



Variation of butyrate production in the gut microbiome in type 2 diabetes patients

Julienne Siptroth¹ · Olga Moskalenko² · Carsten Krumbiegel² · Jörg Ackermann³ · Ina Koch³ · Heike Pospisil¹

Received: 22 September 2022 / Revised: 22 December 2022 / Accepted: 4 January 2023 / Published online: 13 February 2023
© The Author(s) 2023

Abstract

Background Diabetes mellitus type 2 is a common disease that poses a challenge to the healthcare system. The disease is very often diagnosed late. A better understanding of the relationship between the gut microbiome and type 2 diabetes can support early detection and form an approach for therapies. Microbiome analysis offers a potential opportunity to find markers for this disease. Next-generation sequencing methods can be used to identify the bacteria present in the stool sample and to generate a microbiome profile through an analysis pipeline. Statistical analysis, e.g., using Student's t-test, allows the identification of significant differences. The investigations are not only focused on single bacteria, but on the determination of a comprehensive profile. Also, the consideration of the functional microbiome is included in the analyses. The dataset is not from a clinical survey, but very extensive.

Results By examining 946 microbiome profiles of diabetes mellitus type 2 sufferers (272) and healthy control persons (674), a large number of significant genera (25) are revealed. It is possible to identify a large profile for type 2 diabetes disease. Furthermore, it is shown that the diversity of bacteria per taxonomic level in the group of persons with diabetes mellitus type 2 is significantly reduced compared to a healthy control group. In addition, six pathways are determined to be significant for type 2 diabetes describing the fermentation to butyrate. These parameters tend to have high potential for disease detection.

Conclusions With this investigation of the gut microbiome of persons with diabetes type 2 disease, we present significant bacteria and pathways characteristic of this disease.

Keywords Type 2 diabetes · Gut · Butyrate · NGS · Statistical analysis

Background

Diabetes mellitus is one of the most prevalent diseases worldwide. In 2019, the World Health Organization (WHO) included diabetes mellitus in the top 10 leading causes of death (WHO 2020). Diabetes mellitus is a metabolic disease, with some hereditary predispositions. It is divided into two main variants, type 1 and type 2. Type 2 diabetes mellitus (T2D) is the most common form and accounts for over 90% of all cases. In this disease, the pancreas still provides sufficient insulin, but this can be increasingly poorly processed by the cells until a complete insulin resistance is developed (Zaccardi et al. 2016). Lack of

exercise, diet, and obesity are regarded risk factors (Fletcher et al. 2002). In the past, this disease was considered to be a consequence of old age, but nowadays this disease is also increasingly occurring in children and youth. Diabetes mellitus type 2 disease can lead to the development of cardiovascular diseases, diabetic foot and damage to the kidneys and eyes, and can even lead to death (Cannon et al. 2018). Studies have shown that the disease can be pushed back or even defeated completely by changing the way of life (Lean et al. 2018).

The intestinal microbiota is a complex structure of bacteria, fungi and virus. Bacteria account for the largest share, with a total of about 100 trillion bacteria living in human intestines. In addition to digestion, the intestine is involved in many other processes and is also called the control center of the body. For example, it supports the immune system and controls inflammatory processes, among other things. The intestinal microbiome is not only influenced by nutrition, but also by many other factors,

✉ Julienne Siptroth
siptroth@th-wildau.de

such as lifestyle, environment, age and gender. Thus, the composition of the intestinal microbiota varies depending on these factors. With a balanced diet, the diversity of bacteria in the intestine increases, and this leads to a wide formation of various metabolic products (Lozupone et al. 2012; Shreiner et al. 2015). In this context, the short-chain fatty acids (SCFA) are of particular importance (Valdes et al. 2018). An important representative is butyrate, which is produced by a variety of bacteria. Butyrate producers make up about 20% of the total bacterial community. Most frequently occurring representatives are members of the families *Lachnospiraceae* and *Ruminococcaceae* (Vital et al. 2017). Butyrate is formed during the fermentation of carbohydrates via various pathways (pyruvate, succinate, etc.) (Vital et al. 2014). In this process, it represents one of the most important sources of energy for intestinal cells. In addition, butyrate is considered to have a health-promoting effect. It not only ensures the functionality of intestinal cells, but also strengthens the intestinal barrier, regulates immune functions and controls metabolic processes. A reduction of butyrate-producing bacteria is associated with the development of diseases (Mallott and Amato 2022).

With the advent of research into the microbiome, links between gut bacteria and type 2 diabetes mellitus have also emerged. Although the question remains whether causality can be inferred between the correlations found. However, increased research in the field is leading to a better understanding. Existing correlations can be confirmed and deepened, and new ones can form the basis for further studies. Thus, consequences and causes of diseases and changes in the intestinal microbiome can be better understood and possible therapeutic approaches can be developed (Arora and Tremaroli 2021; Li et al. 2020; Sharma and Tripathi 2019; Gurung et al. 2020).

Materials and methods

Data

The available data sets consist of microbiome and individual lifestyle data. These are not from a clinical survey with medical supervision, but are extensive due to the low-threshold manner of the survey. More than 29,000 samples were available, each associated with individual lifestyle information. The data (only with indication of consent for scientific use) is provided by the company BIOMES NGS GmbH as a project partner within the scope of its business-like activity. BIOMES NGS GmbH offers a self-test for the analysis of the intestinal flora. Both the sampling and the answering of the questions about the individual lifestyle were performed by the customers themselves. It is a lifestyle

product without any further verification of the customer's information by a medical doctor.

