

Gut Microbiota Patterns Associated with Colonization of Different *Clostridium difficile* Ribotypes

Jure Skraban^{1,3}, Saso Dzeroski^{3,4,5}, Bernard Zenko^{3,4}, Domen Mongus⁶, Simon Gangl⁶, Maja Rupnik^{1,2,3*}

1 Faculty of Medicine, University of Maribor, Maribor, Slovenia, **2** Centre for Microbiology, Institute of Public Health Maribor, Maribor, Slovenia, **3** Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Ljubljana, Slovenia, **4** Jozef Stefan Institute, Ljubljana, Slovenia, **5** Jozef Stefan International Postgraduate School, Ljubljana, Slovenia, **6** Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

Abstract

C. difficile infection is associated with disturbed gut microbiota and changes in relative frequencies and abundance of individual bacterial taxa have been described. In this study we have analysed bacterial, fungal and archaeal microbiota by denaturing high pressure liquid chromatography (DHPLC) and with machine learning methods in 208 faecal samples from healthy volunteers and in routine samples with requested *C. difficile* testing. The latter were further divided according to stool consistency, *C. difficile* presence or absence and *C. difficile* ribotype (027 or non-027). Lower microbiota diversity was a common trait of all routine samples and not necessarily connected only to *C. difficile* colonisation. Differences between the healthy donors and *C. difficile* positive routine samples were detected in bacterial, fungal and archaeal components. *Bifidobacterium longum* was the single most important species associated with *C. difficile* negative samples. However, by machine learning approaches we have identified patterns of microbiota composition predictive for *C. difficile* colonization. Those patterns also differed between samples with *C. difficile* ribotype 027 and other *C. difficile* ribotypes. The results indicate that not only the presence of a single species/group is important but that certain combinations of gut microbes are associated with *C. difficile* carriage and that some ribotypes (027) might be associated with more disturbed microbiota than the others.

Citation: Skraban J, Dzeroski S, Zenko B, Mongus D, Gangl S, et al. (2013) Gut Microbiota Patterns Associated with Colonization of Different *Clostridium difficile* Ribotypes. PLoS ONE 8(2): e58005. doi:10.1371/journal.pone.0058005

Editor: Michel R. Popoff, Institute Pasteur, France

Received: November 19, 2012; **Accepted:** January 29, 2013; **Published:** February 28, 2013

Copyright: © 2013 Skraban et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Slovenian Research Agency grants 1000-10-310160 and J3-4298 and Operation no. OP13.1.1.2.02.0005 financed by the European Regional Development Fund and the Slovenian Ministry of Higher Education, Science and Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: maja.rupnik@zvv-mb.si

Introduction

Clostridium difficile infection (CDI) is recognised as the leading cause of nosocomial intestinal infections, mainly in elderly hospitalised patients [1]. The onset of CDI has been linked to the use of broad-spectrum antibiotics, historically especially with clindamycin [2,3]. This implicated very early that the healthy gut microbiota has an important role in disease development. The importance of gut microbes is also supported by the high success rate (around 90%) of biotherapy (faecal transplantations) used in patients with severe cases of recurring CDI [4–6].

Several studies on gut microbiota changes associated with *C. difficile* colonization or infection have been published recently. Mice treated with clindamycin lose up to 90% of bacterial taxa in the cecum and are rapidly colonized with *C. difficile* [7]. Other molecular studies on the changes of gut microbiota during CDI have used 16S rRNA clone libraries and showed lower bacterial microbiota diversity in patients with a recurring CDI [8]. The banding patterns on gradient gels connected specific bacterial groups to *C. difficile* colonisation in adults [9] and infants [10]. Another study used 454 sequencing to compare the composition of gut microbiota of *C. difficile* positive (20 asymptomatic and 2 diarrhoeal samples) and negative subjects (252 samples). They found no differences in the composition of faecal microbiota in asymptomatic *C. difficile* carriers and healthy subjects. However, in

C. difficile 027 ribotype positive subjects ($n = 2$) with diarrhoea at the time of sampling faecal microbiota had a lower overall diversity at the genus level [11].

As described, different molecular approaches (including denaturing gradient gel electrophoresis (DGGE), 16S rRNA clone libraries and metagenomics) were used in studies on *C. difficile* and changes in microbiota. Here we have applied a simple and high-throughput DHPLC method (DHPLC - denaturing high pressure liquid chromatography) which separates 16S rDNA amplicons based on fragment size and sequence to analyse human faecal microbiota in the *C. difficile* positive and negative routine samples. It has been shown previously, that DHPLC offers a level of analysis, which is similar to gradient gels (DGGE) in terms of sensitivity and resolution, with the advantage of high automation and repeatability [12,13]. It has been used successfully in community profiling of microbiota of the urogenital tract [14], detecting changes in gut microbiota of patients during antimicrobial therapy [12], following the development of gut microbiota in infants [15], and in profiling various other complex communities from fermenter sludge, compost or soil [13].

Studies on the association of gut microbiota and *C. difficile* mostly describe the bacterial microbiota and very rarely fungi [16], while no studies have included archaea. With DHPLC we have analysed the bacterial, but also the fungal and archaeal gut microbiota in human faecal samples with requested *C. difficile*

testing. With machine learning analysis of the collected data, we were able to show that certain patterns of microbiota composition are associated with *C. difficile* colonization and in particular with the PCR ribotype 027.

