Original research

Gestational diabetes is driven by microbiota-induced inflammation months before diagnosis

Yishay Pinto [ID],[1] Sigal Frishman,[2,3] Sondra Turjeman [ID],[1] Adi Eshel,[1] Meital Nuriel-Ohayon,[1] Oshrit Shrossel,[4] Oren Ziv,[1] William Walters [ID],[5] Julie Parsonnet,[6,7] Catherine Ley,[6] Elizabeth L Johnson [ID],[8] Krithika Kumar,[8] Ron Schweitzer,[9,10] Soliman Khatib [ID],[9,10] Faiga Magzal [ID],[11,12] Efrat Muller,[13] Snait Tamir [ID],[11,12] Kinneret Tenenbaum-Gavish,[2] Samuli Rautava [ID],[14,15] Seppo Salminen,[16] Erika Isolauri [ID],[14] Or Yariv,[17] Yoav Peled,[2,17] Eran Poran,[17] Joseph Pardo,[2,17] Rony Chen,[2] Moshe Hod,[2] Elhanan Borenstein [ID],[13,18,19] Ruth E Ley,[5] Betty Schwartz,[3] Yoram Louzoun,[4] Eran Hadar [ID],[2] Omry Koren [ID][1]

ABSTRACT

Objective Gestational diabetes mellitus (GDM) is a condition in which women without diabetes are diagnosed with glucose intolerance during pregnancy, typically in the second or third trimester. Early diagnosis, along with a better understanding of its pathophysiology during the first trimester of pregnancy, may be effective in reducing incidence and associated short-term and long-term morbidities.

Design We comprehensively profiled the gut microbiome, metabolome, inflammatory cytokines, nutrition and clinical records of 394 women during the first trimester of pregnancy, before GDM diagnosis. We then built a model that can predict GDM onset weeks before it is typically diagnosed. Further, we demonstrated the role of the microbiome in disease using faecal microbiota transplant (FMT) of first trimester samples from pregnant women across three unique cohorts.

Results We found elevated levels of proinflammatory cytokines in women who later developed GDM, decreased faecal short-chain fatty acids and altered microbiome. We next confirmed that differences in GDM-associated microbial composition during the first trimester drove inflammation and insulin resistance more than 10 weeks prior to GDM diagnosis using FMT experiments. Following these observations, we used a machine learning approach to predict GDM based on first trimester clinical, microbial and inflammatory markers with high accuracy.

Conclusion GDM onset can be identified in the first trimester of pregnancy, earlier than currently accepted. Furthermore, the gut microbiome appears to play a role in inflammation-induced GDM pathogenesis, with interleukin-6 as a potential contributor to pathogenesis. Potential GDM markers, including microbiota, can serve as targets for early diagnostics and therapeutic intervention leading to prevention.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The incidence of gestational diabetes mellitus (GDM) is increasing worldwide.
⇒ Early prediction of GDM may reduce short-term and long-term complications to the mother and the offspring.
⇒ At later stages of pregnancy, the gut microbiome of women diagnosed with GDM is different from the microbiome of women without GDM.
⇒ Insulin resistance has been associated with elevated secretion of proinflammatory cytokines.

WHAT THIS STUDY ADDS

⇒ Gut microbiome, metabolome and inflammatory markers were profiled during the first trimester of pregnancy in 394 women.
⇒ Significant differences were found in these markers between women who would and would not later develop GDM.
⇒ The GDM phenotype was transferred to germ-free mice following faecal microbiota transplant from women in their first trimester of pregnancy.
⇒ Accurate prediction of GDM development was made based on first trimester biomarker profiles and clinical data.
⇒ This study suggests diagnosis of GDM/GDM risk can be made earlier allowing for earlier management or even complete prevention.

INTRODUCTION

Gestational diabetes mellitus (GDM), or development of glucose intolerance during pregnancy in women without diabetes, occurs when the pancreas cannot produce enough insulin to balance insulin-inhibiting effects of placental hormones (viz. oestrogen, cortisol and human placental lactogen).[1] Approximately 10% of pregnant women worldwide are diagnosed with GDM. Risk factors include non-white ethnicity, increased maternal age, obesity, family history of diabetes and history of giving birth to large

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Recognition of women at risk of GDM at an early stage of pregnancy, with appropriate risk stratification, may allow specific recommendations for prevention of the disease—currently by lifestyle modification and in the future perhaps by specific pre/pro/postbiotic supplementation.

⇒ If GDM can be prevented, there would be a major reduction in adverse outcomes of GDM, for the mother and offspring, in both the short term and long term.

infants. Consequences of GDM include a wide range of obstetrical and metabolic complications for both the mother (eg, pre-eclampsia, type 2 diabetes and cardiovascular diseases) and the neonate (mainly macrosomia and hypoglycaemia).[2] Many complications are preventable if GDM is detected and appropriately managed and good glycaemic control is achieved by nutrition, exercise and insulin administration, if necessary, along with heightened monitoring during labour and delivery,[3] but earlier detection might allow for complete amelioration of GDM-associated short-term and long-term risks.

The incidence of GDM is increasing worldwide, due primarily to the increase in prevalence of overweight and obesity, advanced maternal age and growth of at-risk populations.[4–6] As such, it is important to expand early-prediction efforts towards reducing its negative consequences. To date, few studies have examined biomarkers of GDM in the first trimester (T1).[7 8] Additionally, while gut microbial dysbiosis has been associated with diabetes,[9] and a recent study has associated gut dysbiosis with GDM in the third trimester (T3),[10] few have focused on T1.[8 11–15]

We sought to identify biomarkers of GDM in T1 of pregnancy. First, we comprehensively profiled the T1 gut microbiome, metabolome and inflammatory cytokine profiles of women who would and would not later be diagnosed with GDM. We then investigated whether the early pregnancy microbiome drove GDM development using germ-free (GF) mice. Finally, we used a machine learning approach to predict GDM based on patient characteristics, T1 microbiome and clinical information, to identify earlier time frames for therapeutic intervention.

METHODS
Pregnant women
Primary prospective cohort
We enrolled a prospective cohort followed throughout pregnancy (online supplemental figure 1). Upon screening for GDM in the second trimester (T2; screening method described in online supplemental methods), women were retroactively classified as 'would go on to develop GDM' and 'would not go on to develop GDM'. This main prospective cohort included 394 pregnant women aged 18–40 years recruited between gestational ages (weeks+days) 11+0–13+6 at women's health centres of Clalit HMO (Dan Petach Tikva District, Israel) during the years 2016–2017. Exclusion criteria included: type 1 or type 2 diabetes mellitus diagnosed before pregnancy (all other chronic diseases were documented in the database); in vitro fertilisation or hormonal therapy in the previous 3 months; use of antibiotics in the previous 3 months and multiple gestation. Initially, 400 women were recruited, but 4 did not provide any samples and 2 did not meet study criteria upon further examination

of medical records (one with antibiotics use, one with type 2 diabetes; online supplemental figure 1). Thus, 394 women were followed through 27–31 weeks of pregnancy; in this study, no women were lost to follow-up as following initial recruitment, all other data (namely GDM diagnosis) could be obtained from digital medical records. Weight and height were assessed at the time of recruitment and blood and faecal samples collected (see online supplemental methods). Dietary consumption (24-hour recall), physical activity (24-hour recall), sleeping hours (3-day recall), stress (validated questionnaire[16]), employment and education details (at recruitment) were recorded. Other maternal demographics, clinical and obstetrical data including pregnancy follow-up and comorbidities were extracted from medical records.

Secondary cohort
Since GDM incidence in Israel is about 10%,[17] a secondary cohort of pregnant women was also recruited. Patients with GDM were enrolled in a cohort study at 24–28 gestational weeks at Rabin Medical Center between the years 2016 and 2017. Exclusion criteria for this cohort were the same as for the main cohort. Medical chart review was performed to identify all demographic and clinical characteristics from T1. Clinical data, but not biological samples, from this secondary cohort are included in the study.