The data on individual lifestyle includes information on age, center of life, height and body weight, as well as information on diet, diseases and medication intake. A total of 99 fields were covered.

The microbiome profile was composed of normalized counts per taxonomic level (kingdom, phylum, class, order, family, genus, species) from sequenced bacterial 16S ribosomal DNA (rDNA). Microbial DNA was analyzed using next-generation sequencing (NGS). This has advantages over classical gut analysis methods because the analysis is more accurate and thus the entirety of the intestinal microbiome can be determined. The 16S rDNA, a gene of the microorganisms, is analyzed.

One subset consisted of clients, who have a diabetes type 2 disease with an age between 18 and 80.

A total of 272 samples met the inclusion criteria, of which 143 were female and 127 were male, and two samples had no gender information.

The diabetes type 2 subgroup was compared with a healthy reference group. Samples for the healthy reference were selected based on the listed parameters.

Inclusion criteria (healthy):

- Age 18-80
- No diseases
- No gastrointestinal complains
- No allergies gluten intolerance
- No medication
- No antibiotic intake <3 months
- No probiotic intake <3 months
- BMI 18.5 to 27.5
- Alcohol intake not daily
- Good wellbeing (> = 5) & health score (>= 6 out of 10)

These criteria resulted in a group size of 674 healthy individuals. Of these, 340 were female, 318 were male, 1 are socially diverse, and 15 samples without gender information. Table 1 shows the distribution of the selected parameters age, gender, BMI, nutrition and sports for the two groups Healthy and T2D. There was no significant difference in nutrition between the two groups. The T2D group did slightly less sports than the Healthy group. This is consistent for the development of diabetes mellitus type 2 disease. For the investigated data set, further adjustment of the parameters in the two groups was not possible, as otherwise unacceptable group sizes would have resulted. The parameters age and BMI are known risk factors and crucial for the development of diabetes type 2. Thus, higher age and higher BMI are characteristic.

Table 1 Distribution of the selected parameters age, gender, BMI, nutrition and sports for the two groups Healthy and T2D

Parameter	Healthy group (n=674)	T2D group (n=272)
Age, years	42,55 ± 12,12	59,71 ± 12,27
Women/Men/Divers, n	340/318/1	143/127/0
BMI, kg/m ²	23,13 ± 2,24	31.05 ± 6.38
Nutrition		
Omnivore	517	245
Vegan	40	2
Vegetarian	74	10
Pescetarian	43	8
Sports		
Daily	22	3
5–6 times per week	61	5
3–4 times per week	199	31
1–2 times per week	172	28
1 time per week	52	17
1 time per 2 weeks	18	7
Less 2 months	2	5
Never	132	107

Methods

Sample preparation and sequencing

Sample storage and lysis

Collected stool samples are stored in 1000 µL DNA-stabilizing buffer at −20 °C until use. For the lysis process, the samples are defrosted and will be centrifuged. Afterwards, warmed up PW buffer from the QIAamp 96 PowerFecal QIAcube HT Kit is added to each sample.

Extraction of stool samples

For the extraction we established the QIAamp 96 PowerFecal QIAcube HT Kit on our liquid handling systems (Hamilton StarLine & Tecan EVO) by using a vacuum chamber as well as a high-pressure chamber. The extracted gDNA is stored at −20 °C until use.

Library preparation for sequencing with the Illumina MiSeq System

The library preparation follows the manual “16S Metagenomic Sequencing Library Preparation- Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq

System”. The mastermix reagents for the target and library amplification are from New England BioLabs, 16S V3V4 primer from Eurofins. For normalization of all samples, a fluorescent dye, and the Biotek Synergy HTX plate reader are used to measure DNA concentrations and to calculate the necessary dilution volume per sample. To ensure a high throughput, all the steps, from the first amplification to the library pooling, are nearly fully automated by using the liquid handling systems (Hamilton StarLine). Hence, we can process between 96 and 192 samples simultaneously and the normalization also works for up to 288 samples. The Library Denaturing and MiSeq Sample Loading is carried out manually.

Processing sequence reads

The determined paired-end reads were filtered in the following. First, the forward/reverse reads were merged using PANDAseq (Masella et al. 2012). This was followed by an alignment using BLASTn (Altschul et al. 1990) against the SILVA rRNA database (version: 138.1) (Quast et al. 2013). Afterwards, filtering was performed. There must be at least 10,000 assigned reads for each sample, for further analysis. Different identity thresholds per taxonomic boundaries (phylum: 75.0%, class: 78.5%, order: 82.0%, family: 86.5%, genus: 94.5%, species: 97.0%) (Yarza et al. 2014) were used. The sequences were clustered according to their similarity (97%) using CD-HIT (Li and Godzik 2006; Fu et al. 2012). Biologically normalized abundances were calculated from the clustered reference sequences using the PICRUSt2 pipeline (Douglas et al. 2020). The output is a table of biologically normalized counts per taxonomic level. The PICRUSt2 pipeline was used to determine the available pathways (MetaCyc Caspi et al. 2016) for each sample applying a predictive model. The abundances of the identified pathways were calculated based on the gene families.

