

Projection of gut microbiome pre and post bariatric surgery to predict surgery outcome

yoram louzoun (✉ louzouy@math.biu.ac.il)

Bar-Ilan University Department of Mathematics <https://orcid.org/0000-0003-1714-6148>

Meirav Ben Izhak

Bar-Ilan University Department of Mathematics

Ruti Cohen

Hi labs

Liora Madar Shapiro

Hi labs

Hamutal Meiri

Hi labs

Chaim Wachtel

Hi labs

Adi Eshel

Bar-Ilan University

Conrad Leung

GeneWize

Edward Messick

GeneWISE

Narisra Jongkam

GeneWISE

Eli Mavor

kaplan hospital

Shimon Sapozhnikov

Rabin Medical Center

Nitsan Maharshak

Tel Aviv Sourasky Medical Center

Subhi Abu-Abeid

Tel Aviv Sourasky Medical Center

Avishai Alis

Rabin Medical Center

Ilanit Mahler

asuta medical center

Aviel Meoded

Tel Aviv Sourasky Medical Center

Shai Eldar

Tel Aviv Sourasky Medical Center

Omry Koren

Research

Keywords: Projection of gut microbiome, pre and post bariatric surgery, to predict surgery outcome

Posted Date: December 13th, 2019

DOI: <https://doi.org/10.21203/rs.2.18879/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Bariatric surgery is often the preferred method to resolve obesity and diabetes, with ~800,000 cases worldwide yearly and high outcome variability. The ability to predict the long-term BMI change following surgery has important implications on individuals and the health care system in general. Given the tight connection between eating habits, sugar consumption, BMI and the gut microbiome, we tested whether the microbiome before any treatment is associated with different treatment outcomes, as well as other intakes (HDL, Triglycerides, etc.).

Results: A projection of the gut microbiome composition of obese (sampled before and after bariatric surgery) and slim patients into principal components was performed and the relation between this projection and surgery outcome was studied. The projection reveals 3 different microbiome profiles belonging to slim, obese and obese who underwent bariatric surgery, with post-surgery more different from the slim than the obese. The same projection allows for a prediction of BMI loss following bariatric surgery, using only the pre-surgery microbiome. A different projection is associated with sugar metabolism and A1C levels.

Conclusions: - the gut microbiome can be naturally decomposed into main components depicting the patient's development and predicting in advance the outcome. Those may be translated into a better clinical management of obese individuals planning to undergo metabolic surgery.

Summary

- Microbiome is known to be associated with BMI and Diabetes.
- Bariatric surgery has large outcome variabilities.
- Microbiome were previously shown to be a good predictor for multiple diseases.
- We here analyze microbiome before and after bariatric surgery and show that
- The microbiome before surgery can be used to predict surgery outcome.
- Post-surgery microbiome drifts farther away from slim microbiome than pre-surgery obese patients.
- The microbiome of the studied slim and obese hosts naturally decomposes to primary components associated with BMI, sugar consumption and post-surgery outcome.
- These results can lead to a microbiome pre-surgery decision whether to perform surgery.

Background

To test the possibility of predicting bariatric surgery outcome, we analyzed 308 fecal sampla

The human body is colonized by a wide variety of micro-organisms, commonly referred to as the human microbiota. The gut microbiota is a complex ecosystem, which provides major functions to the host, such as regulation of metabolism, immune system modulation, and protection against pathogens [1, 2]. The microbiome is strongly associated with weight and sugar consumption, and as such it serves as a proxy for nutrition and life habits and may also influence them. Such life habits may influence the total body mass and BMI in regular conditions, as well as after bariatric surgery.

Obesity and diabetes are world pandemics [3]. Approximately 8-10% of the population develop complications of morbid obesity, (BMI>35), frequently coupled to some form of diabetes. According to the WHO, of the 57 million deaths in 2008 worldwide, 1.3 million were due to metabolic disorders, particularly those associated with obesity [3]. Recently, the gut microbiome of obese individuals has been shown to be very different from the microbiome of slim subjects [4]. Nagpal et al. [5] suggested that some bacteria increase gut permeability and insulin resistance leading to obesity and diabetes. Experimental fecal transplants to mice demonstrated that transplantation of microbiome from obese individuals into slim mice turned them obese [6], showing the importance of the gut microbiome in regulating body weight. Opposite studies of turning obese mice into slim mice have not been successful but one study demonstrated that certain bacteria can prevent weight gain in mice [7].

The introduction of bariatric surgery as a method for losing weight is rapidly adopted as the most efficient method for weight loss and for reducing blood sugar levels [8], however it has drawbacks, including a range of possible complications from nutrition deficiencies to occurrence of life-threatening conditions and a big diversity in the success rates of achieving weight loss and maintaining it [9]. A few studies have shown microbiome changes after bariatric surgery [10]. However, the real potential of the microbiome as a tool for not only monitoring the procedure's outcomes but rather predicting them beforehand has not been explored. An attempt to study and test this potential can result in an essential tool which will assist in the decision whether to consult patient to perform such surgery.

es from patients of 2 main groups: obese who underwent bariatric surgery and naturally slim. For the obese patients (BMI > 35) we sampled the microbiome at five time points (Fig. 1A) – one at enrollment (A, 78 samples), three weeks after a low carbohydrate diet and immediately before the operation (B, 70 samples), and 3 time points following the surgery (two weeks– C, 34 samples; three months – D, 27 samples; and six months E, 16 samples). Not all individuals had been sampled at all time points. This was compared to 83 slim control individuals (BMI 19–25) (For all details, see Supp. Methods). We have collected BMI and sugar A1C information for the same patients in late time points up to a year and a half post-surgery to track their weight loss and the remission of diabetes.

The slim population was younger and had more males compared to the population who underwent surgery (36 +/- 12 vs 48 +/- 12 and 50% males vs 29%). Overall, the patients' mean BMI was reduced from 43.3 +/- 6.8 (Mean +/- SD) to 27.8 +/- 1.5, which represents an average loss of 84.7% overweight (compared to BMI 25), blood sugar levels were reduced from Hemoglobin A1C of 6.5 +/- 0.4 to 5.8 +/- 0.75 or from blood sugar levels of 125 +/- 11 g/dL to 95.4 +/- 10, Triglyceride levels decreased from 183 +/- 20 to 102 +/- 13. All parameters described are significantly ($P < 0.001$) lower from their starting point and not different from the slim control (Fig. 1B-D).

