bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Gut microbiota analyses of Saudi populations for type 2 diabetes-related phenotypes reveals significant association

Fahd A Al-Muhanna[1], Alexa K. Dodwell[2], Abdulmohsen H Al Eleq[1], Waleed I Albaker[1], Andrew W. Brooks[3], Ali I Al-Sultan[1], Abdullah M Al-Rubaish[1], Khaled R Alkharsah[4], Raed M Sulaiman[1], Abdulaziz A Al-Quorain[1], Cyril Cyrus[5], Rudaynah A Alali[1], Chittibabu Vatte, Fred L. Robinson[2], Carrie Nguyen[2], Xin Zhou[3], Michael P. Snyder[3], Afnan F Almuhanna[6], Brendan J Keating[7], Brian D. Piening[2], Amein K Al-Ali[5]

1. Department of Internal Medicine, King Fahd Hospital of the University, Al-Khobar and College of Medicine, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia
2. Earle A Chiles Research Institute, Providence Portland Medical Center, Portland, Oregon, USA
3. Department of Genetics, Stanford University School of Medicine, Stanford, California, USA
4. Department of Microbiology, College of Medicine, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia
5. Department of Clinical Biochemistry, College of Medicine, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia
6. Department of Radiology, King Fahd Hospital of the University, Al-Khobar and College of Medicine, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia
7. Department of Surgery, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

Running Title: Saudi gut microbiota

*** Corresponding Author**

Email: aalali@iau.edu.sa (AKA)

These authors contributed equally to this work.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Large-scale gut microbiome sequencing has revealed key links between microbiome dysfunction and metabolic diseases such as T2D. To date, these efforts have largely focused on Western populations, with few studies assessing T2D microbiota associations in Middle Eastern communities where T2D prevalence is now over 20%. We analyzed the composition of stool 16S rRNA from 461 T2D and 119 non-T2Dparticipants from the Eastern Province of Saudi Arabia. We quantified the abundance of microbial communities to examine any significant differences between subpopulations of samples based on diabetes status and glucose level. We observed overall positive enrichment within diabetics compared to healthy individuals and amongst diabetic participants; those with high glucose levels exhibited slightly more positive enrichment compared to those at lower risk of fasting hyperglycemia. In particular, the genus *Firmicutes* was upregulated in diabetic participants compared to non-diabetic participants, and T2D was associated with an elevated *Firmicutes/Bacteroidetes* ratio, consistent with previous findings. Based on diabetes status and glucose levels of Saudi participants, relatively stable differences in stool composition were perceived by differential abundance and alpha diversity measures.

Author summary

The rates of Type 2 diabetes (T2D) in Saudi Arabia have risen dramatically in the last several decades due to socio-economic changes resulting in changes in dietary and sedentary lifestyles. This emergence has grown more rapidly and affects larger proportions of the population with estimates of T2D prevalence impacting 25% of the population. There is a paucity of microbiome data from Middle Eastern populations, and previous studies have been conducted on

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

small sample sizes. Here we report on the first-ever characterization of gut microbiota T2D versus non-T2D and largest microbiome study ever conducted in a Middle Eastern country. The datasets from this study are important to create a regional reference T2D-microbiome catalogue which will propel the understanding of regional gut flora which are associated with T2D development. Based on T2D status and quantified glucose levels of Middle Eastern participants, relatively stable differences in stool composition were observed by differential abundance and alpha diversity measures. Comparing overlapping and varying patterns in gut microbiota with other studies is critical to assessing novel treatment options in light of a rapidly growing T2D health epidemic.

Introduction

The human gut hosts 100 trillion microorganisms, encompassing thousands of species collectively, weighing an average 1.5 kg per person [1,2]. The human microbiota is important because of its metagenomic repertoire, which is estimated to be 100 times larger than the human genome and encodes a vast array of functionality critical for host physiology and metabolism [2]. The bacterial components responsible for triggering theses physiological functions are currently the subject of intensive research. Differences in human gut microbiome composition have been linked to metabolic diseases such as T2D and obesity [3-7]. Identifying specific bacterial biomarkers within the microbiome could help predict the occurrence of T2D or tailor treatments in high-risk subjects to prevent or delay the onset of metabolic diseases. The molecular mechanisms through which the intestinal microbiota play a key role in metabolic diseases are linked to an increased energy harvesting and the triggering of the low-grade inflammatory status characterizing insulin resistance and obesity [8-9].

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The prevalence of T2D is increasing worldwide, with current data indicating that at least 8.5% of the world's population is affected, with the worldwide prevalence expected to reach 12% by 2025 [10-11]. T2D is mainly caused by insulin resistance and relative insulin deficiency [12]. Saudi Arabia, with a total population of over 20 million, has an estimated T2D constituting 25% of the total population [13]. The rapid rate of increase of T2D disease in some areas of Saudi Arabia, which increased from 16% in 2005 to over 25% in 2011, is thought to be due to rapid lifestyle changes such as diet and sedentary lifestyle, as we;; as adverse environmental factors [13].

We analyzed the composition of 16S rRNA from the stool samples collected from Saudi Arabian participants residing in the Eastern Province and quantified the abundance of microbial communities to determine significant differences between subpopulations of samples based on diabetes status and glucose level. We assessed alpha diversity between the subpopulations to measure species richness and evenness among samples noting that an increased *Firmicutes:Bacteriodetes* ratio has previously been observed in the microbiota of obese/diabetic individuals compared to the microbiota of healthy individuals. Furthermore, individuals with diabetes were tracked for high glucose level (>126 mg/dL) as it is an indicator of fasting hyperglycemia, which could potentially lead to severe long-term complications including cardiovascular disease, neuropathy and kidney failure.