The alpha diversity measures Shannon entropy (Shannon 1948) and inverse Simpson correlation (Simpson 1949) were determined from the rarefied raw counts. The mathematically normalized values were calculated using QIIME2 (Bolyen et al. 2019). Shannon entropy is a measure of diversity and includes both the number of different species and the number of individuals per species. The Shannon index originated in information theory, as a measure of the information content of a message. The inverse Simpson index is also a measurement for describing diversity and, like the Shannon index, belongs to the alpha diversity measures. It is based on the Simpson index. This indicates the probability that two randomly selected individuals belong to different species. The inverse Simpson index is the reciprocal of the Simpson index.

Table 2 Diversity: number of bacteria (at least one relative count) per taxonomic level for the groups Healthy and T2D and in the total of both groups

Group	Phylum	Class	Order	Family	Genus	Species
Total	75	224	490	1045	2393	4326
Healthy	75	217	460	966	2140	3690
T2D	67	184	409	874	1911	3275

Further analysis steps were performed with custom Python scripts using the pandas (McKinney 2010; The pandas development team 2020), NumPy (Harris et al. 2020), scikit-learn (Pedregosa et al. 2011) and SciPy (Virtanen et al. 2020) libraries.

Statistical analysis

Various statistical tests were applied to detect significant differences between the two groups T2D and Healthy. To determine whether the numerical values of microbiome and lifestyle data follow a normal distribution, a test based on D'Agostino and Pearson (D'Agostino 1971; D'Agostino and Pearson 1973) was used. A Student's t-test (Student 1908) was applied to normally distributed samples, and a Mann-Whitney U test (Wilcoxon 1945) was applied to non-normally distributed samples. To determine correlations of individual lifestyle categorical data, the Chi-Square Test of Independence was used. A p-value of > 0.05 was considered significant for all tests used.

Results

Diversity of the intestinal microbiome

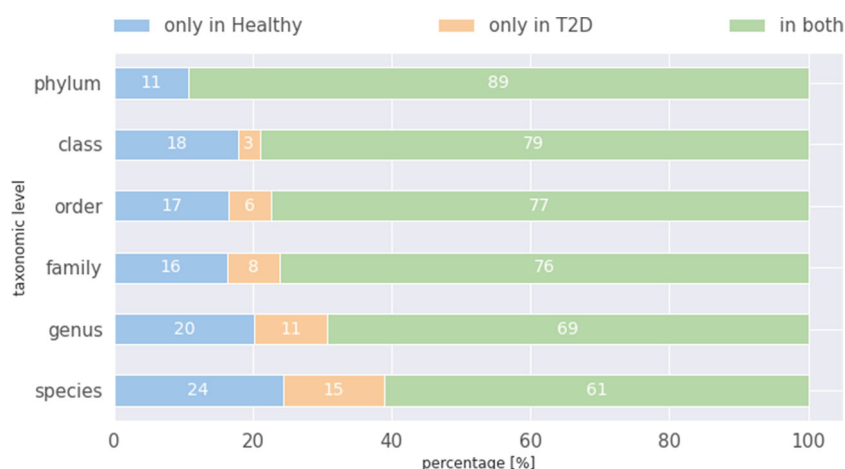
To identify which bacteria/taxa are significant to each group, diversity was examined first. This was done by calculating

the number of bacteria per taxonomic level that had at least one relative count. The quantity of different bacteria per subgroup for each taxonomic level is listed in Table 2. Occurrences over the entire group were also included. It can be seen that the values in the T2D group were about 10% lower than in the Healthy control group across all taxonomic levels.

The uniqueness of bacteria in those groups was investigated afterwards. These percentage values are shown in Fig. 1. The majority of bacteria occur in both groups. In addition, more bacteria occur only in the group Healthy. To determine whether the unique bacteria are characteristic for diabetes, they were tested for significance. For this purpose, different hypothesis tests were applied, Student's t-test for normally distributed samples and Mann-Whitney U test for non-normally distributed samples. A p-value of > 0.05 was assumed as significant. None of the unique bacteria were found to be characteristic of either group.

A hypothesis test was performed for bacteria present in both groups to check for significance. For normally distributed samples, a Student's t-test was applied, and Mann-Whitney U test was used for samples not following a Gaussian distribution. A p-value of 0.05 was assumed as significance threshold. Of 2393 different genera, 25 were identified as significant. Of these, the count values for 4 genera (*Bacteroides*, *Blautia*, *Lachnoclostridium*, and *Prevotella*) were increased in the T2D group in

Fig. 1 Ratio of bacteria present only in the Healthy group (blue), those present only in the T2D group (orange), and in both groups (green), in percent for each taxonomic level



comparison to the Healthy group. The other 19 genera (*Faecalibacterium*, *Lachnospira*, *Roseburia*, *Ruminococcus*, etc.) showed decreased count values in T2D compared to Healthy. Bacteria that could not be precisely assigned (non-specific, etc.) were not included. Table 3 lists significant bacteria (genus level) with respecting p-value and relative abundance in the T2D group.

The significant genera belong to 9 different families (out of 1045 different families). Thereby, the family of *Lachnospiraceae* was the most represented one (16 genera out of 25), followed by *Ruminococcaceae* (3 out of 25). For the remaining 6 genera, each belongs to different families.