Materials and Methods

Collecting and storing of the faecal samples

Faecal samples with requested *C. difficile* testing (n = 171) were obtained from two different routine laboratories. These samples were routinely collected in plastic containers and have been stored at 4°C and shipped to the study site within 48 h, where they have been immediately frozen at -80°C until further processing.

According to the results of the routine *C. difficile* testing, 105 faecal samples were *C. difficile* positive and 66 were *C. difficile* negative. A stool sample was described as non-diarrhoeal or diarrhoeal (has taken the shape of the container). No further clinical information was obtained for the samples.

Additional 37 samples were collected from healthy donors and stored as described above. None of the samples from healthy donors was diarrhoeic. The age of the donors ranged from 14 to 72 years. The gender distribution was 22 females and 15 males.

Ethics Statement

The National Medical Ethics Committee of the Republic of Slovenia (NMEC) has approved the study (number 126/05/12). NMEC has also approved that written or verbal consent is not needed for routine samples and that only verbal informed consent is needed for healthy volunteers. Documenting was not requested. In the case of minor/children participants informed consent was obtained from the parents.

Detection, isolation and ribotyping of *C. difficile* from faecal samples

In each of the two routine laboratories, *C. difficile* was detected by different methods. One laboratory performed the VIDAS ToxA/ToxB test and culturing on CLO selective plates (BioMérieux) after ethanol shock. The second laboratory used the molecular Cepheid Xpert *C. difficile* assay. From the samples tagged as *C. difficile* positive by the Cepheid Xpert *C. difficile* assay, *C. difficile* was isolated by enrichment in Oxoid *C. difficile* broth with Oxoid supplement SR0096E, lysozyme (5 mg/l) and sodium cholate (1%) as described previously [17]. The same enrichment was also used to test the presence of *C. difficile* in the faecal samples of the healthy volunteers.

All isolates were confirmed as *C. difficile* by the amplification of *cdd3*, located downstream from the pathogenicity locus (*PaLoc*) [17] and ribotyped with primers described by Bidet [18]. The profiles were analysed with the BioNumerics software. All isolates were also toxigenic as determined by toxinotyping (data not shown) [19].

Isolation of the total DNA from faecal samples and amplification of the marker genes

The extraction of DNA was performed with the DNA Stool Mini Kit (Qiagen, Germany) after mechanical disruption with the SeptiFast Lyse Kit (Roche) on MagNA Lyser (speed 7000, 70 s; Roche). The DNA concentration and purity were checked with Nanodrop and 60 ng of faecal DNA was used for PCR amplification of bacterial, archaeal and fungal marker genes.

Primers and cycling conditions described by Domann [14] were used for amplification of the variable region V6-V8 of bacterial 16S rRNA genes. The correct amplicon size (470 bp) was checked by agarose gel electrophoresis. Heteroduplex DNA molecules were

reduced by reconditioning as described [20]. Before loading onto the DHPLC column, the PCR products were purified with the PCR Purification Kit (Qiagen).

For the amplification of archaeal 16S rRNA genes, we have targeted the variable region V3 for which we used a modified forward primer PARCH340F 5'-CCCTACGGGGTGCAGCAG-3' previously described by Ovreas [21] and designed a new reverse primer A780R 5'-TACCCGGGTATCTAATCCGGT-3' which anneals to a conserved region about 420 bp downstream. The cycling conditions were 95°C for 7 min, followed by 35 cycles using 95°C for 40 s, 63°C for 30 s, 72°C for 30 s and the final extension at 72°C for 5 min. We have checked the amplicon size, reconditioning and purification as described above.

The fungal internal transcribed spacer region 2 (ITS2) was amplified using a nested PCR. For the first reaction, we have used the primer pair nu-SSU-0817F 5'-TTAGCATGGAATAATRAATAGGA-3' [22] and ITS4R-5' TCCTCCGCTTATTGATATGC-3' [23]. The cycling conditions were 95°C for 2 min, followed by 14 cycles using 95°C for 30 s, 61.3°C for 30 s with a 0.5°C decrement for each following cycle and 72°C for 60 s. This was followed by 19 cycles at 95°C for 30 s, 54.3°C for 30 s and 72°C for 60 s and the final extension at 72°C for 5 min. The correct size of the amplicons (1400–2000 bp) was checked by agarose gel electrophoresis. Before the second PCR reaction, we have degraded any remaining primers from the first PCR reaction. For the degradation, the Shrimp Alkaline Phosphatase and Exonuclease I (both from Fermentas) were used. The reaction mixture of the final volume of 6.5 µl contained 0.15 U/ µl of the Phosphatase, 1.5 U/ µl of the Exonuclease I and 5 µl of the PCR product from the first reaction. The mixture was incubated at 37°C for 15 min and then the enzymes deactivated (85°C, 15 min). Five µl were then used as the template for the second PCR reaction using the primer ITS86F 5'-GTGAATCATC-GAATCTTTGAAC-3' [24] and the ITS4R primer described above. The cycling conditions were 95°C for 2 min, followed by 25 cycles at 95°C for 30 s, 54.3°C for 30 s and 72°C for 40 s and the final extension at 72°C for 5 min. The right size of the amplicons (approximately 200–400 bp) was again checked on agarose gel and the products purified as described above.