RESULTS

Principal coordinate analysis (PCoA) of the generated 16S datasets is shown in Fig S2. The first and second principal coordinated explained 25% and 7%; 29% and 7% and 34% and 6% of the Diabetes Status and Gender variance, respectively. Levels of the 150 most abundance microbial

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

67 genera within T2D and non-T2D participants were observed to differ significantly in stool
68 microbiota abundance derived from 16S sequencing (Figures S3a and S3b).

69 Fig 1 (a and b) shows the rank abundant curve and Permutational Multivariate Analysis of
70 Variance (PERMANOVA) cloud, respectively for Saudi T2D and control 16S stool microbiota
71 datasets. These show that the microbiome communities differ globally between T2D and non-T2D
72 subjects at statistical significance, $p = 0.01$. The abundance of Taxonomic Composition in males
73 and females is clearly evident in both females (Figures S4a and S4b) and in males (Figures S5a
74 and S5b). We also compared Saudi T2D participants with higher glucose >126 mg/dL versus lower
75 glucose strata <=126 mg/dL glucose using the top 150 genera. Amongst the 298 samples with
76 glucose data, n=193 were in the higher glucose strata and n=105 were in the lower strata (Figure
77 S6). Unlike previous studies conducted on Western populations, the Saudi participants with T2D
78 and higher glucose levels showed a trend toward increased diversity, a result that is similar to
79 another recently reported study from a United Arab Emirates (UAE) cohort [14].

80 **Figure 1: Rank abundant curve (a) and permutational multivariate analysis of variance**
81 **cloud (b) for Saudi T2D and control 16S stool microbiota datasets.** This figure shows the rank
82 abundant curve and Permutational Multivariate Analysis of Variance (PERMANOVA) cloud
83 respectively for Saudi T2D and control 16S stool microbiota datasets. These show that the
84 microbiome communities differ globally between T2D and non-T2D subjects at statistical
85 significance, $p = 0.01$.

86

87 Alpha diversity was compared in males versus females (n = 204 and 226, respectively)
88 with no significant differences observed using various different classifications: ACE (Abundance-
89 based Coverage Estimator) and Chao1 indices to estimate richness (measurement of OTUs
90 expected in samples given all the bacterial species identified in the samples); Shannon-Weaver,
91 Simpson and Inverse Simpson to define different levels of resolution (phylum, class, order, family,

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

genus, and species); and Fisher (Fig S7). Alpha diversity of T2D versus non-T2D participants revealed statistically significant enrichment of the Shannon-Weaver and Simpson metrics (Figure S8) ($p < 2.26 \times 10^{-10}$ (CI: -0.392 to -0.718)) and $p < 4.63 \times 10^{-7}$ (CI: -0.049 to -0.108) for Shannon and Simpson diversity, respectively. Saudi T2D cases versus controls showed an association with an elevated *Bacteroidetes/Firmicutes* ratio, $p = 2.2 \times 10^{-5}$ t-test (Fig S9).

We observed an overall positive enrichment of microbiota genus/families for diabetics compared to healthy individuals. In addition, among T2D patients, those with high glucose levels exhibited slightly more positive enrichment compared to those at lower risk of fasting hyperglycemia (Fig 2a and 2b and Table S1). In particular, the *Akkermansia, Acidaminococcus, Megamonas, Dialister, Lactobacillus and Paraprevotella* genus were enriched at $p < 1 \times 10^{-9}$ in T2D versus non-T2D. The *Fusobacterium, Dialister, Akkermansia* and *Prevotella* genus were enriched in low versus high-risk T2D using a fasting glucose cutoff of 126 mg/dL.

Figure 2: Fold Change plots of enriched OTUs for: T2D versus controls (a) and glucose levels for high versus low T2D status (b). An overall positive enrichment of microbiota genus/families for diabetics compared to healthy individuals and amongst diabetic participants was observed. Those with high glucose levels exhibited slightly more positive enrichment compared to those at lower risk of fasting hyperglycemia.

Discussion

In this study we performed the largest microbiome study ever conducted in Saudi Arabia, as well as the first-ever characterization of gut microbiota T2D versus non-T2D in this population. We used shotgun metagenomic sequencing to obtain 16S rRNA reads identifiable down to genus level from the stool samples of 461 T2D and 119 non-T2D Saudi participants from the Eastern Province of Saudi Arabia, a region particularly affected by T2D [15]. We assessed the microbiota

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

abundance based on diabetes status and glucose levels, and examined community diversity patterns to compare with other T2D microbiota studies from around the globe. These efforts are important and warranted given the scarcity of microbiome data in Middle Eastern populations, and these results provide a useful addition to the global microbiome reference dataset in an under-examined community. Saudi Arabian T2D costs have risen over 500% in two decades with 10 million individuals estimated to be diabetic or pre-diabetic, therefore comparing overlapping and varying patterns in gut microbiota with other studies is critical to assessing novel treatment options in light of a rapidly growing T2D health epidemic [15-16].