Quantification by alpha diversity

To describe the diversity of the groups, the Shannon entropy and the inverse Simpson correlation were calculated. Figure 2 shows the results as a violin plot for both groups. Both parameters for alpha diversity are significantly increased in the Healthy group compared to the T2D group. This is particularly visible for the Shannon entropy.

Functional microbiome: consideration of the butyrate production pathways

Furthermore, the identified pathways were analyzed for their significance for type 2 diabetes mellitus. Special attention was paid to pathways in which butyrates are produced. The significant pathways are listed in Table 4. These are six pathways in which fermentation to butanoate occurs. High significance was found for three of six pathways (*PWY-5677*, *P163-PWY*, *CENTFERM-PWY*; cf. Table 4). The three other pathways (*P162-PWY*, *PWY-5676*, *PWY-5022*) showed weak significance.

Discussion

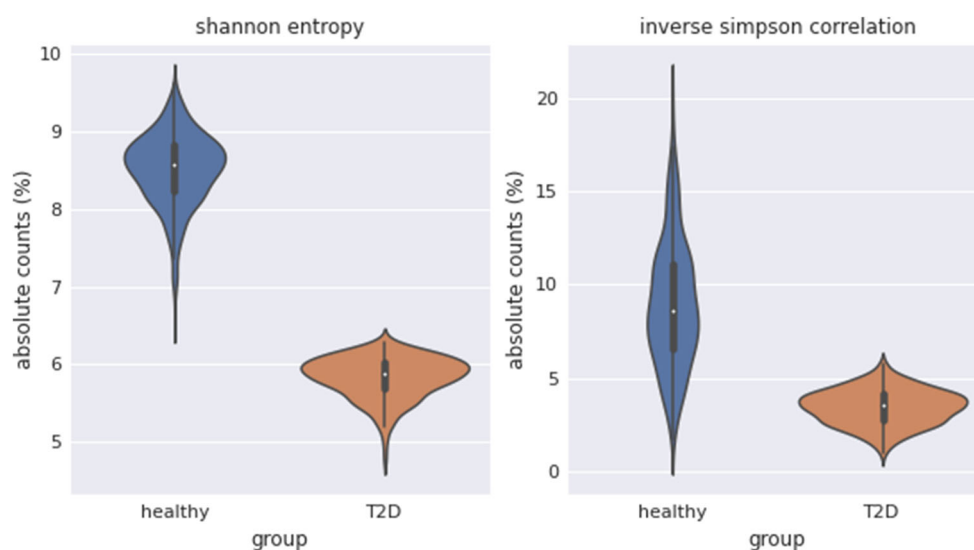
This work analyzed the intestinal microbiome of 272 individuals with diabetes mellitus type 2 disease and 674 healthy control subjects. Significant differences in the alpha diversity measures, Shannon entropy and inverse Simpson correlation, were found between the two groups. It became

Table 3 Bacteria at genus level determined to be significant, with p-value and inclusion of the T2D group relative to the Healthy group (↑ — increases; ↓ — decreases)

Genus	Occurrence in diabetes group	p-value
<i>Lachnospiraceae FCS020 group</i>	↓	2.460e-22
<i>Lachnospiraceae ND3007 group</i>	↓	6.416e-14
<i>Faecalibacterium</i>	↓	8.775e-12
<i>Ruminococcus</i>	↓	7.398e-11
<i>Clostridium sensu stricto 1</i>	↓	3.000e-10*
<i>Lachnospiraceae UCG-001</i>	↓	2.885e-08
<i>Coprococcus</i>	↓	5.091e-07
<i>Subdoligranulum</i>	↓	4.141e-06
<i>Lachnospira</i>	↓	3.639e-06
<i>Lachnospiraceae NC2004 group</i>	↓	2.348e-06
<i>Lachnospiraceae NK4A136 group</i>	↓	2.284e-06
<i>Fusicatenibacter</i>	↓	4.077e-05
<i>Lachnospiraceae UCG-006</i>	↓	2.310e-05
<i>UCG-002</i>	↓	7.378e-03
<i>Roseburia</i>	↓	5.050e-03
<i>Marvinbryantia</i>	↓	3.677e-03
<i>Lachnospiraceae UCG-008</i>	↓	2.881e-03
<i>Agathobacter</i>	↓	2.438e-03*
<i>Alistipes</i>	↓	2.560e-02*
<i>Butyrivibrio</i>	↓	1.747e-02*
<i>Anaerostipes</i>	↓	1.288e-02

*Mann-Whitney U test

Fig. 2 Alpha diversity: violin plot for the group Healthy (blue) and the group T2D (orange) of absolute counts (%) for Shannon entropy (left) and inverse Simpson correlation (right)



clear that the diversity in the diabetes mellitus type 2 group is lower than in the Healthy control group. This was also confirmed by other studies, which show that the microbiome diversity of persons suffering a disease is lower compared to healthy persons (Zhang et al. 2019; Larsen et al. 2010).

Furthermore, the analyses demonstrated that there are bacteria that are only present in one of the two comparison groups. However, the occurrences of these bacteria were not significant for diabetes mellitus type 2 disease. Characteristic bacteria for distinguishing the diabetes microbiome from healthy microbiome were present in both groups. The only significant difference between the genera was in the amount of their occurrence. Four genera (*Bacteroides*, *Blautia*, *Lachnospirillum* and *Prevotella*) showed an increased occurrence in the diabetes mellitus type 2 group than in the Healthy control group. The other significant genera (*Anaerostipes*, *Coprococcus*, *Fusicatenibacter*, *Lachnospira*, *Marvinbryantia*, *Roseburia*, *Faecalibacterium*, *Ruminococcus*, *Subdoligranulum*, *UCG-002*, *Agathobacter*, *Butyricoccus*, *Alistipes*, *Clostridium sensu stricto 1* and all *Lachnospiraceae*) showed the opposite.