Separation of the amplified DNA with DHPLC

To separate the PCR products, we have used denaturing high pressure liquid chromatography (DHPLC; WAVE Microbial Analysis System; Transgenomics, USA). The amplicon volumes loaded on the column were 20 µl for bacteria, 15 µl for archaea and 10 µl for fungi. Conditions for separating the bacterial and fungal amplicons were described previously by Domann [14] and Goldenberg [25], respectively. For archaea, we have optimised the separation conditions which are given in the Supplementary table S1. A UV detector was used in DNA detection and the results were analysed with the Navigator Software version 2.2.0 (Build 25), the peak analysis parameters are given in the Supplementary table S1.

Normalization of retention times and assigning the DHPLC peaks to microbial groups

Before analysis, the retention times of the DHPLC peaks were normalized. A marker designed as a G - C rich oligonucleotide, which eluted after the sample, was loaded onto the column together with the sample. The retention times of the peaks were then adjusted according to the equation $NRT = RT + (RT * (1 - MRT/X))$, where NRT stands for the normalized retention time, RT for the retention time, MRT for

the retention time of the marker and X for an arbitrarily chosen point on the retention time scale.

Identification of the microbial groups

The retention time scale of the DHPLC profiles was divided into 0.2 min intervals. The representative peaks within each 0.2 min interval were collected, sequenced and all peaks within the given 0.2 min interval were then assigned to a bacterial group as described below.

The DHPLC peak fractions were collected from well separated and overlapping peaks and sent for sequencing (Eurofins MWG Operon, Germany). In case of partially overlapping peaks, the collected DHPLC peak fractions sometimes resulted in mixed sequence. In this case the corresponding PCR products were cloned into pGEM[®]-T Easy Vector System (Promega, Germany) following the manufacturers instructions. Ten colonies from each cloned fraction were collected, 16S rRNA gene fragments re-amplified (as above), purified (as above) and screened by DHPLC. Those, which differed in retention times, were selected and sequenced (Eurofins MWG Operon, Germany). The sequences were compared to the databases provided by the National Centre for Biotechnology Information (NCBI) and Ribosomal Database Project-Release 10 and the DHPLC peak fractions named accordingly. All 0.2 min intervals were named after the organisms identified in the fractions belonging to a particular interval. If more than one organism was identified in the DHPLC fractions taken within a particular 0.2 min interval, that interval was named after all the identified organisms. The named time intervals were from now on referred to as microbial groups.

Statistical analysis

The 208 faecal samples were divided into seven sets of samples, based on the following criteria: healthy donors or routine samples, diarrhoea or formed stool, the presence or absence of *C. difficile*, and 027 ribotype or other ribotypes. The names of the sets and the number of samples in each set are given in Tables 1, 2 and 3.

The results were analysed by looking at the differences between the sets of samples in terms of presence or absence of microbial groups or by looking at the percentage of peak area within the

total analysed chromatogram area which was computed by the Navigator Software version 2.2.0 (Build 25) using parameters as described in the Supplementary table S1.

The difference in the number of organism groups was analysed with ANOVA. The equality of the variances in each group was first verified with the Levene test and the pre-planned post-hoc comparisons were carried out with contrast analysis. Since the observed contrasts were not orthogonal, the p-values were adjusted with the Holm method to control the type I error.

Diversity of microbiota was assessed with Reciprocal Simpsons index of diversity and the differences between the compared sets of samples evaluated with Student T-test.

Logistic regression was used to analyse the presence or absence of individual microbial groups. The dependency between the dependent variable (1- individual has an organism group, 0 - individual does not have an organism group) and a microbial group was first tested with the chi - square test. After logistic regression, we used a contrast analysis to carry out the pre-planned comparisons between the sets of samples.

Principal component analysis

Principal Component Analysis (PCA) with respect to Instrumental Variables was performed with R programme (version 2.15.1) and applications ade4 (version 1.5-0) and ade4TkGUI (version 0.2-5). We used the data matrix based on the presence of the peak and percent area of the peaks (representing an individual microbial group) from the DHPLC chromatograms to calculate the spatial coordinates of each faecal sample. We also calculated the centre of gravity and dispersion for each set of samples. Statistical significance of PCA clustering based on microbiota profiles was assessed using a Wilcoxon signed-rank test.