Additional cohorts
In addition to the above cohorts, for faecal microbiome transplant (FMT) experiments in GF models, two additional independent cohorts were included (see online supplemental methods).

Biomarker analysis in the primary cohort
Fasting glucose, liver enzymes and HbA1c were extracted from medical records and serum cytokine and hormone panels performed (online supplemental methods). Bacterial DNA was extracted, amplified (V4 region of the 16S rRNA gene) and sequenced (Illumina MiSeq) from all faecal samples as described in the online supplemental methods. QIIME2 V.2019.4[18] was used for read pre-processing (pipeline in online supplemental methods). Faecal short-chain fatty acid (SCFA) extraction and untargeted metabolomics methods are also described in the online supplemental methods.

FMT into GF mice
Transplantation experiments were performed using faecal samples from the primary prospective cohort and the two additional cohorts (see online supplemental methods).

Prediction
To predict GDM, we developed a prediction model using our prospective cohort (identified T1 biological markers and clinical data) as well as clinical data from our secondary cohort. We checked each combination of the following components: (1) cytokines, (2) microbiome, (3) general clinical information and (4) food questionnaires. The accuracy of the prediction was assessed using the area under the curve of the test set, in a 20%/80% test/training set division and a fivefold cross-validation (see online supplemental methods). To examine generalisability of our model, we applied the classifier to an independently published dataset from a Chinese cohort of 98 pairs of pregnant women with and without GDM (matched) who provided a faecal sample in week 10–15 of pregnancy.[8] We trained the

model on our primary cohort and tested the model performance on the Chinese cohort.

Statistical analysis

Full statistical methods are presented in the online supplemental methods. Briefly, unless otherwise specified, statistical analysis was done using non-parametric Mann-Whitney U tests followed by false discovery rate (FDR) correction. Mantel's correlations between study features were performed. Association of microbial features with GDM was done by Spearman's rank correlations compared with a background distribution followed by a linear model to control for main risk factors. For untargeted metabolomics, the differential abundance of the metabolites between the groups was identified by Student's t-tests and FDR correction. Microbial features of FMT-recipient mice were associated with GDM using MaAsLin2.[19] The MetaCyc pathway abundance in mouse faeces was predicted using PICRUSt2.[20]

Data availability

All sequencing data were submitted to European Bioinformatics Institue (EBI) (project accession number ERP143097). Metabolomics data were deposited at 10.5281/zenodo.6581068.[21] Ethics statement and patient and public involvement are described in online supplemental methods.

RESULTS
Study design

We prospectively recruited 394 women during T1, 44 (11%) of which went on to develop GDM, as diagnosed by glucose tolerance test (GTT) during the second trimester of pregnancy. The other 350 women served as the control group, hereafter 'healthy pregnant women' (online supplemental figure 1). Of the recruited women (regardless of GDM status), 8 suffered spontaneous abortion, 7 delivered preterm and 11 had gestational hypertension or pre-eclampsia. In addition, 4 had polycystic ovary syndrome and 25 had hypothyroidism. These were not exclusion criteria.

Of the 34 women in the GDM group who had blood work on file before pregnancy, 2 had high HbA1c; none had high glucose. Women diagnosed with GDM exhibited other common risk factors (table 1) such as higher maternal age and pre-pregnancy body mass index (BMI). Following pregnancy (6 weeks–6 months), we also examined HbA1c (or glucose) levels of these

women and found one woman with high HbA1c level (out of six who did this blood work) and none with impaired glucose levels (fasting test/75 g oral GTT, out of 22). While beyond the timeline of this T1 study, among women later diagnosed with GDM, dietary consultation/lifestyle change was not sufficient for nine women who therefore received medication to control their GDM.

When examining explained variance between parameters measured (microbiome, SCFA, metabolome, cytokines, hormones, diet and lifestyle; figure 1A), using a Mantel test, we found that the T1 gut microbiome significantly explained the variance of most measurements and was most tightly correlated with the faecal metabolomic profile (figure 1B).

Women with GDM exhibit elevated levels of serum inflammatory cytokines and low levels of SCFAs in T1

Following evidence of elevated inflammatory biomarkers in women diagnosed with GDM,[22] we profiled 10 plasma cytokines, chemokines and hormones in both the GDM (n=35) and control (n=78) groups and found elevated levels of proinflammatory cytokines (interleukin (IL)-4, IL-6, IL-8, granulocyte-macrophage colony-stimulating factor and tumour necrosis factor-α) among the GDM group (figure 1C; $p < 0.05$, FDR-corrected Mann-Whitney U tests) but no differences in leptin and insulin. This result was robust when controlling for BMI and age (see the online supplemental methods; $p < 0.05$, linear model with age and BMI as fixed or random effects).

Another possible early biomarker for GDM are SCFAs, which promote glucose homeostasis and suppress inflammatory response. We found a significant reduction of two branched SCFAs (BSCFAs), isovalerate and isobutyrate, in the GDM group (figure 1D; $p < 0.05$, FDR-corrected Mann-Whitney U tests) and a similar trend for valerate ($p=0.09$).

Gut microbiome is associated with GDM pathogenesis

A number of studies have suggested that the gut microbiome is altered in women with GDM, post-GDM diagnosis. In our study, we did not find differences in T1 gut microbiome α-diversity between women who would and would not develop GDM. Principal coordinate analysis of unweighted UniFrac distances demonstrated that the microbial communities of healthy women and women with GDM trend toward significant differences (figure 2A; $p=0.06$, permutational multivariate analysis of variance; $p=0.23, 0.05, 0.38$ for Bray-Curtis, Jaccard and weighted UniFrac, respectively), supported by results of differential abundance analyses (below). Notably, when fitting a linear model to the distance matrix with GDM outcome and the risk factors age and BMI, widely associated with GDM (sequentially using adonis2, see the online supplemental methods), none of the variables were significant.

We next aimed to characterise the specific subset of differentially abundant bacteria: 1 bacterial species was overrepresented and 16 bacteria under-represented in the GDM group. When repeating this analysis while controlling for age and BMI, we found 15 under-represented species in the GDM group (figure 2B), only 6 of which intersected with the prior, uncontrolled analysis. Controlling for confounding variables allowed us to distinguish between microbial species associated with main risk factors of GDM and the disease itself. We found a lower abundance of *Prevotella* in T1 samples of women who would develop GDM, and this result was replicated in mice (below).

Table 1	Cohort description		
	Control	**GDM**	**P value**
Women	350	44	
BMI	22.6±4.0 (325)	28.2±7.7 (43)	<0.0001
Age (years)	31.2±4.43 (350)	33.1±3.77 (44)	<0.01
Gestation	2.38±1.44 (349)	2.93±1.8 (44)	0.02
Parity	0.97±1.04 (349)	1.2±1.0 (44)	0.16
Fasting glucose (mg/dL)	82.46±7.08 (327)	89.30±14.1 (42)	<0.0001
Delivery week	39.1±1.73(297)	38.7±1.2 (39)	0.14
Newborn weight (kg)	3.2±0.46 (308)	3.24±1.2 (40)	0.66
Stress (rank: 1–4)	1.41±0.59 (232)	1.57±0.63 (30)	0.17
Education (years)	15.36±2.3 (251)	14.75±2.17 (33)	0.15
Current smoker	32	11	<0.01*
Non-smoker	280	28	
Prior smoker	13	1	