The genus *Blautia* was one of the most represented genera in both groups enriched in group T2D. Previous

studies confirm this behavior (Egshatyan et al. 2015; Zhang et al. 2013).

Lachnospirillum were more abundant in the microbiome profiles of diabetic patients. This association has not yet been verified in other studies. However, there are studies in relation to other diseases (Kang et al. 2021). Further research is necessary to identify *Lachnospirillum* as a marker for type 2 diabetes mellitus. There are anomalies in the genera *Bacteroides* and *Prevotella*, both of which are increased in the T2D group, compared to the healthy group. Although similar behavior was detected in other studies, both genera are propionate producers. Propionate, like butyrate, belongs to the short-chain fatty acids. These are associated with a healthy status of the microbiome (Wu et al. 2010; Candela et al. 2016).

A significant reduction of genera *Alistipes*, *Anaerostipes*, *Ruminococcus* was detected in the T2D group. This decrease was associated with diabetes mellitus type 2 and was supported by previous studies. *Ruminococcus* and *Anaerostipes* were associated with a healthy state of the microbiome and showed increased values in the Healthy control group (Gao et al. 2018; Doumatey et al. 2020). Furthermore, various studies also indicated a decrease in genera in the group T2D (Zhang et al. 2019; Gaike et al.

Table 4 Pathways with butyrate as end product

BioCyc ID	Pathways	p-value
PWY-5677	Succinate fermentation to butanoate	1.570e-12
P163-PWY	L-Lysine fermentation to acetate and butanoate	4.804e-10
CENTFERM-PWY	Pyruvate fermentation to butanoate	2.312e-05
P162-PWY	L-Glutamate degradation V (via hydroxyglutarate)	6.015e-02
PWY-5676	Acetyl-CoA fermentation to butanoate	8.599e-02
PWY-5022	4-Aminobutanoate degradation V	2.278e-01

2020; Liu et al. 2020; Salamon et al. 2018; Das et al. 2021).

A decrease of *Alistipes* counts was also identified in the data of group T2D (Thingholm et al. 2019).

Additionally, the analysis pointed out a reduction in the *Lachnospira*, *Roseburia*, *Faecalibacterium* and *Coprococcus* genera in the T2D group compared to the healthy group. These genera were butyrate producers and were associated with healthy gut flora. The decrease of these genera favored obesity and the development of diseases. This fact and other studies supported the assumption that these were genera associated with diabetes mellitus type 2 disease. In these studies, a reduction was also detected in the T2D group (Zhang et al. 2019; Larsen et al. 2010; Kang et al. 2021; Candela et al. 2016; Gao et al. 2018; Liu et al. 2020; Salamon et al. 2018; Das et al. 2021; Anand et al. 2016).

The genus *Subdoligranulum* is closely related to the genera *Faecalibacterium*, both from the family *Ruminococcaceae*. Like *Faecalibacterium*, *Subdoligranulum* is also a butyrate producer. Furthermore, this genus was associated with a healthy metabolic status, but the exact physiological role is yet unknown (Van Hul et al. 2020). The results (cf. Table 3) show that the occurrence of *Subdoligranulum* is reduced in the group with type 2 diabetes mellitus. The reduction of *Subdoligranulum* can also be assumed to have a negative impact and is thus characteristic of type 2 diabetes mellitus disease.

In four other genera (*Fusicatenibacter*, *Agathobacter*, *Butyricoccus*, *Marvinbryantia*), a decrease in the T2D group could also be detected. Those are currently not associated with type 2 diabetes mellitus. All genera produce short-chain fatty acids, e.g., butyrates, and are associated with a healthy intestinal flora (Kang et al. 2021; Lu et al. 2019; Ma et al. 2020; Chen et al. 2021). The reduction of genera leads to a lower proportion of short-chain fatty acids in the organism. This favors obesity and the development of diseases. Thus, these genera may be elements of a profile for the detection of type 2 diabetes mellitus. The exact relationship between the genera *Fusicatenibacter*, *Agathobacter*, *Butyricoccus* and *Marvinbryantia* and type 2 diabetes mellitus disease also needs to be investigated in further trials.

The consideration of the functional microbiome has received greater attention in recent years. It is assumed that not only individual bacteria/taxa, but also their interactions are decisive. So characteristic profiles, e.g., diseases and lifestyle, can be found. Thus, pathway analyses are becoming more and more important. In this work, pathways were investigated in which fermentation to butanoate takes place (Qin et al. 2012; Reichardt et al. 2014). Six different pathways (*PWY-5677*, *P163-PWY*, *CENTFERM-PWY*, *P162-PWY*, *PWY-5676*, *PWY-5022*) could be determined. Through significance analysis,

these were examined with respect to their relevance for the group T2D. Furthermore, significant, determined taxa (e.g., *Lachnospira*, *Roseburia*, *Faecalibacterium*, and *Coprococcus*) are also identified as butyrate producers. These facts support the assumption that the functional microbiome has a high importance for the analysis of the microbiome. Further research must verify the potential found for associations with diabetes mellitus type 2 disease or even other diseases or lifestyles. For this purpose, additional pathways need to be examined. These may also be approaches for early detection and therapies (Arora and Tremaroli 2021).