Machine learning analysis

We further analysed the differences between the sets of samples in terms of presence or absence of microbial groups by using machine learning methods for the induction of decision trees [26], more specifically classification trees [27]. In particular, we used the WEKA J48 implementation [28] of the C4.5 algorithm [26]. The default parameter settings for J48 were applied, which include a

Table 1. Comparison of routine *C. difficile* negative samples (set of samples with formed stool and set with unformed stool) vs. healthy donors by logistic regression model.

| Sample set 1 vs. Sample set 2 ^a | NEG-F (29) vs. Healthy (37) | NEG-D (37) vs. Healthy (37) |
|--|-----------------------------|-----------------------------|
| <i>Peptostreptococcaceae</i> | | 0.21 (0.0028) |
| <i>Ruminococcus bromii</i> | | 0.11 (0.0454) |
| <i>Bifidobacterium longum</i> | 0.27 (0.0288) | |
| <i>Enterobacteriaceae</i> 3 | 0.25 (0.0088) | 0.25 (0.0056) |
| <i>Faecalibacterium</i> sp. | | 0.08 (0.0201) |
| <i>Bacteroides uniformis</i> | | 0.29 (0.0161) |
| <i>Prevotella</i> sp. | 0.23 (0.0256) | 0.18 (0.007) |
| <i>Alistipes</i> sp. | 0.2 (0.0133) | 0.14 (0.0014) |
| <i>Escherichia/Shigella</i> sp. | 0.17 (0.0012) | 0.21 (0.0028) |
| <i>Bacteroides vulgatus</i> | 0.11 (0.0059) | 0.3 (0.0331) |

a - the compared sample sets and the number of samples in each set is specified.

NEG-F: *C. difficile* negative/formed stool; NEG-D: *C. difficile* negative/diarrhoea; HEALTHY: healthy donors.

Results are presented as odds ratios for individual microbial groups between different sets of samples. An odds ratio below 1 indicates that this particular microbial group is less frequent in sample set 1 as compared to sample set 2. The odds ratios are given in the table only for the p-values (shown in parentheses) below 0.05 for the compared sample sets.

doi:10.1371/journal.pone.0058005.t001

Table 2. Comparison of different sets of routine *C. difficile* positive samples vs. healthy donors by logistic regression model.

| Sample set 1 vs. Sample set 2 ^a | OTHER-F (17) vs. Healthy (37) | OTHER-D (31) vs. Healthy (37) | 027-F (16) vs. Healthy (37) | 027-D (41) vs. Healthy (37) |
|--|-------------------------------|-------------------------------|-----------------------------|-----------------------------|
| <i>Peptostreptococcaceae</i> | | | 0.25 (0.0285) | |
| <i>Ruminococcus bromii</i> | | | 0.05 (0.0069) | |
| <i>Bifidobacterium longum</i> | 0.18 (0.0341) | 0.25 (0.0197) | 0.09 (0.0247) | 0.03 (0.0013) |
| <i>Enterobacteriaceae</i> 1 | 0.07 (0.018) | | 0.06 (0.0149) | 0.1 (0.0323) |
| <i>Bacteroides</i> sp. 1 | 0.3 (0.0461) | 0.35 (0.0385) | 0.19 (0.011) | 0.22 (0.0019) |
| <i>Enterobacteriaceae</i> 2 | | 0.26 (0.0116) | 0.15 (0.0036) | 0.13 (0.0001) |
| <i>Enterobacteriaceae</i> 3 | | 0.2 (0.0021) | 0.07 (0.0013) | 0.08 (0.0000) |
| <i>Bacteroides</i> sp. 2 | | 0.28 (0.0326) | 0.1 (0.0322) | 0.12 (0.0017) |
| <i>Faecalibacterium</i> sp. | 0.04 (0.0042) | | | 0.1 (0.0323) |
| <i>Bacteroides uniformis</i> | | 0.35 (0.0403) | 0.12 (0.0106) | 0.27 (0.0084) |
| <i>Prevotella</i> sp. | 0.17 (0.0153) | 0.22 (0.0198) | 0.12 (0.0038) | 0.15 (0.0025) |
| <i>Alistipes</i> sp. | 0.11 (0.0019) | 0.19 (0.0105) | | 0.14 (0.0014) |
| <i>Escherichia/Shigella</i> sp. | 0.13 (0.0022) | 0.09 (0.0000) | 0.19 (0.0105) | 0.15 (0.0002) |
| <i>Bacteroides vulgatus</i> | | 0.1 (0.0044) | | 0.12 (0.0017) |
| <i>Candida albicans</i> | | | | 3.62 (0.0115) |
| <i>Candida glabrata</i> | | | | 10.45 (0.0301) |
| <i>Methanobrevibacter smithii</i> | 0.23 (0.0189) | 0.20 (0.0025) | | 0.25 (0.0041) |

a - the compared sample sets and the number of samples in each set is specified.

OTHER-F: *C. difficile* non 027 ribotype/formed stool; OTHER-D: *C. difficile* non 027 ribotype/diarrhoea; 027-F: *C. difficile* 027 ribotype/formed stool; 027-D: *C. difficile* 027 ribotype/diarrhoea; HEALTHY: healthy donors.

Results are presented as odds ratios for individual microbial groups between different sets of samples. An odds ratio below 1 indicates that this particular microbial group is less frequent in sample set 1 as compared to sample set 2. The odds ratios are given in the table only for the p-values (shown in parentheses) below 0.05 for the compared sample sets.

doi:10.1371/journal.pone.0058005.t002

minimum of two examples per leaf for pre-pruning and a confidence level of 0.25 for post-pruning.

The dependent (class) variable in this analysis was the ribotype of *C. difficile* detected in a sample (none, 027, other). The independent variables (attributes) were the indicators of presence/absence of each of the 34 microbial groups, as well as an indicator of the sample category (healthy donor/routine sample).