Data are presented as mean±SD. P values based on Student's t-test. Parentheses represent sample size.
*P value based on Fisher's exact test.
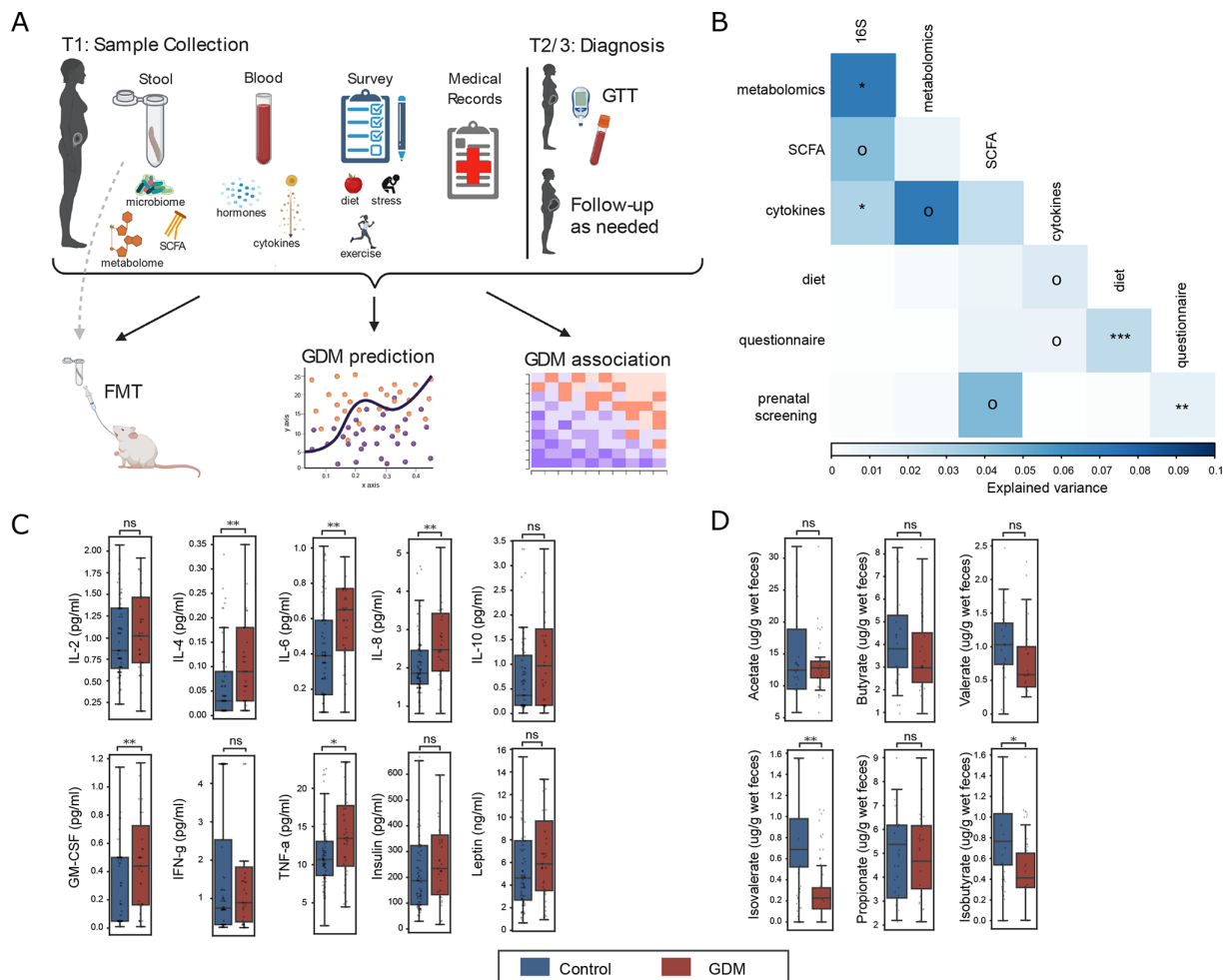BMI, body mass index; GDM, gestational diabetes mellitus.

Figure 1 First trimester blood and faecal biomarkers in women later diagnosed with GDM. (A) Sampling strategy and study design. Samples were collected in first trimester (T1). Stool was collected to profile gut microbiome (GDM: n=28, control: n=236), metabolome (n=15 age/BMI-matched pairs) and SCFAs (n=20 age-matched pairs) and to validate results when transplanted into germ-free mice. Blood samples were used to profile cytokines and hormones (GDM: n=35, control: n=78). Lifestyle surveys and medical records were collected from all participants. (B) Variance explained (square of the Mantel statistic) between all pairs of data types (Mantel test). (C) Serum levels of cytokines and hormones for GDM and control women (false discovery rate (FDR)-corrected Mann-Whitney U tests). (D) Concentration of faecal short-chain fatty acids (FDR-corrected Mann-Whitney U tests). Boxplots indicate the median and IQR; whiskers show IQR×1.5. °P<0.1, *p<0.05, **p<0.01, ***p<0.001. BMI, body mass index; FMT, faecal microbiota transplant; GDM, gestational diabetes mellitus; GM-CSF, granulocyte-macrophage colony-stimulating factor; GTT, glucose tolerance test; IFN, interferon; IL, interleukin; ns, not significant; SCFA, short-chain fatty acid; TNF, tumour necrosis factor.

Glucose impairment and elevated IL-6 levels of women with GDM were phenocopied to mice by FMT

To examine the causal role of the gut microbiome in the pathogenesis of GDM, faecal samples of age-matched and BMI-matched GDM and control samples from the primary cohort were transplanted to GF female mice (figure 3A). Microbiota characterisation was performed 7 and 21 days post-FMT. The GF mice acquired an average of 42 and 48 taxa from donor samples 7 and 21 days post-transfer, respectively. The recipient mice shared ~60% of their taxa with the donor on day 7 and ~55% on day 21 (online supplemental figure 2). On day 7, the microbial communities were significantly different between GDM-recipient and non-GDM recipient mice (figure 3B). Consistent with our observation in women, *P. copri* was found to be reduced in GDM-recipient mice (figure 3C). GTTs revealed GDM-recipient mice exhibited impaired glucose tolerance (figure 3D).

Further, the GDM-recipient mice exhibited elevated levels of both IL-6 (in agreement with our findings in women with GDM)

and IL-10 relative to the control-recipient mice (figure 3E and online supplemental table 1). No differences were found for insulin or leptin levels (online supplemental table 2). We found further support for the role of gut microbiota in GDM pathogenesis with FMT from two additional GDM cohorts (Finnish and American women) (figure 3D; online supplemental figure 3 and online supplemental tables 3–5; combined p=0.15, 0.022, 0.10, 0.24 for time points 0, 30, 60, 120 min, respectively, Fisher's method).