Abbreviations BMI, body mass index; NGS, next-generation sequencing; rDNA, ribosomal deoxyribonucleic acid; SCFA, short-chain fatty acids; T2D, type 2 diabetes; WHO, World Health Organization.

Acknowledgements Thanks to the laboratory team from BIOMES NGS GmbH, especially Philipp Franke. Further thanks go to Christian Rockmann and Prof. Dr. Marcus Frohme for the project acquisition.

Author contribution J.S. and H.P.: conceptualization. J.S.: statistical analysis, writing — original draft. O.M. and C.K.: providing the data (sample preparation and sequencing; processing sequence reads). I.K. and J.A.: writing — review and editing

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Federal Ministry of Education and Research of Germany under grant number 13FH209PX8.

Availability of data and materials The data used are not publicly available because the research project was carried out in collaboration with the company BIOMES NGS GmbH. They can be requested with a reasoned request to the corresponding author.

Declarations

Ethics approval and consent to participate The data are provided by the company BIOMES NGS GmbH. Customers pay for the microbiome test and give their approval for further scientific use in the purchase process. Only data with granted consent was used.

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Anand S, Kaur H, Mande SS (2016) Comparative in silico analysis of butyrate production pathways in gut commensals and pathogens. *Front Microbiol* 7:1945. <https://doi.org/10.3389/fmicb.2016.01945>
- Arora T, Tremaroli V (2021) Therapeutic potential of butyrate for treatment of type 2 diabetes. *Front Endocrinol* 12:761834. <https://doi.org/10.3389/fendo.2021.761834>
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciulek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37(8):852–857. <https://doi.org/10.1038/s41587-019-0209-9>. Number: 8 Publisher: Nature Publishing Group
- Candela M, Biagi E, Soverini M, Consolandi C, Quercia S, Severgnini M, Peano C, Turroni S, Rampelli S, Pozzilli P, Pianesi M, Fallucca F, Brigidi P (2016) Modulation of gut microbiota dysbiosis in type 2 diabetic patients by macrobiotic Ma-Pi 2 diet. *Br J Nutr* 116(1):80–93. <https://doi.org/10.1017/S0007114516001045>
- Cannon A, Handelsman Y, Heile M, Shannon M (2018) Burden of illness in type 2 diabetes Mellitus. *J Manage Care Specialty Pharm* 24(9-a Suppl):5–13. <https://doi.org/10.18553/jmcp.2018.24.9-a.s5>. Publisher: Academy of Managed Care Pharmacy
- Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Subhraveti P, Weaver DS, Karp PD (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 44(Database issue):471–480. <https://doi.org/10.1093/nar/gkv1164>
- Chen Z, Radjabzadeh D, Chen L, Kurilshikov A, Kavousi M, Ahmadizar F, Ikram MA, Uitterlinden AG, Zhernakova A, Fu J, Kraaij R, Voortman T (2021) Association of insulin resistance and type 2 diabetes with gut microbial diversity. *JAMA Netw Open* 4(7):2118811. <https://doi.org/10.1001/jamanetworkopen.2021.18811>
- D'Agostino RB (1971) An omnibus test of normality for moderate and large size samples. *Biometrika* 58(2):341–348. <https://doi.org/10.1093/biomet/58.2.341>
- D'Agostino R, Pearson ES (1973) Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$. *Biometrika* 60(3):613–622. <https://doi.org/10.2307/2335012>. Publisher: [Oxford University Press, Biometrika Trust]
- Das T, Jayasudha R, Chakravarthy S, Prashanthi GS, Bhargava A, Tyagi M, Rani PK, Pappuru RR, Sharma S, Shivaji S (2021) Alterations in the gut bacterial microbiome in people with type 2 diabetes mellitus and diabetic retinopathy. *Sci Rep* 11(1):2738. <https://doi.org/10.1038/s41598-021-82538-0>. Number: 1 Publisher: Nature Publishing Group
- Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI (2020) PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 38(6):685–688. <https://doi.org/10.1038/s41587-020-0548-6>
- Doumatey AP, Adeyemo A, Zhou J, Lei L, Adebamowo SN, Adebamowo C, Rotimi CN (2020) Gut microbiome profiles are associated with type 2 diabetes in urban Africans. *Front Cell Infect Microbiol* 10
- Egshatyan L, Kashtanova D, Popenko A, Tkacheva O, Tyakht A, Alexeev D, Karamnova N, Kostyukova E, Babenko V, Vakhitova M, Boytsov S (2015) Gut microbiota and diet in patients with different glucose tolerance. *Endocr Connect* 5(1):1–9. <https://doi.org/10.1530/EC-15-0094>
- Fletcher B, Gulanick M, Lamendola C (2002) Risk factors for type 2 diabetes mellitus. *J Cardiovasc Nurs* 16(2):17–23
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23):3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Gaike AH, Paul D, Bhute S, Dhotre DP, Pande P, Upadhyaya S, Reddy Y, Sampath R, Ghosh D, Chandraprabha D, Acharya J, Banerjee G, Sinkar VP, Ghaskadbi SS, Shouche YS (2020) The gut microbial diversity of newly diagnosed diabetics but not of prediabetics is significantly different from that of healthy nondiabetics. *mSystems* 5(2):00578–19. <https://doi.org/10.1128/mSystems.00578-19>. Publisher: American Society for Microbiology
- Gao R, Zhu C, Li H, Yin M, Pan C, Huang L, Kong C, Wang X, Zhang Y, Qu S, Qin H (2018) Dysbiosis signatures of gut microbiota along the sequence from healthy, young patients to those with overweight and obesity. *Obesity* 26(2):351–361. <https://doi.org/10.1002/oby.22088>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/oby.22088>
- Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, Shulzhenko N (2020) Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine* 51:102590. <https://doi.org/10.1016/j.ebiom.2019.11.051>
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE (2020) Array programming with NumPy. *Nature* 585(7825):357–362. <https://doi.org/10.1038/s41586-020-2649-2>. Number: 7825 Publisher: Nature Publishing Group
- Kang L, Li P, Wang D, Wang T, Hao D, Qu X (2021) Alterations in intestinal microbiota diversity, composition, and function in patients with sarcopenia. *Sci Rep* 11:4628. <https://doi.org/10.1038/s41598-021-84031-0>
- Larsen N, Vogensen FK, van den Berg FWJ, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sørensen SJ, Hansen LH, Jakobsen M (2010) Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* 5(2):9085. <https://doi.org/10.1371/journal.pone.0009085>
- Lean ME, Leslie WS, Barnes AC, Brosnahan N, Thom G, McCombie L, Peters C, Zhyzhneuskaya S, Al-Mrabeh A, Hollingsworth KG, Rodrigues AM, Rehackova L, Adamson AJ, Snihotta FF, Mathers JC, Ross HM, McIlvenna Y, Stefanetti R, Trenell M, Welsh P, Kean S, Ford I, McConnachie A, Sattar N, Taylor R (2018) Primary care-led weight management

