	<0.0001
Country, *n* (%)					<0.0001
U.S.	2 (4.7)	10 (19.6)	12 (25.0)	20 (87.0)	
Finland	19 (44.2)	16 (31.4)	19 (39.6)	2 (8.7)	
Germany	2 (4.7)	1 (2.0)	5 (10.4)	0 (0.0)	
Sweden	20 (46.5)	24 (47.1)	12 (25.0)	1 (4.3)	
First-degree relatives with T1D, *n* (%)					0.0498
Yes	8 (18.6)	6 (11.8)	16 (33.3)	7 (30.4)	
No	35 (81.4)	45 (88.2)	32 (66.7)	16 (69.6)	
Sex, *n* (%)					
Female	22 (51.2)	29 (56.9)	19 (39.6)	11 (47.8)	0.3829
Male	21 (48.8)	22 (43.1)	29 (60.4)	12 (52.2)	
HLA DR-DQ genotype, *n* (%)					
DR3/DR4	23 (53.5)	23 (45.1)	22 (45.8)	11 (47.8)	0.8541
Other	20 (46.5)	28 (54.9)	26 (54.2)	12 (52.2)	
Type of initial autoantibody, *n* (%)					
IAA	27 (62.8)	30 (58.8)	33 (68.8)	8 (34.8)	0.0519
GADA	16 (37.2)	21 (41.2)	15 (31.3)	15 (65.2)	
Ascorbic acid (μmol/L)	42.99 (20.5)	47.75 (23.5)	63.70 (23.4)	61.96 (22.5)	
25(OH)D (nmol/L)	50.53 (16.4)	53.67 (15.8)	49.98 (20.6)	62.73 (21.5)	

Data are means (SD) unless otherwise indicated. The composition of each fatty acid was the ratio of each fatty acid abundance to the total abundance of all fatty acids. **P* values are from Pearson χ^2 test for categorical variables or Kruskal-Wallis test for continuous variables. Association test on country and clusters was confirmed by exclusion of Germany because of small sample size and sparse cells ($n < 2$) in contingency table.

correlations with ascorbic acid and 25(OH)D, while on-3 fatty acids showed positive connectivity with most CE lipid species and negative connectivity with DG (Supplementary Table 12). The simultaneous higher levels of ascorbic acid and 25(OH)D and reduced n-3 fatty acids observed in Cluster IV was consistent with the correlations among these nutrient biomarkers (Supplementary Table 12).

Cluster-Wise Characteristics Related to IA

The associations of identified clusters with country, family history of T1D, initial autoantibody type, and seroconversion age are shown in Table 1. Most subjects in Cluster I experienced seroconversion after 15 months of age regardless of the type of initial autoantibodies. Cluster II had a relatively lower proportion of first-degree relatives with T1D ($P = 0.0498$). The cluster-wise seroconversion age for each initial autoantibody is presented in Fig. 2*C* and Table 1. Cluster III compared with others was a subgroup with younger seroconversion age ($P < 0.0001$) for either IAA-first or GADA-first. Half of GADA-first children's ($n = 21$) temporal lipidomic profiles in Cluster III were postinfancy, although the lipidomic profiles for half of the IAA-first children ($n = 30$) in this cluster were at infant age. Hence, Cluster III not only represented younger age of lipidome or children but also included the earliest seroconversion of each autoantibody. A subgroup analysis showed that seroconversion among Clusters I, II, and IV was age independent. Relatively higher ascorbic acid and 25(OH)D and lower n-3 fatty acids were found in U.S. participants compared with participants of other countries, while most participants in Cluster IV were from the U.S. ($P = 0.0005$). This explained the nutritional signatures [higher ascorbic acid and 25(OH)D and lower n-3 fatty acids] identified in Cluster IV.