Lower levels of short peptides in stool of women with GDM

We next compared stool metabolome profiles of women who would and would not later develop GDM (n=15 age-matched and BMI-matched pairs). First, we found a significant correlation between the microbiome and metabolome of these women (r=0.26, p=0.02; Mantel test). Although we were limited in sample size, manual exploration of the data revealed that many short peptides had differential concentrations (raw p≤0.05)
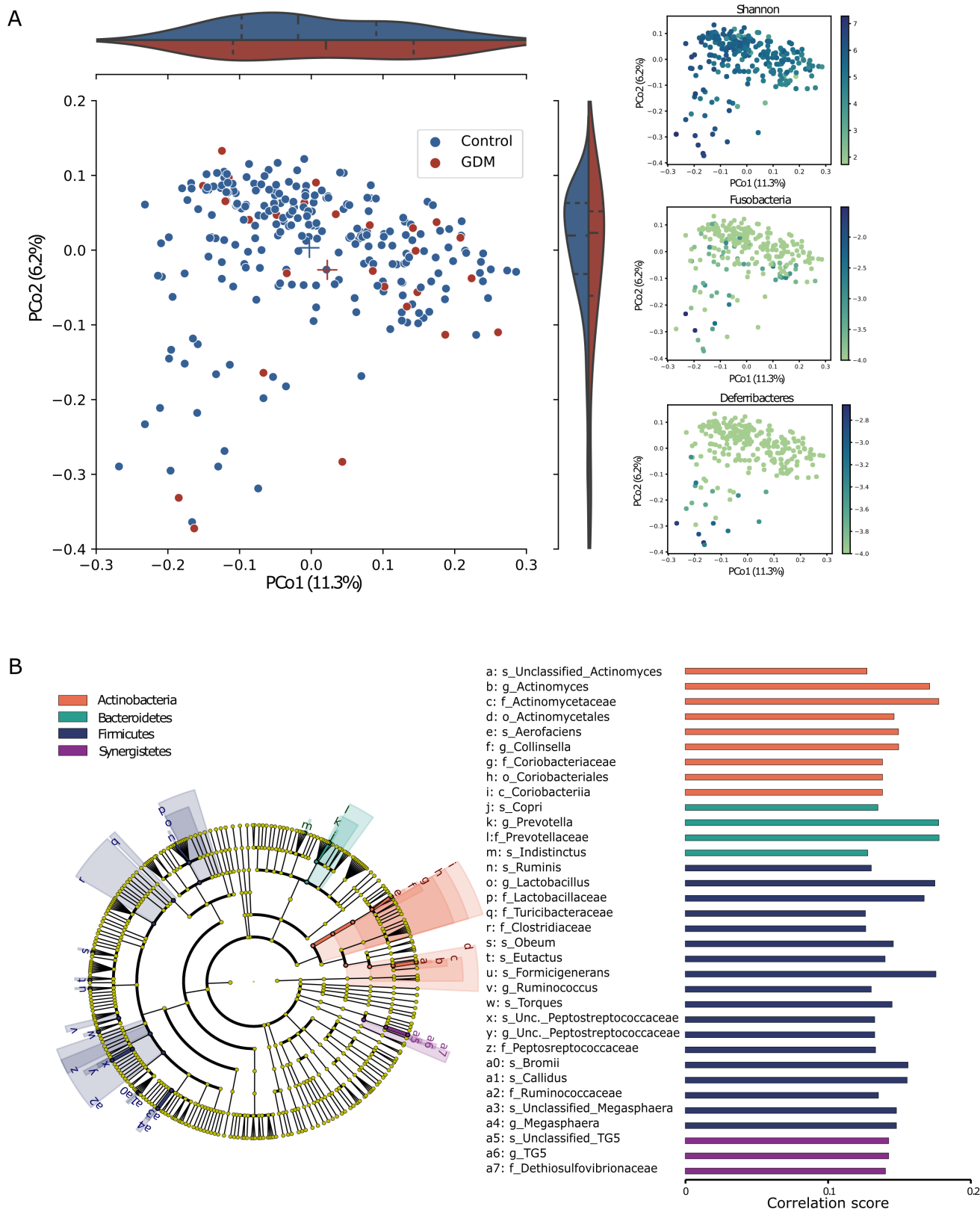
Figure 2 Differences in faecal microbiome composition in first trimester between women who would and would not develop GDM later. (A) Principal coordinate analysis based on 16S rRNA gene sequence profiling of the microbiome (GDM: n=28, control: n=236) using the unweighted UniFrac dissimilarity metric coloured by GDM/control (left; p=0.06, PERMANOVA); violin plots represent the distribution of GDM/control on each axis; Shannon diversity (top right; R^2=0.24 with PCo1) and two phyla that mostly explain the PCo1 and PCo2 variance: Fusobacteria (R^2=0.08 with PCo2) and Deferribacteres (R^2=0.3 with PCo2). (B) The cladogram represents the microbial features associated with the disease state, while controlling for the main risk factors, BMI and age, at all taxonomic ranks. Spearman's rank correlation for each association: a positive association (all associations found), implies over-represented features in the healthy control group. Cladogram and bars are coloured by phylum. BMI, body mass index; GDM, gestational diabetes mellitus; Unc., unclassified; PERMANOVA, permutational multivariate analysis of variance.
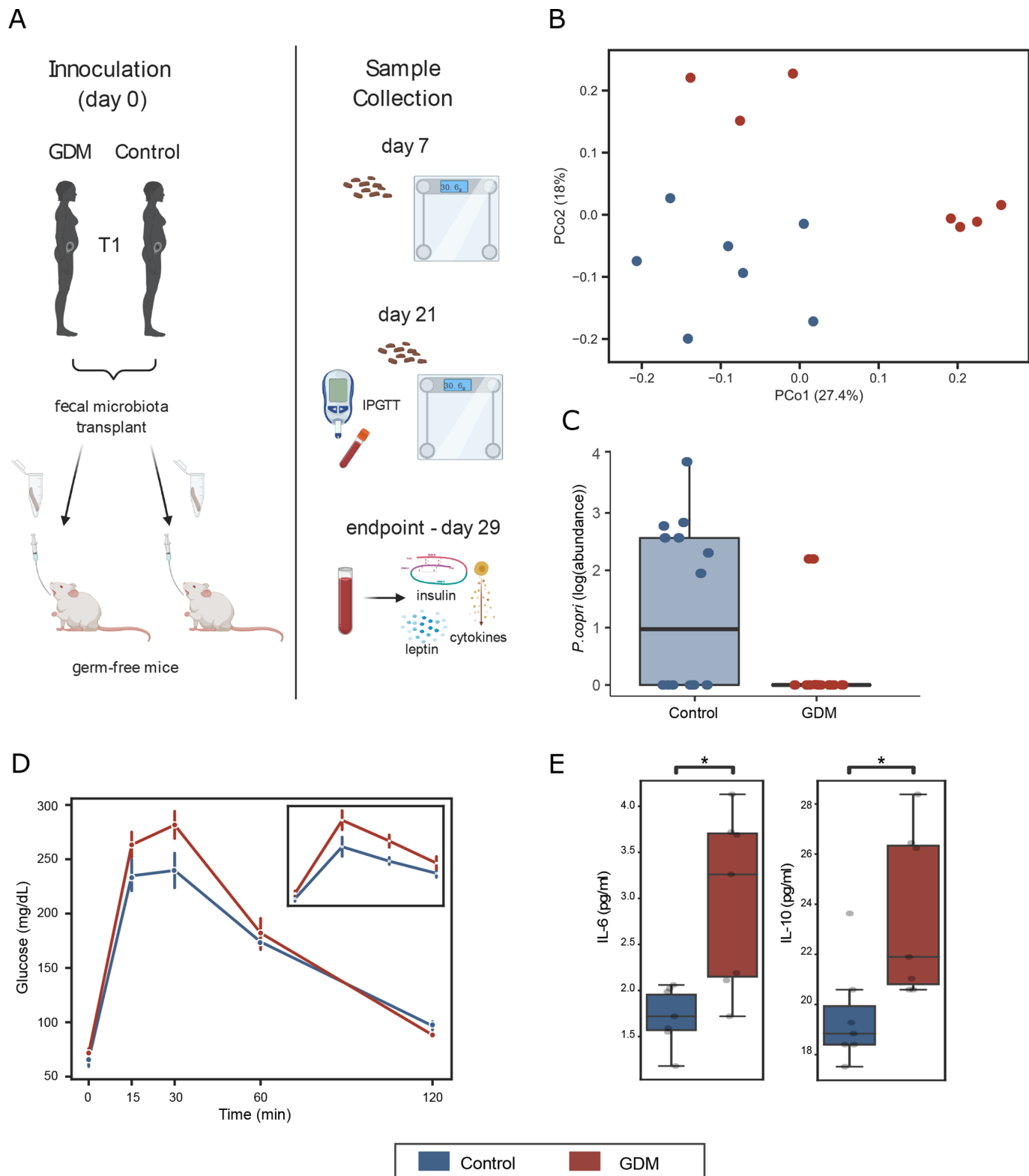