The gut microbiome of all donors was analyzed using 16S rRNA gene sequences (emphasizing the 16SMetaVx.V2) [14], and OTU tables were produced using QIIME2 [15]. The OTU tables were then merged to the genus level. Taxa appearing in only one sample were removed. All samples were log-normalized to highlight the differences in rare bacteria and z-scored. The normalized tables were projected on PCA components. The aim of the first step was to homogenize the description level and reduce the dimension. Since multiple OTUs are associated with the same bacteria, and some OTUs are associated with different levels of classification, we averaged all OTUs associated with the same species in each donor (Fig. 2A-C). Note that while information is lost in the process, such a process is essential for the following machine learning. If an OTU was only present in part of the samples, it was given a value of 0 in all other samples. The resulting values have a scale-free distribution, which often masks large changes in relative frequencies of rare bacteria. To handle that, we log-transformed all OTU values and added a minimal constant value (0.1) to avoid log of zero values. This allows for a narrower distribution of values (Fig. 2D). Finally, given the very high correlation between the relative abundance of different bacteria (Fig. 2E), we projected the bacteria to principal components, which capture most of the variance in the bacterial diversity (Fig. 2F).

The projection on the first principal vector (PC1) clearly delineated axes separating the obese from the slim individuals (Fig. 3A,B). The clear separation of the PC1 projections agrees with observed major differences in the microbiome of slim and obese individuals. The large BMI difference between groups (BMI > 35 in obese versus BMI of < 25 in slim) translates to a large difference in the microbiome. We next tested whether diet or bariatric surgery push back the population toward the slim profile. The results are surprisingly opposite (Fig. 3B). The distance between the projection on the first PC of the post-diet and post-surgery and the slim profile keeps increasing and reaches a maximum after a year. The major difference between these projections allows for a simple classification even with a linear SVM of slim vs obese and pre vs post-surgery samples (Fig. 3C,D ROC curves). The main contributions to the classifiers are from the first two PCs for the healthy (H) vs obese (O) and the same with the 5th PC for the pre vs post-surgery (Fig. 3E). Note that higher test AUC can be obtained by non-linear classifiers. However, the linear classifier gives a clear picture of the contribution of each PC to the microbiome development.

One can then project back the correlations between the PC and the state/BMI to the original OTUs, and find OTUs that are correlated to BMI (PC 1 and 3), the OTUs that are over and underrepresented in obese individuals compared to healthy and the OTUs that change significantly after surgery compared to before surgery (Fig. 4). Interestingly, *Helicobacter* is strongly associated with high BMI (Fig. 4A,B,C) as it has been reported previously in several studies to be associated with inflammation, insulin resistance and BMI [16–19]. The opposite behavior occurs with the bacterial family *Succinivibrionaceae* which are highly associated with healthy individuals and the after-surgery state. Members of this family ferment carbohydrates to acetate and succinate (succinate being an important intermediate for propionate production) [20]. Short-chain fatty acids (SCFAs) such as acetate and propionate have multiple beneficial roles. They are known for their contribution to improved insulin sensitivity and glucose homeostasis [21] and for their protective effect against diet-induced obesity [22]. The inflammatory effects of *Helicobacter* and the SCFA producing capabilities of the *Succinivibrionaceae* help explain our observation the first is higher in the obese state and the latter are associated with a healthy lean state.

To test for possible confounding effects, we tested whether the observed changes in the profile may be the result of age or gender, or whether they are related to the total BMI. There are no significant correlations with gender and a very weak correlation of PC3 and 5 with age (Fig. 5E, F).

Since the first three PC represent the main effect of the current BMI (i.e., BMI at time of sampling), we tested whether the other PCs represent other aspects of the population. Specifically, we tested if other PCs can be used to predict a future change in BMI. We performed an L1 (Lasso) regression of the projection on the first PCs of the A point and the change in BMI between point A and points E (six months after surgery), 1Y and 1.5Y. The prediction was tested using a Leave One Out (LOO) methodology, and the Spearman correlations on the test values between the predicted change and the observed change (on the LOO test), as well as the Area Under Curve of a predictor of whether a patient will have a more/less than average reduction in BMI. The AUCs are significant for all points (Fig. 5B-D). The Spearman correlation is highly significant for point E (where we have more information than other points). The prediction of future BMI change is also determined by a single main PC (the 4th PC Fig. 5A and G), further showing the natural decomposition of the microbiome into elements correlated with the host behavior. When looking at which OTUs are contributing to PC4 (Fig. 5A) we found *Holdemania* to be strongly correlated with future weight loss. *Holdemania* was previously found to be overrepresented in lean individuals [23]. *Coprococcus* which also correlates with greater BMI loss in the future is a known butyrate producer [24]. The beneficial effect of butyrate to weight loss may be due to its role as a substrate in gluconeogenesis [25], and recently, oral butyrate supplementation has even been shown to reduce adiposity and improve insulin sensitivity [26].

Another possible candidate to affect the microbiome is the sugar level. We tested the correlation between the projections on the PCA and the A1C. Indeed, a clear correlation is found between A1C and PC7 (Fig. 6B for correlations

and 6A for the projections). Using the projection, diabetic and non-diabetic patients, as defined by A1C (above 6.5 and below 5.7) can be separated with a high AUC (0.75 average over test $p < 0.01$) (Fig. 6B). This correlation might be used for predicting the disease in healthy subjects both from risk groups and in general. The main OTU negatively affecting PC7 are members of Erysipelotrichaceae which are known to be associated with obesity [27] (Fig. 6A).

Methods

Patients and regulation

Patients were enrolled in the obesity control/bariatric surgery clinics of four medical centers in Israel – Kaplan Medical Center (KMC), Rabin Medical Center (RMC), Tel Aviv Medical Center (TMC) and Poria Medical Center (PMC) during December 2015- November 2018. The ethics committees of each of the respective medical centers approved the study and its amendments and each patient and control signed a written informed consent. Inclusion criteria were: ages 18-70, no antibiotic treatment in the two months prior to the enrolment, no previous bariatric or major gut/stomach operation and filling out a dietary questionnaire. Naturally slim control patients had no diabetes (Hemoglobin A1C<5.0), BMI of 19-25, and had no major medical and endocrine complications. Obese individuals had BMI of 35 and above, with and without active diabetes type 1 or 2 that is treated with medications.

Naturally slim control individuals gave one fecal, blood and urine sample, and completed a dietary and medical questionnaire. Obese/diabetic people gave five samples of each at the following timepoints: time of enrolment (group A), three weeks after low carbohydrate diet and immediately before the operation (Group B), 2 weeks (group C), 3 and 6 months after the operation (groups D&E, respectively). In all visits the patients' weight, blood and urine test results, medications and general health issues were noted along with answers to the dietary questionnaire. In addition, the obese group provided weight values and blood test results at 1, 1.5 and 2 years after the operation. At the time of analysis not all patients completed their course of testing and evaluation.

Blood tests included the standard CBC test list and the blood biochemistry test following 12 hr fasting (for triglycerides, LDL, HDL, blood glucose and many other variables). In addition, patients provide samples for the standard hemoglobin A1 C (HbA1C) test.