Community level differences are evident in the Saudi population between T2D and non-T2D individuals, and diversity patterns appear to vary from well-characterized microbiota from Western cohorts. Indeed, in contrast to Western cohorts that often show associations between decreased gut microbiota diversity and insulin resistance, here we show that Saudi participants with T2D exhibited higher relative diversity in comparison to normal metabolic counterparts [17]. These results are similar to a recent report from Al Bataineh and colleagues who characterized microbiomes in a cohort of 50 T2D and non-T2D individuals from the United Arab Emirates, though higher diversity in that smaller T2D cohort was determined to be insignificant when controlling for age [14]. Sex was not found to play a role in community structural differences, and results were independently validated between females and males. The role of overall community diversity decreasing in T2D populations has been widely cited in early studies on Western populations, yet larger meta-analyses involving global populations have distorted this pattern and highlight the importance of locally representative studies [18].

We observe significant differences between T2D and non-T2D individuals for many microbial taxa, as well as between T2D individuals with high and low fasting blood glucose levels.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Concordant with studies conducted on Western populations is the association of increasing Bacteroidetes/Firmicutes ratio with T2D and in our overweight and obese T2D cohort, increased Bacteroidetes may be functionally related to metabolism of branched chain amino acids which has been linked to obesity-related metabolic phenotypes [3, 19]. Among OTUs assigned at the genus taxonomic level, *Prevotella* and *Bacteroides* OTUs showed some of the most significant log-fold increases in abundance for diabetics (over four-fold increases in abundance), species of which have been functionally associated with the development of insulin resistance and glucose intolerance [20]. Among Firmicutes however, levels of *Acidaminococcus* and *Megasphaera* were positively correlated with T2D, as has been previously observed, and could functionally relate with increases to Bacteroidetes through complementary amino acid metabolism [21-22]. We observed higher levels of *Akkermansia* in the Saudi T2D group, despite potential protective effects for obesity and metabolic disease. Associations of levels of *Akkermansia*, a mucus-consuming taxon, have been observed to be associated with health and with ethnicity in Western populations and may represent an impact of dietary and lifestyle effects on microbiota composition, as this microbe is rarely observed in more traditional cultures across large geographic regions [23]. It should be noted however that *Akkermansia* levels are also often increased in response to metformin intake in T2D subjects (metformin use metadata is not known for the current cohort) [24]. Taxonomic differences associated with T2D likely reflect shared or complementary functional and metabolic traits but may be regionally specific based on dietary and environmental variations known to influence the microbiome [23-25].

Based on diabetes status and quantified glucose levels of Middle Eastern participants, relatively stable differences in stool composition were observed by differential abundance and alpha diversity measures. Many studies have examined T2D associations with gut microbiota in

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

populations around the globe, and while some patterns generally validate across studies such as individual taxon abundance variation, others such as overall community diversity do not replicate consistently. Obesity, diet, lifestyle and ancestry are all factors that influence T2D and each varies significantly from culture to culture around the globe, meaning that the patterns in T2D development and roles of the microbiome likely vary as well. As a rapidly emerging chronic condition in Saudi Arabia and the Middle East, T2D burdens have grown more quickly and affect larger proportions of the population than any other global region, making a regional reference T2D-microbiome dataset critical to understanding the nuances of disease development on a global scale.

Materials and Methods

Study Populations

Between 2015-2019, stool samples and data were collected from 461 consecutive diabetic patients attending the Diabetic Clinics, King Fahd Hospital of the University, Al-Khobar, Saudi Arabia and from 119 healthy controls. Participants ranged in age from 30-75 years and had a body mass index (BMI) ranging from 27 to 40 kg/m^2. The T2D patients had a minimum disease duration of 5 years. Table 1 outlines the patient demographics and clinical characteristics. Baseline measurements included anthropometric measurements, physical examinations and in-person surveys. Participants who had been treated with antibiotics in the previous three months, were pregnant or lactating, or had inflammatory bowel disease were excluded from the study. Blood and stool samples were collected from participants and were stored immediately after collection at −80 °C. Ethical approval of the study was obtained from the local Institutional Review Board

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(IRB) committee and the study was conducted according to the ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines (IRB-2019-01-112). All participants provided written informed consent.

Table 1. Clinical and demographic characteristics for Saudi Arabian T2D cases (n-461) and controls (n=119).

	Ratio	Male	Female
Gender	1: 0.83	54.50%	45.50%
	Mean ± SD		
	Total	**Male**	**Female**
Age (Years)	52.6±8.83	51.82±9.28	53.5±8.25
Glucose(mg/dl)	165.7±68.89	161.45±57.71	166.8±74.09
HBA1c (%)	8.55±1.76	8.45±1.65	8.65±1.85
Duration (Years)	3-25	4-25	3-22
BMI (kg/m^2)	27-40	27-37	30-40

Methods for DNA library preparation and sequencing

Sample collection and microbial DNA extraction were standardized to minimize confounding effects of the technical procedure. Stool samples were taken from T2D (n=461) and from healthy (n=119) participants. Fecal samples were provided by the patients whilst attending the outpatient clinic and immediately stored at −20°C. The samples were subsequently transported on dry ice to the research laboratory where they were stored at −80°C. Bacterial DNA extraction from stool samples was performed using QIAamp Fast DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. In brief, approximately 200 mg of stool was placed in a 2 ml microcentrifuge tube and kept on ice. InhibitEX Buffer (1 ml) was added to each stool

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

sample, homogenized thoroughly by vortexing and incubated at 70°C for 5 minutes. Each sample was centrifuged (20,000 g) for minute and 200 µl supernatant was pipetted into 1.5 ml microcentrifuge tube containing 15 µl proteinase K and 200 µl lysis buffer and incubated at 70°C for 10 minutes. This was followed by the addition of 200 µl of ethanol and mixed by vortexing. The lysate (600 µl) was transferred to the QIAamp spin column and centrifuged (20,000 g) for 1 minute. Finally, the QIAamp spin column was opened and washed twice with two different washing buffers. The DNA was eluted into a new 1.5 ml microcentrifuge tube by adding 200 µl elution buffer. DNA samples were checked for purity using the Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific). Three independent extractions were performed from each sample to ensure robust representation of all microbial content. DNA was stored at −80 °C till the time of processing.