Figure 3 Phenotype transfer via first trimester (T1) FMT to germ-free mice. (A) Study design. (B) PCoA using the unweighted UniFrac metric. Mice receiving FMT from women with GDM exhibit different microbial profiles from mice receiving FMT from the control group (p=0.005, PERMANOVA test, n=7 age/BMI-matched FMT donor pairs). (C) *Prevotella copri,* which was found to be negatively associated with women with GDM, is negatively associated with GDM-transplanted mice as well (p=0.04, linear mixed-effects model). (D) Intraperitoneal glucose tolerance test (ipGTT) revealed impaired glucose sensitivity in mice transplanted with faeces from women with GDM in this study and in the Finnish cohort (insert) (error bars represent ±SEM; *p<0.05 one-tailed Mann-Whitney U test). (E) Serum cytokine level in transplanted mice (*p<0.05 Mann-Whitney U test). Boxplots indicate the median and IQR; whiskers show IQR×1.5. BMI, body mass index; FMT, faecal microbiota transplant; GDM, gestational diabetes mellitus; IL, interleukin; PCoA, principal coordinate analysis; PERMANOVA, permutational multivariate analysis of variance.
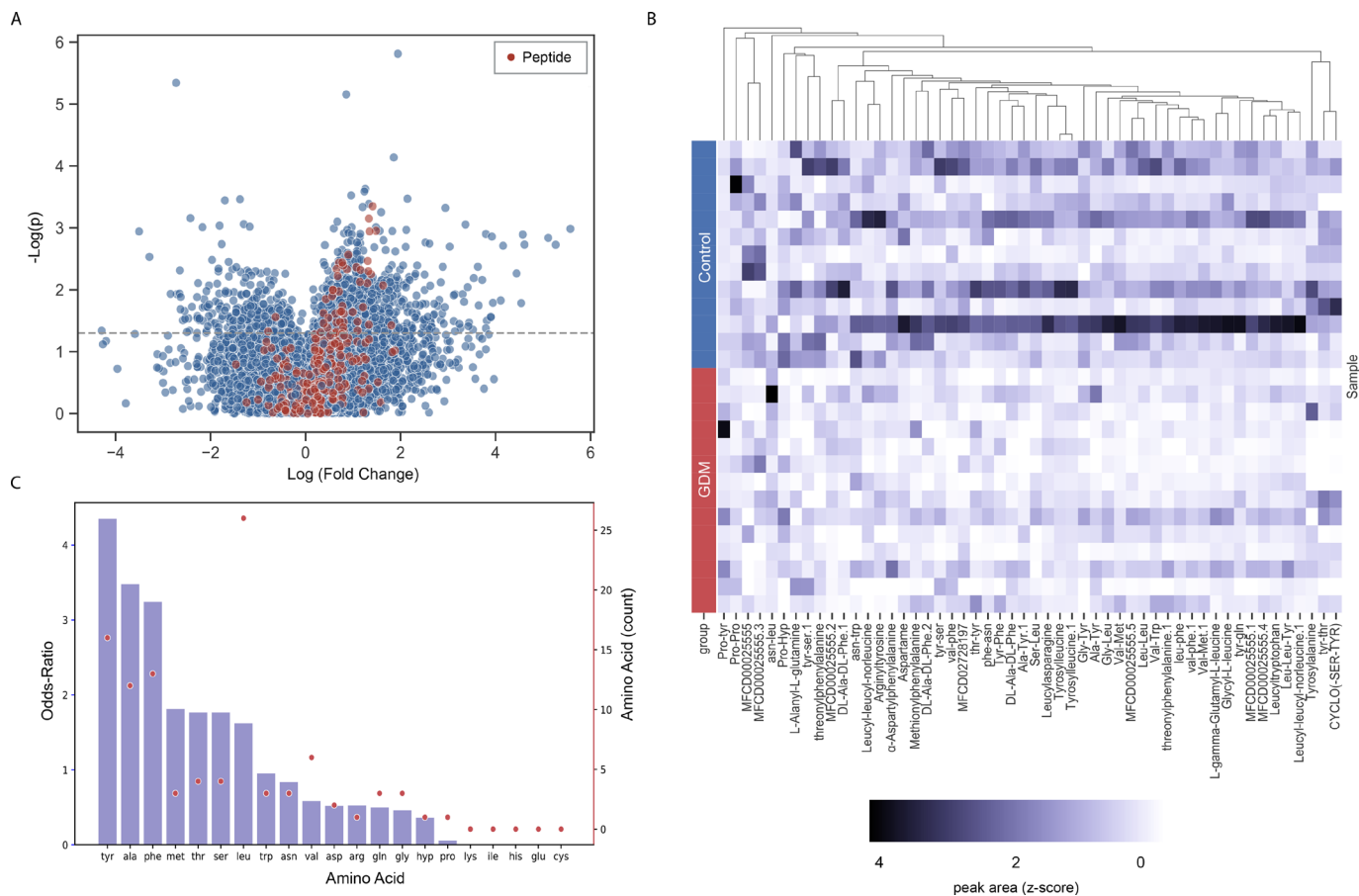
Figure 4 Analysis of first trimester human faecal metabolomics exhibits lower levels of dipeptides for women with GDM. (A) Volcano plot of all metabolites examined in this study, comparing age/BMI-matched metabolite profiles of women who would and would not later develop GDM (n=20 pairs); peptides are coloured in red. (B) Heatmap of the 52 differentially expressed peptides. Each row denotes a sample (grouped by disease state) and each column denotes a peptide. Z-scores were calculated per column. Peptides (columns) were hierarchically clustered based on Euclidean distances. (C) Amino acid composition of the differentially abundant peptides. Bars (left y-axis) represent odds ratios (OR) for each amino acid, and dots (right y-axis) represent the amino acid count in the differentially abundant peptides. BMI, body mass index; GDM, gestational diabetes mellitus.