DNA extraction

Fecal samples were stored in Flora prep tubes (Admera Health, New Jersey, USA) with a proprietary bacterial DNA preservation liquid, allowing for storage at room temperature. The samples were brought into the NGS lab within 1-2 days after collection and were subjected to DNA extraction, purification and cleaning. DNA was extracted from the stool sample using PowerSoil DNA extraction kit (MoBio) according to the manufacturer's instructions. Purified DNA was used for PCR amplification of the variable V3 and V4 regions (Using MetaVx.2™ system developed with GENEWIZ of the 16S rRNA gene as was previously described in US patent 9,745611, and by Caporaso et al[11]. Amplicons were purified using AMPure magnetic beads (Beckman Coulter) and subsequently quantified using Qubit dsDNA quantification kit (Invitrogen) by the Agilent Tap-Station to ensure that the correct size was obtained. Equimolar amounts of DNA from individual samples were pooled and sequenced using the Illumina MiSeq platform and V2 500 cycle kit. Microbial communities were analyzed using QIIME (quantitative insights into microbial ecology)[12] version 1.9.1 applying Usearch to align the sequences to the RDP (ribosomal database project) dataset for 16S rRNA gene sequences.

DNA Amplification

Briefly, the amplification method includes using multiple sets of overlapping primers and generation of multiple frame-shifted amplicons, which increase taxonomically diverse sequence amplification. The primers used in the present work are listed in the Table below and are designed to amplify the variable regions V3 and V4 of the bacterial 16SrRNA (as detailed in paragraphs [0083], [0090], and [0091] of US 9,745,611, using V3-forward and V4-reverse primers to amplify the V3-V4 region).

Table 2: Primers sequences

| Name | Sequence | # |
|------------|--|----|
| U341F-p5 | ACACTCTTTCCCTACACGACGCTCTTCCGATC TNCCTACGGGRSGCAGCA | 1 |
| E343F-p5 | ACACTCTTTCCCTACACGACGCTCTTCCGATC TNTACGGRAGGCAGCAG | 2 |
| E347F-p5 | ACACTCTTTCCCTACACGACGCTCTTCCGATC TNGGAGGCAGCAGTRRGAAT | 3 |
| E347F-p5-n | ACACTCTTTCCCTACACGACGCTCTTCCGATC TNNGGAGGCAGCAGTRRGAAT | 4 |
| A349F-p5 | ACACTCTTTCCCTACACGACGCTCTTCCGATC TNGYGCASCAGKCGMGAA | 5 |
| E802R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNTACNVGGGTATCTAATCC | 6 |
| E803R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNCTACCRGGGTATCTAATCC | 7 |
| P803R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNCTACCRGGGTATCTAAGCC | 8 |
| E806R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNGGACTACHVGGGTWTCTAAT | 9 |
| A806R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNGGACTACVSGGGTATCTAAT | 10 |
| U805R-p7 | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNGACTACHVGGGTATCTAATCC | 11 |
| U805R-p7-n | GACTGGAGTTCAGACGTGTGCTCTTCCGATC TNNGACTACHVGGGTATCTAATCC | 12 |

#: SEQ ID NO; According to IUPAC nucleotide code: K: G/T; M: A/C; R: A/G; Y: C/T; S: C/G; W: A/T; V: A/C/G; H: A/C/T; N:A/G/C/T;

As can be seen from Table 2 of US 9,745,611, these primers are capable of amplifying sequences from bacterial gut microbiome to identify rare species which may be difficult to amplify by similar kits based on 16S rRNA V4 region sequence amplification. In our hands the use of this procedure enables to identify bacterial taxonomic entities that are below 0.2% of the population, and enables the detection of rare species.

Bioinformatics

Paired end sequences were grouped into operational taxonomic units (OTU) using the GreenGene Database 11 (16S RNA database [13]) by grouping sequences with similarity of 97% or greater into the same OTU. Chimeric sequences with similarities were removed and an OTU table has been generated from the remaining sequences resulting in OTU frequencies in each sample.

The following contaminants have been removed from the final table: Thermi, S24-7 and Chloroplast.

Normalization.

OTUs were merged to the genus level by averaging over all OTUs assigned to the same genus. Given the large variation in OTU values, we transformed these values to Z scores by adding a minimal value to each OTU level (0.01) and

calculating the 10-basis log of each value. Statistical Whitening was then performed on the table, by removing the average and dividing by the standard deviation of each OTU.

Machine Learning.

Supervised Learning was performed on the normalized and merged version of the 16S rRNA OTU table in order to recognize patterns in the data. Principal Component Analysis was performed using Python version 3.5 and its package sklearn. A 2-tailed p value of less than 0.05 was considered to indicate statistical significance. A LASSO regression was performed to predict future BMI change. Leave One Out cross-validation method was performed. More complex methods were not used to limit overfitting, given the limited number of samples.

Statistical analysis,

All correlations studied here are Spearman correlations. P values of ROC curves are computed using scrambling the classes (positives or negative) of the samples and computing the AUC of 1,000 scrambles. The real AUC was compared to the 1,000 scrambles. Benjamini Hochberg correction were performed when multiple correlations were computed.

Results And Discussion

To test the possibility of predicting bariatric surgery outcome, we analyzed 308 fecal samples from patients of 2 main groups: obese who underwent bariatric surgery and naturally slim. For the obese patients (BMI>35) we sampled the microbiome at five time points (Fig. 1A) – one at enrollment (A, 78 samples), three weeks after a low carbohydrate diet and immediately before the operation (B, 70 samples), and 3 time points following the surgery (two weeks– C , 34 samples; three months – D, 27 samples; and six months E, 16 samples). Not all individuals had been sampled at all time points. This was compared to 83 slim control individuals (BMI 19-25) (For all details, see Supp. Methods). We have collected BMI and sugar A1C information for the same patients in late time points up to a year and a half post-surgery to track their weight loss and the remission of diabetes.

The slim population was younger and had more males compared to the population who underwent surgery (36 +/- 12 vs 48 +/-12 and 50% males vs 29%). Overall, the patients' mean BMI was reduced from 43.3 +/-6.8 (Mean +/-SD) to 27.8 +/-1.5, which represents an average loss of 84.7% overweight (compared to BMI 25), blood sugar levels were reduced from Hemoglobin A1C of 6.5 +/-0.4 to 5.8 +/-0.75 or from blood sugar levels of 125 +/-11 g/dL to 95.4 +/-10, Triglyceride levels decreased from 183 +/-20 to 102 +/-13. All parameters described are significantly ($P < 0.001$) lower from their starting point and not different from the slim control (Figure 1B-D).