A Kaplan-Meier curve (Fig. 2*D*) showed that subjects in Cluster III compared with other clusters had higher risk of T1D progression from the onset of initial autoantibody, as a result of early seroconversion in Cluster III (2). A subsequent question is whether the subjects in the other Clusters (I, II, IV) had similar risk of progression to T1D from the initial autoantibody, since they experienced seroconversion at similar age. Therefore, we compared the risk of progression from seroconversion to T1D among Clusters I, II, and IV throughout the follow-up from seroconversion to 31 December 2019, adjusting for the age at seroconversion, type of the initial autoantibody, first-degree relatives with T1D, sex, and HLA haplotypes (DR3/DR4 vs. others). Supplementary Table 13 showed that

Cluster IV compared with Cluster I had relatively slower T1D progression after the initial autoantibody (hazard ratio 0.3672 [95% CI 0.1617, 0.8335], $P = 0.0166$). Therefore, we identified distinct clusters of children representing earlier seroconversion of either IAA-first or GADA-first, fewer first-degree relatives with T1D, and slower progression to T1D.

Metabolites and Circulating Vitamins Preceding Early Seroconversion and T1D Progression

Children in Cluster III developed IAA-first at mean age of 12.3 months and GADA-first at mean age of 20.6 months, consistent with the population-wide early onset of either initial autoantibody (1). This subgroup of TEDDY children revealed an analytical definition of early onset of IAA-first and GADA-first. To identify metabolic or nutritional signatures for early onset of an initial autoantibody instead of age, we first defined "early IAA" ($n = 68$) as IAA-first case subjects, "late IAA" ($n = 134$) as those with seroconversion before or after 12 months of age, and GADA-first "early GADA" ($n = 43$) and "late GADA" ($n = 93$) as those with seroconversion before or after 21 months of age. Differential abundance analysis between "early" and "late" onset of each initial autoantibody was performed on analytes measured at the same age during infancy, with country, sex, first-degree relatives with T1D, and HLA haplotype included as covariates. Log fold change in Fig. 3*A* showed that children having reduced plasma ascorbic acid at age 6 months ($P = 0.0362$) or less abundant cholesterol metabolite at age 9 months ($P < 0.0001$, BH-adjusted $P < 0.2$) afterward developed IAA-first earlier (early IAA), while less abundant SM species at age 9 months (BH-adjusted $P < 0.2$) were found to be associated with the following early onset of GADA-first (early GADA). The cholesterol metabolite associated with early IAA was the free form of cholesterol profiled by GC-TOF MS (InChIKey: HVYWMOMLDIMFJA-DPAQBDIFSA-N)—different from CE lipid species profiled by UHPLC-QTOF MS platform.

Differential analysis also found that children who experienced T1D progression ($n = 171$) after either type of initial autoantibody had reduced 25(OH)D ($P = 0.0029$) within 1 year prior to seroconversion compared with the nonprogressors ($n = 167$), while multiple lipid species in lysoPC and DG 34:1 at 6 months prior to GADA-first were less abundant in T1D progressors than nonprogressors (Fig. 3*B*). Age of measurement, country, sex, and first-degree relatives with T1D were included as covariates in the differential analysis of progression. Relatively higher