- for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. *Lancet* 391(10120):541–551. [https://doi.org/10.1016/S0140-6736\(17\)33102-1](https://doi.org/10.1016/S0140-6736(17)33102-1)
- Li W, Godzik A (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* (Oxford, England) 22(13):1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Li W-Z, Stirling K, Yang J-J, Zhang L (2020) Gut microbiota and diabetes: From correlation to causality and mechanism. *World J Diabetes* 11(7):293–308. <https://doi.org/10.4239/wjd.v11.i7.293>
- Liu C, Shao W, Gao M, Liu J, Guo Q, Jin J, Meng F (2020) Changes in intestinal flora in patients with type 2 diabetes on a low-fat diet during 6 months of follow-up. *Exp Ther Med* 20(5):40. <https://doi.org/10.3892/etm.2020.9167>
- Lozupone C, Stombaugh JI, Gordon JI, Jansson JK, Knight R (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489(7415):220–230. <https://doi.org/10.1038/nature11550>
- Lu H-F, Ren Z-G, Li A, Zhang H, Xu S-Y, Jiang J-W, Zhou L, Ling Q, Wang B-H, Cui G-Y, Chen X-H, Zheng S-S, Li L-J (2019) Fecal microbiome data distinguish liver recipients with normal and abnormal liver function from healthy controls. *Front Microbiol* 10:1518. <https://doi.org/10.3389/fmicb.2019.01518>
- Ma Q, Li Y, Wang J, Li P, Duan Y, Dai H, An Y, Cheng L, Wang T, Wang C, Wang T, Zhao B (2020) Investigation of gut microbiome changes in type 1 diabetic mellitus rats based on high-throughput sequencing. *Biomed Pharmacother* 124:109873. <https://doi.org/10.1016/j.biopha.2020.109873>
- Mallott EK, Amato KR (2022) Butyrate production pathway abundances are similar in human and nonhuman primate gut microbiomes. *Mol Biol Evol* 39(1):279. <https://doi.org/10.1093/molbev/msab279>
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinforma* 13(1):31. <https://doi.org/10.1186/1471-2105-13-31>
- McKinney W (2010) Data structures for statistical computing in Python, Austin, Texas, pp 56–61. <https://doi.org/10.25080/Major-a-92bf1922-00a>, <https://conference.scipy.org/proceedings/scipy2010/mckinney.html>
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12(85):2825–2830
- Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto J-M, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490(7418):55–60. <https://doi.org/10.1038/nature11450>. Number: 7418 Publisher: Nature Publishing Group
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41(D1):590–596. <https://doi.org/10.1093/nar/gks1219>
- Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, Flint HJ, Louis P (2014) Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* 8(6):1323–1335. <https://doi.org/10.1038/ismej.2014.14>
- Salamon D, Sroka-Oleksiak A, Kapusta P, Szopa M, Mrozińska S, Ludwig-Słomczyńska AH, Wołkow PP, Bulanda M, Klupa T, Małcki MT, Gosiewski T (2018) Characteristics of gut microbiota in adult patients with type 1 and type 2 diabetes based on next-generation sequencing of the 16S rRNA gene fragment. *Pol Arch Intern Med* 128(6):336–343. <https://doi.org/10.20452/pamw.4246>
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>. Conference Name: The Bell System Technical Journal
- Sharma S, Tripathi P (2019) Gut microbiome and type 2 diabetes: where we are and where to go? *J Nutr Biochem* 63:101–108. <https://doi.org/10.1016/j.jnutbio.2018.10.003>
- Shreiner AB, Kao JY, Young VB (2015) The gut microbiome in health and in disease. *Curr Opin Gastroenterol* 31(1):69–75. <https://doi.org/10.1097/MOG.0000000000000139>
- Simpson EH (1949) Measurement of diversity. *Nature* 163(4148):688–688. <https://doi.org/10.1038/163688a0>. Number: 4148 Publisher: Nature Publishing Group
- Student (1908) The probable error of a mean. *Biometrika* 6(1):1–25. <https://doi.org/10.2307/2331554>. Publisher: [Oxford University Press, Biometrika Trust]
- The pandas development team (2020) pandas-dev/pandas: Pandas 1.0.3. Zenodo. <https://doi.org/10.5281/zenodo.3715232>
- Thingholm LB, Rühlemann MC, Koch M, Fuqua B, Laucke G, Boehm R, Bang C, Franzosa EA, Hübenthal M, Rahnavard A, Frost F, Lloyd-Price J, Schirmer M, Lusi AJ, Vulpe CD, Lerch MM, Homuth G, Kacprowski T, Schmidt CO, Nöthlings U, Karlsten TH, Lieb W, Laudes M, Franke A, Huttenhower C (2019) Obese individuals with and without type 2 diabetes show different gut microbial functional capacity and composition. *Cell Host Microbe* 26(2):252–264. <https://doi.org/10.1016/j.chom.2019.07.004>
- Valdes AM, Walter J, Segal E, Spector TD (2018) Role of the gut microbiota in nutrition and health. *BMJ* 361:2179. <https://doi.org/10.1136/bmj.k2179>. Publisher: British Medical Journal Publishing Group Section: Analysis
- Van Hul M, Le Roy T, Prifti E, Dao MC, Paquot A, Zucker J-D, Delzenne NM, Muccioli G, Clément K, Cani PD (2020) From correlation to causality: the case of Subdoligranulum. *Gut Microbes* 12(1):1–13. <https://doi.org/10.1080/19490976.2020.1849998>
- Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17(3):261–272. <https://doi.org/10.1038/s41592-019-0686-2>. Number: 3 Publisher: Nature Publishing Group
- Vital M, Howe AC, Tiedje JM (2014) Revealing the bacterial butyrate synthesis pathways by analyzing (Meta)genomic Data. *mBio* 5(2):00889–14. <https://doi.org/10.1128/mBio.00889-14>
- Vital M, Karch A, Pieper DH (2017) Colonic butyrate-producing communities in humans: an overview using omics data. *mSystems* 2(6):00130–17. <https://doi.org/10.1128/mSystems.00130-17>
- Wilcoxon F (1945) Individual comparisons by ranking methods. *Biom Bull* 1(6):80–83. <https://doi.org/10.2307/3001968>. Publisher: [International Biometric Society, Wiley]
- Wu X, Ma C, Han L, Nawaz M, Gao F, Zhang X, Yu P, Zhao C, Li L, Zhou A, Wang J, Moore JE, Cherie Millar B, Xu J (2010) Molecular characterisation of the faecal microbiota in patients with type II diabetes. *Curr Microbiol* 61(1):69–78. <https://doi.org/10.1007/s00284-010-9582-9>