between control women and women with GDM (online supplemental table 6). Following curation of all dipeptides and tripeptides, the vast majority of the peptides (50 out of 52) with significant differential concentrations showed a clear tendency of depletion in women with GDM relative to healthy control women (figure 4A,B). These peptides were enriched with the hydrophobic amino acids tyrosine, phenylalanine and alanine (p=8×10^{-4}, 0.01, 0.01, respectively, FDR-corrected Fisher's exact tests; figure 4C). As metabolome profiling in women uncovered important associations to GDM, we decided to use PICRUSt2 to predict metabolic pathways enriched in mice from the faecal microbiota profiles of the mice in our FMT studies. We found 16 differentially abundant metabolic pathways between GDM and control-recipient mice (online supplemental table 7). We observed an enrichment of the mevalonate pathway (PWY-922; online supplemental figure 4), corresponding with evidence of increased IL-6 levels in the GDM group of both our primary cohort and our transplanted mice[23 24] and of the heme pathway (online supplemental figure 4), previously implicated in type 2 diabetes.[25]

Gut microbiome composition improves prediction of GDM early in pregnancy

Finally, we built a machine learning model to predict GDM based on microbiome composition, cytokine profile, medical history and dietary features, all collected during T1 (figure 1). For this aim, we also included T1 clinical data of 66 additional women, recruited retrospectively in later stages of their pregnancy (secondary cohort, see methods). Our Xgboost model predicted GDM with very high accuracy (area under the receiver operating characteristic curve (auROC)=0.83; figure 5). When making predictions based on only a single feature, we found the highest accuracy when using medical records alone (though still 7% lower than our combined model), in agreement with a recent study.[26] Faecal microbiome features resulted in the second highest accuracy (auROC=0.73). Using our two-step method (see online supplemental methods), we improved the odds ratio (OR) from 3.2 to 4, demonstrating the potential for more accurate prediction using the faecal microbiome profile, especially relevant if medical records are incomplete or unavailable. To validate the predictive power of our microbiome model, we used a validation cohort of 98 women who developed GDM with 98 matched healthy controls from a T1 pregnancy study in China.[8] We first built a model based on the validation cohort to test the predictive power for this cohort based only on microbiome data (auROC=0.65). Assuringly, when applying our model (the Israeli cohort learning set) on this cohort, we found a comparable accuracy (auROC=0.6), confirming that despite the striking genetic and lifestyle differences between the cohorts, our findings are, at least partially, generalisable. We also built a model based on 86
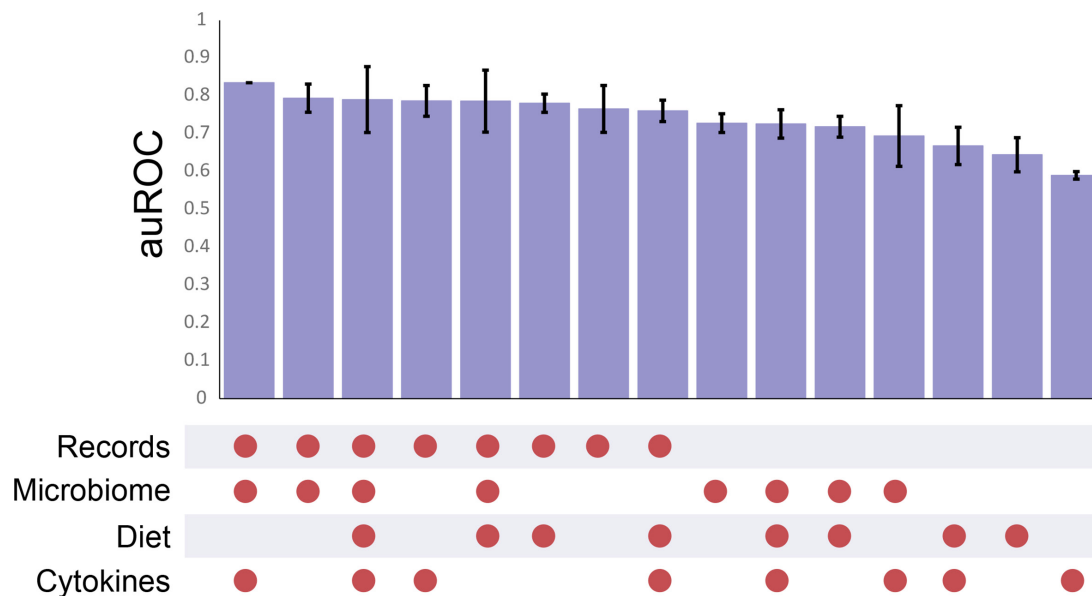
Figure 5 Highly accurate prediction of future disease onset among pregnant women during their first trimester. Area under the receiver operating characteristic curve (auROC) for each combination of features. Error bars represent ±SD.

nutritional characteristics measured in our cohort, which yielded lower predictive accuracy (auROC=0.64; figure 5) than the other features in this study. Further, no differences were found in dietary habits between women in our primary cohort who would and would not later develop GDM (online supplemental tables 8 and 9) suggesting that differences in food consumption during T1 contribute minimally to GDM pathogenesis.

DISCUSSION
GDM biomarkers
Here, using a combination of 'omics' tools, we identify biomarkers of GDM onset as early as the first trimester of pregnancy. Women in T1, who later develop GDM, exhibit gut microbiota dysbiosis as well as increased proinflammatory serum cytokines and lower levels of faecal SCFAs. Further, the specific microbial changes in their microbiota are directly associated with GDM phenotype features (insulin resistance and low-grade inflammation) as revealed by FMT into GF animals. Lastly, we demonstrated that microbiota samples from T1 alone can be used to predict GDM onset and that parameters from patient medical records can improve these predictions, providing a robust tool for early prediction of GDM.

In our primary cohort, women with GDM exhibit elevated levels of serum inflammatory cytokines during T1 of pregnancy. Insulin resistance has been associated with elevated secretion of proinflammatory cytokines,[27] and indeed several studies demonstrated elevated levels of proinflammatory cytokines during T2 and T3.[28–30] These altered cytokine profiles in women with several months prior to a GDM diagnosis suggest that inflammation may be associated with the pathogenesis of GDM and can be used to identify its early onset. This is in line with typical GDM and type 2 diabetes symptomatology. Low-grade chronic inflammation is associated with obesity in general and maternal obesity in particular. But here, we controlled for BMI, suggesting that increased pre-GDM-associated inflammation is beyond that associated with obesity or general pregnancy. This is in line with evidence in the literature of inflammation in cases of type 2 diabetes[31 32] independent of weight. Further, in pregnancy, low-grade inflammation levels can differ among women independent

of BMI, suggesting that some other characteristics like immune–endocrine interactions may also be at play. In our study, we specifically observed higher levels of IL-6 in both women with pre-GDM and in GDM-recipient mice. This suggests that the elevated levels of IL-6 are driven by gut microbes. IL-6 was previously described to play a role in the development of both type 1 and type 2 diabetes[33] and was proposed as a potential biomarker of gestational diabetes in 16 different studies, mostly in later stages of pregnancy.[34] Our findings in T1, both in the focal cohort and in FMT experiments, support inflammation as an early marker of GDM.

Another potential early biomarker for GDM is a decrease in SCFAs, which contribute to the maintenance of glucose homeostasis and suppression of inflammatory response. Hence, SCFAs are thought to play a role in obesity-induced inflammation leading to attenuation of insulin signalling and GDM.[35] We found two BSCFAs reduced in stools of women who later developed GDM. BSCFAs are a product of bacterial fermentation of branched amino acids generated from undigested protein reaching the colon. BSCFAs, proposed markers for protein fermentation, were found to improve insulin sensitivity[36 37] and reduce inflammation.[38] These findings, in line with several studies of later-stage pregnancy[39] (but see findings from Pappa *et al*[40]), suggest faecal BSCFAs could serve as a potential biomarker for GDM in early stages of pregnancy.

The gut microbiome is associated with GDM months before diagnosis
Several studies have found altered gut microbiome composition in women with GDM; most were based on samples collected post-diagnosis.[41 42] Our findings suggest that microbial differences between GDM and control groups, when controlling for confounding variables, exist in T1 and are driven by specific taxa rather than community-wide shifts, leading to subtle differences in composition.

As an illustrative example, *P. copri,* which is known to play a role in glucose homeostasis[43] and has been reported to be more abundant in women diagnosed with GDM,[41 44] was found to be under-represented in women with GDM in our primary cohort,

after controlling for confounders, and also in GDM-recipient mice. A recent study demonstrated that *Prevotella* was a marker of positive glucose metabolism.[45] Kovatcheva-Datchary *et al*. even showed, in a clinical trial, that *Prevotella* protected against *Bacteroides*-induced glucose intolerance and that improvement in glucose metabolism was associated with increased abundance of *Prevotella*.[46] This improved glucose metabolism by presence of *Prevotella* was also demonstrated by supplementing mice with *P. copri*. One possible mechanism, recently proposed in rats, is that *P. copri* improves glucose homeostasis through farnesoid X receptor signalling and increased bile acid metabolism.[47] We chose to discuss *P. copri* specifically as it was found to have lower abundance in both women with GDM and recipient mice and was previously described to play a role in glucose homeostasis. In this study, we also demonstrated the importance of controlling for risk factors. For example, *Akkermansia muciniphila*, which is consistently negatively correlated with obesity,[48] is prima facie negatively associated with GDM when not controlling for the difference in BMI between the groups.

FMT highlights IL-6 as a potential contributor to GDM

Based on our multicohort FMT experiments using T1, pre-GDM-diagnosis samples, we conclude that gut microbes play a causal role in the development of some of the phenotypes of GDM and that their role is likely universal as demonstrated by conservation across cohorts. Increased levels in IL-6, in both women that would develop GDM and transplanted mice that received their microbiota, suggest an important microbiota-related inflammatory mechanism in GDM progression, further supported by functional profile prediction in mice. Two relevant bacterial pathways, the mevalonate pathway and the heme biosynthesis pathway were elevated in GDM-recipient mice. There is evidence that mevalonate can have negative implications on host inflammation—its presence reduces effects of statins in decreasing IL-6 and IL-8,[23] and a kinase deficiency, which increases free mevalonate, leads to autoinflammation.[49] The heme biosynthesis pathway was previously associated with elevated IL-6[50 51] and with type 2 diabetes.[52] We do note, however, that further research is needed to understand if the bacteria themselves, their excreted metabolites or some other factors control this phenotype. Only then can we uncover specific mechanisms of pathogenesis.

