Gut microbiome shifts in people with type 1 diabetes are associated with glycemic control – an INNODIA study

Tommi Vatanen[1,2,3,4,5†], Carine de Beaufort[6,7], M. Loredana Marcovecchio[8], Lut Overbergh[9], Soren Brunak[10], Mark Peakman[11], Chantal Mathieu[12], Mikael Knip[1,13,14†] on behalf of the INNODIA consortium*

*Membership of the INNODIA Consortium is provided in the electronic supplementary material

[1]Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, Finland
[2]Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland
[3]Department of Microbiology, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland
[4]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[5]Liggins Institute, University of Auckland, Auckland, New Zealand
[6]Pediatric Endocrinology and Diabetology (DECCP), Centre Hospitalier de Luxembourg, Luxembourg, Luxembourg
[7]Faculty of Science, Technology and Medicine, University of Luxembourg-Belval Campus, Esch-sur-Alzette, Luxembourg
[8]Department of Paediatrics, University of Cambridge, Cambridge, UK
[9]Katholieke Universiteit Leuven/ Universitaire Ziekenhuizen, Leuven, Belgium
mmunology & Inflammation Research Therapeutic Area, Sanofi, MA, USA
[10]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[11]Immunology & Inflammation Research Therapeutic Area, Sanofi, MA, USA
[12]Department of Chronic Diseases and Metabolism, Endocrinology, Katholieke Universiteit Leuven, Leuven
[13]New Children's Hospital, Helsinki University Hospital, Helsinki, Finland
[14]Tampere Center for Child Health Research, Tampere University Hospital, Tampere, Finland

ORCID identifier
TV   0000-0003-0949-1291
CdB   0000-0003-4310-6799
MLM   0000-0002-4415-316X
LU   0000-0001-7126-356X
SB   0000-0003-0316-5866
MP   0000-0003-3328-3513
CM   0000-0002-4055-5233
MK   0000-0003-0474-0033

†Corresponding authors:
Tommi Vatanen
Department of Microbiology
Faculty of Agriculture and Forestry
FI-00140 University of Helsinki
Helsinki, Finland
E-mail: tommi.vatanen@helsinki.fi

Mikael Knip
Research Programs Unit
Faculty of Medicine
FI-00140 University of Helsinki
Helsinki, Finland
E-mail: mikael.knip@helsinki.fi

Abstract

Aims The gut microbiome is implicated in the disease process leading to clinical type 1 diabetes, but less is known about potential changes in the gut microbiome after the diagnosis of type 1 diabetes and implications in glucose homeostasis. We aimed to analyze potential associations between the gut microbiome composition and clinical and laboratory data during a 2-year follow up of people with newly diagnosed type 1 diabetes, recruited to the INNODIA study. In addition, we analysed the microbiome composition in initially unaffected family members, who progressed to clinical type 1 diabetes during or after their follow-up for 4 years.

Methods We characterized the gut microbiome composition of 98 individuals with newly diagnosed type 1 diabetes (ND cohort) and 194 autoantibody-positive unaffected family members (UFM cohort), representing a subgroup of the INNODIA Natural History study, using metagenomic sequencing. Participants from the ND cohort attended study visits within 6 weeks from the diagnosis and 3, 6, 12 and 24 months later for stool sample collection and laboratory tests (HbA$_{1c}$, C-peptide, diabetes-associated autoantibodies). Participants from the UFM cohort were assessed at baseline, 6, 12, 18, 24 and 36 months later.

Results We observed a longitudinal increase in 21 bacterial species in the ND but not in the UFM cohort. The relative abundance of *Faecalibacterium prausnitzii* was inversely associated with the HbA$_{1c}$ levels at diagnosis (p=0.0019). The rate of the subsequent disease progression in the ND cohort, as assessed by change in HbA$_{1c}$, C-peptide levels and insulin dose, was associated with the abundance of several bacterial species. Individuals with rapid decrease in C-peptide levels in the ND cohort had the lowest gut microbiome diversity. Nineteen individuals who were diagnosed with type 1 diabetes in the UFM cohort had increased abundance of *Sutterella sp. KLE1602* compared to the undiagnosed UFM individuals (p=1.2×10^{-4}).

Conclusions/interpretation Our data revealed associations between the gut microbiome composition and the disease progression in patients with recent onset type 1 diabetes. Future mechanistic studies as well as animal studies and human trials are needed to further validate the significance and causality of these associations.

Keywords Gut microbiome • Newly diagnosed • First-degree relatives • *Faecalibacterium prausnitzii* • HbA$_{1c}$ • C-peptide

Research in context

What is already known about this subject?
- The gut microbiome is implicated in the disease process leading to clinical type 1 diabetes

What is the key question?
- Does the gut microbiome contribute to the post-diagnosis disease progression in type 1 diabetes?

What are the new findings?
- Twelve bacterial species consistently increased during the first 2 years following the diagnosis of type 1 diabetes
- *Faecalibacterium prausnitzii* was inversely associated with the HbA_{1c} levels at the diagnosis of type 1 diabetes
- Individuals with rapidly decreasing C-peptide levels had the lowest gut microbiome diversity

How might this impact on clinical practice in the foreseeable future?
- Modifying the gut microbiome through pre-, pro- and postbiotics may provide means for slowing down the disease process and improving glycemic outcomes in type 1 diabetes

Tweet (X post):
Results from the multicentre INNODIA type 1 diabetes study: Fecal *Feacalibacterium prausnitzii* abundance inversely correlated with HbA1c levels at diagnosis in newly diagnosed patients.

X handles:
@tvatanen (Tommi Vatanen)
@Innodiagroup (INNODIA Study group)
@helsinkiuni (University of Helsinki)
@HelsinkiUniMed (Medicine and Health, University of Helsinki)
@HiLIFE_helsinki (Helsinki Institute of Life Science)

Abbreviations:
AAB	Autoantibody/bodies
CI	Confidence interval
ESM	Electronic supplementary material
FDR	False discovery rate
FMT	Faecal microbiome transplant
GMM	Gut Metabolic Modules
IA	Insulin antibodies
IAA	Insulin autoantibodies
IA-2A	Islet antigen-2 antibodies
*INNODIA	Innovative approaches to understanding and arresting type 1 diabetes
KO	KEGG Orthology identifier
LMM	Linear mixed methods model
MAPQ	Mapping quality
MGS	Metagenomic species
MMTT	Mixed meal tolerance test
ND	Newly diagnosed type 1 diabetes patients
PERMANOVA	Permutational multivariate analysis of variance
SCFA	Short-chain fatty acids
SDS	Standard deviation score
UFM	Unaffected family member

Introduction

Mounting evidence indicates an intricate relationship between the gut microbiome (GM) – a complex consortium of microbes inhabiting the lower gastrointestinal tract – and type 1 diabetes. Both human cohort studies [1–7] and controlled animal experiments [8–12] have demonstrated that the GM harbours both protective and harmful features that may influence disease process leading to type 1 diabetes. For example, microbially produced short chain fatty acids (SCFA) provided protection from type 1 diabetes [1, 7, 12, 13]. Further complicating such analyses, HLA allele combinations providing increased risk for type 1 diabetes also result in changes in the GM composition, potentially through host regulation and selection [14]. Recent and ongoing trials are exploring the possibility to modify the microbiome-linked disease risk through various interventions targeting the GM [13, 15].

The GM is also implicated in the natural course of type 1 diabetes after the diagnosis [13, 16–18]. For example, existing evidence suggests that GM is involved in regulating host glycaemic control [13, 19]. An exploratory faecal microbiome transplant (FMT) trial in patients with recent onset type 1 diabetes found that autologous FMT halted the decline in endogenous insulin secretion [20]. Together, these emerging data suggest that GM modifications could provide possibilities to intervene in the disease process and even slow down its progression. To this end, a combination of exploratory analyses and controlled experiments are needed to identify microbial strains, metabolites and other GM features that are implicated in host physiology and type 1 diabetes-related biomarkers.

Here, we analysed the gut microbiome profiles for associations with clinical and laboratory data from people with newly diagnosed type 1 diabetes and unaffected autoantibody-positive family members participating in the European, multi-centre INNODIA Natural History Study [21] during a follow up period for 2-4 years. INNODIA has collected rich information and clinical data on the participants, including fasting C-peptide and HbA_{1c} measurements, to assess endogenous insulin production and glycaemic control, respectively. We report associations between host glycaemic control, diabetes progression and GM features.

Material and Methods

Study population. This study recruited two cohorts from the large INNODIA Natural History study [21]. The first cohort comprised of individuals with newly diagnosed type 1 diabetes (ND) and the other cohort autoantibody-positive unaffected family members of individuals with type 1 diabetes (UFM) (**Table 1**). INNODIA selected the first 100 ND cohort individuals for in-depth molecular assays, including gut microbiome profiling. Stool samples were available for 98 ND cohort participants. The UFM cohort was included to assess possible GM shifts during the asymptomatic stage of type 1 disease. UFM participants in this study were selected based on stool sample availability. A minority of the UFM participants (10/194; 5.2%) in this study were related to the ND patients. Both ND and UFM participants tested positive for at least one diabetes-associated autoantibody out of the three analysed [GADA, islet antigen-2 antibodies (IA-2A) and ZnT8A]. Insulin antibodies (IA) were measured after the ND patients had already received exogenous insulin which is known to induce insulin antibody production, and these antibodies are not distinguishable from insulin autoantibodies (IAA) by the assay used. BMI was calculated as the weight in kilograms divided by the square of the height in meters. Age- and sex-appropriate standard deviation scores (SDSs) were calculated using World Health Organization 2006 & 2007 data [22]. A harmonized protocol for sample collection and storage was used in the study centres [21]. The study followed the guidelines of the Declaration of Helsinki for research on human participants, and the study protocols were approved by the ethical committees of the

4

participating hospitals. Either the parent or participants themselves gave their written informed consent.

Stool sample collection. Stool samples were collected from the INNODIA ND cohort at baseline (within 6 weeks after the diagnosis of type 1 diabetes) and at 3, 6, 12 and 24 months later. In the INNODIA UFM cohort, stool samples were collected at 6, 12, 18, 24 and 36 months after the screening visit (baseline). The samples were collected in OMNIgene-Gut OMR-200 collection tubes (DNA Genotek, Ottawa, Ontario, Canada), which stabilizes DNA at ambient temperature for up to 60 days. Participants were asked to collect the stool samples at home during the week preceding the next study centre visit and to bring the sample with them to the study centre. The samples were frozen at -80 °C in the local study centre and shipped frozen to the INNODIA Biobank for storage at -80 °C until analysed.

DNA extraction. DNA was extracted from 250 µL aliquots of the fecal samples in OMNIgene-Gut OMR-200 collection tubes using the NucleoSpin® 96 Soil (Macherey-Nagel, Hoerdt, France) kit. Bead beating was done horizontally on a Vortex-Genie 2 at 2700 rpm for 5 min. One negative control and one positive control (ZymoBIOMICS Microbial Community Standard, Zymo Research, Irvine, CA, USA) were included per batch of samples from the DNA extraction and throughout the laboratory process.

DNA sequencing. The quality of extracted DNA was evaluated using agarose gel electrophoresis. The quantity of DNA was measured by Qubit 2.0 fluorometer. DNA was randomly sheared into fragments of 350 bp, on average. Sequencing libraries were constructed using NEBNext Ultra Library Prop Kit (New England Biolabs, Herts, UK). The quality of the DNA libraries was measured using Qubit 2.0 fluorometer and Agilent 2100 Bioanalyzer to assess the fragment size distribution. The DNA concentration of the final library was determined by quantitative real-time PCR (qPCR) prior to sequencing. Metagenomic DNA was sequencing using a 2x150 bp paired-end protocol on an Illumina platform

Sequence data quality control. Quality control of sequencing reads was conducted using KneadData (v0.6.1) to remove low-quality reads and trim low-quality bases. Trimmed reads shorter than 100 bases and reads mapping to human genome GRCh38 were discarded.

Reference gene catalog. The Clinical Microbiomics Human Gut HG04 gene catalog, consisting of 14 355 839 genes, was used as a reference gene catalog. The catalog is based on 12 170 non-public human gut samples, 9428 publicly available metagenomes from 43 countries [24] and 3567 publicly available genome assemblies from isolated microbial strains. Taxonomic abundance profiles were obtained using the Clinical Microbiomics HGMGS version HG4.D.1 set of 2095 metagenomic species (MGS), each represented by a set of genes with highly coherent abundance profiles and base compositions in the 12 170 metagenomes. The metagenomic species concept is described in [25]. Quality controlled reads were mapped to the gene catalog using BWA mem (v. 0.7.16a) algorithm [26] with following criteria: mapping quality ≥ 20, sequence identity ≥ 95% over ≥100bp.

Metagenomic species annotations. An MGS was taxonomically annotated by a BLAST search of its genes against the NCBI RefSeq genome database (2020-01-27) combined with rank-specific annotation criteria. An MGS was assigned to a taxon if at least M % of its genes were mapped to the taxon and no more than D % of its genes were mapped to a different taxon. We only considered blast hits with an alignment length ≥ 100 bp, ≥ 50 % query coverage and % identity ≥ PID. Parameters M, D and PID were defined for subspecies, species, genus, family, order, class, phylum, and superkingdom as follows: PID = (95, 95,

85, 75, 65, 55, 50, 45); M = (75, 75, 60, 50, 40, 30, 25, 20); and D = (10, 10, 10, 20, 20, 20, 20, 15), respectively. Finally, each MGS was processed with CheckM [27], and the annotation was updated with the CheckM result if this resulted in a lower taxonomic rank.

MGS abundance calculation. MGS Abundances were calculated as described in [25]. Briefly, an abundance of each MGS for estimated using gene abundances of 100 signature genes optimized for accurate abundance profiling. MGS abundances were normalized to account for length of the signature genes and the relative abundance of the data.

Functional annotation and profiling. We measured microbiome functional modules using The Gut Metabolic Modules (GMMs) which consist of 103 conserved metabolic pathways, each defined as a series of enzymatic steps represented by KEGG Orthology (KO) identifiers [28]. Abundance of a functional module was calculated as the proportion of all mapped reads that mapped to a KOs belonging to the given module.

Diversity estimation. Alpha and beta diversities were estimated using rarefied MGS abundances. We used the number of entities detected (richness) and Shannon's index to measure alpha diversity. Beta diversity was calculated using the Bray–Curtis dissimilarity.

Statistical analysis. Associations between microbial alpha-diversities and clinical covariates were tested using linear models. Associations between microbial beta-diversities and clinical covariates were tested with permutational multivariate analysis of variance (PERMANOVA) as implemented in vegan R package. Associations between individual microbial features (species, functional modules) were tested by linear mixed effect models in MaAsLin2 R package [29], assuming normally distributed data and normally distributed random intercepts per study subject. The statistical models included covariates to correct for effects related to the clinical site, gender, and age at diagnosis (age at recruitment for the UFM cohort). Statistical models involving longitudinal analyses additionally included covariates to correct for the time from baseline and subject specific random intercepts. All data analyses were conducted in R v4.0.0 in the INNODIA Cloud environment. False discovery rate (FDR) correction with all tests including multiple features was performed using Benjamini-Hochberg procedure and the resulting q-values (i.e. FDR corrected p-values) are reported when appropriate.

Results

We analysed 368 faecal samples collected from 98 individuals newly diagnosed with type 1 diabetes (ND cohort) and 492 faecal samples collected from 194 UFM participants (**Fig. 1, Table 1**). Participants were recruited across 25 clinical centres in 13 European countries (**Table 2**). The age of the ND participants varied from one to 38 years (mean 12.3, SD 8.64) and the cohort comprised 48 female and 50 male study participants. The average age at the diagnosis of was 12.3 years (SD 8.6) and the mean diabetes duration was 3.7 weeks (SD 1.6, minimum 0.7, maximum 6.3 weeks) at the first study visit. At baseline, the average total daily insulin dose was 0.52 IU/kg (SD 0.26), HbA$_{1c}$ 75.3 mmol/mol [SD 23.2; 9.0 (4.0) %], and the mean fasting C-peptide level 272 pmol/l (IQR 236). The mean plasma glucose reading at baseline was 7.73 mmol/l (IQR 2.77). The mean BMI SDS was 0.41 (SD 1.11). Following the results of a previous INNODIA study, we divided the ND cohort participants into three groups based on the change of C-peptide levels over time, depicting the rate of disease progression: rapid decline, slow decline and increasing [30]. The age of 194 UFM participants varied from one to 44 years (mean 21.2, SD 14.1). Clinical and demographic data of both cohorts are presented in **Table 1**. Stool samples were metagenomically

sequenced with an average depth of 9 Gb per sample corresponding to 30.0 million (M) read pairs per sample (Illumina 2×150 paired-end). On average 87.6 % of the high-quality microbiome reads from a sample were mapped to the gut gene catalogue.

We first analysed the metagenomes from the ND cohort to investigate links between disease biomarkers and the GM following the diagnosis of type 1 diabetes. The most prevalent and abundant GM species included several *Bifidobacterium* and *Bacteroides* species as well as *Faecalibacterium prausnitzii*, all common human GM members (**Fig. 2**). We observed a stark dichotomy in the relative abundance of *Prevotella copri* which was highly abundant in a subset of the metagenomes (N=88, 23.8%) and missing in others. *P. copri* presence was associated with change in abundance of 42 microbiome functional modules (Wilcoxon test, FDR corrected p<0.05, **ESM Table 1**).

We tested for associations between microbial alpha- and beta-diversities and demographic/clinical/biochemical data at baseline in the ND cohort. We detected shifts in microbial profiles (beta-diversities) between the clinical centres (PERMANOVA test, R^2=0.183, p=0.007). There were no detectable shifts or associations between microbial alpha- or beta-diversities and participants' gender, age, BMI SDS, insulin dose, HbA_{1c}, fasting C-peptide measurement or C-peptide/glucose ratio at baseline (alpha-diversities, linear mixed model, p>0.1; beta-diversities, PERMANOVA test, p>0.05). We observed associations between the abundance of 60 bacterial species and the clinical centres (q<0.20, **ESM Table 2**).

We next pooled data from the follow-up period to assess longitudinal microbiome changes over the visit schedule in the ND cohort. Twenty-one microbial species and eight functional modules showed longitudinal trends (linear mixed effects model, q<0.20, **Fig. 3a-b, ESM Table 2**). All these species had an increasing trend over time, and they included both common and rare species with high and low average relative abundances.

We tested for association between HbA_{1c} values at the time of the diagnosis of type 1 diabetes and the microbiome at the baseline visit (within 6 weeks after the diagnosis) in the ND cohort. The relative abundance of *Faecalibacterium prausnitzii* was inversely correlated with HbA_{1c} [linear mixed effects model, beta-coefficient=-0.15 (95% CI: -0.24, -0.058), q=0.19, nominal p=0.0019, **Fig. 4a, ESM Table 3**] and microbial hydrogen metabolism (Functional module MF0098) was positively associated with HbA_{1c} (beta-coefficient=2.43 (95% CI: 0.92, 3.94), q=0.17, nominal p=0.0020, **ESM Table 3**). Since patients with ketoacidosis at the diagnosis of type 1 diabetes had significantly higher HbA_{1c} (p=5.2×10⁻⁵; the average HbA_{1c} in patients with ketoacidosis 117.0 mmol/mol, in patients with no ketoacidosis 93.3 mmol/mol), we conducted sensitivity analysis controlling for ketoacidosis. Both associations above became less statistically significant (*F. prausnitzii*, q=0.29, nominal p=0.0031; hydrogen metabolism, q=0.35, nominal p=0.0077), indicating that ketoacidosis explained away a fraction of the associations. Ketoacidosis status alone was not associated to any GM features at baseline in the ND cohort.