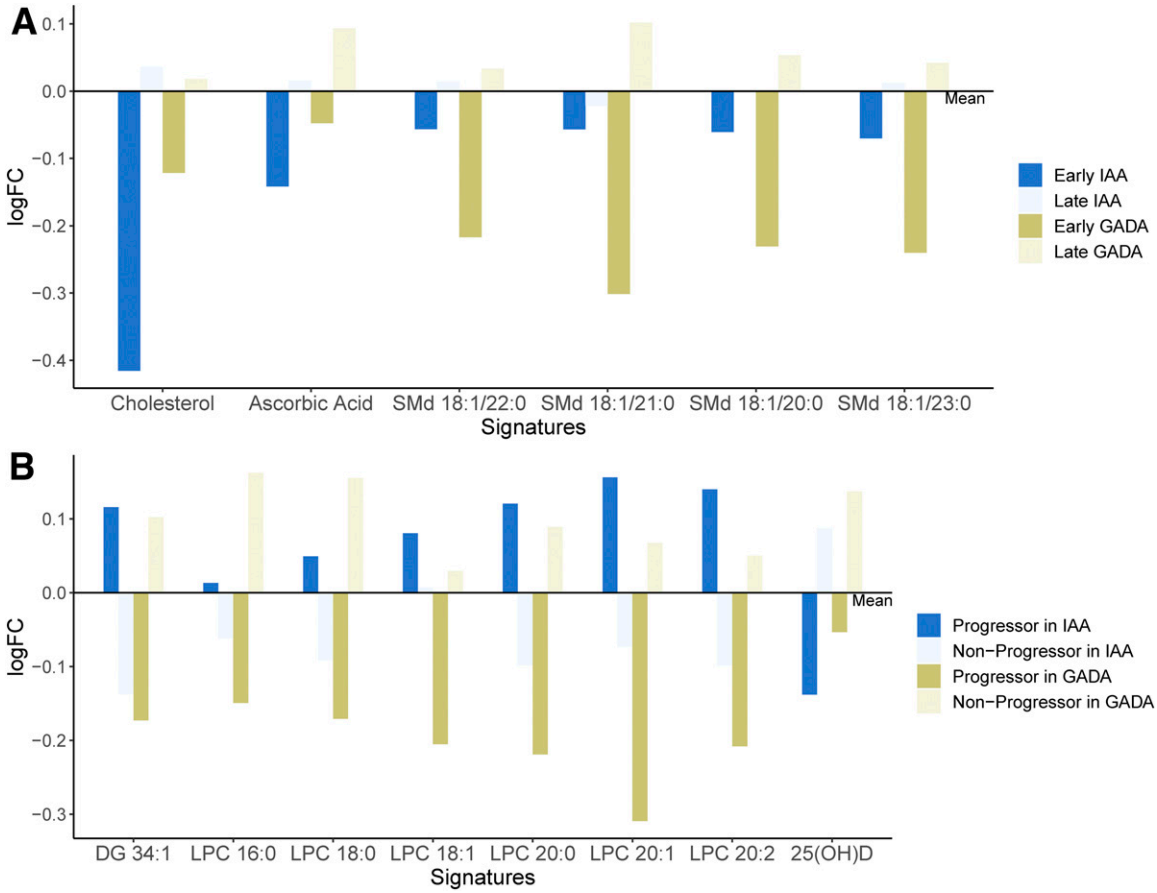

Figure 3—Metabolites and circulating vitamins preceding early seroconversion (*A*) or further progression to T1D (*B*). FC, fold change; SMd, sphingomyelins (SM).