Sequencing was performed using either the Swift Amplicon 16S panel (Swift Biosciences) or a custom protocol. For the Swift protocol, 20 ng of stool-derived DNA was used for 16S sequencing library preparation using the 16S Primer Panel v2, the Swift Normalase Amplicon Panels (SNAP) Core Kit, and the SNAP Combinatorial Dual Index Primer Kit (Sets 1A and 1B) (Swift Biosciences, CA). The indexed libraries were on average 620 base pairs (bp) in length, and individual DNA libraries were diluted to 2.5 nM, pooled in equimolar proportion, and sequenced on a NovaSeq 6000 SP flow cell (Illumina, CA) using 250 bp paired-end reads. For the custom approach, PCR was performed on each sample using the 515F primer (forward primer) and one of the 100 806rcbc primers (reverse primer). These primers contained: sequence homologous to region V4 of the 16S rRNA in forward and reverse; Illumina adaptors; and the reverse primers contained indexing sequences. Taq PCR Master Mix from Qiagen was used to prepare the PCR master mix. A PCR reaction was performed on each extracted DNA sample, i.e. each stool sample

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

had three PCR reactions. The PCR product was run on 1% agarose gel. The band of expected size (381bp) was excised from gel and purified with gel purification kit from Qiagen. The three PCR products from each sample were pooled together. The pooled and purified PCR product was quantified with NanoDrop 2000 (Thermo Sciences, USA).

Equal concentrations of DNA from each sample (5ng of DNA) were pooled together. For each sequencing run, DNA from 50 samples was pooled to make the DNA library for each batch. The final concentration of the DNA library was quantified with real time PCR using the Kapa library quantification kit (Roche, USA) according to the manufacturer's instructions. The DNA library of each batch was sequenced using the MiSeq platform from Illumina (Illumina, USA) using the MiSeq reagent V2 500cycles Kit from Illumina and the custom read1 (TATGGTAATTGTGTGCCAGCMGCCGCGGTAA), read2 (AGTCAGTCAGCCGGACTACH VGGGTWTCTAAT) and index (ATTAGAWACCCBDGTAGTCCGGCTGACTGACT) sequencing primers. PhiX DNA (Illumina, USA) was used as a control library.

Analyses

Figure S1 overviews the analytical pipeline and workflow employed for these analyses. 16S rRNA (V4 region) sequences were used in this study and sequenced with Illumina software which handled the initial primer and barcode processing of all raw sequences. Raw sequences were demultiplexed with Illumina's bcl2fastq2 v2.20 [26]. FastQC was then used for further processing to remove samples with low quality scores across the majority of bases [27]. After de-multiplexing the raw sequences and screening via FastQC, the majority of data processing was executed in QIIME2 with custom scripts. Paired-end reads were joined using VSEARCH. Chimera amplicon removal and abundance filtering were processed using Deblur [28]. Amplicon sequences were

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

clustered and assembled into Operational Taxonomical Units (OTUs) using closed reference clustering against the Greengenes 13_8 database via VESEARCH. Taxonomic assignment was performed using a pre-trained Naïve Bayes classifier with Greengenes OTU database. The abundance tables and data obtained from QIIME2 were combined into a Phyloseq object and further analyzed in R with custom scripts [29].

Acknowledgement

The authors would like to acknowledge the financial supported extended by King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia, grant numbers 12-MED2799-46 and 13-MED1881-46. We are also grateful to the nurses and technical staff for their work and dedication.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285):59-65.
2. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449(7164):804-10.
3. Larsen N, Vogensen FK, van den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, et al. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. PLoS One. 2010;5(2):e9085.
4. Greenhill C. Obesity: Gut microbiota, host genetics and diet interact to affect the risk of developing obesity and the metabolic syndrome. Nat Rev Endocrinol. 2015;11(11):630.
5. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-7.
6. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. Gut. 2016;65(3):426-36.
7. Piening BD, Zhou W, Contrepois K, Röst H, Gu Urban GJ, Mishra T, et al. Integrative Personal Omics Profiles during Periods of Weight Gain and Loss. Cell Syst. 2018;6(2):157-70.e8.
8. Shoelson SE, Lee J, Goldfine AB. Inflammation and insulin resistance. J Clin Invest. 2006;116(7):1793-801.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