We analysed possible associations between the clinical and laboratory data at the baseline visit (within 6 weeks from the diagnosis) and the baseline microbiome composition in the ND cohort. Tested variables included HbA_{1c} value, insulin dose per kg, fasting C-peptide concentration, fasting C-peptide/glucose ratio and BMI SDS. We did not observe any associations between baseline measurements and microbiome features. We also tested for associations between the number of detected autoantibodies (AABs) at baseline and the baseline microbiome composition. Insulin antibodies (IA) were excluded from the AAB count since the patients had already received exogenous insulin which is known to induce insulin

antibody production, and these antibodies are not distinguishable from IAA by the assay applied. We observed a positive correlation between *Coprococcus eutactus* and the number of detectable AABs at the baseline visit (beta-coefficient=0.41 (95% CI: 0.16, 0.66), q=0.16, nominal p=0.0014, **Fig. 4b**, **ESM Table 4**) and an inverse correlation between microbial tyrosine degradation and the number of AABs (beta-coefficient=-0.48 (95% CI: -0.77, -0.19), q=0.17, nominal p=0.0014, **ESM Table 4**). In addition, we analysed the data for associations between the change in clinical parameters (HbA$_{1c}$, insulin dose, fasting C-peptide) during the 2-year follow-up and the baseline microbiome composition. Relative abundance of *Blautia obeum* was inversely associated with change in fasting C-peptide between baseline and visit 5 (beta-coefficient=-0.25 (95% CI: -0.38, -0.12), q=0.07, **Fig. 5, ESM Table 5**).

We compared the microbiome composition between the three groups of participants in the ND cohort based on the the change of C-peptide levels over time and observed differences in *Streptococcus salivarius* (q=0.12), *Cambylobacter concisus* (q=0.15) and *Veillonella atypica* (q=0.19) abundances between the groups (**ESM Table 6**). Bacterial alpha-diversity, measured by richness (number of observed species), differed between the groups (repeated measures analysis of variance, corrected for clinical centres, age groups and an interaction between the age groups and progression groups, p=0.021) such that the individuals with rapid C-peptide decline had the lowest bacterial richness, on average (**Fig. 6a**). As there was a wide age range (1-38 years) within the ND cohort potentially affecting the gut microbiome composition, we further divided the patients in three age categories, age <7, age 7-12, and age 13 or older, based on distinct immunohistological profiles [31] and clinical characteristics [32]. Baseline clinical characteristics per age category are shown in **ESM Table 7**. We found that the association with C-peptide decline was most apparent in the group who were diagnosed with type 1 diabetes before the age of seven (**Fig. 6b-d**), although the interaction term between the age groups and the progression groups was not significant (p=0.77).

Nineteen individuals in the UFM cohort progressed to clinical type 1 diabetes during or after the follow up. We compared the GMs of these individuals to the rest of the UFM cohort. Individuals diagnosed with diabetes had increased abundance of *Sutterella sp. KLE1602* (longitudinal analysis, beta-coefficient=1.20 (95% CI: 0.59, 1.80), q=0.033, nominal p=1.2×10^{-4}, **ESM Table 7**) which was also more prevalent in these individuals: ten out of 19 patients (53%) diagnosed with type 1 diabetes had *Sutterella sp. KLE1602* in at least one stool sample compared to 27% (48 of 175) in non-diabetic UFM cohort participants. The progressors also had increased abundance of the functional module *pyruvate:ferredoxin oxidoreductase* (MF0073, beta-coefficient=0.018 (95% CI: 0.0092, 0.028), q=0.059, nominal p=9.8×10^{-5}, **ESM Table 7**) in their GMs. We did not observe any longitudinal GM changes or associations between HbA$_{1c}$ levels and GM features, nor associations between the stages of type 1 diabetes and microbial alpha diversities in the UFM cohort.

Discussion and Conclusions

Baseline relative abundance of *Faecalibacterium prausnitzii* was inversely correlated with HbA$_{1c}$ at the diagnosis of type 1 diabetes. *F. prausnitzii* is a beneficial commensal bacterium with important health-promoting functions [33] by, for example, production of butyrate which serves as an energy source for colonocytes. Microbial butyrate production is among the strongest mechanisms through which the GM mediates protection against islet autoimmunity and type 1 diabetes [7, 12, 34]. In a pilot trial supplementing resistant starch modified with SCFAs releasing resistant starch to humans with type 1 diabetes, subjects with the highest

SCFA concentrations exhibited the best glycaemic control [13]. Future human cohort studies with faecal and/or serum butyrate quantification could further investigate the role of butyrate in glycaemic control in type 1 diabetes.

We observed differences in microbial richness – the number of observed taxa – according to the disease progression in the ND cohort. Individuals with the fastest decline in C-peptide levels, reflecting decline in endogenous insulin production, had the lowest GM richness. We also observed an inverse association between *Blautia obeum* and the change in fasting C-peptide levels between baseline and 2 years after the diagnosis. The difference in microbial richness is compatible with the observations from the DIABIMMUNE cohort where the decline in microbial diversity appeared prior to the diagnosis of type 1 diabetes [6]. However, such pattern had been absent in most other human cohort studies suggesting that microbial diversity alone has a poor predictive value for type 1 diabetes. Together, these observations paint a picture where a multitude of microbial features may be weakly involved in the host-GM interplay in established disease.

UFM cohort members who had been diagnosed with type 1 diabetes during or after the follow-up had higher relative abundance and prevalence of *Sutterella sp. KLE1602* compared to other UFM cohort members. While *Sutterella spp.* are not as extensively studied as many other prevalent gut bacteria, a study in mice found that *Sutterella spp.* degraded immunoglobulin A (IgA) in the gut [35]. Authors showed that mice with *Sutterella spp.* and resulting low gut IgA levels had increased Dextran Sulfate Sodium (DSS)-induced colon ulceration compared to mice with normal IgA levels and no *Sutterella*. A randomized double-blind placebo-controlled clinical trial in humans testing the efficacy of fecal microbiome transplant in ulcerative colitis found that *Sutterella spp.* were enriched in patients who failed to stay in remission following the transplant [36]. Speculatively, Sutterella could disrupt the intestinal antibacterial immune response and indirectly induce local and systemic inflammation by increase in intracellular pathogens and pathobionts [37]. Potential roles of *Sutterella spp.* and colonic IgA in type 1 diabetes could be further investigated by, for example, including faecal IgA assays and sorting strategies that identify gut bacterial IgA coating in current and upcoming type 1 diabetes cohort studies.

We observed longitudinal GM changes over the 2-year observation period in the ND cohort. These changes included increases in known oral species *Bifidobacterium dentium* and *Streptococcus salivarius*, as well as several butyrate producing *Ruminococcaceae* species and *Coprococcus catus*. Such longitudinal shifts were absent in the UFM cohort, suggesting the changes in the ND cohort could be related to the manifestation of type 1 diabetes and subsequent lifestyle changes. For example, patients with newly diagnosed type 1 diabetes are recommended to adapt their dietary patterns and contents of their diet which might be reflected in the GM structure.

We found *Prevotella copri* in 88 (23.8%) metagenomes. Presence of *P. copri* tended to have significant effects in many functional modules encoded by the entire microbial community. *Prevotella*-rich microbiomes are more common in non-Western populations, while the GMs in Western countries are more commonly dominated by *Bacteroides* spp. [38]. Generally, high abundance of *Prevotella* is associated with plant-rich, high-fiber diet and overweight adults consuming a whole-grain diet for 6-weeks lost more weight if their GM harboured *Prevotella* [39]. Here, the observed dichotomy in *Prevotella* abundance could reflect long-term dietary habits and other external factors affecting the GM composition. While we did not observe any associations with *P. copri* (or other *Prevotella* spp.) and clinical data in our analyses, *Prevotella*-rich microbiomes could still provide means for stratifying the metagenomes to different community types in further exploratory analyses.

These data included prominent differences in the GM composition between recruiting clinical sites, reflected in beta-diversities and individual microbial species and functional modules. Such shifts are likely related to geographic, lifestyle, ethnic and other differences between the clinical sites distributed across 13 European countries. These factors are known confounders in microbiome studies since they often prominently correlate with GM structure and therefore pose a challenge to designing large GM studies in rare or otherwise difficult-to-study human conditions. Nonetheless, these data further underscore the necessity of controlling for these and other known confounding factors in GM studies. Specifically, these factors should be considered in trial design for best reproducibility and reliability.

Strengths This study investigated the GM of a relatively large cohort of 98 newly-diagnosed type 1 diabetes patients with a two-year longitudinal follow-up involving five visits to the clinic. On all clinic visits, harmonized clinical data on the participants, including fasting C-peptide and HbA_{1c} measurements, were collected. Additionally, stool samples were collected using a harmonized process, enabling robust comparisons between the study centres.

Limitations The GM data included prominent differences between the clinical sites distributed across 13 European countries. Such shifts are likely related to lifestyle, dietary, ethnic, and other differences between the populations but since this information has not been collected in INNODIA we were unable to further dissect the source of the observed variation. Among these, the diet is a major GM modulator and dietary changes following type 1 diabetes could plausibly induce shifts in the GM composition. This hypothesis cannot be investigated here due to the lack of information on participant diets. Stool samples and microbiome data were not available at all time points from all individuals. Even though we are not aware of any systematic stool sampling bias, in theory, any such systematic bias could skew the results. Stimulated peak C-peptide level or area under curve from MMTT are more accurate measures of beta cell function compared to fasting C-peptide used here, as stimulated C-peptide data was not available at baseline for all participants.

Conclusions This study reports associations between the GM composition and type 1 diabetes in an INNODIA multi-centre human cohort study. Albeit these associations were statistically weak, potential mechanisms behind these associations could be further investigated using animal models and immunological assays. Rigorous, controlled trials are needed to assess the roles of identified bacteria and their potential causality in type 1 diabetes.

Data Availability The data generated and analysed is person-sensitive and can be accessed in secure environments only. Access can be provided by application to the INNODIA Data Access Committee.

References

1. de Goffau MC, Luopajärvi K, Knip M, et al (2013) Fecal microbiota composition differs between children with β-cell autoimmunity and those without. Diabetes 62(4):1238–1244

2. Needell JC, Zipris D (2016) The role of the intestinal microbiome in type 1 diabetes pathogenesis. Curr Diab Rep 16(10):89

3. Davis-Richardson AG, Ardissone AN, Dias R, et al (2014) Bacteroides dorei dominates gut microbiome prior to autoimmunity in Finnish children at high risk for type 1 diabetes. Front Microbiol 5:678

4. Endesfelder D, zu Castell W, Ardissone A, et al (2014) Compromised gut microbiota networks in children with anti-islet cell autoimmunity. Diabetes 63(6):2006–2014

5. Maffeis C, Martina A, Corradi M, et al (2016) Association between intestinal permeability and fecal microbiota composition in Italian children with beta cell autoimmunity at risk for type 1 diabetes. Diabetes Metab Res Rev. https://doi.org/10.1002/dmrr.2790

6. Kostic AD, Gevers D, Siljander H, et al (2015) The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. Cell Host Microbe 17(2). https://doi.org/10.1016/j.chom.2015.01.001

7. Vatanen T, Franzosa EA, Schwager R, et al (2018) The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. Nature 562(7728):589–594

8. Wen L, Ley RE, Volchkov PY, et al (2008) Innate immunity and intestinal microbiota in the development of Type 1 diabetes. Nature 455(7216):1109–1113

9. Burrows MP, Volchkov P, Kobayashi KS, Chervonsky AV (2015) Microbiota regulates type 1 diabetes through Toll-like receptors. Proc Natl Acad Sci U S A 112(32):9973–9977

10. Girdhar K, Huang Q, Chow I-T, et al (2022) A gut microbial peptide and molecular mimicry in the pathogenesis of type 1 diabetes. Proc Natl Acad Sci U S A 119(31):e2120028119

11. Lv W, Graves DT, He L, et al (2020) Depletion of the diabetic gut microbiota resistance enhances stem cells therapy in type 1 diabetes mellitus. Theranostics 10(14):6500–6516

12. Mariño E, Richards JL, McLeod KH, et al (2017) Gut microbial metabolites limit the frequency of autoimmune T cells and protect against type 1 diabetes. Nat Immunol. https://doi.org/10.1038/ni.3713

13. Bell KJ, Saad S, Tillett BJ, et al (2022) Metabolite-based dietary supplementation in human type 1 diabetes is associated with microbiota and immune modulation. Microbiome 10(1):9

14. Russell JT, Roesch LFW, Ördberg M, et al (2019) Genetic risk for autoimmunity is associated with distinct changes in the human gut microbiome. Nat Commun 10(1):3621

15. Ziegler A-G, Arnolds S, Kölln A, et al (2021) Supplementation with Bifidobacterium longum subspecies infantis EVC001 for mitigation of type 1 diabetes autoimmunity: the GPPAD-SINT1A randomised controlled trial protocol. BMJ Open 11(11):e052449

16. Yuan X, Wang R, Han B, et al (2022) Functional and metabolic alterations of gut microbiota in children with new-onset type 1 diabetes. Nat Commun 13(1):6356

17. Luo S, Yue T, Liu Z, et al (2022) Gut microbiome and metabolic activity in type 1 diabetes: An analysis based on the presence of GADA. Front Endocrinol 13:938358

18. Mejía-León ME, Petrosino JF, Ajami NJ, Domínguez-Bello MG, de la Barca AMC (2014) Fecal microbiota imbalance in Mexican children with type 1 diabetes. Sci Rep 4(1):3814

19. Shilo S, Godneva A, Rachmiel M, et al (2022) The gut microbiome of adults with type 1 diabetes and Its association with the host glycemic control. Diabetes Care. https://doi.org/10.2337/dc21-1656

20. de Groot P, Nikolic T, Pellegrini S, et al (2020) Faecal microbiota transplantation halts progression of human new-onset type 1 diabetes in a randomised controlled trial. Gut. https://doi.org/10.1136/gutjnl-2020-322630

21. Dunger DB, Bruggraber SFA, Mander AP, et al (2022) INNODIA Master Protocol for the evaluation of investigational medicinal products in children, adolescents and adults with newly diagnosed type 1 diabetes. Trials 23(1):414

22. Cole TJ (1990) The LMS method for constructing normalized growth standards. Eur J Clin Nutr 44(1):45–60

23. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359

24. Pasolli E, Asnicar F, Manara S, et al (2019) Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell 176(3):649-662.e20

25. Nielsen HB, Almeida M, Juncker AS, et al (2014) Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol 32(8):822–828

26. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760

27. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25(7):1043–1055

28. Vieira-Silva S, Falony G, Darzi Y, et al (2016) Species-function relationships shape ecological properties of the human gut microbiome. Nat Microbiol 1(8):16088

29. Mallick H, Rahnavard A, McIver LJ, et al (2021) Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol 17(11):e1009442

30. Armenteros JJA, Brorsson C, Johansen CH, et al (2023) Multi-omics analysis reveals drivers of loss of β-cell function after newly diagnosed autoimmune type 1 diabetes: An INNODIA multicenter study. Preprint at https://www.medrxiv.org/content/10.1101/2023.03.22.23287261v1

31. Arif S, Leete P, Nguyen V, et al (2014) Blood and islet phenotypes indicate immunological heterogeneity in type 1 diabetes. Diabetes 63(11):3835–3845

32. Parviainen A, Härkönen T, Ilonen J, But A, Knip M, Finnish Pediatric Diabetes Register (2022) Heterogeneity of type 1 diabetes at diagnosis supports existence of age-related endotypes. Diabetes Care 45(4):871–879

33. Lopez-Siles M, Duncan SH, Garcia-Gil LJ, Martinez-Medina M (2017) Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J 11(4):841–852

34. Jacob N, Jaiswal S, Maheshwari D, et al (2020) Butyrate induced Tregs are capable of migration from the GALT to the pancreas to restore immunological tolerance during type-1 diabetes. Sci Rep 10(1):19120

35. Moon C, Baldridge MT, Wallace MA, et al (2015) Vertically transmitted faecal IgA levels determine extra-chromosomal phenotypic variation. Nature 521(7550):90–93

36. Paramsothy S, Kamm MA, Kaakoush NO, et al (2017) Multidonor intensive faecal microbiota transplantation for active ulcerative colitis: a randomised placebo-controlled trial. Lancet 389(10075):1218–1228

37. Hansen IS, Baeten DLP, den Dunnen J (2019) The inflammatory function of human IgA. Cell Mol Life Sci 76(6):1041–1055

38. Prasoodanan P K V, Sharma AK, Mahajan S, et al (2021) Western and non-western gut microbiomes reveal new roles of Prevotella in carbohydrate metabolism and mouth-gut axis. NPJ Biofilms Microbiomes 7(1):77

39. Christensen L, Vuholm S, Roager HM, et al (2019) Prevotella abundance predicts weight loss success in healthy, overweight adults consuming a whole-grain diet ad libitum: a post hoc analysis of a 6-wk randomized controlled trial. J Nutr 149(12):2174–2181

Tables

Table 1. Clinical and demographic data of the study participants from the ND and UFM cohorts. Insulin antibodies (IA) were measured after the patients had already received exogenous insulin which is known to induce insulin antibody production, and these antibodies are not distinguishable from IAA by the assay used.

	ND	UFM
Total participants (N)	98	194
Sex		
Female	48 (49%)	107 (55%)
Male	50 (51%)	87 (45%)
Age at diagnosis/recruitment		
Mean (SD)	12.3 (8.64)	21.2 (14.1)
Median (IQR) [min, max]	11 (8) [1, 38]	16 (26) [1, 44]
BMI SDS at baseline		
Mean (SD)	0.41 (1.11)	N/A
Median (IQR) [min, max]	0.4 (1.61) [-2.32, 2.72]	N/A
Glucose reading at baseline (mmol/mol)		
Mean (SD)	7.73 (4.68)	N/A
Median (IQR) [min, max]	6.4 (2.77) [3.6, 31.5]	N/A
HbA$_{1c}$ at diagnosis (mmol/mol)		
Mean (SD)	101 (26.9)	N/A
Median (IQR) [min, max]	104 (35.5) [4.2, 150]	N/A
HbA$_{1c}$ at diagnosis (%)		
Mean (SD)	11.4 (4.6)	N/A
Median (IQR)	11.7 (5.4)	N/A
HbA$_{1c}$ at baseline (mmol/mol)		
Mean (SD)	75.3 (23.2)	33.7 (4.61)
Median (IQR) [min, max]	78.5 (25.5) [8.7, 130]	34 (4.89) [5.2, 48]
HbA$_{1c}$ at baseline (%)		
Mean (SD)	9.0 (4.3)	5.3
Median (IQR)	9.3 (4.5)	5.3
Insulin dose at baseline (IU/kg)		
Mean (SD)	0.52 (0.26)	N/A
Median (IQR) [min, max]	0.53 (0.32) [0.036, 1.2]	N/A
Fasted C-peptide at baseline (pmol / l)		
Mean (SD)	272 (233)	N/A
Median (IQR) [min, max]	211 (236) [25.8, 1290]	N/A
GADA		
Negative	26 (27%)	74 (38%)
Positive	72 (73%)	120 (62%)
IA-2A		
Negative	30 (31%)	161 (83%)
Positive	68 (69%)	33 (17%)
IA		
Negative	24 (24%)	102 (53%)
Positive	74 (76%)	92 (47%)
ZnT8A		
Negative	35 (36%)	150 (77%)
Positive	63 (64%)	44 (33%)
Detectable autoantibodies		
Mean (SD)	2.83 (0.93)	1.49 (0.95)
Median (IQR) [min, max]	3 (2) [1, 4]	1 (1) [0, 4]

Table 2. Number of participants (N) per clinical centres and study cohorts.