We next reversed the direction of analysis and analysed the overall microbiota composition depending on the detected ribotype of *C. difficile* and other sample properties. To this end, we used predictive clustering trees - PCTs [29] for multi-target classification, as implemented in the Clus data mining suite (<http://clus.sourceforge.net>). Each of the indicators of presence/absence of the 34 microbial groups was considered as a target (class) variable, all of them to be predicted simultaneously by a

Table 3. Comparisons of all routine *C. difficile* positive vs. all routine *C. difficile* negative samples by logistic regression model.

| Sample set 1 vs. Sample set 2 ^a | OTHER-D (31) vs. NEG-D (37) | 027-D (41) vs. NEG-D (37) | 027-D (41)/OTHER-D (31) vs. NEG-D (37) |
|---|-----------------------------|---------------------------|--|
| <i>Streptococcus</i> sp./ <i>Enterococcus</i> sp. 2 | | 3.7 (0.0471) | 3.23 (0.0461) |
| <i>Peptostreptococcaceae</i> | | 2.9 (0.0259) | 2.63 (0.0232) |
| <i>Ruminococcus bromii</i> | | 4.8 (0.0491) | 3.33 (0.0453) |
| <i>Bifidobacterium longum</i> | | 0.04 (0.0031) | 0.12 (0.0012) |
| <i>Enterobacteriaceae</i> 1 | | 0.1 (0.037) | |
| <i>Bacteroides</i> sp. 1 | | 0.39 (0.0465) | |
| <i>Enterobacteriaceae</i> 2 | | 0.35 (0.0289) | |
| <i>Enterobacteriaceae</i> 3 | | 0.33 (0.0498) | |
| <i>Methanobrevibacter smithii</i> | 0.39 (0.0491) | 0.31 (0.0268) | 0.35 (0.0138) |

a - the compared sample sets and the number of samples in each set is specified.

NEG-D: *C. difficile* negative/diarrhoea; OTHER-D: *C. difficile* non 027 ribotype/diarrhoea; 027-D: *C. difficile* 027 ribotype/diarrhoea.

Results are presented as odds ratios for individual microbial groups between different sets of samples. An odds ratio below 1 indicates that this particular microbial group is less frequent in sample set 1 as compared to sample set 2. The odds ratios are given in the table only for the p-values (shown in parentheses) below 0.05 for the compared sample sets.

doi:10.1371/journal.pone.0058005.t003

single tree. As independent variables (attributes), we considered the sample category (healthy donor/routine sample), the consistency of the stool (formed/diarrhoea) and the ribotype of *C. difficile* detected in a sample (none, 027, other).

Results

The aim of this study was to compare the gut microbiota in faecal samples from healthy donors and from routine samples negative or positive for *C. difficile* and to detect possible differences in dominant microbial groups (bacteria, archaea, fungi).

We have analysed 208 faecal samples, of which 171 were routine samples and 37 were from healthy volunteers. All faecal samples from healthy volunteers were *C. difficile* negative. Of the 171 routine samples, 105 were *C. difficile* positive and 66 were *C. difficile* negative. From all 105 positive faecal samples *C. difficile* was isolated and strains were assigned to 22 different *C. difficile* PCR ribotypes. The five most frequent ribotypes were 027 (57), 014/020 (6), 081 (6), 002 (5) and 023 (5).

DHPLC analysis of archaeal, fungal and bacterial microbiota

The representative DHPLC profiles in Fig. 1 show different complexity of archaeal, fungal and bacterial microbiota in human faecal samples. This section will only briefly list archaeal, fungal and bacterial groups detected by the DHPLC method, while the differences in faecal microbiota between samples (healthy, *C.*

difficile positive and negative samples) will be described in more detail in further sections.

The archaeal microbiota was simple in all 208 analysed samples and was composed maximally of two methanogenic archaea, *Methanosphaera stadtmanae* and *Methanobrevibacter smithii* (Fig. 1a).

The fungal microbiota was more diverse (Fig. 1b, Fig. 2). We have found nine different fungi altogether. The most frequent among them were *Saccharomyces cerevisiae* (in 57% of samples) and *Candida albicans* (in 27% of samples). *Candida glabrata*, *Candida diversa*, *Clavispora lusitanae*, *Pichia burtonii*, *Lenoniera sp.*, *Eurotium sp.* and *Sclerotium sp.* were found only sporadically.