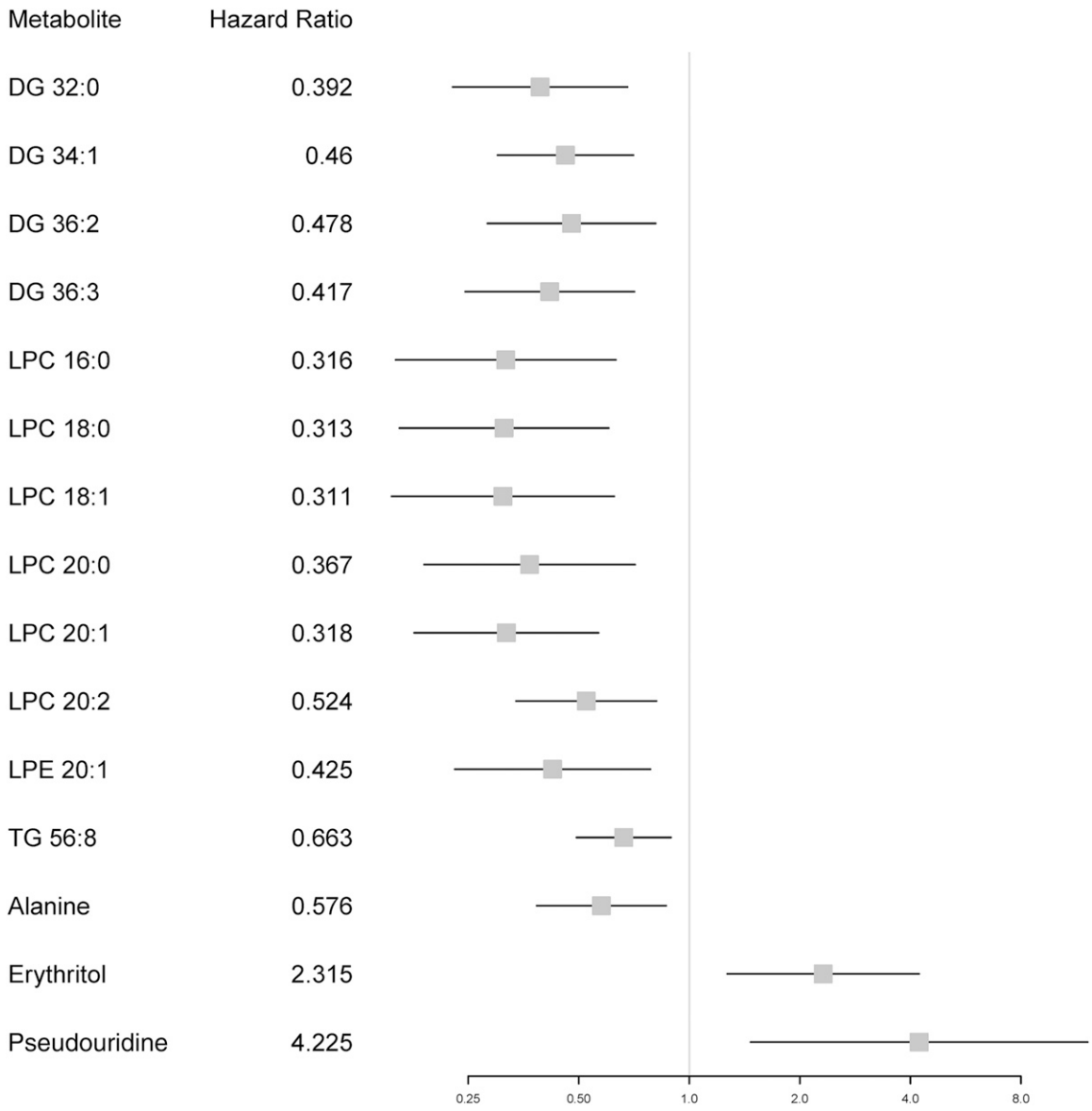

Figure 4—Hazard ratio for metabolites associated with progression risk from GADA-first seroconversion to T1D. LPC, lysoPC; LPE, lysoPE.

abundance (log fold change in Fig. 3*B*) of lipid species before IAA-first onset in T1D progressors compared with nonprogressors was not confirmed with statistical significance. Reduced abundance of other lipid species in DG, TG 56:8, and amino acid alanine at 6 months before GADA-first onset was also found signaling afterward T1D progression by Cox regression analysis (BH-adjusted $P < 0.2$), with hazard ratio presented in Fig. 4. The prognostic signatures for T1D progression following GADA-first (lysoPC and DG 34:1) identified in differential analysis were also confirmed by Cox regression. Our analysis did not find metabolic signatures for T1D progression following IAA-first. These results were in agreement with the higher 25(OH)D and slower T1D progression observed in Cluster IV.

A further analysis on the analytes measured at infant age among IAA-first and GADA-first case subjects revealed that HLA-DR3/DR4 haplotype and the corresponding alleles (at HLA genes DRB1, DQA1, DQB1) were associated with reduced 25(OH)D at age of 6 months (Table 2), although ascorbic acid at the same age was only associated with HLA-DR3/DR4 haplotype. Association between HLA-DR3/DR4 and circulating vitamins was not observed in matched control subjects.