Clinical centre	Country	N (ND)	N (UFM)
UH - University of Helsinki	Finland	29	22
ULI - University of Ljubljana	Slovenia	18	7
HH-RH - Herlev University Hospital	Denmark	17	6
UCAM - University of Cambridge	UK	8	10
MUG - Medical University of Graz	Austria	5	11
SUM - Slaski Uniwersytet Medyczny w Katowicach	Poland	4	9
ULB - Universite Libre de Bruxelles	Belgium	3	19
UOUL - Oulun Yliopisto	Finland	2	13
HKA - Hannoversche Kinderheilanstalt	Germany	2	12
CHL - Centre Hospitalier de Luxembourg	Luxembourg	2	9
UNISI - Universita degli Studi di Siena	Italy	2	6
UK - Norfolk and Norwich	UK	2	5
UK - Barts Health NHS Trust	UK	2	
KU Leuven	Belgium	1	21
UULM - Universitat Ulm	Germany	1	1
OUS - Oslo Universitetssykehus HF	Norway		13
Paris - Hospital Robert Debre	France		8
Lund University	Sweden		5
INSERM - Institut National de la Sante et de la Recherche Medicale	France		4
Paris - Centre Hospitalier Sud-Francilien	France		3
B - ZH Geel	Belgium		2
Paris - Hospital Jean-Verdier	France		2
UK - Northampton General Hospital	UK		2
UK-Birmingham Childrens Hospital	UK		2
IT- Ospedale Pediatrico Bambino Gesu	Italy		1
MUW - Medical University Vienna	Austria		1

Figure captions

Figure 1. Stool sample collection and metagenomes in the ND (N=98 patients) and UFM (N=194 individuals) cohorts within the INNODIA study. Stool samples were collected at home during the week preceding the next study centre visit. N shows the number of metagenomes generated per time point.

Figure 2. Heatmap displaying relative abundances of the 30 most abundant bacterial species in N=370 gut microbiome profiles from the participants in the INNODIA ND cohort. Row and column orders were determined by hierarchical clustering using Ward's clustering criterion.

Figure 3. Longitudinal changes in the ND cohort during the follow up. N=21 bacterial species had statistically significant longitudinal trends (linear mixed effects model, q<0.10).

Figure 4. Associations between gut bacteria at baseline and clinical covariates in the ND cohort.
a) Relative abundance of *Faecalibacterium prausnitzii* was inversely associated with HbA_{1c} values at the diagnosis of type 1 diabetes (linear model, q=0.19). **b**) Relative abundance of *Coprococcus eutactus* correlated with the number of AABs at the baseline visit (linear model, q=0.16).

Figure 5. Relative abundance of *Blautia obeum* was inversely associated with the change in fasting C-peptide concentration during the first two years of clinical type 1 diabetes (linear model, q=0.07) in the ND cohort.

Figure 6. Difference in microbial richness according to different C-peptide profiles during the follow up **a**) in all ND cohort participants and in three age groups, **b**) under 7, **c**) 7-12 and **d**) over 12 years old patients, within the ND cohort. Individuals were divided into three groups based on the decline of C-peptide levels over time, depicting the rate of disease progression: rapid and slow progression and increasing C-peptide levels [30].

	ND	UFM
Total participants (N)	98	194
Sex		
Female	48 (49%)	107 (55%)
Male	50 (51%)	87 (45%)
Age at diagnosis/recruitment		
Mean (SD)	12.3 (8.64)	21.2 (14.1)
Median (IQR) [min, max]	11 (8) [1, 38]	16 (26) [1, 44]
BMI SDS at baseline		
Mean (SD)	0.41 (1.11)	N/A
Median (IQR) [min, max]	0.4 (1.61) [-2.32, 2.72]	N/A
Glucose reading at baseline (mmol/mol)		
Mean (SD)	7.73 (4.68)	N/A
Median (IQR) [min, max]	6.4 (2.77) [3.6, 31.5]	N/A
HbA1c at diagnosis (mmol/mol)		
Mean (SD)	101 (26.9)	N/A
Median (IQR) [min, max]	104 (35.5) [4.2, 150]	N/A
HbA1c at diagnosis (%)		
Mean (SD)	11.4 (4.6)	N/A
Median (IQR)	11.7 (5.4)	N/A
HbA1c at baseline (mmol/mol)		
Mean (SD)	75.3 (23.2)	33.7 (4.61)
Median (IQR) [min, max]	78.5 (25.5) [8.7, 130]	34 (4.89) [5.2, 48]
HbA1c at baseline (%)		
Mean (SD)	9.0 (4.3)	5.3
Median (IQR)	9.3 (4.5)	5.3
Insulin dose at baseline (IU/kg)		
Mean (SD)	0.52 (0.26)	N/A
Median (IQR) [min, max]	0.53 (0.32) [0.036, 1.2]	N/A
Fasted C-peptide at baseline (pmol / l)		
Mean (SD)	272 (233)	N/A
Median (IQR) [min, max]	211 (236) [25.8, 1290]	N/A
GADA		
Negative	26 (27%)	74 (38%)
Positive	72 (73%)	120 (62%)
IA-2A		
Negative	30 (31%)	161 (83%)
Positive	68 (69%)	33 (17%)
IA		
Negative	24 (24%)	102 (53%)
Positive	74 (76%)	92 (47%)
ZnT8A		
Negative	35 (36%)	150 (77%)
Positive	63 (64%)	44 (33%)
Detectable autoantibodies		
Mean (SD)	2.83 (0.93)	1.49 (0.95)
Median (IQR) [min, max]	3 (2) [1, 4]	1 (1) [0, 4]

Clinical center	Country	N (ND)	N (UFM)
UH - University of Helsinki	Finland	29	22
ULI - University of Ljubljana	Slovenia	18	7
HH-RH - Herlev University Hospital	Denmark	17	6
UCAM - University of Cambridge	UK	8	10
MUG - Medical University of Graz	Austria	5	11
SUM - Slaski Uniwersytet Medyczny w Katowicach	Poland	4	9
ULB - Universite Libre de Bruxelles	Belgium	3	19
UOUL - Oulun Yliopisto	Finland	2	13
HKA - Hannoversche Kinderheilanstalt	Germany	2	12
CHL - Centre Hospitalier de Luxembourg	Luxembourg	2	9
UNISI - Universita degli Studi di Siena	Italy	2	6
UK - Norfolk and Norwich	UK	2	5
UK - Barts Health NHS Trust	UK	2	
KU Leuven	Belgium	1	21
UULM - Universitat Ulm	Germany	1	1
OUS - Oslo Universitetssykehus HF	Norway		13
Paris - Hospital Robert Debre	France		8
Lund University	Sweden		5
INSERM - Institut National de la Sante et de la Recherche Medical	France		4
Paris - Centre Hospitalier Sud-Francilien	France		3
B - ZH Geel	Belgium		2
Paris - Hospital Jean-Verdier	France		2
UK - Northampton General Hospital	UK		2
UK-Birmingham Childrens Hospital	UK		2
IT- Ospedale Pediatrico Bambino Gesu	Italy		1
MUW - Medical University Vienna	Austria		1

Figure 1
Page 21 of 74   Page 20 of 74

Newly Diagnosed (ND) cohort:

| T1D Diagnosis | Visit 1, baseline N=89 | Visit 2, 3 months N=85 | Visit 3, 6 months N=77 | Visit 4, 12 months N=61 | Visit 5, 24 months N=56 |

0-6 weeks

Unaffected Family Members (UFM) cohort:

| Visit 1, baseline N=165 | Visit 2, 6 months N=117 | Visit 3, 12 months N=88 | Visit 4, 18 months N=68 | Visit 5, 24 months N=42 | Visit 6, 36 months N=12 |

Bacteroides uniformis
Bacteroides vulgatus
Faecalibacterium prausnitzii
Subdoligranulum sp. APC924/74
[Eubacterium] rectale
Bacteroides ovatus
Fusicatenibacter saccharivorans
Blautia obeum
Clostridiaceae bacterium
Bifidobacterium longum subsp. longum
Clostridium sp. AF34−10BH
Ruminococcaceae bacterium
Prevotella copri
Bacteroides caccae
Alistipes shahii
Bacteroidaceae bacterium
Alistipes onderdonkii
Alistipes finegoldii
Alistipes putredinis
Parabacteroides distasonis
Gemmiger formicilis
Ruminococcaceae bacterium TF06−43
Roseburia faecis
Ruminococcus bromii
Bifidobacterium adolescentis
Ruminococcus bicirculans
Bacteroides dorei
Akkermansia muciniphila
Bacteroides fragilis
Dialister invisus

N=370 gut metagenomes from the INNODIA ND cohort

Bacterial species

- *Alistipes ihumii*
- *Alistipes indistinctus*
- *Alistipes shahii*
- *Bacteroidaceae bacterium*
- *Bacteroides massiliensis*
- *Bacteroides nordii*
- *Bifidobacterium dentium*
- *Blautia sp. SG772*
- *Coprococcus catus*
- *Coprococcus comes*
- *Dorea sp. AF24 7LB*
- *Eubacteriaceae bacterium*
- *Holdemania sp. Marseille P2844*
- *Odoribacter splanchnicus*
- Paraprevotella.clara
- Ruminococcaceae.bacterium.AM28.23LB
- Ruminococcaceae.bacterium.D5
- Ruminococcaceae.bacterium.TF06.43
- Ruminococcus.sp..AF41.9
- Streptococcus.salivarius
- Victivallales.bacterium.CCUG.44730

a

b

Electronic supplementary material for

Gut microbiome shifts in people with type 1 diabetes are associated with glycemic control – an INNODIA study

Members of the INNODIA and INNODIA HARVEST consortia

Consortium members

Mathieu	Chantal	KU Leuven, Belgium
Gillard	Pieter	KU Leuven, Belgium
Casteels	Kristina	KU Leuven, Belgium
Overbergh	Lutgart	KU Leuven, Belgium
Dunger	David	University of Cambridge, United Kingdom
Wallace	Chris	University of Cambridge, United Kingdom
Evans	Mark	University of Cambridge, United Kingdom
Thankamony	Ajay	University of Cambridge, United Kingdom
Hendriks	Emile	University of Cambridge, United Kingdom
Bruggraber	Sylvaine	University of Cambridge, United Kingdom
Marcoveccchio	Loredana	University of Cambridge, United Kingdom
Peakman	Mark	King's College London, United Kingdom
Tree	Timothy	King's College London, United Kingdom
Morgan	Noel G.	University of Exeter, United Kingdom
Richardson	Sarah	University of Exeter, United Kingdom
Todd	John A	University of Oxford, United Kingdom
Wicker	Linda	University of Oxford, United Kingdom
Mander	Adrian	Cardiff University, United Kingdom
Dayan	Colin	Cardiff University, United Kingdom
Alhadj Ali	Mohammad	Cardiff University, United Kingdom
Pieber	Thomas	Medical University of Graz, Austria
Eizirik	Decio L.	Universite Libre de Bruxelles, Belgium
Cnop	Miriam	Universite Libre de Bruxelles, Belgium
Brunak	Søren	University of Copenhagen, Denmark
Pociot	Flemming	Herlev University Hospital, Region Hovedstaden, Denmark
Johannesen	Jesper	Herlev University Hospital, Region Hovedstaden, Denmark
Rossing	Peter	Herlev University Hospital, Region Hovedstaden, Denmark
Legido Quigley	Cristina	Herlev University Hospital, Region Hovedstaden, Denmark
Mallone	Roberto	Cochin Institute Paris, France
Scharfmann	Raphael	Cochin Institute Paris, France
Boitard	Christian	Cochin Institute Paris, France
Knip	Mikael	University of Helsinki, Finland
Otonkoski	Timo	University of Helsinki, Finland
Veijola	Riitta	University of Oulu, Finland
Lahesmaa	Riitta	University of Turku, Finland
Oresic	Matej	University of Turku, Finland
Toppari	Jorma	University of Turku, Finland

Danne	Thomas	Children's and Youth Hospital Hannover, Germany
Ziegler	Anette G.	Helmholtz Zentrum Muenchen, Germany
Achenbach	Peter	Helmholtz Zentrum Muenchen, Germany
Rodriguez-Calvo	Teresa	Helmholtz Zentrum Muenchen, Germany
Solimena	Michele	TU Dresden, Germany
Bonifacio	Ezio E.	TU Dresden, Germany
Speier	Stephan	TU Dresden, Germany
Holl	Reinhard	University of Ulm, Germany
Dotta	Francesco	University of Siena, Italy
Chiarelli	Francesco	University of Chieti, Italy
Marchetti	Piero	University of Pisa, Italy
Bosi	Emanuele	University Vita-Salute San Raffaele, Italy
Cianfarani	Stefano	Bambino Gesù Children's Hospital, Italy
Ciampalini	Paolo	Bambino Gesù Children's Hospital, Italy
De Beaufort	Carine	Centre Hospitalier de Luxembourg, Luxemburg
Dahl-Jørgensen	Knut	Oslo University Hospital, Norway
Skrivarhaug	Torild	Oslo University Hospital, Norway
Joner	Geir	Oslo University Hospital, Norway
Krogvold	Lars	Oslo University Hospital, Norway
Jarosz-Chobot	Przemka	Medical University of Silesia, Poland
Battelino	Tadej	University of Ljubljana, Slovenia
Thorens	Bernard	University of Lausanne, Switzerland
Gotthardt	Martin	Radboud University Medical Center, The Netherlands
Roep	Bart O.	Leiden University Medical Center, The Netherlands
Nikolic	Tanja	Leiden University Medical Center, The Netherlands
Zaldumbide	Arnaud	Leiden University Medical Center, The Netherlands
Lernmark	Ake	Lund University, Sweden
Lundgren	Marcus	Lund University, Sweden
Costacalde	Guillaume	Univercell-Biosolutions, France
Strube	Thorsten	Sanofi, Germany
Schulte	Anke M.	Sanofi, Germany
Nitsche	Almut	Sanofi, Germany
Peakman	Mark	Sanofi, United States
Vela	Jose	Sanofi, United States
Von Herrath	Matthias	Novo Nordisk, Denmark
Wesley	Johnna	Novo Nordisk, Denmark
Napolitano-Rosen	Antonella	GlaxoSmithKline, United Kingdom
Thomas	Melissa	Eli Lilly, United Kingdom
Schloot	Nanette	Eli Lilly, United Kingdom
Goldfine	Allison	Novartis Pharma AG, Switzerland
Waldron-Lynch	Frank	Novartis Pharma AG, Switzerland
Kompa	Jill	Novartis Pharma AG, Switzerland
Vedala	Aruna	Novartis Pharma AG, Switzerland
Hartmann	Nicole	Novartis Pharma AG, Switzerland
Nicolas	Gwenaelle	Novartis Pharma AG, Switzerland
van Rampelbergh	Jean	Imcyse SA, Belgium
Bovy	Nicolas	Imcyse SA, Belgium
Dutta	Sanjoy	Juvenile Diabetes Research Foundation, United States
Soderberg	Jeannette	Juvenile Diabetes Research Foundation, United States
Ahmed	Simi	Juvenile Diabetes Research Foundation, United States
Martin	Frank	Juvenile Diabetes Research Foundation, United States
Latres	Esther	Juvenile Diabetes Research Foundation, United States

| Agiostratidou | Gina | The Leona M. and Harry B. Helmsley Charitable Trust, United States |
| Koralova | Anne | The Leona M. and Harry B. Helmsley Charitable Trust, United States |

2. Associated clinical sites

Willemsen	Ruben	Barts Health NHS Trust, United Kingdom
Smith	Anne	Northampton General Hospital NHS Trust, United Kingdom
Anand	Binu	West Suffolk NHS FT, United Kingdom
Datta	Vipan	Norfolk & Norwich University NHS FT, United Kingdom
Puthi	Vijith	North West Anglia NHS FT, United Kingdom
Zac-Varghese	Sagen	East & North Hertfordshire NHS Trust, United Kingdom
Dias	Renuka	Birmingham Women's and Children's NHS FT, United Kingdom
Sundaram	Premkumar	University Hospitals of Leicester NHS Trust, United Kingdom
Vaidya	Bijay	Royal Devon & Exeter NHS FT, United Kingdom
Patterson	Catherine	NHS Fife, United Kingdom
Owen	Katharine	Oxford University Hospitals NHS FT, United Kingdom
Dayan	Colin	Cardiff & Vale University Health Board, United Kingdom
Piel	Barbara	Queen Elizabeth Hospital, King's Lynn FT, United Kingdom
Heller	Simon	Sheffield Teaching Hospitals NHS FT, United Kingdom
Randell	Tabitha	Nottingham University Hospitals NHS Trust, United Kingdom
Gazis	Tasso	Nottingham University Hospitals NHS Trust, United Kingdom
Bismuth Reisman	Elise	Hospital Robert Debre, France
Carel	Jean-Claude	Hospital Robert Debre, France
Riveline	Jean-Pierre	Hospital Lariboisiere, France
Gautier	Jean-Francois	Hospital Lariboisiere, France
Andreelli	Fabrizio	Hospital Lapitie-Salpetriere, France
Travert	Florence	Hospital Bichat Claude Bernard, France
Cosson	Emmanuel	Hospital Jean-Verdier, France
Penfornis	Alfred	Centre Hospitalier Sud-Francilien, France
Petit	Catherine	Centre Hospitalier Sud-Francilien, France
Feve	Bruno	Hospital St Antoine, France
Lucidarme	Nadine	Hospital Jean-Verdier Pediatrie, France
Cosson	Emmanuel	Hospital Avicenne, France
Beressi	Jean-Paul	Hospital Andre Mignot, France
Ajzenman	Catherina	Hospital Andre Mignot Pediatrie, France
Radu	Alina	Hospital Europeen Georges-Pompidou, France
Greteau-Hamoumou	Stephanie	Hospital Louis Mourier, France
Bibal	Cecile	Hospital Kremlin Bicetre, France
Meissner	Thomas	Universitatsklinikum der Heinrich-Heine-Univeritat Dusseldorf, Germany
Heidtmann	Bettina	Katholisches Kinderkrankenhaus Wilhelmstift, Germany
Toni	Sonia	AOU Meyer, Italy
Rami-Merhar	Birgit	Medical University of Vienna, Austria
Eeckhout	Bart	Algemeen Ziekenhuis Geel Sint-Dimpna Geel, Belgium
Peene	Bernard	Algemeen Ziekenhuis Geel Sint-Dimpna Geel, Belgium
Vantongerloo	N	Algemeen Ziekenhuis Geel Sint-Dimpna Geel, Belgium
Maes	Toon	Imeldziekenhuis Bonheiden, Belgium
Gommers	Leen	Imeldziekenhuis Bonheiden, Belgium

ESM Table 1. Associations between *Prevotella copri* presence and the abundance of microbiome functional modules in the ND cohort. Associations were tested using Wilcoxon-test between samples with or without *Prevotella copri*. Nominal p-values were corrected for multiple testing using Holm's method.