Bacterial microbiota was most complex (Fig. 1c, Fig. 3). Altogether we have identified 23 different bacterial groups. Some chromatographic peaks were assigned to more than one genus because they either contained mixed sequences, or because the short sequences used in the DHPLC analysis did not allow distinction between them (e.g. *Clostridium difficile*/*Sporacetigenium sp.*; *Citrobacter freundii*/*Enterobacter sp.*; *Streptococcus sp.*/*Enterococcus sp.* 1 and 2). Bacterial groups that were identified to the same genus or family but had different retention times were designated by a number (*Enterobacteriaceae* 1, 2 and 3; *Bacteroides sp.* 1 and 2, *Streptococcus sp.*/*Enterococcus sp.* 1 and 2). Because short fragments are used in the DHPLC analysis further identification of particular amplified fraction was not possible. The remaining identified groups were *Clostridium sp.*, *Peptostreptococcaceae*, *Blautia sp.*, *Ruminococcus bromii*, *Ruminococcus sp.*, *Faecalibacterium sp.*, *Klebsiella pneumoniae*; *Escherichia/Shigella sp.*, *Prevotella sp.*, *Alistipes sp.*, *Bacteroides uniformis*, *Bacteroides vulgatus*, *Bifidobacterium longum* and *Bifidobacterium sp.*

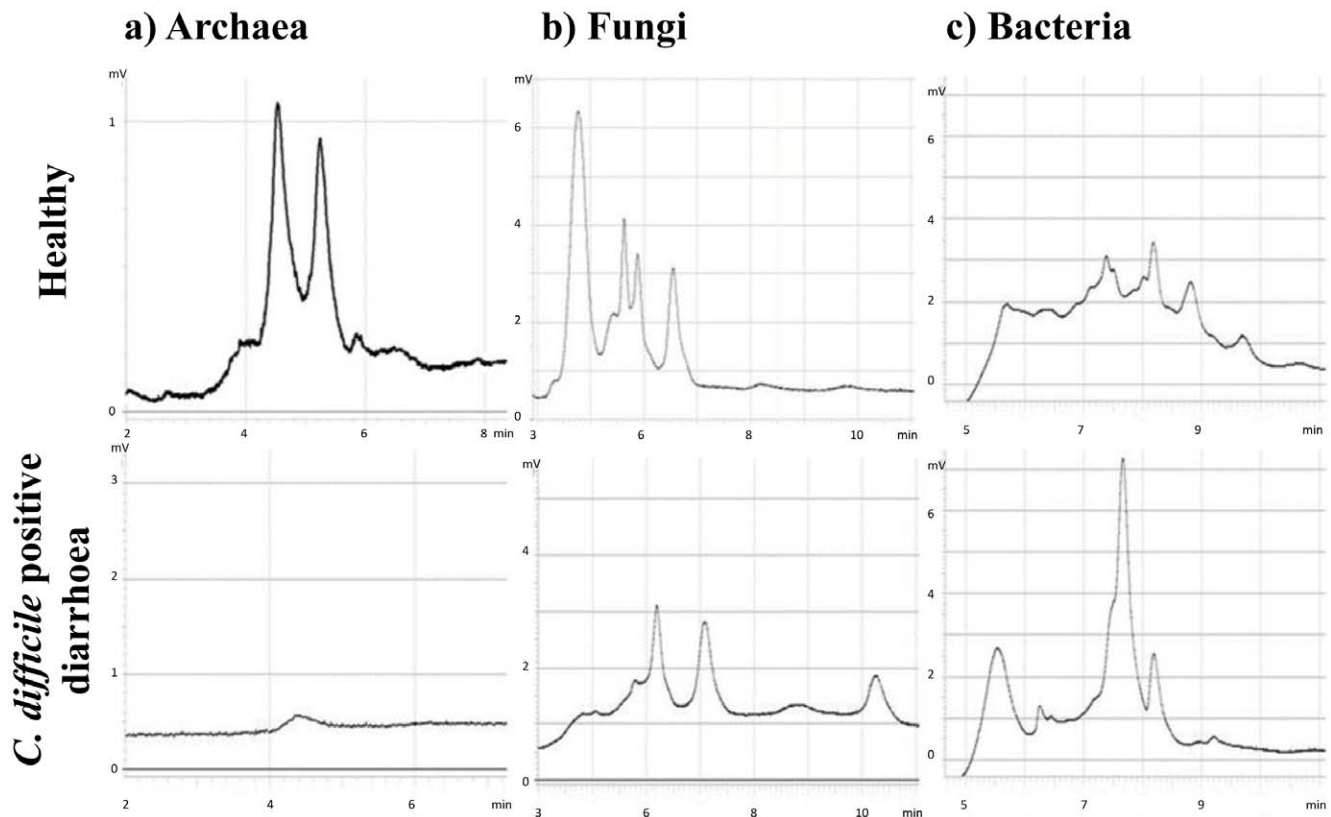


Figure 1. Representative DHPLC profiles of faecal microbiota. The profiles show archaeal (a), fungal (b) and bacterial (c) faecal microbiota. The faecal profiles of healthy volunteers are in the upper row and the profiles from the diarrhoeal patients colonised by *Clostridium difficile* 027 ribotypes are in the lower row. The selected profiles show typical chromatograms of the three studied microbial groups. The chromatograms show increasing complexities from archaeal to fungal and bacterial microbiota, respectively.
doi:10.1371/journal.pone.0058005.g001