DISCUSSION

Our analysis in the current study revealed the relationship among HLA DR3/DR4 genotypes, circulating vitamins, metabolome, islet autoimmunity, and further progression to T1D in early childhood. We provided an analytical definition for early seroconversion of IAA-first (12 months) and GADA-first (21 months), respectively. We identified reduced ascorbic acid, cholesterol metabolite, and SM species as early-seroconversion signatures for IAA-first and GADA-first (Fig. 3), while 25(OH)D and multiple

Table 2—Association of HLA haplotypes/genotypes with plasma 25(OH)D and ascorbic acid at age of 6 months for IAA-first and GADA-first children

	Plasma 25(OH)D (nmol/L)		Plasma ascorbic acid (μmol/L)	
HLA haplotypes or genotypes (%)	Mean (SD)	*P*	Mean (SD)	*P*
DR3/DR4 (51%)	49.62 (21.18)	0.018	55.51 (27.99)	0.04[+]
Others (49%)	56.62 (24.88)		62.37 (28.54)	
DRB1*0301 (64%)	51.23 (20.49)	0.023	58.90 (29.00)	0.845
Others (36%)	56.24 (27.34)		58.90 (27.54)	
DQA1*0501 (64%)	51.43 (20.56)	0.033	58.92 (28.80)	0.856
Others (36%)	56.00 (27.43)		58.85 (27.87)	
DQB1*201 (64%)	51.43 (20.56)	0.033	58.92 (28.80)	0.856
Others (36%)	56.00 (27.43)		58.85 (27.87)	

Association analysis per circulating nutrient was performed by linear regression on each genetic factor with covariates of country, sex, and family history of T1D. [+]Association between HLA DR3/DR4 and ascorbic acid was $P = 0.001$ for GADA-first children.

metabolites before either initial autoantibody were predictive of further progression to T1D.

Multiple metabolites near the onset of GADA-first were found as prognostic signature for T1D (Fig. 4). The association between reduced amino acid alanine and T1D progression was consistent with the negative association between alanine and GADA-first risk previously reported (20). Our previous analysis on primary metabolites in TEDDY (20) also found alanine to be negatively associated with the risk of IAA-first. Higher pseudouridine was found to be associated with T1D progression following GADA-first, which was also a risk marker for development of type 2 diabetes or renal dysfunction in T1D or type 2 diabetes (37,38). Pseudouridine is typically the most abundant marker of modified RNA bases in human plasma. Erythritol, a critical sugar alcohol, was also found to be positively associated with T1D progression after GADA-first. Furthermore, results in Table 2 not only confirmed previous findings that lower circulating vitamins were associated with higher IA risk (8,25) but also implied that HLA-DR3/DR4 potentially linked to the gene expression dysregulation of these circulating vitamins.

Our analysis found reduced ascorbic acid only preceding early seroconversion of IAA-first, consistent with the conclusion that ascorbic acid only reduced the risk of IAA-first from a recent analysis in TEDDY (25). Previous results about the role of ascorbic acid in T1D were unclear. A previous study (39) showed that ascorbic acid level was significantly increased in plasma and spleen of the diabetes-prone rats compared with the control group. Higher ascorbate concentration in mouse was also found to have immediate inhibition on insulin secretion, but was completely reversible (40), which was associated with hyperpolarization of the pancreatic β-cell. On the other hand, oral ascorbic acid supplementation was reported to provide an effective prophylaxis against exercise-induced free radical–mediated lipid peroxidation in T1D (7).

A recent study in TEDDY provided evidence for the association between 25(OH)D and islet autoimmunity (8) without analyzing the risk of T1D. The protective effect of 25(OH)D on T1D was reported in a Norwegian birth control study (41) and in a non-Hispanic White population in the U.S. (42) but not in other studies (43–47). Our present study sheds light on the potential causes for previous inconsistent conclusions about 25(OH)D and the overall risk for T1D, that is, the time between 25(OH)D measurement and clinical onset of T1D. Our findings underscore the importance of when to relate 25(OH)D to T1D, i.e., within 1 year prior to seroconversion, at (or immediately after) seroconversion, at a second IA (48), or before the onset of multiple positive autoantibodies (49,50).