Funtional module	Nominal p-value	FDR corrected p-value
lactaldehyde degradation	6.04161E-33	6.16245E-31
Entner-Doudoroff pathway	3.3406E-31	3.374E-29
maltose degradation	6.60922E-30	6.60922E-28
xylose degradation	7.96503E-29	7.88538E-27
glutamate degradation II	1.00707E-28	9.86929E-27
nitrate reduction (dissimilatory)	2.6296E-28	2.55071E-26
glutamine degradation I	2.0182E-26	1.93747E-24
pentose phosphate pathway (oxidative phas	4.81301E-26	4.57235E-24
threonine degradation II	7.14526E-26	6.71654E-24
butyrate production I	9.25355E-26	8.6058E-24
Bifidobacterium shunt	2.92281E-25	2.68899E-23
lactate production	3.63801E-25	3.31058E-23
proline degradation	7.99054E-25	7.19149E-23
histidine degradation	1.23247E-24	1.0969E-22
asparagine degradation	1.51143E-24	1.33006E-22
arabinose degradation	1.09322E-21	9.51105E-20
pectin degradation I	1.51829E-16	1.30573E-14
tryptophan degradation	4.60567E-16	3.91482E-14
lysine degradation II	8.51059E-10	7.1489E-08
fructose degradation	5.63414E-09	4.67634E-07
urea degradation	8.5332E-09	6.99722E-07
serine degradation	8.95387E-09	7.25264E-07
methanogenesis from carbon dioxide	7.49908E-07	5.99927E-05
methanol conversion	9.06862E-07	7.16421E-05
methanogenesis - methyl-coM	9.06862E-07	7.16421E-05
sucrose degradation II	1.70735E-06	0.000131466
glycerol degradation III	2.49388E-06	0.000189535
homoacetogenesis	4.30778E-06	0.000323083
arginine degradation IV	7.76152E-06	0.000574353
ethanol production I	8.15004E-06	0.000594953
methionine degradation I	0.00001259	0.000906478
arginine degradation V	2.33648E-05	0.001658901
arabinoxylan degradation	3.05085E-05	0.002135593
glycolysis (pay-off phase)	4.13257E-05	0.002851474
sulfate reduction (dissimilatory)	7.22929E-05	0.004915915
aspartate degradation II	0.000122585	0.008213174
pyruvate:formate lyase	0.0002114	0.013952388
threonine degradation I	0.000382816	0.024883053
galactose degradation	0.000421728	0.026990622
tyrosine degradation I	0.000493461	0.03108803
lactose degradation	0.000716412	0.044417527
glutamate degradation III	0.000725538	0.044417527

ESM Table 2. Associations between bacterial species, clinical sites, and visit schedule from linear mixed effects models in MaAsLin2 in the ND cohort. The analysis included microbiome data from all study visits. The model controlled for the age at the diagnosis of type 1 diabetes. More details on the tested variables

Bacterial species	Variable	Value
Catabacter.hongkongensis	site	UCAM - University of Cambridge
Clostridiaceae.bacterium	site	KU Leuven
Lactobacillus.sakei.subsp..sakei	site	UK - Norfolk and Norwich
Coprococcus.catus	time_from_baseline	time_from_baseline
Flavonifractor.plautii	site	HH-RH - Herlev University Hospital
Firmicutes.bacterium.OM04.13BH	site	HKA - Hannoversche Kinderheilanstalt
Lactococcus.lactis.subsp..cremoris	site	HKA - Hannoversche Kinderheilanstalt
Ruminococcaceae.bacterium.AM28.23LB	time_from_baseline	time_from_baseline
Paraprevotella.clara	time_from_baseline	time_from_baseline
Bifidobacterium.dentium	time_from_baseline	time_from_baseline
Dysosmobacter.welbionis	site	HH-RH - Herlev University Hospital
Eubacteriaceae.bacterium	time_from_baseline	time_from_baseline
Odoribacter.splanchnicus	time_from_baseline	time_from_baseline
Ruminococcus.sp..AF46.10NS	site	CHL - Centre Hospitalier de Luxembourg
Sellimonas.intestinalis	site	HH-RH - Herlev University Hospital
Tidjanibacter.massiliensis	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Ruminococcus.sp..AF41.9	site	HH-RH - Herlev University Hospital
Massilicoli.timonensis	site	UCAM - University of Cambridge
Eggerthella.lenta	site	HH-RH - Herlev University Hospital
Firmicutes.bacterium.AM31.12AC	site	HH-RH - Herlev University Hospital
Streptococcus.salivarius	time_from_baseline	time_from_baseline
Dorea.sp..AF24.7LB	time_from_baseline	time_from_baseline
Anaerostipes.hadrus	site	UK - Barts Health NHS Trust
X.Clostridium..bolteae	site	HH-RH - Herlev University Hospital
Ruminococcaceae.bacterium.D5	site	ULB - Univer Libre de Bruxelles
Clostridiales.bacterium	site	ULB - Univer Libre de Bruxelles
Bacteroides.clarus	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Bacteroides.massiliensis	site	MUG - Medical University of Graz
Victivallales.bacterium.CCUG.44730	time_from_baseline	time_from_baseline
X.Clostridium..symbiosum	site	HH-RH - Herlev University Hospital
Clostridiales.bacterium.CCNA10	site	UNISI - Università degli Studi di Siena
Bifidobacterium.longum.subsp..longum	site	KU Leuven
Blautia.sp..SG.772	time_from_baseline	time_from_baseline
Eisenbergiella.massiliensis	site	ULI - University of Ljubljana

Haemophilus.pittmaniae	site	HKA - Hannoversche Kinderheilanstalt
X.Ruminococcus..gnavus	site	HH-RH - Herlev University Hospital
Bacteroidaceae.bacterium	time_from_baseline	time_from_baseline
Clostridium.sp..AT4	site	HH-RH - Herlev University Hospital
Bacteroides.faecis	site	UNISI - Università degli Studi di Siena
Flavonifractor.plautii	site	ULB - Univer Libre de Bruxelles
Alistipes.shahii	time_from_baseline	time_from_baseline
Ruthenibacterium.lactatiformans	site	ULI - University of Ljubljana
Ruthenibacterium.lactatiformans	site	HH-RH - Herlev University Hospital
Clostridiales.bacterium.CCNA10	site	HH-RH - Herlev University Hospital
Eisenbergiella.massiliensis	site	HH-RH - Herlev University Hospital
Lactococcus.lactis.subsp..cremoris	site	UOUL - Oulun Yliopisto
Hungatella.hathewayi	site	HH-RH - Herlev University Hospital
Intestinibacter.bartlettii	site	UULM - Universität Ulm
Alistipes.indistinctus	time_from_baseline	time_from_baseline
Bacteroides.ovatus	site	HKA - Hannoversche Kinderheilanstalt
Alistipes.timonensis	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Campylobacter.concisus	site	UOUL - Oulun Yliopisto
Clostridium.sp..SN20	site	UNISI - Università degli Studi di Siena
Holdemania.sp..Marseille.P2844	time_from_baseline	time_from_baseline
Ruminococcus.sp..AM36.2AA	site	ULI - University of Ljubljana
Ruminococcus.sp..AM36.2AA	site	HH-RH - Herlev University Hospital
Clostridium.sp..AF15.31	site	UK - Barts Health NHS Trust
Lachnoclostridium.sp..An138	site	HH-RH - Herlev University Hospital
X.Clostridium..innocuum	site	HH-RH - Herlev University Hospital
Anaerotruncus.colihominis	site	HH-RH - Herlev University Hospital
Bacteroides.caccae	site	HH-RH - Herlev University Hospital
Eisenbergiella.tayi	site	UCAM - University of Cambridge
X.Clostridium..aldenense	site	HH-RH - Herlev University Hospital
Clostridium.sp..TM06.18	site	MUG - Medical University of Graz
Haemophilus.parainfluenzae	site	MUG - Medical University of Graz
Lactobacillus.sakei.subsp..sakei	site	UNISI - Università degli Studi di Siena
Bacteroides.massiliensis	time_from_baseline	time_from_baseline
Methanobrevibacter.smithii	site	ULB - Univer Libre de Bruxelles
Veillonella.parvula	site	MUG - Medical University of Graz
Lachnospira.pectinoschiza	site	UCAM - University of Cambridge
Lactobacillus.sakei.subsp..sakei	site	HKA - Hannoversche Kinderheilanstalt
Alistipes.ihumii	time_from_baseline	time_from_baseline
Clostridium.sp..OF03.18AA	site	MUG - Medical University of Graz
Odoribacter.splanchnicus	site	HH-RH - Herlev University Hospital
Coprobacillus.cateniformis	site	HH-RH - Herlev University Hospital
Victivallis.vadensis	site	UK - Barts Health NHS Trust
Tyzzerella.nexilis	site	ULB - Univer Libre de Bruxelles
Blautia.producta	site	HH-RH - Herlev University Hospital
Lachnospiraceae.bacterium.OF11.28	site	KU Leuven
Bacteroides.nordii	time_from_baseline	time_from_baseline
Bifidobacterium.longum.subsp..longum	site	ULI - University of Ljubljana
Blautia.sp..SG.772	site	HH-RH - Herlev University Hospital
Coprococcus.comes	time_from_baseline	time_from_baseline
Eubacterium.ramulus	site	ULI - University of Ljubljana
Ruminococcus.sp..AF41.9	time_from_baseline	time_from_baseline
Bifidobacterium.breve	site	HH-RH - Herlev University Hospital

Ruminococcaceae.bacterium.D5	time_from_baseline	time_from_baseline
Streptococcus.gordonii	site	UK - Norfolk and Norwich
Ellagibacter.isourolithinifaciens	site	ULB - Univer Libre de Bruxelles
Ruminococcaceae.bacterium.TF06.43	time_from_baseline	time_from_baseline
Odoribacter.splanchnicus	site	UCAM - University of Cambridge

Beta-coefficient	Standard error	p-value	q-value	prevalence
0.392413145	0.068805957	1.85008E-07	0.000683789	0.127717391
-1.899703832	0.374137486	8.17406E-07	0.001510566	0.994565217
0.769954732	0.156884149	1.81639E-06	0.002237788	0.10326087
0.086261548	0.018266908	0.000003717	0.003252187	0.866847826
-0.872113016	0.176938128	4.3996E-06	0.003252187	0.980978261
1.343898855	0.278262787	6.54552E-06	0.00403204	0.138586957
0.510311701	0.113225941	8.96952E-06	0.004630632	0.125
0.109131423	0.024257534	0.000010023	0.004630632	0.214673913
0.107929181	0.024664146	1.72406E-05	0.007080123	0.211956522
0.060775185	0.014563881	3.97201E-05	0.009672244	0.119565217
-0.768035441	0.175744224	0.00003695	0.009672244	0.961956522
0.111615794	0.026811512	4.18712E-05	0.009672244	0.940217391
0.127128583	0.030127226	3.32202E-05	0.009672244	0.817934783
1.474868992	0.332579997	2.95391E-05	0.009672244	0.100543478
-0.706845187	0.161156101	3.46334E-05	0.009672244	0.317934783
1.967870184	0.45465561	4.09282E-05	0.009672244	0.154891304
0.940639071	0.21849716	4.64524E-05	0.010099288	0.266304348
0.541979856	0.126320152	5.06364E-05	0.010397341	0.152173913
-0.824322169	0.195859751	6.55576E-05	0.012752671	0.850543478
0.593474283	0.142742979	8.04487E-05	0.01486692	0.279891304
0.124576957	0.031375224	9.04748E-05	0.015923569	0.89673913
0.112209491	0.028354929	9.64908E-05	0.016210449	0.714673913
-0.939315147	0.235168999	0.000126621	0.020347461	0.983695652
-0.672301407	0.167867858	0.000139183	0.021434216	0.758152174
1.293359713	0.326124209	0.000155572	0.022999816	0.206521739
-0.879248432	0.223451364	0.000174543	0.024812021	0.991847826
1.645423405	0.419464961	0.000181738	0.024877882	0.157608696
1.870000455	0.482766181	0.000210556	0.027793373	0.192934783
0.080263808	0.021641289	0.000252227	0.031074359	0.258152174
-0.653925161	0.17049876	0.000252119	0.031074359	0.570652174
1.923825918	0.504255426	0.000266983	0.03183126	0.252717391
-2.813942745	0.764721396	0.000317806	0.0326564	0.953804348
0.102969113	0.028095427	0.000296233	0.0326564	0.758152174
-0.495990834	0.131218241	0.000300888	0.0326564	0.296195652

0.904852213	0.240027879	0.000318082	0.0326564	0.116847826
-0.837943865	0.221724507	0.000300216	0.0326564	0.567934783
0.150473477	0.041423527	0.000334856	0.033449428	0.720108696
-0.506165935	0.135297203	0.000345042	0.033559849	0.236413043
2.914788862	0.784752913	0.0003734	0.034767332	0.184782609
-1.343343111	0.361887569	0.00037627	0.034767332	0.980978261
0.140475317	0.039087525	0.000385837	0.034781798	0.779891304
-0.489567812	0.133334789	0.000447206	0.03903527	0.97826087
-0.484458207	0.132002441	0.000454144	0.03903527	0.97826087
-0.761041661	0.210616454	0.000523795	0.042085803	0.252717391
-0.472104295	0.130132332	0.000501721	0.042085803	0.296195652
0.444198719	0.126782701	0.000518349	0.042085803	0.125
-0.609709109	0.169305007	0.00055821	0.042982132	0.717391304
-1.722394029	0.476156776	0.000555911	0.042982132	0.929347826
0.097673475	0.02805212	0.000579885	0.043057223	0.470108696
-1.419201995	0.39595683	0.000582484	0.043057223	0.967391304
1.093653144	0.309303931	0.000663719	0.043805423	0.111413043
0.423801704	0.119442694	0.00064144	0.043805423	0.141304348
1.254470481	0.352700716	0.000634131	0.043805423	0.404891304
0.057524306	0.016690382	0.000657191	0.043805423	0.241847826
0.88464311	0.248777786	0.000643021	0.043805423	0.83423913
0.881098777	0.246703489	0.000611741	0.043805423	0.83423913
2.555523785	0.73207635	0.000762033	0.048602153	0.233695652
-0.405846351	0.115649288	0.000766214	0.048602153	0.336956522
-0.514025706	0.146743375	0.000775846	0.048602153	0.758152174
-0.470820359	0.135357007	0.000832151	0.04881952	0.8125
1.427693183	0.410973842	0.000825571	0.04881952	0.695652174
0.654195291	0.188591495	0.000827945	0.04881952	0.432065217
-0.526864781	0.151065534	0.000802951	0.04881952	0.513586957
1.080504497	0.315138698	0.000916734	0.052941412	0.804347826
-1.09562912	0.323778513	0.001057731	0.060144196	0.869565217
0.401570367	0.118352719	0.001098409	0.061510878	0.10326087
0.082727217	0.025211861	0.001168345	0.063680992	0.192934783
1.37784283	0.40977627	0.00117162	0.063680992	0.157608696
-0.952874834	0.285867727	0.001230158	0.065893705	0.701086957
1.420897317	0.425524408	0.001249773	0.065987993	0.529891304
0.386346597	0.114892956	0.00127059	0.066142251	0.10326087
0.106748985	0.032851357	0.001301222	0.066796053	0.494565217
-1.026086616	0.309891138	0.00131994	0.066828768	0.907608696
0.8981843	0.27099918	0.001378859	0.068868434	0.817934783
-0.480768858	0.144965297	0.001427963	0.069444109	0.377717391
1.733749593	0.525078274	0.00142004	0.069444109	0.192934783
0.994307294	0.301062681	0.001486646	0.071358991	0.14673913
-0.394751546	0.120314766	0.001517951	0.071927512	0.222826087
-2.028811046	0.638047079	0.001683817	0.078777074	0.970108696
0.062136879	0.019746501	0.001833375	0.08175593	0.203804348
-0.61468335	0.190817086	0.001838782	0.08175593	0.953804348
0.744842725	0.230450256	0.001779582	0.08175593	0.758152174
0.078878067	0.025041283	0.001810788	0.08175593	0.880434783
0.714961004	0.222226127	0.001858089	0.08175593	0.836956522
0.073111436	0.023596079	0.002146127	0.093318648	0.266304348
-0.337352198	0.10669506	0.002214841	0.095032328	0.105978261

0.047592083	0.01544049	0.002262674	0.095032328	0.206521739
0.408624821	0.131403368	0.002261255	0.095032328	0.130434783
0.852571592	0.271131711	0.002341794	0.097062715	0.192934783
0.109772313	0.035770332	0.00236354	0.097062715	0.79076087
1.118460377	0.35707151	0.002401308	0.097530052	0.817934783

ESM Table 3. Associations between baseline bacterial abundances and functions, and clinical measurements at the time of the diagnosis of type 1 diabetes from linear mixed effects models in MaAsLin2 in the ND cohort. Associations with variables included for statistical corrections (clinical site, age at diagnosis) are also shown. The analysis included microbiome data from the baseline visit. More details on the tested variables are given below.

Variable	Value	Explanation
site	categorical	Clinical site. Comparisons to University of Helsink
age_at_diagnosis	continuous	Age at diagnosis in years
hba1c_at_diagnosis	continuous	HbA1c measurement at the time of T1D diagnosi

Other columns

Variable		metadata variable that was associated with the s
Value		value of the metadata which has the given associ
Beta-coefficient		the fixed effect coefficient of the given value
Standard error		Standard error of the beta-coefficient
p-value		nominal p-value for the given association
q-value		corrected p-value, fdr q-value, for the given assoc
prevalence		Prevalence (proportion of samples with presence

Bacterial species	Variable	Value
Lactobacillus.sakei.subsp..sakei	site	UK - Norfolk and Norwich
Alistipes.timonensis	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Lactococcus.lactis.subsp..cremoris	site	HKA - Hannoversche Kinderheilanstalt
Bacteroides.clarus	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Bacteroides.massiliensis	site	MUG - Medical University of Graz
Haemophilus.pittmaniae	site	HKA - Hannoversche Kinderheilanstalt
Bacteroides.faecis	site	UNISI - Università degli Studi di Siena
Eisenbergiella.tayi	site	UCAM - University of Cambridge
Haemophilus.pittmaniae	site	UOUL - Oulun Yliopisto
Bacteroides.eggerthii	site	UCAM - University of Cambridge
Clostridium.celatum	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Firmicutes.bacterium.OM04.13BH	site	HKA - Hannoversche Kinderheilanstalt
Firmicutes.bacterium.AM31.12AC	site	HH-RH - Herlev University Hospital
Lactobacillus.sakei.subsp..sakei	site	HKA - Hannoversche Kinderheilanstalt
Paraprevotella.clara	site	UULM - Universität Ulm
Ellagibacter.isourolithinifaciens	site	ULB - Univer Libre de Bruxelles
Holdemania.filiformis	site	HH-RH - Herlev University Hospital
Haemophilus.sputorum	age_at_diagnosis	age_at_diagnosis
Schaalia.odontolytica	site	UNISI - Università degli Studi di Siena
Bacteroides.stercoris	site	KU Leuven
Christensenella.minuta	site	UCAM - University of Cambridge
Clostridiaceae.bacterium	site	KU Leuven
Firmicutes.bacterium.OM04.13BH	site	CHL - Centre Hospitalier de Luxembourg
Lachnotalea.sp..AF33.28	site	UCAM - University of Cambridge
Massilioclostridium.coli	site	MUG - Medical University of Graz
Neisseria.subflava	site	UOUL - Oulun Yliopisto
Ruminococcus.sp..AF41.9	site	HH-RH - Herlev University Hospital
Veillonella.parvula	site	ULI - University of Ljubljana
X.Clostridium..aldenense	site	KU Leuven
Campylobacter.concisus	site	UOUL - Oulun Yliopisto
Catabacter.hongkongensis	site	UCAM - University of Cambridge

Dysosmobacter.welbionis	site	HH-RH - Herlev University Hospital
Faecalibacterium.prausnitzii	hba1c_at_diagnosis	hba1c_at_diagnosis
Flavonifractor.plautii	site	HH-RH - Herlev University Hospital