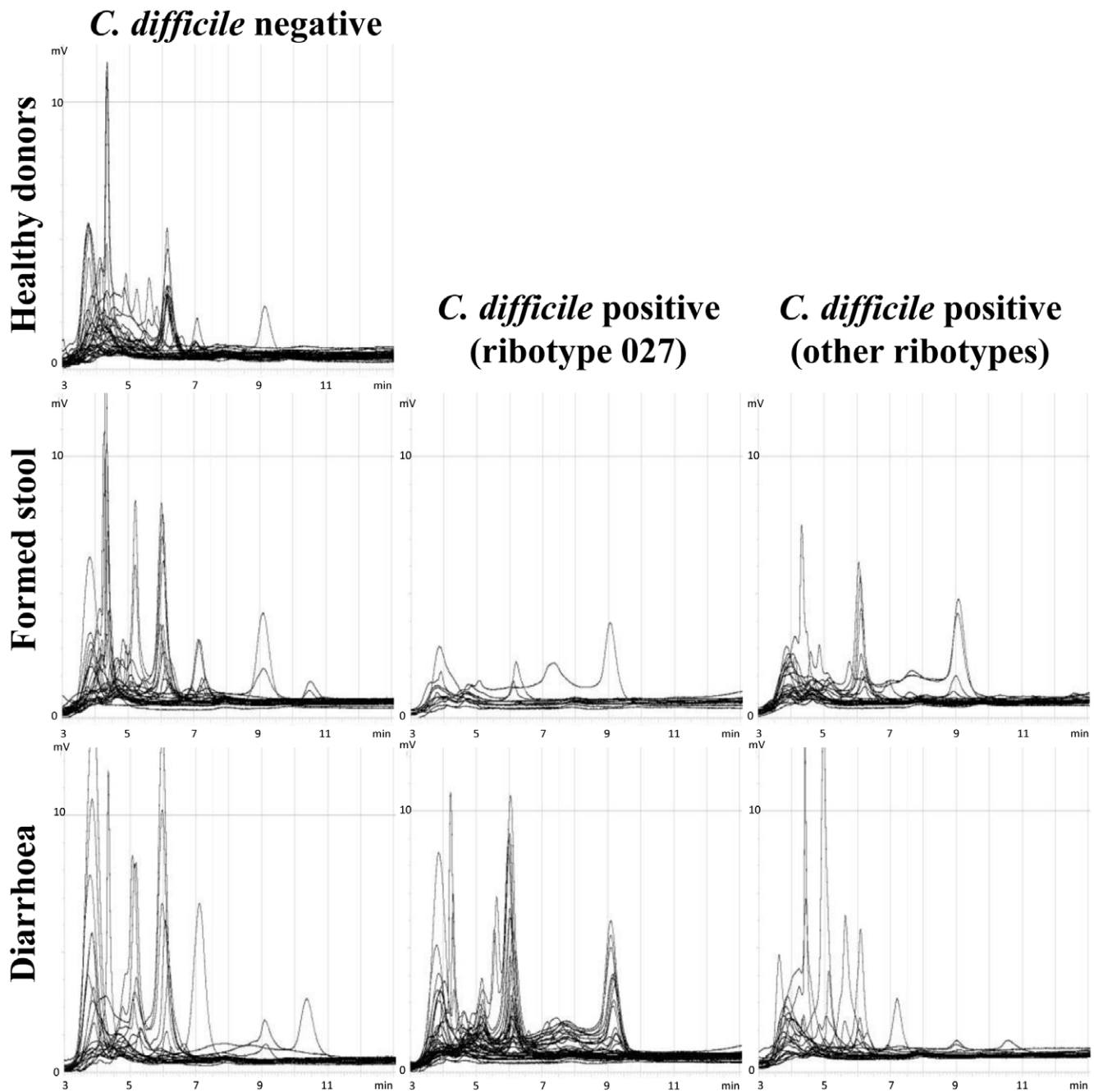


Figure 2. Fungal DHPLC profiles of all seven sets of samples. The samples were distributed into sets according to consistency, source (healthy/routine) and *C. difficile* presence. The peaks with different retention times (x - axis) represents different fungal groups. Fungal microbiota in healthy volunteers and *C. difficile* positive samples differs in the composition (position of the peaks) but not in diversity (number of the peaks). doi:10.1371/journal.pone.0058005.g002

DHPLC profiles and principal component analysis

All 208 faecal samples were partitioned into seven sets according to origin (routine sample/healthy), appearance (diarrhoea/formed) and *C. difficile* colonization status (negative/ribotype 027/any other ribotype). We have named the sample sets accordingly (diarrhoeal/*C. difficile* negative samples; diarrhoeal/*C. difficile* ribotype 027 positive samples; diarrhoeal/*C. difficile* other ribotypes positive samples; formed stool/*C. difficile* negative samples; formed stool/*C. difficile* ribotype 027 positive samples; formed stool/*C. difficile* other ribotypes positive samples; healthy donors) (Fig. 2 and 3). Because of the large number of samples

colonised with the 027 ribotype (57 out of 105 colonised samples), we were able to separate those from other ribotypes and treat it as a separate set of samples.