- WHO, WHO (2020) WHO reveals leading causes of death and disability worldwide: 2000-2019. fact sheet. <https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019>
- Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12(9):635–645. <https://doi.org/10.1038/nrmicro3330>. Number: 9 Publisher: Nature Publishing Group
- Zaccardi F, Webb DR, Yates T, Davies MJ (2016) Pathophysiology of type 1 and type 2 diabetes mellitus: a 90-year perspective. *Postgrad Med J* 92(1084):63–69. <https://doi.org/10.1136/postgradmedj-2015-133281>
- Zhang X, Shen D, Fang Z, Jie Z, Qiu X, Zhang C, Chen Y, Ji L (2013) Human gut microbiota changes reveal the progression of glucose intolerance. *PLoS ONE* 8(8):71108. <https://doi.org/10.1371/journal.pone.0071108>
- Zhang F, Wang M, Yang J, Xu Q, Liang C, Chen B, Zhang J, Yang Y, Wang H, Shang Y, Wang Y, Mu X, Zhu D, Zhang C, Yao M, Zhang L (2019) Response of gut microbiota in type 2 diabetes to hypoglycemic agents. *Endocrine* 66(3):485–493. <https://doi.org/10.1007/s12020-019-02041-5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Julienne Siptroth¹ · Olga Moskalenko² · Carsten Krumbiegel² · Jörg Ackermann³ · Ina Koch³ · Heike Pospisil¹

Olga Moskalenko
olga.moskalenko@biomes.world

Carsten Krumbiegel
carsten.krumbiegel@biomes.world

Jörg Ackermann
j.ackermann@bioinformatik.uni-frankfurt.de

Ina Koch
ina.koch@bioinformatik.uni-frankfurt.de

Heike Pospisil
heike.pospisil@th-wildau.de

- ¹ High Performance Computing in Life Sciences, Technical University of Applied Sciences Wildau, Wildau, Germany
- ² BIOMES NGS GmbH, Schwartzkopffstraße 1, 15745 Wildau, Germany
- ³ Department of Molecular Bioinformatics, Institute of Computer Science, Goethe University Frankfurt, 60325 Frankfurt am Main, Germany