Module	Name	Variable
MF0073	pyruvate:ferredoxin	site
MF0091	ethanol production	site
MF0030	glutamate degradat	site
MF0098	hydrogen metaboli	site
MF0102	sulfate reduction (d	site
MF0060	glycerol degradatio	site
MF0032	glutamate degradat	age_at_diagnosis
MF0024	phenylalanine degr	site
MF0042	asparagine degrada	site
MF0004	pectine degradatio	site
MF0030	glutamate degradat	site
MF0073	pyruvate:ferredoxin	site
MF0014	arabinose degradat	site
MF0042	asparagine degrada	site
MF0081	methanol conversio	site
MF0099	methanogenesis - m	site
MF0100	methanogenesis fro	site
MF0029	aspartate degradati	site
MF0036	isoleucine degradat	site
MF0050	threonine degradat	site
MF0070	pentose phosphate	site
MF0098	hydrogen metaboli	hba1c_at_diagnosis
MF0102	sulfate reduction (d	site

Beta-coefficie	Standard error	p-value	q-value	prevalence
1.399816747	0.200852291	1.83609E-09	6.84496E-06	0.120481928
2.210811235	0.336371982	9.25734E-09	1.22344E-05	0.108433735
1.348111386	0.205588026	9.84525E-09	1.22344E-05	0.120481928
1.966948238	0.335829713	1.63449E-07	0.000152335	0.13253012
1.733827749	0.32567917	1.31076E-06	0.000977305	0.13253012
1.073685764	0.221449227	7.90726E-06	0.004913043	0.108433735
1.851881253	0.400833236	1.83075E-05	0.007717416	0.144578313
1.23018646	0.26654746	1.86311E-05	0.007717416	0.361445783
1.701314479	0.366076374	1.65713E-05	0.007717416	0.108433735
1.84249993	0.415992565	3.64331E-05	0.013582257	0.204819277
0.880376778	0.214283465	0.00011222	0.038032458	0.108433735
1.430228379	0.36123343	0.000186673	0.057992958	0.13253012
0.601900532	0.153556621	0.000213296	0.061166723	0.240963855
0.807666795	0.209183361	0.000259603	0.069128468	0.120481928
1.535043031	0.404289089	0.000321221	0.079834127	0.156626506
0.788283982	0.212570153	0.000429629	0.095500885	0.156626506
-0.72192731	0.194894672	0.000435492	0.095500885	0.734939759
-0.37114067	0.103157297	0.00061438	0.127244994	0.253012048
0.633132778	0.179768282	0.00078261	0.153556369	0.144578313
3.017190995	0.889113677	0.001171041	0.162256164	0.240963855
0.74107777	0.215679914	0.001025753	0.162256164	0.168674699
-1.19669237	0.346715506	0.000977162	0.162256164	0.987951807
1.166950852	0.342250124	0.001113671	0.162256164	0.13253012
0.748228804	0.216862145	0.000981029	0.162256164	0.265060241
0.616436784	0.1776864	0.00092427	0.162256164	0.156626506
1.297787939	0.382562413	0.001175139	0.162256164	0.144578313
0.882320414	0.258298942	0.001092222	0.162256164	0.228915663
-0.8108047	0.240673529	0.001263686	0.162448971	0.746987952
2.068695687	0.612400284	0.00122856	0.162448971	0.56626506
1.457404353	0.441335616	0.001550852	0.18915906	0.204819277
0.391397185	0.119792944	0.001725163	0.18915906	0.144578313

-0.79095182	0.241836027	0.0017078	0.18915906	0.927710843	
-0.14768593	0.045070393	0.001676065	0.18915906	1	
-0.83348186	0.25339133	0.001613367	0.18915906	0.975903614	

Value	Beta-coefficie	Standard error	p-value	q-value	prevalence
HKA - Hannove	-36.1893951	4.92220573	3.81594E-10	6.10551E-07	1
UOUL - Oulun Y	0.087679354	0.014740128	1.1377E-07	9.10161E-05	0.12048193
MUG - Medical	0.263623338	0.050479285	1.93263E-06	0.001030734	0.1686747
ULB - Univer Li	17.93341393	4.058745967	3.78514E-05	0.013699153	1
UCAM - Univer	0.265460977	0.060553896	4.28099E-05	0.013699153	0.19277108
ULB - Univer Li	8.639505491	2.057870033	8.22334E-05	0.021928902	1
age_at_diagno	0.839856528	0.216033166	0.000237544	0.0542958	0.95180723
UOUL - Oulun Y	0.189334095	0.049690653	0.000307332	0.055340801	0.28915663
MUG - Medical	28.67187847	7.532567701	0.000311292	0.055340801	1
CHL - Centre H	34.63723318	9.583773869	0.000582957	0.088095264	1
HKA - Hannove	0.159638957	0.044652829	0.000660714	0.088095264	0.1686747
UOUL - Oulun Y	-29.1342679	8.136868453	0.000649375	0.088095264	1
HH-RH - Herlev	-6.52378522	1.92165928	0.001166002	0.143507893	1
KU Leuven	29.63031048	9.063139982	0.001714511	0.161365775	1
ULB - Univer Li	0.123671499	0.037752581	0.00168087	0.161365775	0.1686747
ULB - Univer Li	0.123671499	0.037752581	0.00168087	0.161365775	0.1686747
ULB - Univer Li	0.123671499	0.037752581	0.00168087	0.161365775	0.1686747
MUG - Medical	37.42035018	11.71545243	0.00215113	0.166468308	1
UULM - Univer	4.327322492	1.365231904	0.002314122	0.166468308	1
HKA - Hannove	33.90225791	10.66417614	0.002249992	0.166468308	1
MUG - Medical	22.55150936	7.140146901	0.002392982	0.166468308	1
hba1c_at_diag	2.433531985	0.754268009	0.001952513	0.166468308	1
ULB - Univer Li	0.292502718	0.091018934	0.0020286	0.166468308	0.19277108

ESM Table 4. Associations between baseline bacterial abundances and functions, and the number of diabetes-associated autoantibodies at baseline visit from linear mixed effects models in MaAsLin2 in the ND cohort. Associations with variables included for statistical corrections (clinical site, age at diagnosis) are also shown. The analysis included microbiome data from the baseline visit. More details on the tested variables are given below.

Variable	Value	Explanation
site	categorical	Clinical site. Comparisons to University of Helsink
age_at_diagnosis	continuous	Age at diagnosis in years
num_aab	continuous	Number of positive AABs at baseline visit

Other columns

Variable		metadata variable that was associated with the s
Value		value of the metadata which has the given associ
Beta-coefficient		the fixed effect coefficient of the given value
Standard error		Standard error of the beta-coefficient
p-value		nominal p-value for the given association
q-value		corrected p-value, fdr q-value, for the given assoc
prevalence		Prevalence (proportion of samples with presence

Bacterial species	Variable	Value
Lactobacillus.sakei.subsp..sakei	site	UK - Norfolk and Norwich
Alistipes.timonensis	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Ruminococcus.sp..AF46.10NS	site	CHL - Centre Hospitalier de Luxembourg
Lactococcus.lactis.subsp..cremoris	site	HKA - Hannoversche Kinderheilanstalt
Bacteroides.clarus	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Clostridiaceae.bacterium	site	KU Leuven
Haemophilus.pittmaniae	site	HKA - Hannoversche Kinderheilanstalt
Bacteroides.massiliensis	site	MUG - Medical University of Graz
Bacteroides.faecis	site	UNISI - Università degli Studi di Siena
Haemophilus.pittmaniae	site	UOUL - Oulun Yliopisto
Eisenbergiella.tayi	site	UCAM - University of Cambridge
Lactobacillus.sakei.subsp..sakei	site	HKA - Hannoversche Kinderheilanstalt
Clostridium.sp..AF15.31	site	UK - Barts Health NHS Trust
Firmicutes.bacterium.AM31.12AC	site	HH-RH - Herlev University Hospital
Flavonifractor.plautii	site	HH-RH - Herlev University Hospital
Lachnoclostridium.sp..An131	age_at_diagnosi	age_at_diagnosis
Clostridium.celatum	site	SUM - Śląski Uniwersytet Medyczny w Katowicac
Firmicutes.bacterium.OM04.13BH	site	HKA - Hannoversche Kinderheilanstalt
Ruminococcus.sp..AF41.9	site	HH-RH - Herlev University Hospital
Schaalia.odontolytica	site	UNISI - Università degli Studi di Siena
X.Clostridium..bolteae	site	HH-RH - Herlev University Hospital
Veillonella.parvula	site	ULI - University of Ljubljana
Clostridiaceae.bacterium	site	UK - Barts Health NHS Trust
Ellagibacter.isourolithinifaciens	site	ULB - Univer Libre de Bruxelles
Anaerostipes.hadrus	site	UK - Barts Health NHS Trust
Bacteroides.eggerthii	site	UCAM - University of Cambridge
Holdemania.filiformis	site	HH-RH - Herlev University Hospital
Bifidobacterium.dentium	site	KU Leuven
Clostridium.paraputrificum	site	UK - Barts Health NHS Trust
Coprococcus.eutactus	num_aab	num_aab
Dysosmobacter.welbionis	site	HH-RH - Herlev University Hospital

Firmicutes.bacterium.OM04.13BH	site	CHL - Centre Hospitalier de Luxembourg
Odoribacter.splanchnicus	site	ULI - University of Ljubljana
Pseudoflavonifractor.sp..An184	age_at_diagnosi	age_at_diagnosis
Roseburia.sp..AM16.25	site	UK - Norfolk and Norwich
X.Ruminococcus..gnavus	site	HH-RH - Herlev University Hospital
Actinomyces.graevenitzii	site	UOUL - Oulun Yliopisto
Anaerotruncus.colihominis	site	HH-RH - Herlev University Hospital
Bacteroides.plebeius	site	CHL - Centre Hospitalier de Luxembourg
Clostridiales.bacterium.CCNA10	site	HH-RH - Herlev University Hospital
Clostridiales.bacterium.CCNA10	site	UNISI - Università degli Studi di Siena
Collinsella.aerofaciens	site	HH-RH - Herlev University Hospital
Eisenbergiella.massiliensis	site	HH-RH - Herlev University Hospital
Eisenbergiella.tayi	site	KU Leuven
Massilioclostridium.coli	site	MUG - Medical University of Graz
Prevotella.corporis	site	UK - Barts Health NHS Trust
Ruminococcus.sp..AF17.22AC	site	HKA - Hannoversche Kinderheilanstalt
Eisenbergiella.massiliensis	site	ULI - University of Ljubljana
X.Clostridium..symbiosum	site	HH-RH - Herlev University Hospital
Bacteroides.stercoris	site	KU Leuven
Christensenella.minuta	site	MUG - Medical University of Graz
Haemophilus.sputorum	age_at_diagnosi	age_at_diagnosis

Module	Name	Variable
MF0073	pyruvate:ferredo	site
MF0091	ethanol producti	site
MF0030	glutamate degra	site
MF0073	pyruvate:ferredo	site
MF0083	succinate consu	site
MF0060	glycerol degrada	site
MF0098	hydrogen metab	site
MF0102	sulfate reduction	site
MF0102	sulfate reduction	site
MF0014	arabinose degrad	site
MF0026	tyrosine degrada	num_aab
MF0036	isoleucine degra	site
MF0081	methanol conver	site
MF0099	methanogenesis	site
MF0100	methanogenesis	site
MF0074	pyruvate:format	site

ki (highest n)

pecies in question
iation

ciation
) of the given bacterial species

Beta-coefficie	Standard err	p-value	q-value	prevalence
1.404828187	0.19055099	2.4247E-10	9.934E-07	0.11235955
2.274273771	0.33440406	2.7152E-09	5.56208E-06	0.1011236
1.981761847	0.29986772	6.0756E-09	8.2972E-06	0.1011236
1.319376202	0.20958438	2.238E-08	2.29228E-05	0.12359551
2.286124487	0.38484085	9.6008E-08	7.86689E-05	0.13483146
-1.97121586	0.37168484	1.2283E-06	0.000838706	0.98876404
1.088212813	0.20752884	1.5521E-06	0.000908446	0.1011236
2.058430058	0.4018136	2.4829E-06	0.001271547	0.14606742
1.859617599	0.384077	7.2785E-06	0.003313327	0.14606742
1.401524421	0.3173927	3.5273E-05	0.014451345	0.1011236
1.121563564	0.2637928	6.3567E-05	0.023083064	0.38202247
0.823579622	0.19450071	0.00006761	0.023083064	0.11235955
2.191695683	0.53107337	9.8732E-05	0.03111565	0.20224719
0.575034658	0.14825826	0.00023228	0.067974027	0.23595506
-0.95271471	0.24957199	0.00028553	0.07568118	0.97752809
0.247475482	0.06500356	0.00029556	0.07568118	0.12359551
0.841457991	0.22947066	0.00047053	0.092832446	0.11235955
1.307556824	0.35523147	0.00044951	0.092832446	0.13483146
0.918236366	0.25064217	0.00047583	0.092832446	0.23595506
0.625869056	0.16839056	0.00039929	0.092832446	0.13483146
-0.87430906	0.23655919	0.00042773	0.092832446	0.7752809
-0.84785169	0.23237582	0.00049966	0.093050293	0.73033708
-0.96301742	0.26536566	0.00053271	0.09489261	0.98876404
0.91831975	0.25638003	0.00062095	0.106001627	0.16853933
-1.05295122	0.3032343	0.00088233	0.139034929	0.97752809
1.342979714	0.38642291	0.00087392	0.139034929	0.19101124
-0.65308786	0.19000192	0.00098629	0.149660453	0.70786517
0.945082276	0.28207858	0.00129517	0.156561315	0.11235955
1.018334201	0.30352405	0.00127669	0.156561315	0.13483146
0.410202945	0.12314658	0.00137569	0.156561315	0.61797753
-0.78146125	0.23346376	0.00130803	0.156561315	0.93258427

1.18772598	0.35385961	0.00127093	0.156561315	0.13483146
0.79453189	0.23545918	0.00120185	0.156561315	0.78651685
0.319538199	0.09548531	0.00131122	0.156561315	0.2247191
1.038936461	0.30583997	0.0011197	0.156561315	0.1011236
-0.9549683	0.28650045	0.00136632	0.156561315	0.56179775
1.02809919	0.31896796	0.00191575	0.170664463	0.2247191
-0.65672658	0.20120236	0.00169138	0.170664463	0.73033708
0.950542401	0.28883219	0.00155688	0.170664463	0.1011236
-0.66917096	0.20388778	0.00160025	0.170664463	0.28089888
1.669713796	0.51696066	0.00187731	0.170664463	0.28089888
0.676162144	0.20827022	0.00178412	0.170664463	0.76404494
-0.52533633	0.16157734	0.00175848	0.170664463	0.28089888
1.999531688	0.61782799	0.00184037	0.170664463	0.38202247
0.518165693	0.15838828	0.00165303	0.170664463	0.16853933
0.86760953	0.26977451	0.00195783	0.170664463	0.23595506
1.225617189	0.38094495	0.00195039	0.170664463	0.28089888
-0.51390493	0.16046627	0.00203954	0.174083267	0.28089888
-0.69039769	0.21687575	0.0021614	0.180719297	0.59550562
3.030039635	0.95432851	0.00221638	0.181609872	0.24719101
1.145849063	0.36525136	0.00248342	0.195806086	0.17977528
-0.29993889	0.09561603	0.00248521	0.195806086	0.25842697

Value	Beta-coeffici	Standard err	p-value	q-value	prevalence
HKA - Hannove	-34.567115	4.64228832	1.77276E-10	3.0137E-07	1
UOUL - Oulun Y	0.08848979	0.0126587	1.22268E-09	1.0393E-06	0.11235955
MUG - Medical	0.81602494	0.16041515	2.85235E-06	0.00161633	0.19101124
UOUL - Oulun Y	-31.558804	7.09987282	3.17216E-05	0.01078536	1
UK - Barts Heal	1.79573682	0.40308718	3.05911E-05	0.01078536	1
ULB - Univer Li	8.35195883	1.91376908	4.25024E-05	0.01204233	1
ULB - Univer Li	14.9051018	3.99462978	0.000380558	0.09242114	1
UCAM - Univer	0.20031984	0.05489096	0.000498367	0.10590306	0.19101124
ULB - Univer Li	0.29954547	0.08571341	0.000821736	0.15521677	0.19101124
HH-RH - Herlev	-6.5647836	1.93799596	0.001153873	0.17262894	1
num_aab	-0.4810412	0.14507887	0.001442428	0.17262894	1
UULM - Univer	4.33752914	1.2667649	0.001028168	0.17262894	1
ULB - Univer Li	0.11452881	0.03472588	0.001523197	0.17262894	0.17977528
ULB - Univer Li	0.11452881	0.03472588	0.001523197	0.17262894	0.17977528
ULB - Univer Li	0.11452881	0.03472588	0.001523197	0.17262894	0.17977528
HKA - Hannove	-28.909881	8.95082564	0.001877365	0.19947	1

**ESM Table 5. Associations between baseline bacterial abundances
insulin dose in the ND cohort during the follow-up from linear mixe**
between bacterial species and HbA1c, fasting C-peptide and insulin d
5 – baseline) year follow up are shown. Associations with variables in
age at diagnosis) are also shown. The analysis included microbiome d

Variable	Value
site	categorical
age_at_diagnosis	continuous
hba1c_v5_minus_v1	continuous
insulin_dose_per_kg_v4_minus_v	continuous
fasted_c_pep_result_v5_minus_v	continuous

Other columns

Variable
Value
Beta-coefficient
Standard error
p-value
q-value
prevalence

Change during year one:

Bacterial species	Variable
Lactobacillus.sakei.subsp..sakei	site
Lactococcus.lactis.subsp..cremoris	site
Ruminococcus.sp..AF46.10NS	site
Alistipes.timonensis	site
Bacteroides.clarus	site
Bacteroides.massiliensis	site
Clostridiaceae.bacterium	site
Ellagibacter.isourolithinifaciens	site
Haemophilus.pittmaniae	site
Bacteroides.faecis	site
Clostridium.sp..AF15.31	site
Haemophilus.pittmaniae	site
Lactobacillus.sakei.subsp..sakei	site
Pseudomonas.aeruginosa	site
Clostridiales.bacterium.CCNA10	site
Bacteroides.intestinalis	site
Alistipes.timonensis	site
Anaerostipes.hadrus	site
Dorea.sp..AF24.7LB	site
Firmicutes.bacterium.OM04.13BH	site
Tyzzerella.nexilis	site
Bacteroides.plebeius	site
Clostridiaceae.bacterium	site
Eisenbergiella.tayi	site
Firmicutes.bacterium.OM04.13BH	site
Lachnoclostridium.sp..An131	age_at_diagnosis
Lachnoclostridium.sp..An138	site
Sutterella.seckii	site

Lactococcus.lactis.subsp..cremoris	site
Roseburia.sp..AM16.25	insulin_dose_per_kg_v4_minus_v
Schaalia.odontolytica	site

Change during two years:

Bacterial species	Variable
Ruminococcus.sp..AF46.10NS	site
Alistipes.timonensis	site
Clostridium.paraputrificum	site
Clostridiaceae.bacterium	site
Ruminococcus.sp..AF46.10NS	site
Bacteroides.clarus	site
Ellagibacter.isourolithinifaciens	site
Clostridiaceae.bacterium	site
Olsenella.sp..GAM18	site
Clostridiales.bacterium.CCNA10	site
Clostridium.celatum	site
Bacteroides.massiliensis	site
Klebsiella.pneumoniae.subsp..pne	hba1c_v5_minus_v1
X.Clostridium..aldenense	site
Pseudomonas.aeruginosa	site
Bacteroides.faecis	site
Alistipes.timonensis	site
Sutterella.seckii	site
Bacteroides.plebeius	site
X.Ruminococcus..gnavus	age_at_diagnosis
Haemophilus.pittmaniae	site
Blautia.obeum	fasted_c_pep_result_v5_minus_v
Klebsiella.pneumoniae.subsp..pne	site
Firmicutes.bacterium.OM04.13BH	site
Ruminococcus.sp..AF21.42	site
Clostridium.sp..SN20	site
Prevotella.copri	site
Bacteroides.intestinalis	site
Ruminococcus.sp..AF46.10NS	site
Anaerostipes.hadrus	site
Bacteroides.plebeius	site
Bifidobacterium.longum.subsp..lo	site
Streptococcus.australis	site
Klebsiella.pneumoniae.subsp..pne	site
Tyzzerella.nexilis	site

and change in HbA1c, fasting C-peptide and
d effects models in MaAsLin2. Associations
dose during one (visit 4 – baseline) and two (visit
ncluded for statistical corrections (clinical site,
data from the baseline visit. More details on the

Explanation

Clinical site. Comparisons to University of Helsinki (highest n)
Age at diagnosis in years
Change in HbA1c measurement between baseline visit and visit 5 (during two year follow up)
Change in insulin dose between baseline visit and visit 4 (during the first year of follow up)
Change in fasting C-peptide measurement between baseline visit and visit 5 (during two year follow u

metadata variable that was associated with the species in question
value of the metadata which has the given association
the fixed effect coefficient of the given value
Standard error of the beta-coefficient
nominal p-value for the given association
corrected p-value, fdr q-value, for the given association
Prevalence (proportion of samples with presence) of the given bacterial species

Value	Beta-coefficient	Standard error	p-value
UK - Norfolk and Norwich	1.438255465	0.208180653	3.57977E-09
HKA - Hannoversche Kinderheilanstalt	1.417753993	0.212132924	8.65494E-09
CHL - Centre Hospitalier de Luxembourg	1.97207268	0.312648928	3.74238E-08
SUM - Śląski Uniwersytet Medyczny w Katowicac	1.616412111	0.268831955	1.16999E-07
SUM - Śląski Uniwersytet Medyczny w Katowicac	2.678170945	0.497273603	1.26773E-06
MUG - Medical University of Graz	2.241234396	0.427301113	2.14207E-06
KU Leuven	-1.953234296	0.384339386	3.91374E-06
ULB - Univer Libre de Bruxelles	1.320106215	0.268910232	7.36585E-06
HKA - Hannoversche Kinderheilanstalt	1.105151932	0.228245094	9.39556E-06
UNISI - Università degli Studi di Siena	1.838375187	0.424347896	0.000057069
UK - Barts Health NHS Trust	2.278129312	0.536872051	7.74885E-05
UOUL - Oulun Yliopisto	1.458609336	0.346418093	8.66808E-05
HKA - Hannoversche Kinderheilanstalt	0.867519207	0.218207615	0.000191193
CHL - Centre Hospitalier de Luxembourg	0.538646953	0.136271059	0.000206279
UNISI - Università degli Studi di Siena	1.991686602	0.507349965	0.000225614
ULB - Univer Libre de Bruxelles	2.482881045	0.642626541	0.00027665
MUG - Medical University of Graz	0.777555814	0.214228684	0.000588405
UK - Barts Health NHS Trust	-1.113085006	0.307513571	0.000607218
HH-RH - Herlev University Hospital	1.120874818	0.307630411	0.000562798
HKA - Hannoversche Kinderheilanstalt	1.301553818	0.364743255	0.000713704
UCAM - University of Cambridge	1.17583386	0.330364878	0.000734657
CHL - Centre Hospitalier de Luxembourg	1.0012275	0.294726535	0.001214008
UK - Barts Health NHS Trust	-0.936554652	0.271260313	0.00102374
UCAM - University of Cambridge	1.087470832	0.315956329	0.001058306
CHL - Centre Hospitalier de Luxembourg	1.196240958	0.347058327	0.001042239
age_at_diagnosis	0.226637274	0.066395367	0.001154838
HH-RH - Herlev University Hospital	-0.405521259	0.118846909	0.001159519
ULB - Univer Libre de Bruxelles	1.390365441	0.408045504	0.001176539

Value			
UOUL - Oulun Yliopisto	1.089128223	0.32196391	0.001268533
insulin_dose_per_kg_v4_minus_v1	-0.153189501	0.047293401	0.00198621
UNISI - Università degli Studi di Siena	0.622167561	0.186251664	0.001442711

Value	Beta-coefficient	Standard error	p-value
CHL - Centre Hospitalier de Luxembourg	1.992336218	0.263979371	1.21599E-09
SUM - Śląski Uniwersytet Medyczny w Katowicac	1.655431267	0.284005441	4.8821E-07
UK - Barts Health NHS Trust	2.055690912	0.369484826	1.2235E-06
KU Leuven	-1.842040029	0.359419644	5.50498E-06
UK - Barts Health NHS Trust	1.891816571	0.367699542	5.14345E-06
SUM - Śląski Uniwersytet Medyczny w Katowicac	2.606871749	0.523056121	8.88195E-06
ULB - Univer Libre de Bruxelles	1.41057104	0.281384087	8.05103E-06
UK - Barts Health NHS Trust	-1.659979071	0.346045308	1.66526E-05
UK - Barts Health NHS Trust	2.068268491	0.438838198	2.20371E-05
UNISI - Università degli Studi di Siena	2.470420215	0.546512373	4.16958E-05
UK - Barts Health NHS Trust	1.297566223	0.286309556	4.01215E-05
MUG - Medical University of Graz	2.399186917	0.550648625	0.000071104
hba1c_v5_minus_v1	0.125835972	0.029482691	9.48059E-05
KU Leuven	2.42357964	0.565172766	8.88568E-05
CHL - Centre Hospitalier de Luxembourg	0.534748202	0.128151923	0.00012878
UNISI - Università degli Studi di Siena	1.878240082	0.464687624	0.000195204
MUG - Medical University of Graz	1.163426222	0.290633734	0.000220666
ULB - Univer Libre de Bruxelles	1.46830375	0.368533965	0.000234174
CHL - Centre Hospitalier de Luxembourg	1.083141172	0.277886041	0.000306827
age_at_diagnosis	-0.871598948	0.222910426	0.000295293
UOUL - Oulun Yliopisto	1.346721455	0.354669982	0.000419098
fasted_c_pep_result_v5_minus_v1	-0.2505483	0.06326155	0.00031743
CHL - Centre Hospitalier de Luxembourg	0.606630215	0.161527423	0.000476136
CHL - Centre Hospitalier de Luxembourg	1.159516698	0.321622365	0.000751885
UK - Barts Health NHS Trust	2.456197761	0.697828669	0.000971058
UNISI - Università degli Studi di Siena	1.794314882	0.51223996	0.001021088
UK - Barts Health NHS Trust	2.805047296	0.805910583	0.001090867
ULB - Univer Libre de Bruxelles	2.398802869	0.695098606	0.001190496
ULB - Univer Libre de Bruxelles	0.980483026	0.285878279	0.00126759
UK - Barts Health NHS Trust	-1.504691121	0.445458595	0.001475694
UK - Barts Health NHS Trust	1.310514554	0.387070282	0.001442112
KU Leuven	-3.11591371	0.913513235	0.001339592
UNISI - Università degli Studi di Siena	1.715678877	0.507511923	0.001464014
UOUL - Oulun Yliopisto	0.950849382	0.28645011	0.001748881
KU Leuven	2.092685329	0.632859254	0.001814355

p)

q-value	prevalence
1.46663E-05	0.112359551
1.77297E-05	0.123595506
5.11084E-05	0.101123596
0.000119837	0.101123596
0.001038776	0.134831461
0.001462678	0.146067416
0.002290656	0.988764045
0.003772234	0.168539326
0.004277066	0.101123596
0.023381179	0.146067416
0.028860946	0.202247191
0.029594264	0.101123596
0.060255064	0.112359551
0.060366197	0.101123596
0.061622597	0.280898876
0.070839566	0.112359551
0.130935272	0.101123596
0.130935272	0.97752809
0.130935272	0.651685393
0.143328133	0.134831461
0.143328133	0.146067416
0.177635437	0.101123596
0.177635437	0.988764045
0.177635437	0.382022472
0.177635437	0.134831461
0.177635437	0.123595506
0.177635437	0.303370787
0.177635437	0.146067416

0.17921308	0.123595506
0.180211087	0.101123596
0.197026245	0.134831461

q-value	prevalence
4.10276E-06	0.101123596
0.00082361	0.101123596
0.001376029	0.134831461
0.003714758	0.988764045
0.003714758	0.101123596
0.004281099	0.134831461
0.004281099	0.168539326
0.007023233	0.988764045
0.008261475	0.101123596
0.012789232	0.280898876
0.012789232	0.112359551
0.019992084	0.146067416
0.022848224	0.101123596
0.022848224	0.573033708
0.028966925	0.101123596
0.041163677	0.146067416
0.043795711	0.101123596
0.043894538	0.146067416
0.051761779	0.101123596
0.051761779	0.561797753
0.067335058	0.101123596
0.0717199	0.97752809
0.07302188	0.101123596
0.110298254	0.134831461
0.136514608	0.258426966
0.137806038	0.393258427
0.141560956	0.224719101
0.148767895	0.112359551
0.152744576	0.101123596
0.155593494	0.97752809
0.155593494	0.101123596
0.155593494	0.91011236
0.155593494	0.651685393
0.178809833	0.101123596
0.180048059	0.146067416

ESM Table 6. Associations between bacterial species and diseas in the ND cohort from linear mixed effects models in MaAsLin2.
statistical corrections (clinical site, age, gender, time from baselin microbiome data from all study visits. More details on the tested

Variable	Value
site	categorical
age	categorical
time_from_baseline	continuous
gender	categorical
nd_progression_group	categorical

Other columns
Variable
Value
Beta-coefficient
Standard error
p-value
q-value
prevalence

Species	Variable
Clostridiaceae.bacterium	site
Firmicutes.bacterium.OM04.13BH	site
Lactobacillus.sakei.subsp..sakei	site
Coprococcus.catus	time_from_baseline
Catabacter.hongkongensis	site
Dysosmobacter.welbionis	site
Ruminococcaceae.bacterium.AM28.23LB	time_from_baseline
Paraprevotella.clara	time_from_baseline
Flavonifractor.plautii	site
Bifidobacterium.dentium	time_from_baseline
Odoribacter.splanchnicus	time_from_baseline
Lactococcus.lactis.subsp..cremoris	site
Tidjanibacter.massiliensis	site
Eubacteriaceae.bacterium	time_from_baseline
Streptococcus.salivarius	time_from_baseline
Ruminococcus.sp..AF46.10NS	site
Dorea.sp..AF24.7LB	time_from_baseline
Haemophilus.pittmaniae	site
Bacteroides.clarus	site
Blautia.sp..SG.772	time_from_baseline
Victivallales.bacterium.CCUG.44730	time_from_baseline
Ruminococcus.sp..AF41.9	site
Alistipes.shahii	time_from_baseline
Eisenbergiella.tayi	site
Anaerostipes.hadrus	site
Clostridiales.bacterium.CCNA10	site
Ruthenibacterium.lactatiformans	site
Tyzzerella.nexilis	site
Haemophilus.parainfluenzae	age_at_diagnosis

Sellimonas.intestinalis	site
Alistipes.timonensis	site
Bacteroides.faecis	site
Bacteroides.ovatus	site
Firmicutes.bacterium.AM31.12AC	site
Alistipes.indistinctus	time_from_baseline
Bacteroidaceae.bacterium	time_from_baseline
Eggerthella.lenta	site
Holdemania.sp..Marseille.P2844	time_from_baseline
Campylobacter.concisus	site
Eisenbergiella.massiliensis	site
Ruthenibacterium.lactatiformans	site
Clostridium.sp..AT4	site
Coprococcus.comes	time_from_baseline
Lactobacillus.sakei.subsp..sakei	site
Victivallis.vadensis	site
Bifidobacterium.longum.subsp..longum	site
Clostridium.sp..SN20	site
Tyzzerella.nexilis	site
Alistipes.ihumii	time_from_baseline
X.Clostridium..aldenense	site
Bacteroides.caccae	site
Lactobacillus.sakei.subsp..sakei	site
Clostridiales.bacterium.CCNA10	site
Lachnoclostridium.sp..An138	site
Eisenbergiella.massiliensis	site
Anaerotruncus.colihominis	site
Neisseria.subflava	nd_progression_group
Clostridiales.bacterium	site
Clostridium.sp..AF15.31	site
Lachnotalea.sp..AF33.28	site
Massilicoli.timonensis	site
Methanobrevibacter.smithii	site
Ruminococcaceae.bacterium.AM28.23LB	site
Ruminococcaceae.bacterium.D5	site
Streptococcus.salivarius	site
X.Clostridium..symbiosum	site
Ruminococcaceae.bacterium.D5	time_from_baseline
Ruminococcus.sp..AF41.9	time_from_baseline
Streptococcus.salivarius	nd_progression_group
X.Clostridium..bolteae	site
Ruminococcaceae.bacterium.TF06.43	time_from_baseline
Subdoligranulum.sp..APC924.74	time_from_baseline
Bifidobacterium.animalis.subsp..lactis	site
Campylobacter.concisus	nd_progression_group
Flavonifractor.plautii	site
Firmicutes.bacterium.AF19.2LB	time_from_baseline
Bifidobacterium.longum.subsp..longum	site
Ruminococcaceae.bacterium.TF06.43	site
Bacteroides.massiliensis	site
Erysipelatoclostridium.ramosum	site
Gordonibacter.urolithinfaciens	gender

Eisenbergiella.tayi	site
Eubacterium.ramulus	site
Firmicutes.bacterium.OM04.13BH	site
X.Clostridium..spiroforme	site
Blautia.sp..AF19.10LB	time_from_baseline
Parasutterella.excrementihominis	time_from_baseline
Anaeromassilibacillus.sp..An250	site
Bacteroides.fragilis	site
Firmicutes.bacterium.OM04.13BH	site
Lachnoclostridium.sp..An131	site
Lachnospiraceae.bacterium.OF11.28	site
Ruminococcus.sp..AM36.2AA	site
Veillonella.parvula	site
Asaccharobacter.celatus	time_from_baseline
Clostridium.sp..AT4	time_from_baseline
Streptococcus.gordonii	site
X.Clostridium..leptum	site
X.Ruminococcus..gnavus	site
Collinsella.aerofaciens	site
Faecalibacterium.sp..OF04.11AC	time_from_baseline
Haemophilus.parainfluenzae	site
Veillonella.atypica	nd_progression_group
Lactococcus.lactis.subsp..cremoris	site
Senegalimassilia.anaerobia	site
Blautia.massiliensis	time_from_baseline
Haemophilus.sputorum	age_at_diagnosis
Bacteroides.stercoris	site
Ruminococcus.sp..AM36.2AA	site
Alistipes.putredinis	time_from_baseline

Value	Beta-coefficient	Standard error	p-value
KU Leuven	-1.924737182	0.378067357	7.8859E-07
HKA - Hannoversche Kinderheilanstalt	1.386253628	0.275630115	3.17955E-06
UK - Norfolk and Norwich	0.774214195	0.157047112	1.6676E-06
time_from_baseline	0.087893993	0.018510623	3.32574E-06
UCAM - University of Cambridge	0.367957756	0.076204551	6.42659E-06
HH-RH - Herlev University Hospital	-0.868980178	0.182140105	8.47637E-06
time_from_baseline	0.110668508	0.024627604	1.03197E-05
time_from_baseline	0.10947944	0.025030437	1.74904E-05
HH-RH - Herlev University Hospital	-0.846500092	0.188703048	0.00002495
time_from_baseline	0.062822394	0.01477942	2.88533E-05
time_from_baseline	0.129578497	0.030558244	0.000030608
HKA - Hannoversche Kinderheilanstalt	0.477537973	0.114614996	3.91733E-05
SUM - Śląski Uniwersytet Medyczny w Katowicac	2.073044719	0.4822433	4.76007E-05
time_from_baseline	0.112627691	0.02720197	0.000046134
time_from_baseline	0.130900442	0.03166263	4.68469E-05
CHL - Centre Hospitalier de Luxembourg	1.437194271	0.335707956	5.35067E-05
time_from_baseline	0.114866116	0.028764465	8.39124E-05
HKA - Hannoversche Kinderheilanstalt	0.946796152	0.236707141	0.000148773
SUM - Śląski Uniwersytet Medyczny w Katowicac	1.758244654	0.442920456	0.000159025
time_from_baseline	0.106452449	0.028468896	0.000224841
time_from_baseline	0.082108723	0.02196913	0.000227122
HH-RH - Herlev University Hospital	0.856118373	0.223692444	0.000260483
time_from_baseline	0.146584397	0.039619561	0.000261258
UCAM - University of Cambridge	0.781903088	0.206527373	0.000288227
UK - Barts Health NHS Trust	-0.889405171	0.237795655	0.00031734
UNISI - Università degli Studi di Siena	2.072467919	0.55160508	0.000330415
ULI - University of Ljubljana	-0.527485021	0.140298136	0.000340393
ULB - Univer Libre de Bruxelles	1.216011856	0.321475121	0.000335325
age_at_diagnosis	-0.295543532	0.080587571	0.000440584

HH-RH - Herlev University Hospital	-0.625160248	0.169805584	0.000426049
SUM - Śląski Uniwersytet Medyczny w Katowicac	1.166241946	0.322076338	0.000512261
UNISI - Università degli Studi di Siena	3.107962372	0.860850549	0.000537325
HKA - Hannoversche Kinderheilanstalt	-1.446345647	0.399459517	0.000527127
HH-RH - Herlev University Hospital	0.51263077	0.142137531	0.000549089
time_from_baseline	0.099693705	0.028476899	0.000543033
time_from_baseline	0.148347605	0.042010887	0.000486042
HH-RH - Herlev University Hospital	-0.702934494	0.197251721	0.000625851
time_from_baseline	0.058661762	0.016915214	0.000610226
UOUL - Oulun Yliopisto	0.419231047	0.118897334	0.00068334
ULI - University of Ljubljana	-0.492387	0.138978378	0.000677233
HH-RH - Herlev University Hospital	-0.497680796	0.140212826	0.00068406
HH-RH - Herlev University Hospital	-0.506351734	0.143381854	0.00070287
time_from_baseline	0.086118618	0.025205351	0.000730039
UNISI - Università degli Studi di Siena	0.446414723	0.129388044	0.000898588
UK - Barts Health NHS Trust	1.84758495	0.537025735	0.000927238
KU Leuven	-2.518141479	0.760469532	0.001154099
UNISI - Università degli Studi di Siena	1.329368413	0.393020599	0.001123048
UCAM - University of Cambridge	0.722559008	0.214065501	0.001141373
time_from_baseline	0.10902856	0.033330139	0.001210784
KU Leuven	2.057703421	0.627176177	0.001273094
HH-RH - Herlev University Hospital	1.458919849	0.438518074	0.001339705
HKA - Hannoversche Kinderheilanstalt	0.380547627	0.113442057	0.001318188
HH-RH - Herlev University Hospital	-0.732466438	0.220793928	0.00138141
HH-RH - Herlev University Hospital	-0.407833649	0.122964227	0.001430538
HH-RH - Herlev University Hospital	-0.457445047	0.138904827	0.001496527
HH-RH - Herlev University Hospital	-0.458954433	0.139739331	0.00156402
Rapid	0.101194901	0.032380701	0.001928263
ULB - Univer Libre de Bruxelles	-0.783486938	0.244316891	0.001972622
UK - Barts Health NHS Trust	2.42871327	0.749483476	0.001718312
UCAM - University of Cambridge	0.623106065	0.192802328	0.001807378
UCAM - University of Cambridge	0.435663575	0.136062707	0.001993856
UK - Barts Health NHS Trust	1.555268315	0.487275389	0.002014303
SUM - Śląski Uniwersytet Medyczny w Katowicac	1.05556634	0.329311505	0.001906724
ULB - Univer Libre de Bruxelles	1.13656677	0.353625839	0.001908787
ULI - University of Ljubljana	-0.528519458	0.163177446	0.001782928
HH-RH - Herlev University Hospital	-0.575796694	0.179317663	0.001938013
time_from_baseline	0.048935481	0.015679691	0.001997902
time_from_baseline	0.075439301	0.023947714	0.001813971
Slow	0.475288947	0.149740472	0.002154466
HH-RH - Herlev University Hospital	-0.546869393	0.172062543	0.002150343
time_from_baseline	0.112375745	0.036315768	0.002178466
time_from_baseline	0.106123859	0.034413903	0.002251405
ULI - University of Ljubljana	-0.442364629	0.142867354	0.00276067
Rapid	0.123264521	0.039726727	0.002823156
ULB - Univer Libre de Bruxelles	-1.23628911	0.402005521	0.002920284
time_from_baseline	0.114824565	0.038228379	0.002913634
ULI - University of Ljubljana	-0.591265907	0.192931212	0.003003785
UNISI - Università degli Studi di Siena	-2.453678344	0.801101134	0.003035353
MUG - Medical University of Graz	1.953147259	0.640200135	0.003095693
KU Leuven	1.780079923	0.595351705	0.003135128
MALE	-0.32030196	0.105735127	0.003307902