9. Hotamisligil GS. Inflammation and metabolic disorders. Nature. 2006;444(7121):860-7.
10. Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, Di Angelantonio E, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010;375(9733):2215-22.
11. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. Nature. 2001;414(6865):782-7.
12. Association AD. 2. Classification and Diagnosis of Diabetes. Diabetes Care. 2017;40(Suppl 1):S11-S24.
13. Alotaibi A, Perry L, Gholizadeh L, Al-Ganmi A. Incidence and prevalence rates of diabetes mellitus in Saudi Arabia: An overview. J Epidemiol Glob Health. 2017;7(4):211-8.
14. Al Bataineh MT, Dash NR, Bel Lassen P, Banimfreg BH, Nada AM, Belda E, et al. Revealing links between gut microbiome and its fungal community in Type 2 Diabetes Mellitus among Emirati subjects: A pilot study. Sci Rep. 2020;10(1):9624.
15. Alshayban D, Joseph R. Health-related quality of life among patients with type 2 diabetes mellitus in Eastern Province, Saudi Arabia: A cross-sectional study. PLoS One. 2020;15(1):e0227573.
16. Robert AA, Al Dawish MA, Braham R, Musallam MA, Al Hayek AA, Al Kahtany NH. Type 2 Diabetes Mellitus in Saudi Arabia: Major Challenges and Possible Solutions. Curr Diabetes Rev. 2017;13(1):59-64.
17. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. Nature. 2013;500(7464): 541-6.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

18. Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, et al. Role of gut microbiota in type 2 diabetes pathophysiology. EBioMedicine. 2020;51:102590.
19. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science. 2013;341(6150):1241214.
20. Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BA, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature. 2016;535(7612):376-81.
21. Gaike AH, Paul D, Bhute S, Dhotre DP, Pande P, Upadhyaya S, et al. The Gut Microbial Diversity of Newly Diagnosed Diabetics but Not of Prediabetics Is Significantly Different from That of Healthy Nondiabetics. mSystems. 2020;5(2).
22. Yang Q, Lin SL, Kwok MK, Leung GM, Schooling CM. The Roles of 27 Genera of Human Gut Microbiota in Ischemic Heart Disease, Type 2 Diabetes Mellitus, and Their Risk Factors: A Mendelian Randomization Study. Am J Epidemiol. 2018;187(9):1916-22.
23. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science. 2017;357(6353):802-6.
24. de la Cuesta-Zuluaga J, Mueller NT, Corrales-Agudelo V, Velásquez-Mejía EP, Carmona JA, Abad JM, et al. Metformin Is Associated with Higher Relative Abundance of Mucin-Degrading Akkermansia muciniphila and Several Short-Chain Fatty Acid-Producing Microbiota in the Gut. Diabetes Care. 2017;40(1):54-62.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

25. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.

26. H. Li, Seqtk: a fast and light weight tool for processing FASTA or FASTQ sequences, 2013. https://github.com/lh3/seqtk

27. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

28. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems. 2017;2(2).

29. McMurdie, P.J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One8, (2013).

Supporting information

Fig S1: Data processing and analyses pipeline for Saudi T2D 16S microbiota study.

Fig S2: Principal coordinate analyses of 16S microbiota data from Saudi T2D and control participants using: (a) sex and (b) T2D status.

Fig S3a: Heatmap of top 150 genus for (a) non-T2D and (b) T2D (OTU abundance based on BrayCurtis dissimilarity).

Fig S3b: Heatmap of top 50 genus for (a) non-T2D and (b) T2D individuals listed, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig S4a: Heatmap of top 150 gut microbiota 16S genus for (a) T2D and (b) T2D in Saudi females (OTU abundance based on BrayCurtis dissimilarity.

Fig S4b: Abundance of gut microbiota 16s taxonomic composition of: a) non-T2D vs (b) T2D in Saudi females

Fig S5a: Heatmap of top 150 gut microbiota 16S genus: (a) non-T2D (b) T2D in Saudi males (OTU abundance based on BrayCurtis dissimilarity).

Fig S5b: Abundance of gut microbiota 16S taxonomic composition of: (a) non-T2D versus (b) T2D in Saudi males.

Fig S6: Heatmap of top 150 genus for Saudi 16S gut microbiota for individuals with: (a) < 126 mg/dL and (b) >126 mg/dL (OTU abundance based on BrayCurtis dissimilarity).

Fig S7: Alpha diversity 16S gut microbiota assessment in Saudi males and females using: Chao1, ACE, Shannon-Weaver, Simpson, Inverse Simpson and Fisher indices.

Fig S8: Shannon and Simpson Alpha diversity: (a) T2D versus (b) non-T2D status.

Fig S9: *Bacteroidetes-Firmicutes* ratio in Saudi non-T2D cases and controls using 16S gut microbiota data.

Table S1: The most divergent microbiota genus between Saudi T2D cases and controls (a) and between T2D cases with high (> 126 mg/ dL) and low (< 126 mg/ dL) glucose (b). Positive 16S fold change indicates upregulation in diabetics.

Fig 1: Rank abundant curve (a) and permutational multivariate analysis of variance (PERMANOVA) cloud (b) for Saudi T2Ds and Control 16S stool microbiota datasets. This figure shows the rank abundant curve and Permutational Multivariate Analysis of Variance (PERMANOVA) cloud respectively for Saudi T2D and control 16S stool microbiota datasets. These show that the microbiome communities differ globally between T2D and non-T2D subjects at statistical significance, $p = 0.01$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.25.465666; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a