The gut microbiome of all donors was analyzed using 16S rRNA gene sequences (emphasizing the 16S MetaVx.V2) [14], and OTU tables were produced using QIIME2 [15]. The OTU tables were then merged to the genus level. Taxa appearing in only one sample were removed. All samples were log-normalized to highlight the differences in rare bacteria and z-scored. The normalized tables were projected on PCA components. The aim of the first step was to homogenize the description level and reduce the dimension. Since multiple OTUs are associated with the same bacteria, and some OTUs are associated with different levels of classification, we averaged all OTUs associated with the same species in each donor (Fig. 2A-C). Note that while information is lost in the process, such a process is essential for the following machine learning. If an OTU was only present in part of the samples, it was given a value of 0 in all other samples. The resulting values have a scale-free distribution, which often masks large changes in relative frequencies of rare bacteria. To handle that, we log-transformed all OTU values and added a minimal constant value (0.1) to avoid log of zero values. This allows for a narrower distribution of values (Fig. 2D). Finally, given the very high

correlation between the relative abundance of different bacteria (Fig. 2E), we projected the bacteria to principal components, which capture most of the variance in the bacterial diversity (Fig. 2F).

The projection on the first principal vector (PC1) clearly delineated axes separating the obese from the slim individuals (Fig. 3A,B). The clear separation of the PC1 projections agrees with observed major differences in the microbiome of slim and obese individuals. The large BMI difference between groups (BMI>35 in obese versus BMI of <25 in slim) translates to a large difference in the microbiome. We next tested whether diet or bariatric surgery push back the population toward the slim profile. The results are surprisingly opposite (Fig. 3B). The distance between the projection on the first PC of the post-diet and post-surgery and the slim profile keeps increasing and reaches a maximum after a year. The major difference between these projections allows for a simple classification even with a linear SVM of slim vs obese and pre vs post-surgery samples (Fig. 3C,D ROC curves). The main contributions to the classifiers are from the first two PCs for the healthy (H) vs obese (O) and the same with the 5th PC for the pre vs post-surgery (Fig. 3E). Note that higher test AUC can be obtained by non-linear classifiers. However, the linear classifier gives a clear picture of the contribution of each PC to the microbiome development.

One can then project back the correlations between the PC and the state/BMI to the original OTUs, and find OTUs that are correlated to BMI (PC 1 and 3), the OTUs that are over and underrepresented in obese individuals compared to healthy and the OTUs that change significantly after surgery compared to before surgery (Fig. 4). Interestingly, *Helicobacter* is strongly associated with high BMI (Fig. 4A,B,C) as it has been reported previously in several studies to be associated with inflammation, insulin resistance and BMI [16-19]. The opposite behavior occurs with the bacterial family *Succinivibrionaceae* which are highly associated with healthy individuals and the after-surgery state. Members of this family ferment carbohydrates to acetate and succinate (succinate being an important intermediate for propionate production) [20]. Short-chain fatty acids (SCFAs) such as acetate and propionate have multiple beneficial roles. They are known for their contribution to improved insulin sensitivity and glucose homeostasis [21] and for their protective effect against diet-induced obesity [22]. The inflammatory effects of *Helicobacter* and the SCFA producing capabilities of the *Succinivibrionaceae* help explain our observation the first is higher in the obese state and the latter are associated with a healthy lean state.

To test for possible confounding effects, we tested whether the observed changes in the profile may be the result of age or gender, or whether they are related to the total BMI. There are no significant correlations with gender and a very weak correlation of PC3 and 5 with age (Fig. 5E, F).

Since the first three PC represent the main effect of the current BMI (i.e., BMI at time of sampling), we tested whether the other PCs represent other aspects of the population. Specifically, we tested if other PCs can be used to predict a future change in BMI. We performed an L1 (Lasso) regression of the projection on the first PCs of the A point and the change in BMI between point A and points E (six months after surgery), 1Y and 1.5Y. The prediction was tested using a Leave One Out (LOO) methodology, and the Spearman correlations on the test values between the predicted change and the observed change (on the LOO test), as well as the Area Under Curve of a predictor of whether a patient will have a more/less than average reduction in BMI. The AUCs are significant for all points (Fig. 5B-D). The Spearman correlation is highly significant for point E (where we have more information than other points). The prediction of future BMI change is also determined by a single main PC (the 4th PC Fig. 5A and G), further showing the natural decomposition of the microbiome into elements correlated with the host behavior. When looking at which OTUs are contributing to PC4 (Fig. 5A) we found *Holdemania* to be strongly correlated with future weight loss. *Holdemania* was previously found to be overrepresented in lean individuals [23]. *Coprococcus* which also correlates with greater BMI loss in the future is a known butyrate producer [24]. The beneficial effect of butyrate to weight loss may be due to its

role as a substrate in gluconeogenesis [25], and recently, oral butyrate supplementation has even been shown to reduce adiposity and improve insulin sensitivity [26].

Another possible candidate to affect the microbiome is the sugar level. We tested the correlation between the projections on the PCA and the A1C. Indeed, a clear correlation is found between A1C and PC7 (Fig. 6B for correlations and 6A for the projections). Using the projection, diabetic and non-diabetic patients, as defined by A1C (above 6.5 and below 5.7) can be separated with a high AUC (0.75 average over test $p < 0.01$) (Fig. 6B). This correlation might be used for predicting the disease in healthy subjects both from risk groups and in general. The main OTU negatively affecting PC7 are members of Erysipelotrichaceae which are known to be associated with obesity [27] (Fig. 6A).

Conclusions

To summarize, we have shown that the decomposition of the relative bacteria frequency (as represented by log value) naturally represents the different aspects of the donors. This decomposition highlights that people who lost weight after bariatric surgery have a very different microbiome composition compared to people who are "naturally" slim. Furthermore, the more weight they lose, the more their microbiome profile differs not only from their starting profile as obese but also from naturally slim people. Moreover, one can predict in advance whether surgery will succeed in reducing BMI and whether a subject has diabetes. The PCs are determined by the composition of the studied populations, and the analysis of different populations may highlight different possible projections of the microbiome composition. Interestingly, 2 main PCs remain with no clear correlation to the phenotypes studied here. Those may represent other important elements affecting and affected by the microbiome not tested in this study.