One of our aims in the present analyses was to identify potential metabolic and nutritional markers for the subtypes of islet autoimmunity. Hence, we focused on TEDDY children who developed GADA only (33%) or IAA only (49%) as the initial autoantibody to ensure heterogeneity among subjects in clustering, without considering the simultaneous onset of GADA and IAA (14%) or other autoantibodies at seroconversion. Future studies on lipidome and dietary biomarkers under the TEDDY NCC design should include all subjects having positively confirmed IA and their matched control subjects for sufficient statistical power. The TEDDY full cohort is dominated by a White non-Hispanic population (>90%); hence, sample size in other race or ethnicity groups was too small to assess whether our findings were relevant to races. Another limitation of the present study existed in the prognostic analysis on either metabolome or nutrient biomarkers, which was conducted within the TEDDY NCC design for IA (at 1:3 case:control ratio). The population-wide risk of progression from initial autoantibody to T1D might differ from that in the TEDDY NCC, as it only represents a subgroup of the full TEDDY cohort who seroconverted at an early age (<5 years). Future prognostic analysis on metabolome or nutrient biomarkers should be extended to a larger cohort for reproducibility and validation.

Conclusion

Infants having reduced plasma ascorbic acid and cholesterol experienced seroconversion of IAA-first earlier, while early onset of GADA-first was preceded by reduced SM

at infancy. Both IAA-first and GADA-first children with further T1D progression compared with nonprogressors had lower plasma 25(OH)D prior to the onset of each autoantibody. Plasma ascorbic acid and 25(OH)D were negatively associated with HLA-DR3/DR4 among infants who later experienced IAA-first and GADA-first, implying dysregulation of gene expression potentially linked to reduced circulating vitamins.

Funding. The TEDDY study is funded by National Institute of Diabetes and Digestive and Kidney Diseases (U01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865, U01 DK63863, U01 DK63836, U01 DK63790, UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865, UC4 DK63863, UC4 DK63836, UC4 DK95300, UC4 DK100238, UC4 DK106955, UC4 DK112243, and UC4 DK117483) and contract no. HHSN267200700014C from the National Institute of Allergy and Infectious Diseases, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institute of Environmental Health Sciences, Centers for Disease Control and Prevention, and JDRF. This work supported in part by the National Institutes of Health/National Center for Advancing Translational Sciences Clinical and Translational Science Awards to the University of Florida (UL1 TR000064) and the University of Colorado (UL1 TR002535).

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

Author Contributions. Å.L., W.H., M.R., J.-X.S., J.T., A.-G.Z., B.A., and J.P.K. contributed to conception and design of the study. Q.L., X.L., and J.P.K. developed analytical plans and interpreted the results. Q.L. conducted analysis programming and drafted the manuscript with X.L. J.Y. interpreted results, provided input in discussion, and edited the manuscript. Å.L., M.R., J.-X.S., J.T., and A.-G.Z. facilitated physical data collection from patients. Q.L., X.L., J.Y., I.E., Å.L., W.H., M.R., J.-X.S., J.T., A.-G.Z., B.A., and J.P.K. reviewed, edited, and approved the final draft of the manuscript. Q.L. 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

1. Krischer JP, Lynch KF, Schatz DA, et al.; TEDDY Study Group. The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study. Diabetologia 2015;58:980–987

2. Liu X, Vehik K, Huang Y, et al.; TEDDY Study Group. Distinct growth phases in early life associated with the risk of type 1 diabetes: the TEDDY study. Diabetes Care 2020;43:556–562

3. Pundziute-Lyckå A, Persson L-A, Cedermark G, et al. Diet, growth, and the risk for type 1 diabetes in childhood: a matched case-referent study. Diabetes Care 2004;27:2784–2789

4. Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. Lancet 2016;387:2340–2348

5. Hummel S, Beyerlein A, Tamura R, et al.; TEDDY Study Group. First infant formula type and risk of islet autoimmunity in The Environmental Determinants of Diabetes in the Young (TEDDY) study. Diabetes Care 2017;40:398–404

6. Lund-Blix NA, Dong F, Mårild K, et al. Gluten intake and risk of islet autoimmunity and progression to type 1 diabetes in children at increased risk of the disease: the Diabetes Autoimmunity Study in the Young (DAISY). Diabetes Care 2019;42:789–796

7. Davison GW, Ashton T, George L, et al. Molecular detection of exercise-induced free radicals following ascorbate prophylaxis in type 1 diabetes mellitus: a randomised controlled trial. Diabetologia 2008;51:2049–2059