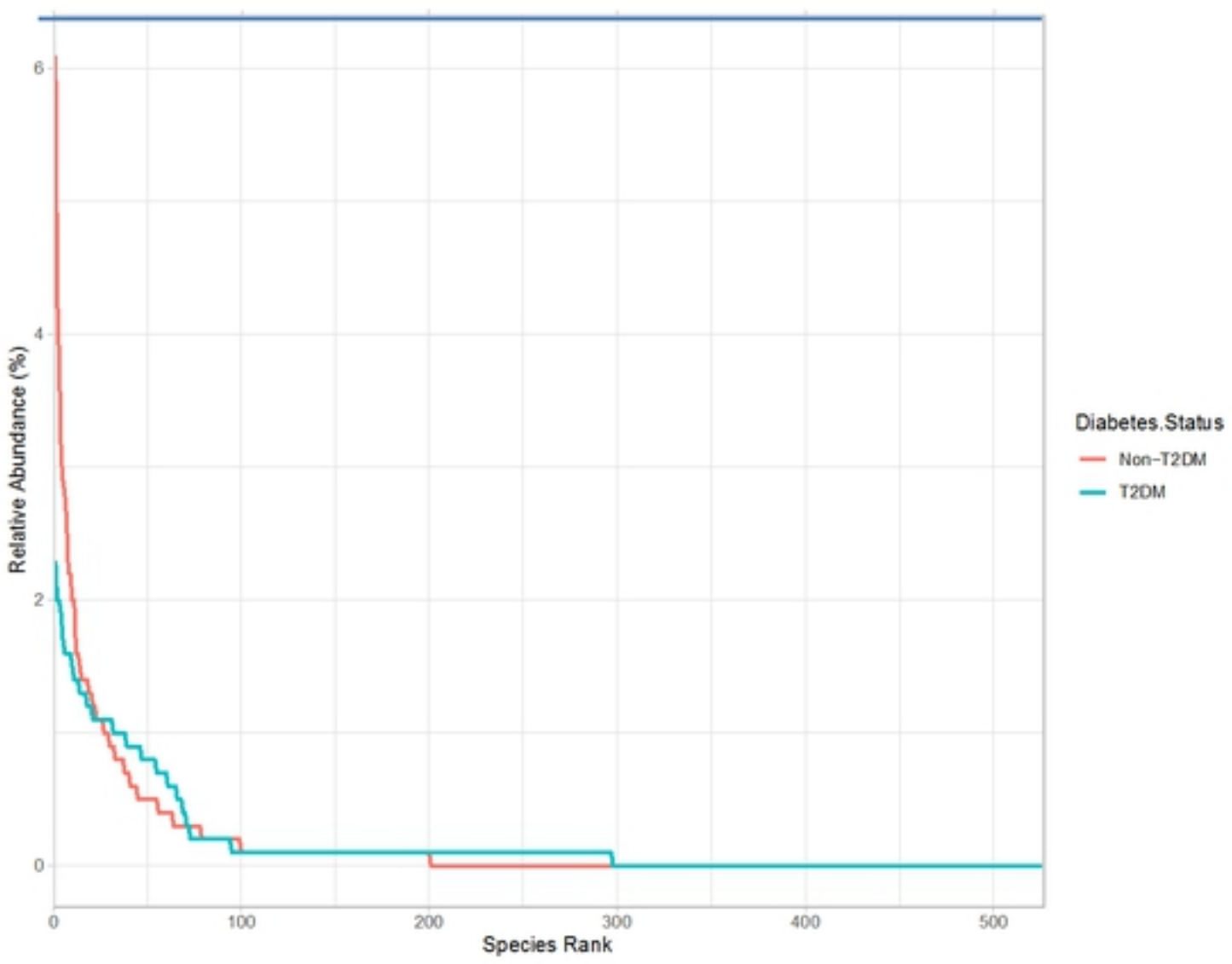

Figure 1a

b

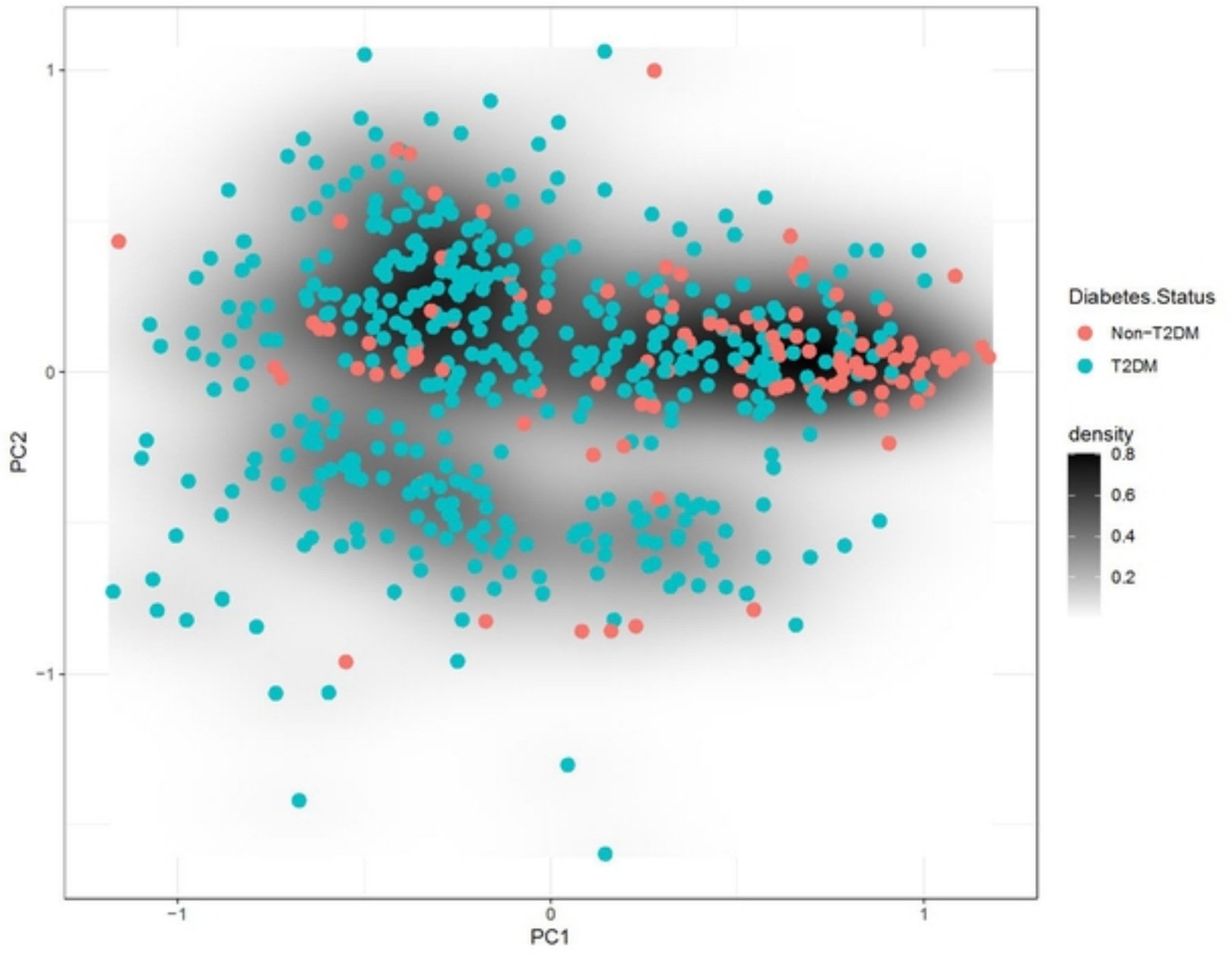
1
0
−1
PC2
−1
0
1
PC1
Diabetes.Status
Non−T2DM
T2DM
density
0.8
0.6
0.4
0.2

Figure 1b

Fig 2: Fold change plots of enriched OTUs for: T2D vs controls (a) and glucose levels for high vs low T2D status (b). An overall positive enrichment of microbiota genus/families for diabetics compared to healthy individuals and amongst diabetic participants was observed. Those with high glucose levels exhibited slightly more positive enrichment compared to those at lower risk of fasting hyperglycemia.