KU Leuven	1.897985955	0.637546942	0.003268857
ULI - University of Ljubljana	0.671969892	0.221485405	0.003275821
UK - Barts Health NHS Trust	0.852019351	0.281747173	0.003320057
KU Leuven	1.855364178	0.626975067	0.003653718
time_from_baseline	0.102911687	0.03519804	0.003751203
time_from_baseline	0.078713694	0.026917472	0.003746002
MUG - Medical University of Graz	-1.042945834	0.354118326	0.004143277
ULI - University of Ljubljana	-1.242296377	0.420315	0.004124498
HH-RH - Herlev University Hospital	0.352394139	0.119292509	0.00412658
HH-RH - Herlev University Hospital	-0.202346187	0.067961906	0.004084607
KU Leuven	-1.881915242	0.645758843	0.003944735
ULI - University of Ljubljana	0.766155253	0.258785569	0.004112385
ULI - University of Ljubljana	-0.525106635	0.17666191	0.003995995
time_from_baseline	0.085269006	0.029343919	0.00396363
time_from_baseline	-0.050104717	0.017203339	0.003881906
UK - Norfolk and Norwich	0.372692097	0.128461903	0.004308823
ULI - University of Ljubljana	-0.506323922	0.171875268	0.004306454
HH-RH - Herlev University Hospital	-0.651048854	0.221231518	0.004289372
HH-RH - Herlev University Hospital	0.936799624	0.319045353	0.004355722
time_from_baseline	0.089316875	0.03117549	0.004497052
UOUL - Oulun Yliopisto	1.293616407	0.443172025	0.004586861
Slow	0.353584772	0.12137396	0.004659653
HH-RH - Herlev University Hospital	-0.144811113	0.051214649	0.004965406
ULI - University of Ljubljana	0.736222362	0.254469889	0.00492569
time_from_baseline	0.085671162	0.030261754	0.004969822
age_at_diagnosis	-0.13589337	0.047317334	0.005229592
KU Leuven	4.108149631	1.435975641	0.005263794
HH-RH - Herlev University Hospital	0.74361229	0.258733757	0.005268237
time_from_baseline	0.10686635	0.03796962	0.005244754

q-value	prevalence
0.003278959	0.994490358
0.003457109	0.137741047
0.003457109	0.104683196
0.003457109	0.865013774
0.005344348	0.129476584
0.005874125	0.961432507
0.006129878	0.217630854
0.009090641	0.201101928
0.011526906	0.980716253
0.011569827	0.121212121
0.011569827	0.815426997
0.013194927	0.123966942
0.013194927	0.157024793
0.013194927	0.939393939
0.013194927	0.898071625
0.013905052	0.101928375
0.020523978	0.724517906
0.034366522	0.115702479
0.034801323	0.159779614
0.04497011	0.754820937
0.04497011	0.261707989
0.047230908	0.269972452
0.047230908	0.776859504

0.049935297	0.424242424
0.050548381	0.983471074
0.050548381	0.242424242
0.050548381	0.977961433
0.050548381	0.148760331
0.061064947	0.876033058

0.061064947	0.308539945
0.063419829	0.104683196
0.063419829	0.187327824
0.063419829	0.966942149
0.063419829	0.283746556
0.063419829	0.476584022
0.063419829	0.716253444
0.068481265	0.848484848
0.068481265	0.231404959
0.069373681	0.143250689
0.069373681	0.300275482
0.069373681	0.977961433
0.069584169	0.225895317
0.070593071	0.878787879
0.084916587	0.104683196
0.085676788	0.195592287
0.099973851	0.953168044
0.099973851	0.396694215
0.099973851	0.148760331
0.102743703	0.50137741
0.105870472	0.506887052
0.107124842	0.691460055
0.107124842	0.104683196
0.108375522	0.242424242
0.110151455	0.341597796
0.113137419	0.300275482
0.116128476	0.809917355
0.123168722	0.110192837
0.123168722	0.991735537
0.123168722	0.236914601
0.123168722	0.303030303
0.123168722	0.154269972
0.123168722	0.159779614
0.123168722	0.217630854
0.123168722	0.209366391
0.123168722	0.898071625
0.123168722	0.564738292
0.123168722	0.209366391
0.123168722	0.269972452
0.127578317	0.898071625
0.127578317	0.754820937
0.127578317	0.801652893
0.130018616	0.909090909
0.157244746	0.220385675
0.158630857	0.143250689
0.159770257	0.980716253
0.159770257	0.757575758
0.16180765	0.953168044
0.16180765	0.801652893
0.16293534	0.192837466
0.162948259	0.292011019
0.164342798	0.380165289

0.164342798	0.424242424
0.164342798	0.840220386
0.164342798	0.137741047
0.178731311	0.239669421
0.179281646	0.611570248
0.179281646	0.826446281
0.179455695	0.738292011
0.179455695	0.694214876
0.179455695	0.137741047
0.179455695	0.123966942
0.179455695	0.96969697
0.179455695	0.845730028
0.179455695	0.707988981
0.179455695	0.85399449
0.179455695	0.225895317
0.180970576	0.132231405
0.180970576	0.820936639
0.180970576	0.570247934
0.181110903	0.809917355
0.18513607	0.727272727
0.186982039	0.876033058
0.188105223	0.443526171
0.194948306	0.123966942
0.194948306	0.300275482
0.194948306	0.964187328
0.19913936	0.236914601
0.19913936	0.256198347
0.19913936	0.845730028
0.19913936	0.804407713

ESM Table 7. Clinical and demographic data of the study participants from the N the age categories (<7 years, 7-12 years, >=13 years).

	<7 years	7-12 yeras
Total participants (N)	23	40
Sex		
Female	11	18
Male	12	22
BMI SDS at baseline		
Mean (SD)	0.41 (1.11)	0.24 (0.94)
Median (IQR) [min, max]	0.68 (1.23) [-2.32, 1.78]	0.17 (1.23) [-1.18, 2.28]
Glucose reading at baseline (mmol/mol)		
Mean (SD)	6.11 (3.10)	7.92 (5.58)
Median (IQR) [min, max]	5.4 (1.9) [3.6, 19.1]	6.3 (2.2) [3.8, 31.5]
HbA1c at diagnosis (mmol/mol)		
Mean (SD)	86.8 (26.3)	105 (20.3)
Median (IQR) [min, max]	85.8 (32) [4.2, 150]	105 (30.7) [51, 131]
HbA1c at baseline (mmol/mol)		
Mean (SD)	62.5 (21.3)	78.9 (25.2)
Median (IQR) [min, max]	66 (21) [8.7, 85.8]	80.3 (30.1) [13.4, 130]
Insulin dose at baseline (IU/kg)		
Mean (SD)	0.40 (0.18)	0.53 (0.25)
Median (IQR) [min, max]	0.45 (0.24) [0.11, 0.74]	0.54 (0.25) [0.036, 1.2]
Fasted C-peptide at baseline (pmol / l)		
Mean (SD)	123 (62.9)	227 (169)
Median (IQR) [min, max]	104 (39) [65.4, 280]	199 (199) [25.8, 659]
GADA		
Negative	7	8
Positive	6	23
IA-2A		
Negative	3	9
Positive	10	22
IAA		
Negative	2	6
Positive	11	25
ZnT8A		
Negative	7	8
Positive	6	23
Detectable autoantibodies		
Mean (SD)	2.57 (0.99)	3 (0.91)
Median (IQR) [min, max]	2 (1) [1, 4]	3 (2) [1, 4]

ND cohort stratified by

>=13 years
35
19
16
0.60 (1.30)
0.72 (1.66) [-2.0, 2,72]
8.57 (4.25)
7.4 (3.1) [3.7, 20.9]
107 (30.8)
111 (32) [4.2, 150]
77.5 (19.8)
81 (24.8) [40, 122]
0.59 (0.28)
0.57 (0.40) [0.093, 1.2]
399 (289)
316 (271) [87.6, 1290]
2
23
9
16
10
15
6
19
2.8 (0.90)
3 (1) [1, 4]

ESM Table 8. Associations between bacterial abundances and functions, and diag effects models in MaAsLin2. Associations with variables included for statistical corr shown. The analysis included microbiome data from all study visits. More details on

Variable	Value
site	categorical
age_at_consent	continuous
diabetes	binary
visit	categorical

Other columns

Variable
Value
Beta-coefficient
Standard error
p-value
q-value
prevalence

Species	Variable
Lawsonibacter.asaccharolyticus	age_at_consent
Massilioclostridium.coli	age_at_consent
Negativibacillus.massiliensis	site
Clostridiales.bacterium.VE202.01	age_at_consent
Bifidobacterium.catenulatum.subsp..kashiwanohens	age_at_consent
Clostridiales.bacterium.Choco116	site
Clostridiales.bacterium.1_7_47FAA	site
Anaerotruncus.massiliensis	site
Clostridiales.bacterium.VE202.01	site
Haemophilus.parainfluenzae	age_at_consent
Anaerostipes.caccae	site
Eubacterium.sp..AM49.13BH	age_at_consent
Bifidobacterium.animalis.subsp..lactis	site
Clostridiales.bacterium.1_7_47FAA	site
Massilimicrobiota.timonensis	site
Clostridium.sp..AF34.10BH	site
Prevotella.corporis	site
Roseburia.inulinivorans	site
Candidatus.Borkfalkia.ceftriaxoniphila	site
Clostridiales.bacterium.VE202.01	site
Faecalibacterium.prausnitzii	visit
Lachnoclostridium.edouardi	site
Lactobacillus.paracasei.subsp..paracasei	site
Intestinimonas.butyriciproducens	site
Massilioclostridium.coli	site
Bifidobacterium.pseudocatenulatum	site
Sutterella.sp..KLE1602	diabetes
Butyricimonas.virosa	age_at_consent
Gordonibacter.urolithinfaciens	site
Firmicutes.bacterium.AF25.13AC	site
Clostridiales.bacterium.VE202.21	age_at_consent

Lactobacillus.paracasei.subsp..paracasei	site
Paraprevotella.xylaniphila	site
Coprococcus.catus	site
Bacteroides.massiliensis	site
Lawsonibacter.asaccharolyticus	site
X.Clostridium..innocuum	site
X.Clostridium..leptum	site
Bifidobacterium.animalis.subsp..lactis	site
Clostridiales.bacterium.1_7_47FAA	site
Clostridium.sp..TF06.15AC	visit
Gordonibacter.pamelaeae	age_at_consent
Bifidobacterium.animalis.subsp..lactis	site
Bifidobacterium.pseudocatenulatum	site
Clostridiaceae.bacterium.OF09.1	site
Clostridium.sp..TF06.15AC	site
Odoribacter.splanchnicus	site
Bifidobacterium.animalis.subsp..lactis	site
Clostridiales.bacterium.VE202.01	site
Propionibacterium.freudenreichii.subsp..shermanii	site
Bacteroides.fragilis	site
Clostridium.sp..TM06.18	site
Clostridium.sp..OF03.18AA	site
X.Clostridium..leptum	site
Clostridiales.bacterium.VE202.01	site
Alistipes.sp..5CPEGH6	age_at_consent
Bifidobacterium.animalis.subsp..lactis	site
Clostridiales.bacterium	site
Intestinimonas.butyriciproducens	site
Lactobacillus.paracasei.subsp..paracasei	site
Merdimonas.faecis	site
Parabacteroides.johnsonii	site
Ruminococcus.lactaris	visit
Ruthenibacterium.lactatiformans	gender
Sutterella.sp..KLE1602	site
Tyzzerella.nexilis	site
Alistipes.timonensis	site
Tyzzerella.nexilis	site
Bacteroides.fragilis	site
Clostridiales.bacterium.VE202.01	site
Clostridium.sp..AT4	site
Eubacterium.sp..AM49.13BH	site
Ruminococcus.sp..AF17.22AC	site
Veillonella.atypica	age_at_consent
Blautia.obeum	site
Clostridiales.bacterium.Choco116	site
X.Clostridium..leptum	site
Anaerostipes.hadrus	age_at_consent
Dorea.longicatena	site
Bifidobacterium.adolescentis	site
Bifidobacterium.animalis.subsp..lactis	site
Blautia.massiliensis	site
Coprobacter.secundus	visit

Romboutsia.timonensis	site
Ruminococcus.sp..AF17.22AC	visit
Ruminococcus.sp..AF21.42	age_at_consent
Streptococcus.parasanguinis	gender
Tidjanibacter.massiliensis	visit
Clostridiales.bacterium.VE202.01	site
Veillonella.dispar	age_at_consent
Alistipes.timonensis	site
Phascolarctobacterium.faecium	gender
Agathobaculum.butyriciproducens	site
Asaccharobacter.celatus	site
Bacteroides.vulgatus	site
Victivallales.bacterium.CCUG.44730	site
Bifidobacterium.animalis.subsp..lactis	site
Gordonibacter.pamelaeae	site
Bacteroides.fragilis	site
Bacteroides.massiliensis	site
Propionibacterium.freudenreichii.subsp..shermanii	site
Alistipes.finegoldii	site
Bacteroides.fragilis	site
Collinsella.aerofaciens	site
Lactobacillus.sakei.subsp..sakei	site
Phascolarctobacterium.faecium	visit
Pseudoflavonifractor.sp..An184	site
Bifidobacterium.pseudocatenulatum	site
Bacteroides.fragilis	site
Intestinibacter.bartlettii	age_at_consent
Gordonibacter.urolithinfaciens	site
Clostridiales.bacterium.VE202.01	site
Lactococcus.lactis.subsp..lactis	site
Massilioclostridium.coli	site
Massilioclostridium.coli	site
Roseburia.hominis	site
Schaalia.odontolytica	site
X.Clostridium..asparagiforme	site
Bacteroidaceae.bacterium	diabetes
Escherichia.coli	site
Lachnospiraceae.bacterium	site
Lawsonibacter.asaccharolyticus	site
Olsenella.sp..GAM18	site
Coprobacter.fastidiosus	site
Alistipes.finegoldii	site
Anaerobutyricum.hallii	site
Bacteroides.coprocola	site
Bifidobacterium.animalis.subsp..lactis	site
Senegalimassilia.anaerobia	site
X.Eubacterium..rectale	site
Anaerostipes.hadrus	site
Ruminococcus.sp..AF46.10NS	site
Bacteroides.thetaiotaomicron	site
Bifidobacterium.dentium	visit
Bifidobacterium.animalis.subsp..lactis	site

Clostridiaceae.bacterium.OF09.1	site
Clostridiales.bacterium.VE202.01	site
Clostridium.paraputrificum	site
Coprobacillus.cateniformis	age_at_consent
Lawsonibacter.asaccharolyticus	site
Odoribacter.splanchnicus	site
Parabacteroides.johnsonii	site
Blautia.hydrogenotrophica	site
Erysipelatoclostridium.ramosum	site
Streptococcus.salivarius	age_at_consent
Ruminococcaceae.bacterium.D5	site
Methanobrevibacter.smithii	visit
Anaerostipes.hadrus	site
Gordonibacter.pamelaeae	site
Pseudoflavonifractor.sp..BIOML.A3	visit
Senegalimassilia.anaerobia	visit
Slackia.isoflavoniconvertens	site
Bacteroides.dorei	visit
Blautia.hydrogenotrophica	visit
X.Clostridium..spiroforme	site
Allisonella.histaminiformans	site
Clostridium.sp..AM49.4BH	site
Candidatus.Borkfalkia.ceftriaxoniphila	site
Catabacter.hongkongensis	site
Coprobacter.secundus	visit
Eisenbergiella.tayi	gender
Methanobrevibacter.smithii	site
Lachnoclostridium.edouardi	visit

Module	Name
MF0035	glycine degradation
MF0102	sulfate reduction (dissimilator
MF0025	tryptophan degradation
MF0038	methionine degradation I
MF0073	pyruvate:ferredoxin oxidoredu
MF0059	anaerobic fatty acid beta-oxid
MF0082	putrescine degradation
MF0063	glyoxylate bypass
MF0013	allose degradation
MF0095	propionate production III
MF0088	butyrate production I
MF0035	glycine degradation
MF0038	methionine degradation I
MF0040	proline degradation
MF0084	succinate conversion to propio
MF0027	tyrosine degradation II
MF0041	valine degradation I
MF0008	maltose degradation
MF0030	glutamate degradation I
MF0041	valine degradation I
MF0076	4-aminobutyrate degradation
MF0080	lactate consumption II

nosis of type 1 diabetes in the UFM cohort from linear mixed

rections (clinical site, age, gender, time from baseline) are also

n the tested variables are given below.