Declarations

Ethics approval and consent to participate

Kaplan Medical center 0068-15-KMC

Rabin Medical Center 0088-16-RMC

Tel Aviv Medical Center 0548-16-TLV

Poria Medical center 0057-18-POR

Consent for publication

All coauthors have agreed to the publication

Availability of data and material

The Sequence files are now uploaded to the EBI

• Competing interests

We have no competing interests

Funding

US-Israel bi-national Research and Development Fund (FIRD-F) project # 1459 (RC, LMS, HM, CL)

Authors' contributions

| <u>contributions</u> | Ethics and protocol | Data and sample collection | NGS testing | Database | Analysis | Manuscript Drafting | Manuscript editing | Funding |
|--|----------------------------|-----------------------------------|--------------------|-----------------|-----------------|----------------------------|---------------------------|----------------|
| Meirav Ben Izhak , | | | | X | X | X | X | |
| Dr. Hamutal Meiri, PhD, MBA (hylabs) | X | X | | X | X | X | X | X |
| Dr. Liora Madar Shapiro , PhD, MS (Hylabs), | | X | X | X | X | X | | X |
| Dr. Ruti Cohen, PhD (Hylabs), | X | X | | X | X | X | | X |
| Dr. Chaim Wechtel, PhD (Hylabs) | X | X | X | X | X | X | | |
| Conrad Leung, PhD (GENEWIZ, | | | X | X | X | X | X | X |
| Edward Messick, MS(GENEWIZ) | | | | X | X | X | | |
| Narisra Jongkam, MS (GENEWIZ | | | X | X | X | X | | |
| Omry Koren, PhD, Bar Ilan | X | X | X | X | X | X | X | X |
| Yoram Louzoun, PhD, Bar Ilan | | X | X | X | X | X | X | |
| Eli Mavor, MD, Kaplan Medical Center | X | X | | X | X | X | | |
| Shimon Sapoznikov, MD, Kaplan Medical Center | X | X | | X | X | X | | |
| Prof. Nitsan Maharshak, MD | X | X | | X | X | X | | |
| Sobchi AbuAvid, MD | X | X | | X | X | X | | |
| Prof. Avishai Alis, MD | X | X | | X | X | | | |
| Dr. Ilanit Mahaler, MD | | X | | X | X | X | | |

Ethics and protocol

Data and sample collection

NGS testing

Database

Analysis

Manuscript writing

Funding

Funding

Acknowledgements

Diana Bluvshstein, Manar Hadad, Natali Ogo for their help as clinical nurse involved in data collection and patient enrolment.

References

1. Spor A, Koren O, Ley R: **Unravelling the effects of the environment and host genotype on the gut microbiome.** *Nat Rev Microbiol* 2011, **9**(4):279-290.
2. Korem T, Zeevi D, Zmora N, Weissbrod O, Bar N, Lotan-Pompan M, Avnit-Sagi T, Kosower N, Malka G, Rein M: **Bread affects clinical parameters and induces gut microbiome-associated personal glycemic responses.** *Cell metabolism* 2017, **25**(6):1243-1253. e1245.
3. Organization WH: **Global status report on noncommunicable diseases 2014.** In.: World Health Organization; 2014.
4. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI: **An obesity-associated gut microbiome with increased capacity for energy harvest.** *Nature* 2006, **444**(7122):1027-1031.
5. Nagpal R, Kumar A, Kumar M, Behare PV, Jain S, Yadav H: **Probiotics, their health benefits and applications for developing healthier foods: a review.** *FEMS Microbiol Lett* 2012, **334**(1):1-15.
6. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR *et al*: **Gut microbiota from twins discordant for obesity modulate metabolism in mice.** *Science* 2013, **341**(6150):1241-1244.
7. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blehman R, Beaumont M, Van Treuren W, Knight R, Bell JT *et al*: **Human genetics shape the gut microbiome** *Cell* 2014, **159**(4).
8. Brethauer SA, Aminian A, Romero-Talamas H, Batayyah E, Mackey J, Kennedy L, Kashyap SR, Kirwan JP, Rogula T, Kroh M *et al*: **Can diabetes be surgically cured? Long-term metabolic effects of bariatric surgery in obese patients with type 2 diabetes mellitus.** *Ann Surg* 2013, **258**(4):628-636; discussion 636-627.
9. Ma IT, Madura JA: **Gastrointestinal complications after bariatric surgery.** *Gastroenterology & hepatology* 2015, **11**(8):526.
10. Graessler J, Qin Y, Zhong H, Zhang J, Licinio J, Wong ML, Xu A, Chavakis T, Bornstein AB, Ehrhart-Bornstein M *et al*: **Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters.** *Pharmacogenomics J* 2013, **13**(6):514-522.
11. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M: **Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.** *The ISME journal* 2012, **6**(8):1621.
12. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI *et al*: **QIIME allows analysis of high-throughput community sequencing data.** *Nat Methods* 2010, **7**(5):335-336.
13. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL: **Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.** *Applied and*

environmental microbiology 2006, **72**(7):5069-5072.

14. Huang J, Zhenzhen Z, Hairong D, Wu X, Li S: **Methods and kits for identifying microorganisms in a sample**. In.: Google Patents; 2017.
15. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F *et al*: **Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2**. *Nat Biotechnol* 2019, **37**(8):852-857.
16. Chen LW, Kuo SF, Chen CH, Chien CH, Lin CL, Chien RN: **A community-based study on the association between Helicobacter pylori Infection and obesity**. *Sci Rep* 2018, **8**(1):10746.
17. Chen LW, Chien CY, Yang KJ, Kuo SF, Chen CH, Chien RN: **Helicobacter pylori Infection Increases Insulin Resistance and Metabolic Syndrome in Residents Younger than 50 Years Old: A Community-Based Study**. *PLoS One* 2015, **10**(5):e0128671.
18. Polyzos SA, Kountouras J, Zavos C, Deretzi G: **The association between Helicobacter pylori infection and insulin resistance: a systematic review**. *Helicobacter* 2011, **16**(2):79-88.
19. Wu MS, Lee WJ, Wang HH, Huang SP, Lin JT: **A case-control study of association of Helicobacter pylori infection with morbid obesity in Taiwan**. *Arch Intern Med* 2005, **165**(13):1552-1555.
20. Santos EdO, Thompson F: **The Family Succinivibrionaceae**. In: *The Prokaryotes: Gammaproteobacteria*. Edited by Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014: 639-648.
21. Canfora EE, Jocken JW, Blaak EE: **Short-chain fatty acids in control of body weight and insulin sensitivity**. *Nat Rev Endocrinol* 2015, **11**(10):577-591.
22. Rios-Covian D, Ruas-Madiedo P, Margolles A, Gueimonde M, de Los Reyes-Gavilan CG, Salazar N: **Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health**. *Front Microbiol* 2016, **7**:185.
23. Andoh A, Nishida A, Takahashi K, Inatomi O, Imaeda H, Bamba S, Kito K, Sugimoto M, Kobayashi T: **Comparison of the gut microbial community between obese and lean peoples using 16S gene sequencing in a Japanese population**. *J Clin Biochem Nutr* 2016, **59**(1):65-70.
24. Anand S, Kaur H, Mande SS: **Comparative In silico Analysis of Butyrate Production Pathways in Gut Commensals and Pathogens**. *Front Microbiol* 2016, **7**:1945.
25. Arora T, Backhed F: **The gut microbiota and metabolic disease: current understanding and future perspectives**. *J Intern Med* 2016, **280**(4):339-349.
26. McNabney SM, Henagan TM: **Short Chain Fatty Acids in the Colon and Peripheral Tissues: A Focus on Butyrate, Colon Cancer, Obesity and Insulin Resistance**. *Nutrients* 2017, **9**(12).
27. Zhou W, Sailani MR, Contrepois K, Zhou Y, Ahadi S, Leopold SR, Zhang MJ, Rao V, Avina M, Mishra T *et al*: **Longitudinal multi-omics of host-microbe dynamics in prediabetes**. *Nature* 2019, **569**(7758):663-671.