a

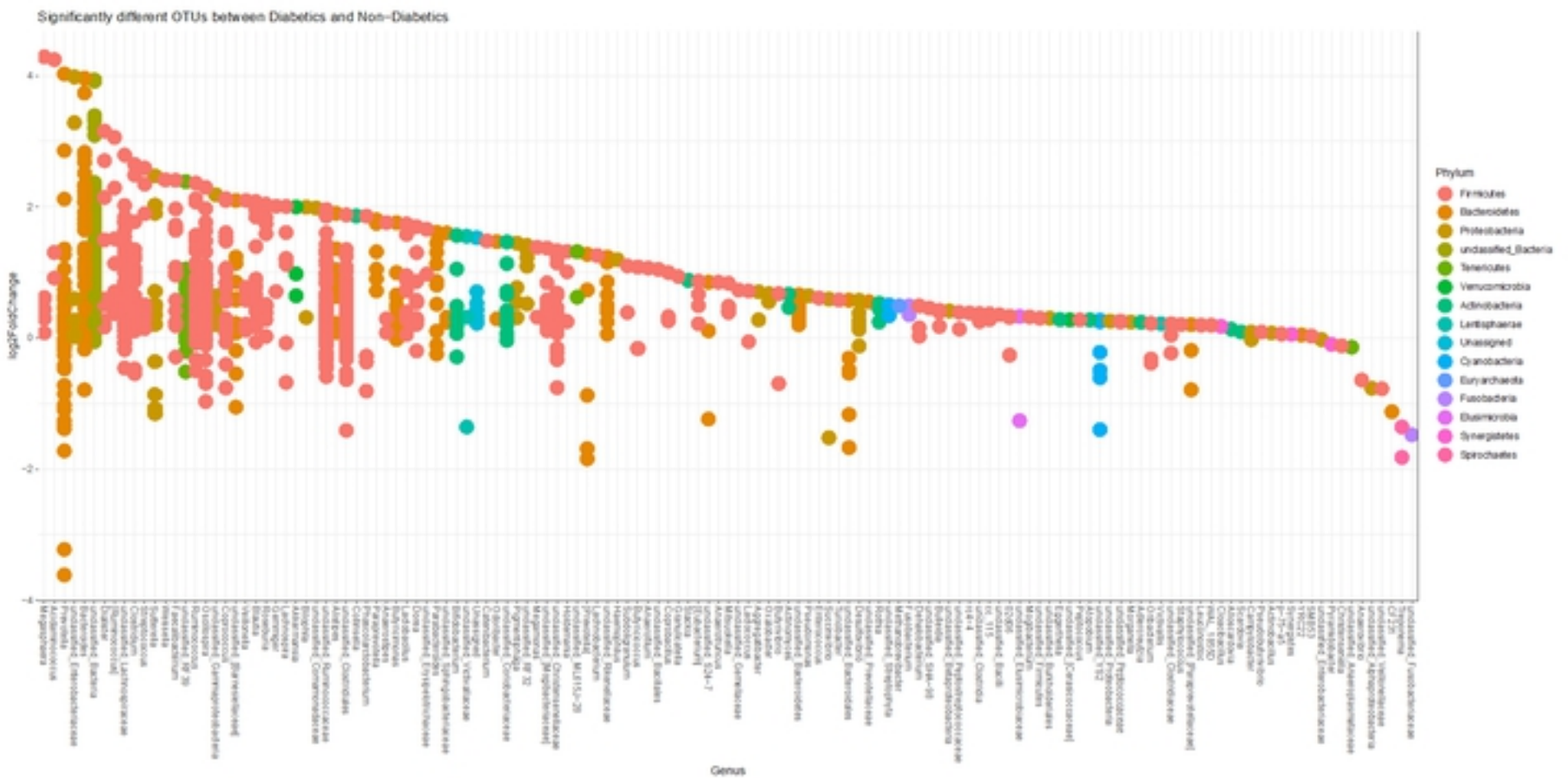

Figure 2a

b

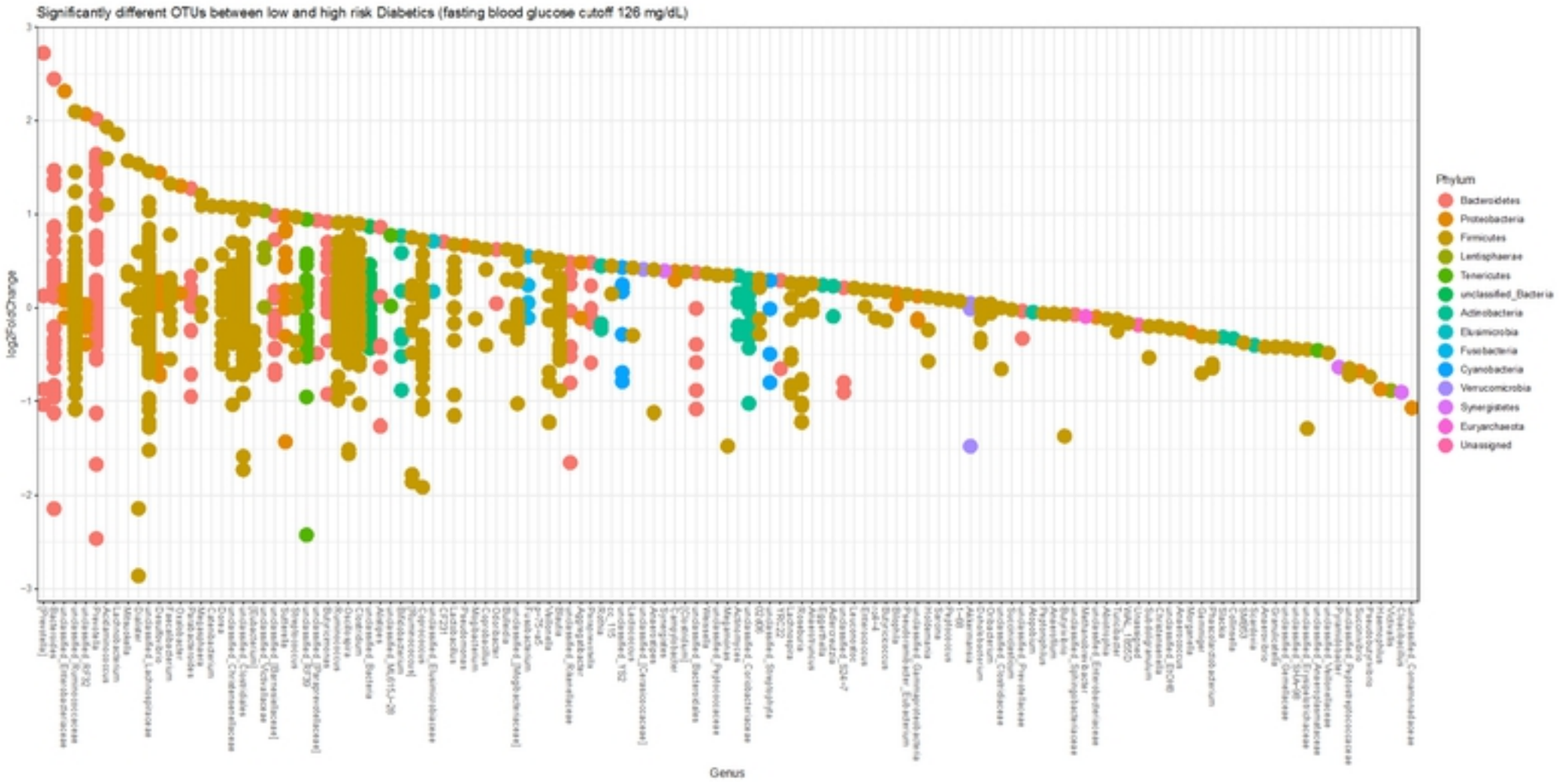
Significantly different OTUs between low and high risk Diabetics (fasting blood glucose cutoff 126 mg/dL)
log2FoldChange
Genus
Phylum
Bacteroidetes
Proteobacteria
Firmicutes
Lentisphaerae
Tenericutes
unclassified_Bacteria
Actinobacteria
Elusimicrobia
Fusobacteria
Cyanobacteria
Verrucomicrobia
Synergistetes
Euryarchaeota
Unassigned

Figure 2b