8. Norris JM, Lee H-S, Frederiksen B, et al.; TEDDY Study Group. Plasma 25-hydroxyvitamin D concentration and risk of islet autoimmunity. Diabetes 2018; 67:146–154

9. Uusitalo U, Liu X, Yang J, et al.; TEDDY Study Group. Association of early exposure of probiotics and islet autoimmunity in the TEDDY study. JAMA Pediatr 2016;170:20–28

10. Miller MR, Yin X, Seifert J, et al. Erythrocyte membrane omega-3 fatty acid levels and omega-3 fatty acid intake are not associated with conversion to type 1 diabetes in children with islet autoimmunity: the Diabetes Autoimmunity Study in the Young (DAISY). Pediatr Diabetes 2011;12:669–675

11. Niinistö S, Takkinen H-M, Erlund I, et al. Fatty acid status in infancy is associated with the risk of type 1 diabetes-associated autoimmunity. Diabetologia 2017;60:1223–1233

12. Virtanen SM, Niinistö S, Nevalainen J, et al. Serum fatty acids and risk of advanced beta-cell autoimmunity: a nested case-control study among children with HLA-conferred susceptibility to type I diabetes. Eur J Clin Nutr 2010;64:792–799

13. Lequin RM. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clin Chem 2005;51:2415–2418

14. Salminen I, Alfthan G. Plasma ascorbic acid preparation and storage for epidemiological studies using TCA precipitation. Clin Biochem 2008;41:723–727

15. Elorinne A-L, Alfthan G, Erlund I, et al. Food and nutrient intake and nutritional status of Finnish vegans and non-vegetarians [published correction appears in PLoS One 2016;11:e0151296]. PLoS One 2016;11:e0148235

16. van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol 2008;9:112–124

17. Orešič M, Simell S, Sysi-Aho M, et al. Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. J Exp Med 2008;205:2975–2984

18. Pflueger M, Seppänen-Laakso T, Suortti T, et al. Age- and islet autoimmunity-associated differences in amino acid and lipid metabolites in children at risk for type 1 diabetes. Diabetes 2011;60:2740–2747

19. Kyle JE, Burnum-Johnson KE, Wendler JP, et al. Plasma lipidome reveals critical illness and recovery from human Ebola virus disease. Proc Natl Acad Sci U S A 2019;116:3919–3928

20. Li Q, Parikh H, Butterworth MD, et al.; TEDDY Study Group. Longitudinal metabolome-wide signals prior to the appearance of a first islet autoantibody in children participating in the TEDDY study. Diabetes 2020;69:465–476

21. TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) study: study design. Pediatr Diabetes 2007;8:286–298

22. TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) study. Ann N Y Acad Sci 2008;1150:1–13

23. American Diabetes Association. Standards of medical care in diabetes—2014. In Clinical Practice Recommendations, 2014. Diabetes Care 2014;37(Suppl. 1):S14–S80

24. Lee H-S, Burkhardt BR, McLeod W, et al.; TEDDY Study Group. Biomarker discovery study design for type 1 diabetes in The Environmental Determinants of Diabetes in the Young (TEDDY) study. Diabetes Metab Res Rev 2014;30:424–434

25. Mattila M, Erlund I, Lee H-S, et al.; TEDDY Study Group. Plasma ascorbic acid and the risk of islet autoimmunity and type 1 diabetes: the TEDDY study. Diabetologia 2020;63:278–286

26. Skogerson K, Wohlgemuth G, Barupal DK, Fiehn O. The volatile compound BinBase mass spectral database. BMC Bioinformatics 2011;12:321

27. Kind T, Liu K-H, Lee DY, DeFelice B, Meissen JK, Fiehn O. LipidBlast in silico tandem mass spectrometry database for lipid identification. Nat Methods 2013;10: 755–758

28. Fan S, Kind T, Cajka T, et al. Systematic error removal using random forest for normalizing large-scale untargeted lipidomics data. Anal Chem 2019;91:3590–3596

29. Li Q, Fisher K, Meng W, et al. GMSimpute: a generalized two-step Lasso approach to impute missing values in label-free mass spectrum analysis. Bioinformatics 2020;36:257–263