Explanation

Clinical site. Comparisons to University of Helsinki (highest n)

Age at recruitment to INNODIA

Patient was diagnosed with type 1 diabetes after joining the INNODIA study

Study visit

metadata variable that was associated with the species in question

value of the metadata which has the given association

the fixed effect coefficient of the given value

Standard error of the beta-coefficient

nominal p-value for the given association

corrected p-value, fdr q-value, for the given association

Prevalence (proportion of samples with presence) of the given bacterial species

Value	Beta-coefficie	Standard error	p-value
age_at_consent	0.335122489	0.042221402	7.25059E-13
age_at_consent	0.252584745	0.031981081	4.52139E-13
INSERM - Institut National de la Sante et de la Recherche Medi	1.434019933	0.19397604	9.76269E-12
age_at_consent	-0.2862844	0.04416239	1.24266E-09
age_at_consent	-0.56238034	0.092237415	8.0644E-09
MUW - Medical University Vienna	2.610924275	0.467466551	5.50395E-08
Paris - Centre Hospitalier Sud-Francilien	1.002192412	0.186437734	1.50082E-07
IT- Ospedale Pediatrico Bambino Gesu	1.695618513	0.339960766	1.25544E-06
ULB - Univer Libre de Bruxelles	-0.81356042	0.166127654	2.47628E-06
age_at_consent	-0.24326682	0.050129391	3.20037E-06
INSERM - Institut National de la Sante et de la Recherche Medi	0.93523367	0.1998324	6.44374E-06
age_at_consent	-0.33991703	0.075898624	1.50225E-05
ULB - Univer Libre de Bruxelles	-0.68030134	0.151627893	1.70174E-05
B - ZH Geel	0.964565502	0.224243417	2.24927E-05
Paris - Hospital Jean-Verdier	0.941462967	0.218193781	3.29372E-05
B - ZH Geel	-2.79424661	0.671698886	4.09898E-05
MUW - Medical University Vienna	1.650984844	0.39584589	3.83566E-05
INSERM - Institut National de la Sante et de la Recherche Medi	-1.4370059	0.337612102	4.16535E-05
ULB - Univer Libre de Bruxelles	-1.12888634	0.273854191	6.03131E-05
KU Leuven	-0.65308971	0.160557107	7.65562E-05
UFM Visit 6 (36 months)	-0.52634553	0.13214515	8.41948E-05
Paris - Centre Hospitalier Sud-Francilien	0.983721703	0.244255284	7.70059E-05
UK - Northampton General Hospital	0.907886287	0.224139866	8.15729E-05
ULB - Univer Libre de Bruxelles	-0.77453186	0.19469134	0.000106954
ULI - University of Ljubljana	-0.67436475	0.169247738	0.000102761
KU Leuven	-1.2938457	0.327141036	0.000117018
TRUE	1.195743432	0.30327905	0.000117538
age_at_consent	0.272629518	0.069409728	0.000126385
UCAM - University of Cambridge	0.696487423	0.179755565	0.000152134
Paris - Centre Hospitalier Sud-Francilien	-1.52188011	0.400926313	0.000179186
age_at_consent	-0.14109178	0.036982511	0.00019979

OUS - Oslo Univertssykehus HF	0.427223928	0.113659189	0.000220612
UK - Northampton General Hospital	1.626221165	0.429446255	0.000214716
INSERM - Institut National de la Sante et de la Recherche Medi	-1.52199328	0.403278318	0.000231103
UK-Birmingham Childrens Hospital	3.286014616	0.879882111	0.000250875
ULB - Univer Libre de Bruxelles	-0.59614286	0.159043631	0.000264025
INSERM - Institut National de la Sante et de la Recherche Medi	1.077953829	0.286924293	0.00026538
ULB - Univer Libre de Bruxelles	-0.57559761	0.15296211	0.000264261
Paris - Hospital Robert Debre	-0.75946786	0.20320323	0.000279351
Paris - Hospital Robert Debre	0.372182815	0.100119921	0.000305624
UFM Visit 6 (36 months)	-0.5488645	0.151216537	0.000331074
age_at_consent	-0.17818567	0.048654495	0.00034271
CHL - Centre Hospitalier de Luxembourg	-0.67639359	0.183721909	0.000382217
ULB - Univer Libre de Bruxelles	-1.2286484	0.338445103	0.00038534
INSERM - Institut National de la Sante et de la Recherche Medi	-1.2524533	0.344735833	0.00038616
INSERM - Institut National de la Sante et de la Recherche Medi	-1.44055365	0.395869152	0.000369477
UK-Birmingham Childrens Hospital	-1.68268684	0.465059527	0.000366978
KU Leuven	-0.53423079	0.14665135	0.000398918
Paris - Centre Hospitalier Sud-Francilien	-1.27836899	0.355986475	0.000409245
MUW - Medical University Vienna	1.509398647	0.425719179	0.000438462
KU Leuven	-1.32145045	0.370264856	0.000473899
Paris - Hospital Jean-Verdier	-1.64983951	0.463493159	0.000508089
UK - Northampton General Hospital	-2.12869029	0.608738965	0.000608771
ULI - University of Ljubljana	-0.75558849	0.213857597	0.000602152
UCAM - University of Cambridge	-0.70812813	0.204486201	0.000689705
age_at_consent	0.226988423	0.066576158	0.000821241
SUM - Slaski Uniwersytet Medyczny w Katowicach	-0.66784432	0.194732693	0.000793913
ULB - Univer Libre de Bruxelles	-0.44655153	0.129994678	0.000775088
UCAM - University of Cambridge	-0.82438493	0.241121197	0.000797722
UULM - Universitat Ulm	0.973596737	0.282239962	0.000720227
IT- Ospedale Pediatrico Bambino Gesu	1.92757387	0.568555614	0.000820162
MUW - Medical University Vienna	2.487786762	0.727281607	0.00076588
UFM Visit 6 (36 months)	-0.83494527	0.247448558	0.000833197
MALE	-0.23405795	0.068453441	0.000826173
UK-Birmingham Childrens Hospital	3.063285868	0.90025879	0.000819376
INSERM - Institut National de la Sante et de la Recherche Medi	1.128606393	0.328840972	0.000798675
MUW - Medical University Vienna	1.598495476	0.474322001	0.000887431
Paris - Hospital Robert Debre	0.859726495	0.254529099	0.000942992
ULB - Univer Libre de Bruxelles	-1.27824242	0.382972321	0.001049873
INSERM - Institut National de la Sante et de la Recherche Medi	-0.9621442	0.287571846	0.001042919
HKA - Hannoversche Kinderheilanstalt	0.766334214	0.228609049	0.000990175
HKA - Hannoversche Kinderheilanstalt	-1.12055102	0.33543278	0.001042859
CHL - Centre Hospitalier de Luxembourg	0.714267252	0.213808941	0.001048147
age_at_consent	-0.13593587	0.040685869	0.001051886
MUG - Medical University of Graz	-0.44229186	0.13268714	0.001081883
Paris - Hospital Jean-Verdier	0.903751598	0.270990142	0.001109434
Paris - Hospital Jean-Verdier	-1.10014774	0.327484081	0.001101183
age_at_consent	-0.10530584	0.031674708	0.001133728
Paris - Hospital Jean-Verdier	-0.88404703	0.266347658	0.001165441
Paris - Hospital Robert Debre	-1.40840945	0.428800426	0.001261184
UOUL - Oulun Yliopisto	0.53648404	0.16227529	0.00127899
UNISI - Universita degli Studi di Siena	-1.06029434	0.323510906	0.001273778
UFM Visit 3 (12 months)	0.168184139	0.051677957	0.001262591

UCAM - University of Cambridge	0.713152662	0.217186253	0.001298874
UFM Visit 4 (18 months)	0.194103037	0.05993656	0.0013218
age_at_consent	0.198988632	0.060442632	0.001226134
MALE	-0.26804304	0.081783885	0.001314794
UFM Visit 6 (36 months)	-0.6494746	0.199319751	0.001249969
UK - Norfolk and Norwich	-0.89017473	0.273993373	0.001386911
age_at_consent	-0.14015444	0.043118994	0.001414058
B - ZH Geel	1.088195264	0.338579391	0.001506453
MALE	0.677354929	0.209764908	0.001496496
INSERM - Institut National de la Sante et de la Recherche Medi	-0.95021605	0.293721252	0.001539974
ULB - Univer Libre de Bruxelles	-0.84845347	0.263100352	0.001545463
UNISI - Universita degli Studi di Siena	-0.72099534	0.224754322	0.001590743
Lund University	1.525157453	0.475717114	0.001579693
HH-RH - Herlev University Hospital	-0.6529951	0.200907119	0.001621956
UNISI - Universita degli Studi di Siena	-0.88702649	0.27736126	0.001640611
CHL - Centre Hospitalier de Luxembourg	-1.53615374	0.480319021	0.001672996
UK - Norfolk and Norwich	1.866445234	0.584787909	0.001680156
INSERM - Institut National de la Sante et de la Recherche Medi	0.514592782	0.16027099	0.001692446
KU Leuven	-0.9953773	0.31291836	0.001784138
UK - Norfolk and Norwich	-1.99537903	0.627717371	0.001737634
INSERM - Institut National de la Sante et de la Recherche Medi	-1.3814987	0.434167775	0.001762801
HKA - Hannoversche Kinderheilanstalt	0.291913434	0.09192508	0.001748956
UFM Visit 6 (36 months)	-0.47524768	0.150829766	0.001793769
INSERM - Institut National de la Sante et de la Recherche Medi	1.046803112	0.329244011	0.00178504
UCAM - University of Cambridge	-1.31853222	0.416372511	0.001850237
ULI - University of Ljubljana	-1.69901886	0.53789045	0.001892286
age_at_consent	-0.16103859	0.051117009	0.001970133
INSERM - Institut National de la Sante et de la Recherche Medi	0.788706773	0.251030624	0.00200583
UNISI - Universita degli Studi di Siena	-0.78367608	0.250914963	0.002105593
MUG - Medical University of Graz	0.512952356	0.164157749	0.002082468
UCAM - University of Cambridge	-0.46452847	0.148859261	0.002129486
UK - Norfolk and Norwich	-0.63353729	0.202803033	0.002048685
ULB - Univer Libre de Bruxelles	-0.65751411	0.209676922	0.002136233
ULB - Univer Libre de Bruxelles	0.138638845	0.043861693	0.002062291
SUM - Slaski Uniwersytet Medyczny w Katowicach	0.639796857	0.204338665	0.002085637
TRUE	0.932102468	0.300115241	0.002217548
UNISI - Universita degli Studi di Siena	1.277540765	0.41160631	0.002245537
Paris - Hospital Jean-Verdier	-0.92901443	0.29660965	0.002192846
Paris - Hospital Jean-Verdier	-1.08970396	0.34818008	0.002232631
Lund University	0.755983608	0.244111462	0.002259593
SUM - Slaski Uniwersytet Medyczny w Katowicach	1.28451484	0.415466099	0.002342225
OUS - Oslo Univertssykehus HF	-1.14971449	0.373138991	0.002419216
Paris - Hospital Jean-Verdier	-1.28095409	0.413649319	0.002433006
ULI - University of Ljubljana	0.92875317	0.302181031	0.002486292
HKA - Hannoversche Kinderheilanstalt	-0.57205464	0.185575529	0.002427319
CHL - Centre Hospitalier de Luxembourg	0.973414331	0.316473734	0.002470912
Paris - Hospital Jean-Verdier	-1.42840766	0.460811711	0.002470377
MUG - Medical University of Graz	-0.45163235	0.147302961	0.002532747
SUM - Slaski Uniwersytet Medyczny w Katowicach	0.579501363	0.189110964	0.002545235
UULM - Universitat Ulm	-2.46176375	0.804734726	0.002632633
UFM Visit 6 (36 months)	0.356317288	0.117904286	0.002670028
INSERM - Institut National de la Sante et de la Recherche Medi	-0.78869566	0.258572295	0.002844093

Paris - Hospital Jean-Verdier	-1.36719797	0.449540046	0.002837022
OUS - Oslo Univertssykehus HF	-0.58066103	0.191870574	0.00287315
SUM - Slaski Uniwersytet Medyczny w Katowicach	0.366123335	0.121140415	0.002893546
age_at_consent	-0.13310297	0.043718719	0.002785996
KU Leuven	-0.46813322	0.153509894	0.002764312
Paris - Hospital Jean-Verdier	-1.22502803	0.403152951	0.002880815
KU Leuven	0.628409967	0.207741274	0.002883748
INSERM - Institut National de la Sante et de la Recherche Medi	0.546005233	0.180654711	0.002935712
Paris - Hospital Jean-Verdier	0.936065173	0.307608473	0.002921964
age_at_consent	-0.13950772	0.046060214	0.002960404
IT- Ospedale Pediatrico Bambino Gesu	1.978327687	0.659617463	0.003010464
UFM Visit 4 (18 months)	0.174860346	0.058558902	0.003052507
ULB - Univer Libre de Bruxelles	-0.35843947	0.119257252	0.003156639
KU Leuven	-0.53066984	0.17690428	0.003148533
UFM Visit 6 (36 months)	-0.27427943	0.092116295	0.003129448
UFM Visit 6 (36 months)	-0.36257497	0.121667889	0.003118064
UNISI - Universita degli Studi di Siena	1.240678864	0.417774854	0.003407772
UFM Visit 6 (36 months)	0.570214376	0.193512917	0.003467485
UFM Visit 4 (18 months)	0.109155306	0.037160924	0.003537734
INSERM - Institut National de la Sante et de la Recherche Medi	1.003272333	0.339141913	0.003572653
ULB - Univer Libre de Bruxelles	0.447245771	0.151342055	0.00359971
ULB - Univer Libre de Bruxelles	-0.89756661	0.305461192	0.003830315
HKA - Hannoversche Kinderheilanstalt	-0.92959309	0.3190278	0.004072266
IT- Ospedale Pediatrico Bambino Gesu	0.984941351	0.339080673	0.004084253
UFM Visit 4 (18 months)	0.164882977	0.056994633	0.00408904
MALE	-0.25459963	0.087253704	0.004102857
IT- Ospedale Pediatrico Bambino Gesu	2.56055167	0.883404312	0.004186284
UFM Visit 4 (18 months)	-0.11571835	0.04017317	0.004231676

Variable	Value	Beta-coefficie	Standard error
gender	MALE	-0.12434623	0.025594236
visit	UFM Visit 6 (36	0.616305806	0.136464052
gender	MALE	-0.14470089	0.033045732
gender	MALE	-0.10962925	0.027134238
diabetes	TRUE	0.018461388	0.004635053
site	UNISI - Univers	1.514868185	0.399790861
site	UNISI - Univers	1.516847017	0.400215668
site	UNISI - Univers	1.463845832	0.391594492
site	UK - Northamp	-1.4272442	0.386188366
site	MUW - Medica	1.509398647	0.425719179
gender	MALE	-0.09929977	0.027847578
site	Paris - Hospital	-0.38228244	0.109403646
site	Paris - Hospital	-0.40889524	0.116284826
gender	MALE	-0.10435321	0.030032693
site	UNISI - Univers	1.419968497	0.407011352
site	UK - Northamp	-0.84798931	0.2467083
site	INSERM - Instit	0.903511367	0.265126679
gender	MALE	-0.10004867	0.030464185
age_at_consent	age_at_consen	0.233738937	0.070835759
site	UNISI - Univers	0.817853119	0.248180043
site	UNISI - Univers	0.869721663	0.265135939
site	UNISI - Univers	1.42552588	0.42671997

q-value	prevalence
2.74942E-09	0.947046843
2.74942E-09	0.405295316
2.46801E-08	0.146639511
2.35608E-06	0.376782077
1.22321E-05	0.549898167
6.95699E-05	0.122199593
0.000162603	0.156822811
0.00119016	0.130346232
0.002086674	0.376782077
0.002427163	0.798370672
0.004442669	0.177189409
0.009494237	0.619144603
0.009927696	0.228105906
0.012184604	0.156822811
0.016653063	0.16496945
0.017550028	0.959266802
0.017550028	0.17107943
0.017550028	0.953156823
0.02407443	0.533604888
0.027762317	0.376782077
0.027762317	0.99185336
0.027762317	0.181262729
0.027762317	0.120162933
0.032445455	0.83503055
0.032445455	0.405295316
0.033015063	0.464358452
0.033015063	0.252545825
0.034232411	0.386965377
0.039785563	0.393075356
0.045298308	0.945010183
0.048877579	0.684317719

0.050700633	0.120162933
0.050700633	0.140529532
0.051549519	0.867617108
0.052964198	0.33808554
0.052964198	0.947046843
0.052964198	0.727087576
0.052964198	0.890020367
0.054323097	0.228105906
0.057946375	0.156822811
0.061240595	0.896130346
0.061883674	0.617107943
0.062311448	0.228105906
0.062311448	0.464358452
0.062311448	0.902240326
0.062311448	0.896130346
0.062311448	0.918533605
0.063029054	0.228105906
0.063341049	0.376782077
0.066505884	0.114052953
0.070471537	0.637474542
0.074102824	0.855397149
0.085498522	0.853360489
0.085498522	0.890020367
0.095104061	0.376782077
0.095741872	0.189409369
0.095741872	0.228105906
0.095741872	0.99592668
0.095741872	0.83503055
0.095741872	0.120162933
0.095741872	0.209775967
0.095741872	0.120162933
0.095741872	0.66395112
0.095741872	0.987780041
0.095741872	0.252545825
0.095741872	0.201629328
0.10045191	0.103869654
0.10517135	0.201629328
0.10780414	0.637474542
0.10780414	0.376782077
0.10780414	0.193482688
0.10780414	0.619144603
0.10780414	0.437881874
0.10780414	0.419551935
0.109272022	0.99592668
0.109272022	0.122199593
0.109272022	0.890020367
0.110233255	0.977596741
0.11188231	0.957230143
0.113915161	0.759674134
0.113915161	0.228105906
0.113915161	0.949083503
0.113915161	0.342158859

0.113915161	0.949083503
0.113915161	0.437881874
0.113915161	0.439918534
0.113915161	0.788187373
0.113915161	0.279022403
0.118183526	0.376782077
0.119157926	0.49898167
0.12418416	0.103869654
0.12418416	0.393075356
0.124689251	0.969450102
0.124689251	0.879837067
0.125668684	0.99389002
0.125668684	0.323828921
0.126813575	0.228105906
0.126963234	0.617107943
0.127084245	0.637474542
0.127084245	0.33808554
0.127084245	0.114052953
0.127139688	0.881873727
0.127139688	0.637474542
0.127139688	0.869653768
0.127139688	0.114052953
0.127139688	0.393075356
0.127139688	0.33401222
0.129927731	0.464358452
0.131661453	0.637474542
0.135831711	0.885947047
0.137046981	0.393075356
0.137298238	0.376782077
0.137298238	0.317718941
0.137298238	0.405295316
0.137298238	0.405295316
0.137298238	0.914460285
0.137298238	0.128309572
0.137298238	0.344195519
0.139323226	0.871690428
0.139323226	0.83910387
0.139323226	0.983706721
0.139323226	0.947046843
0.139323226	0.101832994
0.143253518	0.543788187
0.145046444	0.881873727
0.145046444	0.926680244
0.145046444	0.105906314
0.145046444	0.228105906
0.145046444	0.364562118
0.145046444	0.973523422
0.146235293	0.977596741
0.146235293	0.122199593
0.15011948	0.916496945
0.151115594	0.118126273
0.154539799	0.228105906

0.154539799	0.902240326
0.154539799	0.376782077
0.154539799	0.105906314
0.154539799	0.266802444
0.154539799	0.947046843
0.154539799	0.918533605
0.154539799	0.120162933
0.154614156	0.183299389
0.154614156	0.242362525
0.154839327	0.940936864
0.156379195	0.297352342
0.15748443	0.236252546
0.158542727	0.977596741
0.158542727	0.617107943
0.158542727	0.140529532
0.158542727	0.364562118
0.170029893	0.299389002
0.171878449	0.702647658
0.174221928	0.183299389
0.174806441	0.389002037
0.175001276	0.134419552
0.185026162	0.759674134
0.193267514	0.533604888
0.193267514	0.156822811
0.193267514	0.342158859
0.193267514	0.409368635
0.195980102	0.236252546
0.196889768	0.181262729

p-value	q-value	prevalence
3.13212E-06	0.009922544	1
9.09731E-06	0.014410142	0.24439919
2.29805E-05	0.024267358	1
8.51637E-05	0.059443639	1
9.38189E-05	0.059443639	1
0.000211631	0.095777949	0.86761711
0.000210661	0.095777949	0.86761711
0.000256879	0.101724239	0.88187373
0.000298058	0.104916393	0.99592668
0.000438462	0.138904695	0.11405295
0.000482798	0.139045756	1
0.000690745	0.145885442	1
0.000625022	0.145885442	1
0.000660127	0.145885442	1
0.000621131	0.145885442	0.85132383
0.000749306	0.148362601	0.99796334
0.000895896	0.166952827	0.49898167
0.001266083	0.182315985	1
0.001187439	0.182315985	0.26680244
0.001237816	0.182315985	0.49898167
0.00125303	0.182315985	0.99389002
0.001036205	0.182315985	0.85132383