Short peptides association with GDM

We found lower levels of short peptides in T1 stools of women with pre-GDM. These peptides are enriched with the amino acids phenylalanine, alanine and tyrosine. Previously, plasma levels of these three hydrophobic amino acids have also been reported to be significantly associated with diabetes.[53] One study found a link between their elevated blood levels and decreased insulin secretion.[54] Interestingly, Jiang *et al*. recently found elevated levels of alanine and tyrosine in maternal blood at 12–16 gestational weeks in women later diagnosed with GDM[55]; alanine is also used by the liver for gluconeogenesis.[56] Increased amino acid levels in the blood may result in lower levels excreted in stool,[57] though this requires further study.

Prediction of GDM

We were able to accurately predict future GDM onset in T1, weeks before the complication is typically diagnosed. Our combined model predicts GDM with very high accuracy, and even a microbiota-centric model could predict disease onset in two geographically diverse cohorts. This tool allows for accurate early prediction, care plans and potential prevention of this disease, improving both maternal and fetal outcomes. This is further supported by phenotype transfer in samples originating from cohorts in three different continents. On the whole, prediction could (and likely should) be improved using local microbiota characteristics, but genus-level differences in the microbiome can be used as general predictors in the absence of local data.

CONCLUSIONS

In summary, we found broad and consistent evidence that GDM pathology begins as early as T1 in a large prospective cohort of pregnant women. Additionally, we successfully demonstrated that the precursors of GDM originate in the gut microbiota and that early-onset GDM has a bacterial signature at least partially responsible for the GDM phenotype, evident from phenotype transfer following FMT. Our findings suggest that GDM is induced through heightened inflammation, initiated by microbial dysbiosis. Future research based on our findings can help unravel the underlying mechanisms.

This study has several limitations. Bacterial dysbiosis could be a first response to disease onset rather than a cause. Additionally, the phenotype transfer we observed may be caused by other faecal material including metabolites, eukaryotic microorganisms, human viruses and bacteriophages, though in this case as well, the bacterial biomarkers identified can be relevant for diagnostics. Lastly, throughout this study, we treated the major risk factors of GDM, BMI and age, using either matching or relevant statistical methods. We cannot exclude the effect of other clinical or demographic features on our results and also wish to highlight the potentially important contribution of these two 'confounding' risk factors. Despite limitations, addition of microbiome data to a machine learning model improved our ability to predict GDM and can even serve as a standalone snapshot predictor. These results may be of use in the future when exploring preventive measures for GDM.

Author affiliations
[1]Azrieli Faculty of Medicine, Bar-Ilan University, Safed, Israel
[2]Helen Schneider Hospital for Women, Rabin Medical Center and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[3]Institute of Biochemistry, School of Nutritional Sciences Food Science and Nutrition, The School of Nutritional Sciences, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel
[4]Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel
[5]Department of Microbiome Science, Max Planck Institute for Developmental Biology, Tubingen, Germany
[6]Department of Medicine, Stanford University, Stanford, California, USA
[7]Department of Epidemiology and Population Health, Stanford University, Stanford, California, USA
[8]Division of Nutritional Sciences, Cornell University, Ithaca, New York, USA
[9]Department of Natural Compounds and Analytical Chemistry, Migal-Galilee Research Institute, Kiryat Shmona, Israel
[10]Analytical Chemistry Laboratory, Tel-Hai College, Upper Galilee, Israel
[11]Laboratory of Human Health and Nutrition Sciences, Migal-Galilee Technology Center, Kiryat Shmona, Israel
[12]Nutritional Science Department, Tel Hai College, Upper Galilee, Israel
[13]The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel
[14]Department of Pediatrics, University of Turku and Turku University Hospital, Turku, Finland
[15]University of Helsinki & Helsinki University Hospital, New Children's Hospital, Pediatric Research Center, Helsinki, Finland
[16]Functional Foods Forum, University of Turku, Turku, Finland
[17]Clalit Health Services, Tel Aviv, Israel
[18]Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[19]Santa Fe Institute, Santa Fe, New Mexico, USA

Twitter Sondra Turjeman @SondraTurjeman, Soliman Khatib @Soliman72227832 and Omry Koren @OmryKoren

ORCID iDs
Yishay Pinto http://orcid.org/0000-0002-5280-1040
Sondra Turjeman http://orcid.org/0000-0002-5224-1215
William Walters http://orcid.org/0000-0003-1716-3984
Elizabeth L Johnson http://orcid.org/0000-0002-8290-1122
Soliman Khatib http://orcid.org/0000-0002-6762-2408
Faiga Magzal http://orcid.org/0000-0003-0975-1858
Snait Tamir http://orcid.org/0000-0002-7473-7369
Samuli Rautava http://orcid.org/0000-0003-0935-3593
Erika Isolauri http://orcid.org/0000-0002-1635-8002
Elhanan Borenstein http://orcid.org/0000-0003-3002-0945
Eran Hadar http://orcid.org/0000-0002-3101-2827
Omry Koren http://orcid.org/0000-0002-7738-1337