30. Stewart CJ, Ajami NJ, O'Brien JL, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature 2018;562:583–588

31. La Torre D, Seppänen-Laakso T, Larsson HE, et al.; DiPiS Study Group. Decreased cord-blood phospholipids in young age-at-onset type 1 diabetes. Diabetes 2013;62:3951–3956

32. Kulkarni H, Mamtani M, Wong G, et al. Genetic correlation of the plasma lipidome with type 2 diabetes, prediabetes and insulin resistance in Mexican American families. BMC Genet 2017;18:48

33. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J 2016;8:289–317

34. Smyth GK. limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, Eds. New York, Springer, 2005, pp. 397–420

35. Barupal DK, Fiehn O. Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets. Sci Rep 2017;7:14567

36. Benjamini Y, Yekutieli D. False discovery rate–adjusted multiple confidence intervals for selected parameters. J Am Stat Assoc 2005;100:71–81

37. Niewczas MA, Mathew AV, Croall S, et al. Circulating modified metabolites and a risk of ESRD in patients with type 1 diabetes and chronic kidney disease. Diabetes Care 2017;40:383–390

38. Solini A, Manca ML, Penno G, Pugliese G, Cobb JE, Ferrannini E. Prediction of declining renal function and albuminuria in patients with type 2 diabetes by metabolomics. J Clin Endocrinol Metab 2016;101:696–704

39. Behrens WA, Madere R. Vitamin C and vitamin E status in the spontaneously diabetic BB rat before the onset of diabetes. Metabolism 1991;40:72–76

40. Bergsten P, Moura AS, Atwater I, Levine M. Ascorbic acid and insulin secretion in pancreatic islets. J Biol Chem 1994;269:1041–1045

41. Tapia G, Mårild K, Dahl SR, et al. Maternal and newborn vitamin D-binding protein, vitamin D levels, vitamin D receptor genotype, and childhood type 1 diabetes. Diabetes Care 2019;42:553–559

42. Munger KL, Levin LI, Massa J, Horst R, Orban T, Ascherio A. Preclinical serum 25-hydroxyvitamin D levels and risk of type 1 diabetes in a cohort of US military personnel. Am J Epidemiol 2013;177:411–419

43. Simpson M, Brady H, Yin X, et al. No association of vitamin D intake or 25-hydroxyvitamin D levels in childhood with risk of islet autoimmunity and type 1 diabetes: the Diabetes Autoimmunity Study in the Young (DAISY). Diabetologia 2011;54:2779–2788

44. Mäkinen M, Mykkänen J, Koskinen M, et al. Serum 25-hydroxyvitamin D concentrations in children progressing to autoimmunity and clinical type 1 diabetes. J Clin Endocrinol Metab 2016;101:723–729

45. Thorsen SU, Mårild K, Olsen SF, et al. Lack of association between maternal or neonatal vitamin D status and risk of childhood type 1 diabetes: a Scandinavian case-cohort study. Am J Epidemiol 2018;187:1174–1181

46. Thorsen SU, Pipper CB, Johannesen J, Mortensen HB, Pociot F, Svensson J. '25-Hydroxyvitamin D, autoantigenic and total antibody concentrations: results from a Danish case-control study of newly diagnosed patients with childhood type 1 diabetes and their healthy siblings'. Scand J Immunol 2018;87:46–53

47. Mäkinen M, Löyttyniemi E, Koskinen M, et al. Serum 25-hydroxyvitamin D concentrations at birth in children screened for HLA-DQB1 conferred risk for type 1 diabetes. J Clin Endocrinol Metab 2019;104:2277–2285

48. Vehik K, Lynch KF, Schatz DA, et al.; TEDDY Study Group. Reversion of β-cell autoimmunity changes risk of type 1 diabetes: TEDDY study. Diabetes Care 2016;39:1535–1542

49. Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. JAMA 2013;309:2473–2479

50. Krischer JP, Liu X, Vehik K, et al.; TEDDY Study Group. Predicting islet cell autoimmunity and type 1 diabetes: an 8-year TEDDY study progress report. Diabetes Care 2019;42:1051–1060