Figures

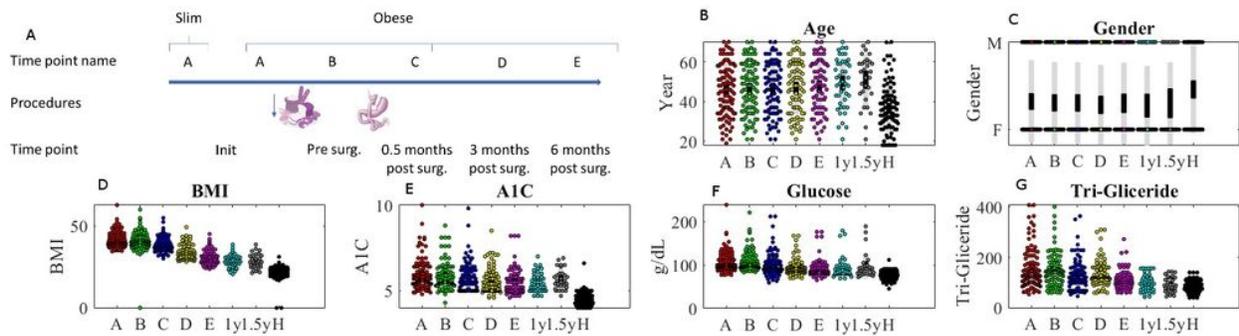


Figure 1

1A – Experimental setup. The experiment tested the microbiome and different intakes (HDL, LDL, Triglycerides, BMI and A1C) of patients from two groups – obese who underwent Bariatric surgery and slim individuals. The slim individuals have been sampled once while the obese patients were sampled in 5 time points – before the entire process (A), after low carbohydrate diet and before surgery (B), 2-3 weeks after surgery (C), 3 months after surgery (D) and 6 months after surgery (E). Fig. 1B-E Distribution of Age, Gender, BMI, A1C, Glucose and Tri-glyceride levels at the samples taken in each group and time point. The black boxes represent the 25th-75th percentile. The gray boxes represent the 10th-90th percentile. The gender distribution is plotted similarly to the other continuous values.

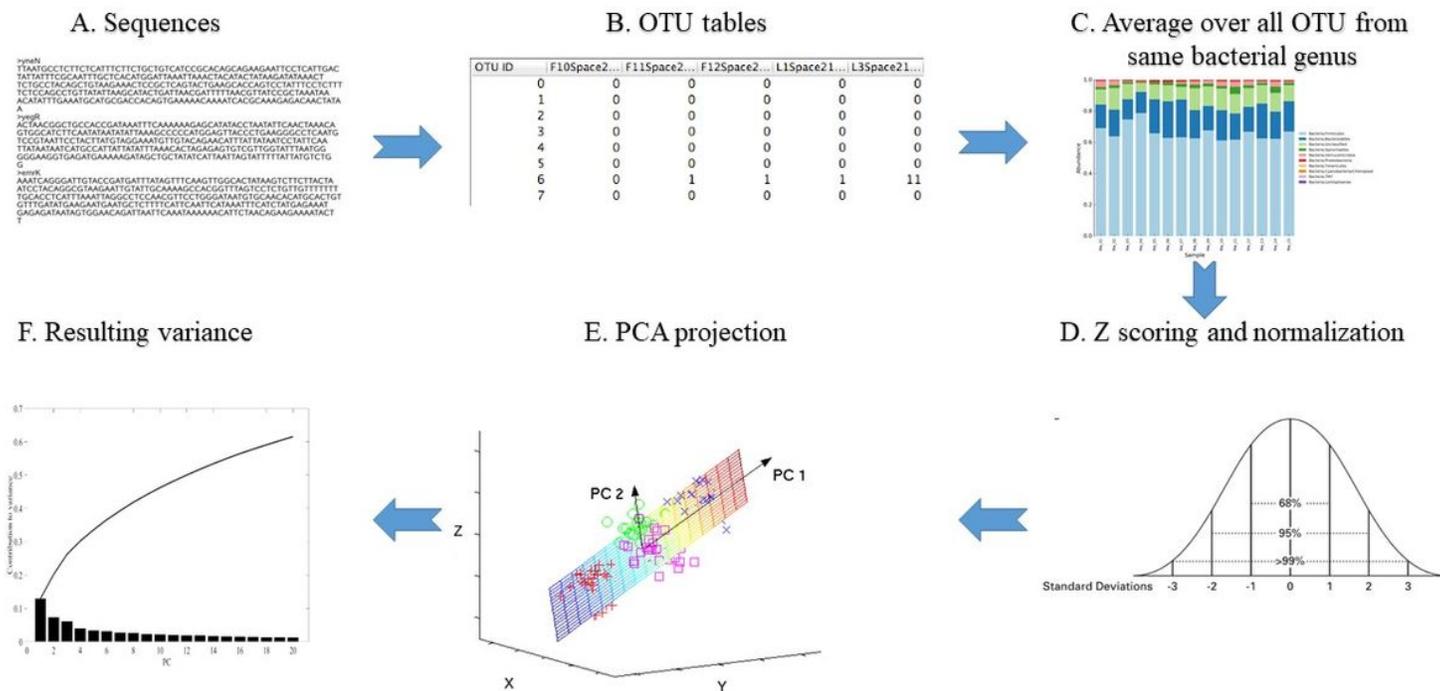


Figure 2

Outline of analysis (from upper left to upper right and then from lower right to lower left). First, the Fastq sequences are quality controlled. The good quality sequences are translated to OTUs using QIIME2. In order to homogenize the description level, the OTU levels belonging to the same genus in a given sample are averaged to the genus level. (Lower level). The sample distribution is heavy-tailed. It is thus log-transformed with a minimal value (0.1) added to each OTU level to avoid log of zero values. The results are then z scored by removing the average and dividing by the standard

deviation of each sample. The dimension of the z-scores is further reduced using PCA. The first 8 PC explain approximately 40 % of the total variance (lower left figure).