REFERENCES

1   Baz B, Riveline J-P, Gautier J-F. Endocrinology of pregnancy: gestational diabetes mellitus: definition, aetiological and clinical aspects. *Eur J Endocrinol* 2016;174:R43–51.
2   Plows JF, Stanley JL, Baker PN, *et al*. The pathophysiology of gestational diabetes mellitus. *Int J Mol Sci* 2018;19. doi:10.3390/ijms19113342. [Epub ahead of print: 26 Oct 2018].
3   Lende M, Rijhsinghani A. Gestational diabetes: overview with emphasis on medical management. *Int J Environ Res Public Health* 2020;17. doi:10.3390/ijerph17249573. [Epub ahead of print: 21 Dec 2020].
4   Zhu Y, Zhang C. Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. *Curr Diab Rep* 2016;16:7.
5   Eades CE, Cameron DM, Evans JMM. Prevalence of gestational diabetes mellitus in Europe: a meta-analysis. *Diabetes Res Clin Pract* 2017;129:173–81.
6   Casagrande SS, Linder B, Cowie CC. Prevalence of gestational diabetes and subsequent type 2 diabetes among U.S. women. *Diabetes Res Clin Pract* 2018;141:200–8.
7   Tenenbaum-Gavish K, Sharabi-Nov A, Binyamin D, *et al*. First trimester biomarkers for prediction of gestational diabetes mellitus. *Placenta* 2020;101:80–9.
8   Ma S, You Y, Huang L, *et al*. Alterations in gut microbiota of gestational diabetes patients during the first trimester of pregnancy. *Front Cell Infect Microbiol* 2020;10:58.
9   Qin J, Li Y, Cai Z, *et al*. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 2012;490:55–60.
10 Crusell MKW, Hansen TH, Nielsen T, *et al*. Gestational diabetes is associated with change in the gut microbiota composition in third trimester of pregnancy and postpartum. *Microbiome* 2018;6:89.
11 Liu Y, Qin S, Feng Y, *et al*. Perturbations of gut microbiota in gestational diabetes mellitus patients induce hyperglycemia in germ-free mice. *J Dev Orig Health Dis* 2020;11:580–8.
12 Mokkala K, Paulin N, Houttu N, *et al*. Metagenomics analysis of gut microbiota in response to diet intervention and gestational diabetes in overweight and obese women: a randomised, double-blind, placebo-controlled clinical trial. *Gut* 2021;70:309–18.
13 Mokkala K, Houttu N, Vahlberg T, *et al*. Gut microbiota aberrations precede diagnosis of gestational diabetes mellitus. *Acta Diabetol* 2017;54:1147–9.
14 Gomez-Arango LF, Barrett HL, McIntyre HD, *et al*. Connections between the gut microbiome and metabolic hormones in early pregnancy in overweight and obese women. *Diabetes* 2016;65:2214–23.
15 Zheng W, Xu Q, Huang W, *et al*. Gestational diabetes mellitus is associated with reduced dynamics of gut microbiota during the first half of pregnancy. *mSystems* 2020;5. doi:10.1128/mSystems.00109-20. [Epub ahead of print: 24 03 2020].
16 Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:385–96.
17 Kalter-Leibovici O, Freedman LS, Olmer L, *et al*. Screening and diagnosis of gestational diabetes mellitus: critical appraisal of the new international association of diabetes in pregnancy Study Group recommendations on a national level. *Diabetes Care* 2012;35:1894–6.
18 Bolyen E, Rideout JR, Dillon MR, *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–7.
19 Mallick H, Rahnavard A, McIver LJ, *et al*. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol* 2021;17:e1009442.
20 Douglas GM, Maffei VJ, Zaneveld JR, *et al*. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020;38:685–8.
21 Pinto Y. Months before diagnosis, gestational diabetes is characterized by microbiota-induced inflammation. 2022.
22 Chen X, Stein TP, Steer RA, *et al*. Individual free fatty acids have unique associations with inflammatory biomarkers, insulin resistance and insulin secretion in healthy and gestational diabetic pregnant women. *BMJ Open Diabetes Res Care* 2019;7:e000632.
23 Sakoda K, Yamamoto M, Negishi Y, *et al*. Simvastatin decreases IL-6 and IL-8 production in epithelial cells. *J Dent Res* 2006;85:520–3.
24 Caesar R, Nygren H, Orešič M, *et al*. Interaction between dietary lipids and gut microbiota regulates hepatic cholesterol metabolism. *J Lipid Res* 2016;57:474–81.
25 Kwan S-Y, Sabotta CM, Joon A, *et al*. Gut microbiome alterations associated with diabetes in Mexican Americans in South Texas. *mSystems* 2022;7:e0003322.
26 Artzi NS, Shilo S, Hadar E, *et al*. Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med* 2020;26:71–6.
27 Esser N, Legrand-Poels S, Piette J, *et al*. Inflammation as a link between obesity, metabolic syndrome and type 2 diabetes. *Diabetes Res Clin Pract* 2014;105:141–50.
28 Ross KM, Miller G, Culhane J, *et al*. Patterns of peripheral cytokine expression during pregnancy in two cohorts and associations with inflammatory markers in cord blood. *Am J Reprod Immunol* 2016;76:406–14.
29 Žák P, Souček M. Correlation of tumor necrosis factor alpha, interleukin 6 and interleukin 10 with blood pressure, risk of preeclampsia and low birth weight in gestational diabetes. *Physiol Res* 2019;68:395–408.
30 Sudharshana Murthy KA, Bhandiwada A, Chandan SL, *et al*. Evaluation of oxidative stress and proinflammatory cytokines in gestational diabetes mellitus and their correlation with pregnancy outcome. *Indian J Endocrinol Metab* 2018;22:79–84.
31 Pickup JC, Crook MA. Is type II diabetes mellitus a disease of the innate immune system? *Diabetologia* 1998;41:1241–8.
32 Pradhan AD, Manson JE, Rifai N, *et al*. C-Reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *JAMA* 2001;286:327.
33 Kristiansen OP, Mandrup-Poulsen T. Interleukin-6 and diabetes: the good, the bad, or the indifferent? *Diabetes* 2005;54 Suppl 2:S114–24.
34 Amirian A, Mahani MB, Abdi F. Role of interleukin-6 (IL-6) in predicting gestational diabetes mellitus. *Obstet Gynecol Sci* 2020;63:407–16.
35 Chambers ES, Preston T, Frost G, *et al*. Role of gut Microbiota-Generated short-chain fatty acids in metabolic and cardiovascular health. *Curr Nutr Rep* 2018;7:198–206.
36 Heimann E, Nyman M, Pålbrink A-K, *et al*. Branched short-chain fatty acids modulate glucose and lipid metabolism in primary adipocytes. *Adipocyte* 2016;5:359–68.
37 Haufe S, Engeli S, Kaminski J, *et al*. Branched-Chain amino acid catabolism rather than amino acids plasma concentrations is associated with diet-induced changes in insulin resistance in overweight to obese individuals. *Nutr Metab Cardiovasc Dis* 2017;27:858–64.

38 Yao CK, Muir JG, Gibson PR. Review article: insights into colonic protein fermentation, its modulation and potential health implications. *Aliment Pharmacol Ther* 2016;43:181–96.

39 Chorell E, Hall UA, Gustavsson C, *et al*. Pregnancy to postpartum transition of serum metabolites in women with gestational diabetes. *Metabolism* 2017;72:27–36.

40 Pappa KI, Vlachos G, Theodora M, *et al*. Intermediate metabolism in association with the amino acid profile during the third trimester of normal pregnancy and diet-controlled gestational diabetes. *Am J Obstet Gynecol* 2007;196:65.e1–65.e5.

41 Ponzo V, Fedele D, Goitre I, *et al*. Diet-Gut microbiota interactions and gestational diabetes mellitus (GDM). *Nutrients* 2019;11. doi:10.3390/nu11020330. [Epub ahead of print: 03 Feb 2019].

42 Ferrocino I, Ponzo V, Gambino R, *et al*. Changes in the gut microbiota composition during pregnancy in patients with gestational diabetes mellitus (GDM). *Sci Rep* 2018;8:12216.

43 Cani PD. Human gut microbiome: hopes, threats and promises. *Gut* 2018;67:1716–25.

44 Hasain Z, Mokhtar NM, Kamaruddin NA, *et al*. Gut microbiota and gestational diabetes mellitus: a review of Host-Gut microbiota interactions and their therapeutic potential. *Front Cell Infect Microbiol* 2020;10:188.

45 Asnicar F, Berry SE, Valdes AM, *et al*. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat Med* 2021;27:321–32.

46 Kovatcheva-Datchary P, Nilsson A, Akrami R, *et al*. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. *Cell Metab* 2015;22:971–82.

47 Péan N, Le Lay A, Brial F, *et al*. Dominant gut Prevotella copri in gastrectomised non-obese diabetic Goto-Kakizaki rats improves glucose homeostasis through enhanced FXR signalling. *Diabetologia* 2020;63:1223–35.

48 Xu Y, Wang N, Tan H-Y, *et al*. Function of *Akkermansia muciniphila* in Obesity: Interactions With Lipid Metabolism, Immune Response and Gut Systems. *Front Microbiol* 2020;11:219.

49 Politiek FA, Waterham HR. Compromised protein prenylation as pathogenic mechanism in mevalonate kinase deficiency. *Front Immunol* 2021;12:724991.

50 Gbotosho OT, Kapetanaki MG, Ghosh S, *et al*. Heme induces IL-6 and cardiac hypertrophy genes transcripts in sickle cell mice. *Front Immunol* 2020;11:1910.

51 Kang DY, Sp N, Jo ES, *et al*. New insights into the pivotal role of iron/heme metabolism in TLR4/NF-κB signaling-mediated inflammatory responses in human monocytes. *Cells* 2021;10. doi:10.3390/cells10102549. [Epub ahead of print: 27 Sep 2021].

52 Fernandez-Cao JC, Arija V, Aranda N, *et al*. Heme iron intake and risk of new-onset diabetes in a Mediterranean population at high risk of cardiovascular disease: an observational cohort analysis. *BMC Public Health* 2013;13:1042.

53 Wang TJ, Larson MG, Vasan RS, *et al*. Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;17:448–53.

54 Vangipurapu J, Stancáková A, Smith U, *et al*. Nine amino acids are associated with decreased insulin secretion and elevated glucose levels in a 7.4-Year follow-up study of 5,181 Finnish men. *Diabetes* 2019;68:1353–8.

55 Jiang R, Wu S, Fang C, *et al*. Amino acids levels in early pregnancy predict subsequent gestational diabetes. *J Diabetes* 2020;12:503–11.

56 Bloomgarden Z. Diabetes and branched-chain amino acids: what is the link? *J Diabetes* 2018;10:350–2.

57 West KA, Kanu C, Maric T, *et al*. Longitudinal metabolic and gut bacterial profiling of pregnant women with previous bariatric surgery. *Gut* 2020;69:1452–9.