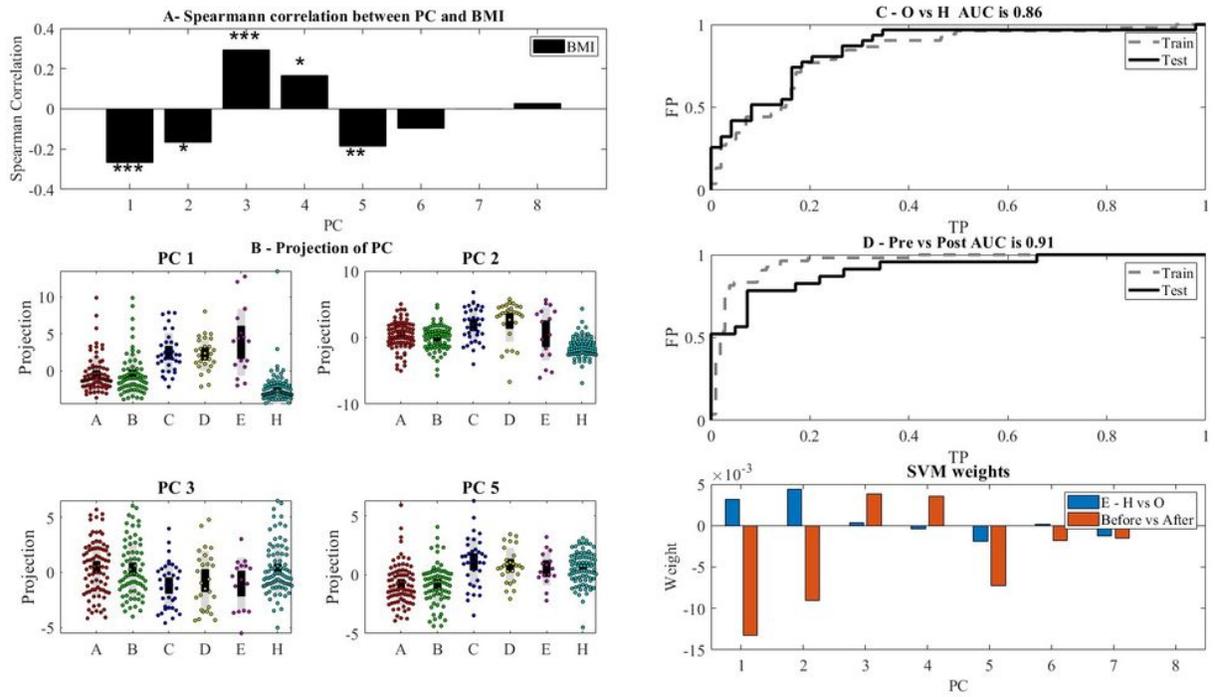


Figure 3

Upper left plot. Spearman correlation between BMI of samples and the eight highest variance PC. ***,** and * represent significance level of 0.001,0.01 and 0,05 respectively (in this and all following figures). PC1 and 3 are the most correlated with BMI, but some correlation is also observed between PC 2 and 4. Lower left plots projection of the significant PC on the different stages. One can see in PC1 a clear difference between the H and the obese states. Following surgery, the projection is farther away from the H state than before. The same can be seen to a lesser degree in PC3. Note that we do not have microbiome samples from the latest time points. Upper right plots ROC curves of linear SVM classification using the projections on the first 8 PC of the H vs O and within the O group before vs after surgery, and the resulting weights (lower right plot), with the main contribution again from PC1,2 and 5.

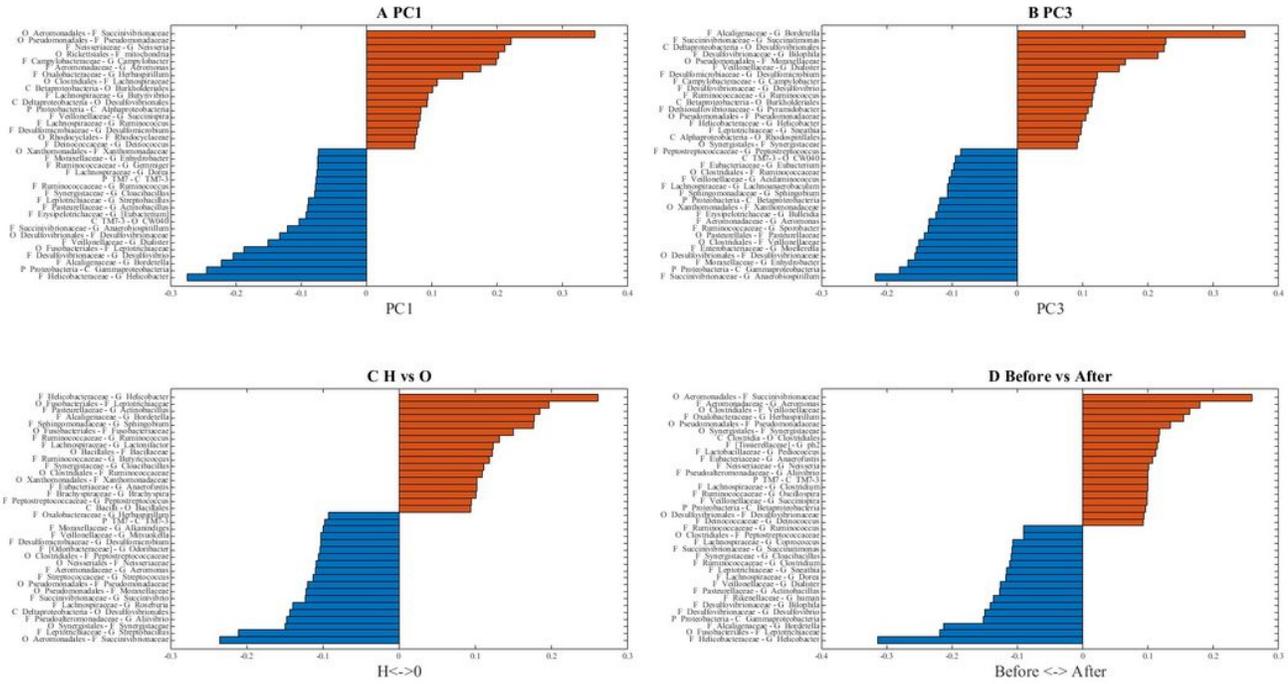


Figure 4

Upper (A, B) plots - Bacterial composition of PC1 and PC3. Lower plots (C,D) – linear SVM classifier for H vs O and before vs after surgery. Only the top 15 % of OTUs are presented.

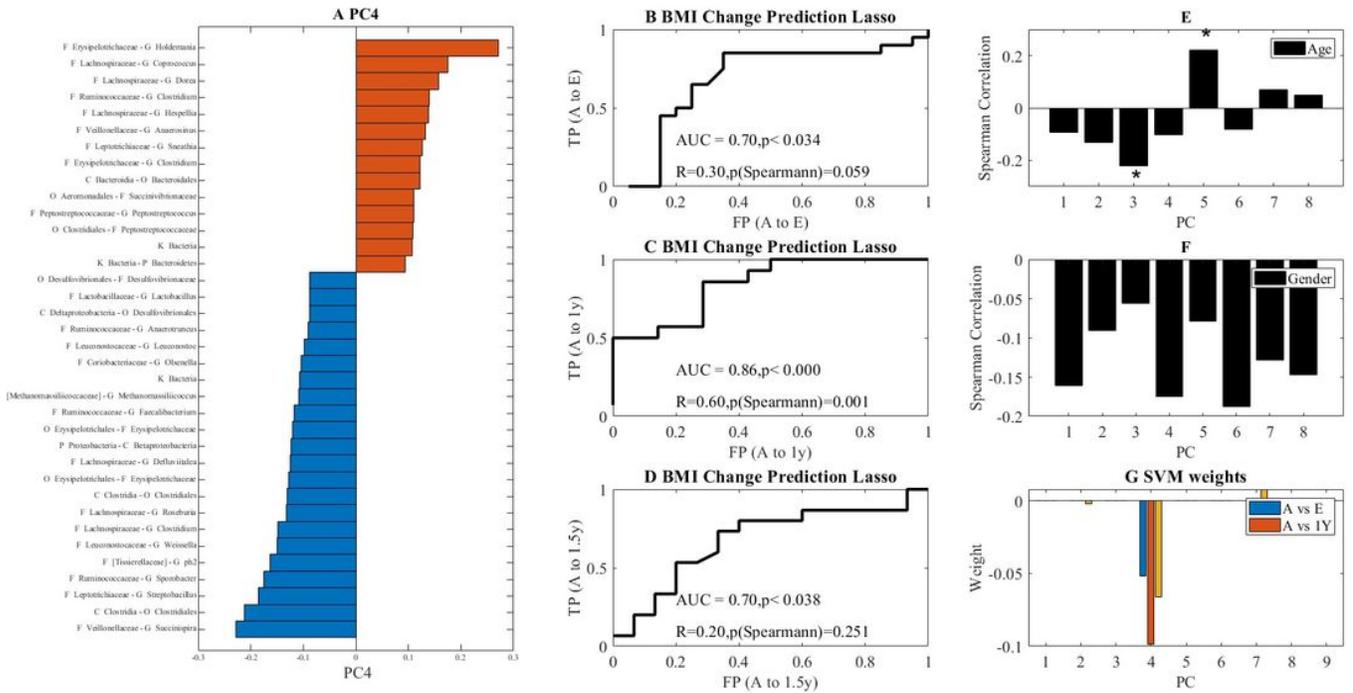


Figure 5

A Composition of PC4 (which is most correlated with future BMI change). B-D correlation of Age, Gender (represented as a one-hot) and BMI (reproduced from Figure 3) with the PC. The correlation with Age is borderline and does not pass a multiple measurement correction. E-F ROC curves of future BMI change based only on the PC at point A. The ROC curves for above and below median change in BMI, using a LASSO regression and a LOO validation. E is for six months and F is for one-year change. The first p-value is the ROC curve p-value compared with ROC curves of 1000 scrambling of the predicted values. The second p is the Spearman correlation p-value between predicted and actual change. G – average regression weights over all LOO learning sessions. One can see that prediction is based on the 4th PC.

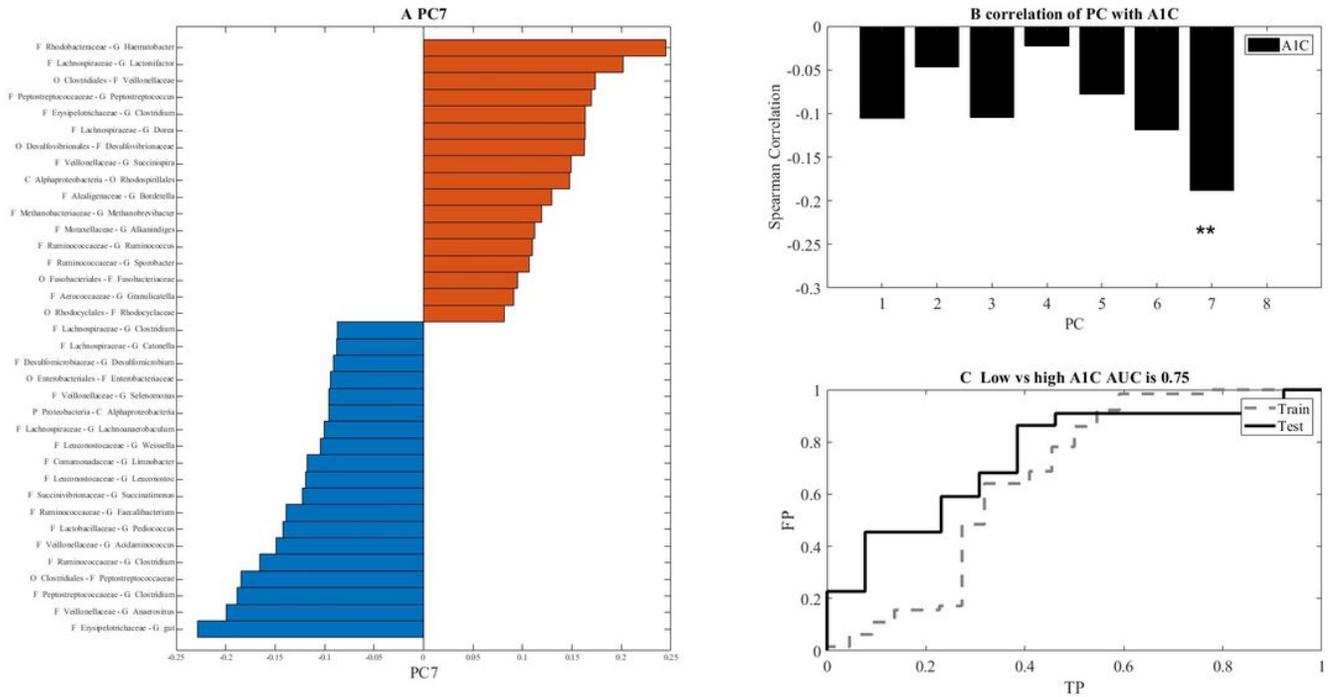


Figure 6

B correlation of the different projection with A1C. The only significant correlations are for PC7. C. ROC curve of SVM classifier of A1C with a test AUC of 0.75. A is the projection of PC7.