To compare the faecal profiles of the dominant microbiota in the seven sets of samples, we first used PCA to calculate the spatial coordinates of each sample (Fig. 4). All sets of samples showed a large dispersion, particularly samples sent for routine *C. difficile* testing (Fig. 4a). Although the centres of gravity were different for samples from healthy persons and for routine samples, they also substantially overlapped (Fig. 4a). The result was similar when we

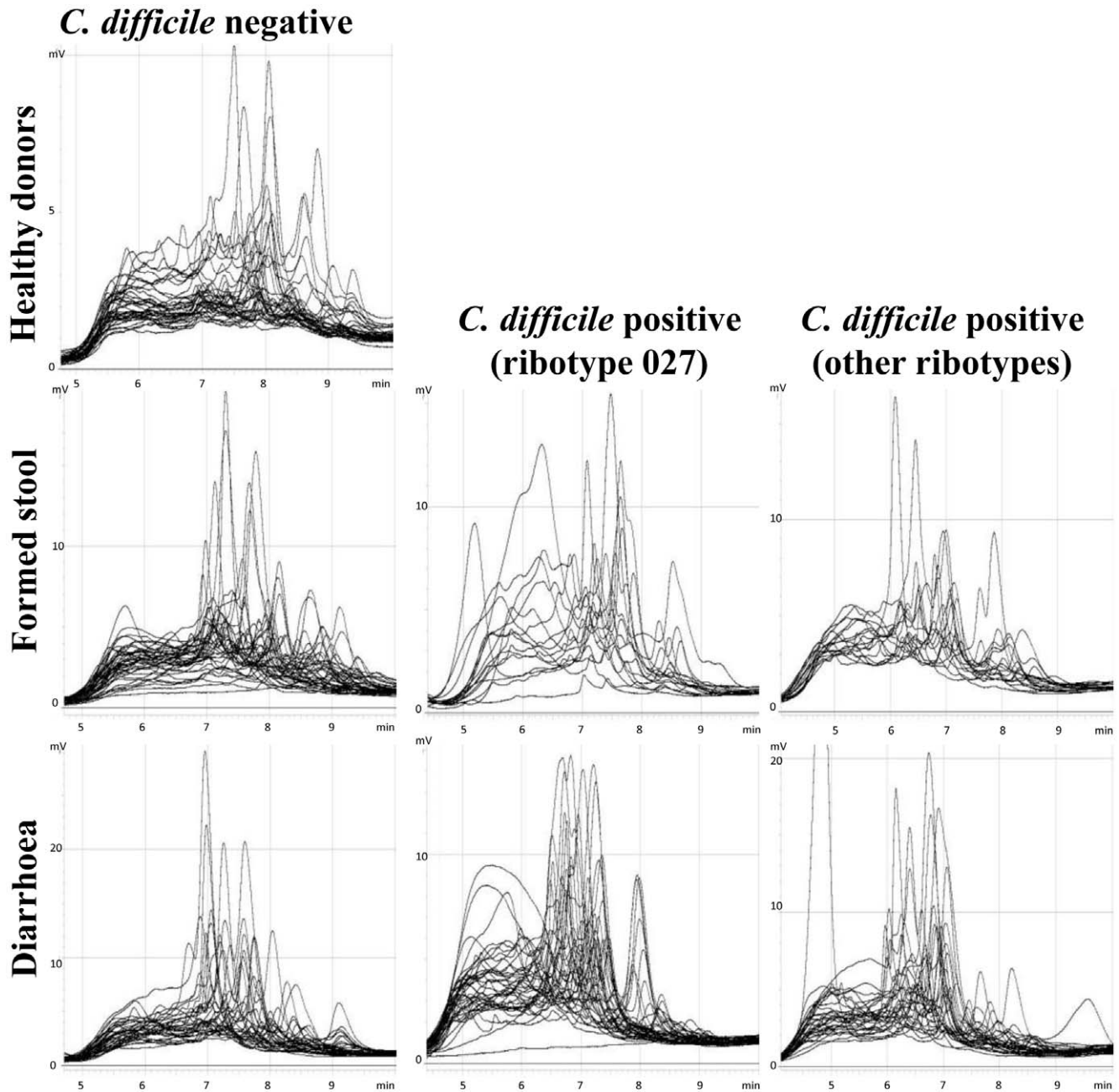


Figure 3. Bacterial DHPLC profiles of all seven sets of samples. Bacterial microbiota in healthy individuals is composed of a large number of species, hence the peaks are numerous but the area below each peak is small. The presence of *C. difficile* correlates with decreased diversity reflected by a lower number of peaks which are, however, more prominent.
doi:10.1371/journal.pone.0058005.g003

compared the samples based on appearance (diarrhoea/formed) and *C. difficile* colonisation status (Fig. 4b, c and d).

Regardless of the overlap between the sets of samples, Wilcoxon signed-rank test performed on each pair of the sets from Fig. 4 (healthy donors or routine samples, diarrhoea or formed stool, the presence or absence of *C. difficile*, and 027 ribotype or other ribotypes) show a clear distinction between them. In all the cases, location shift greater than 0.5 can be shown with 95% confidence, while the probability of sets being different is larger than 99.9%. This confirms that the partition of the samples into the sets at hand was appropriate and we kept this grouping to further analyse the

differences in microbial groups using statistical analyses and machine learning methods.

Comparison of microbiota present in routine samples and in samples from healthy volunteers

Within each of the seven sample sets, we have determined the average number of bacterial or fungal groups (Fig. 5). The difference in the average number of present bacterial groups between the healthy volunteers and each of the six different routine sample sets was statistically significant ($P < 0.0001$). The comparison of the average number of bacterial and fungal groups

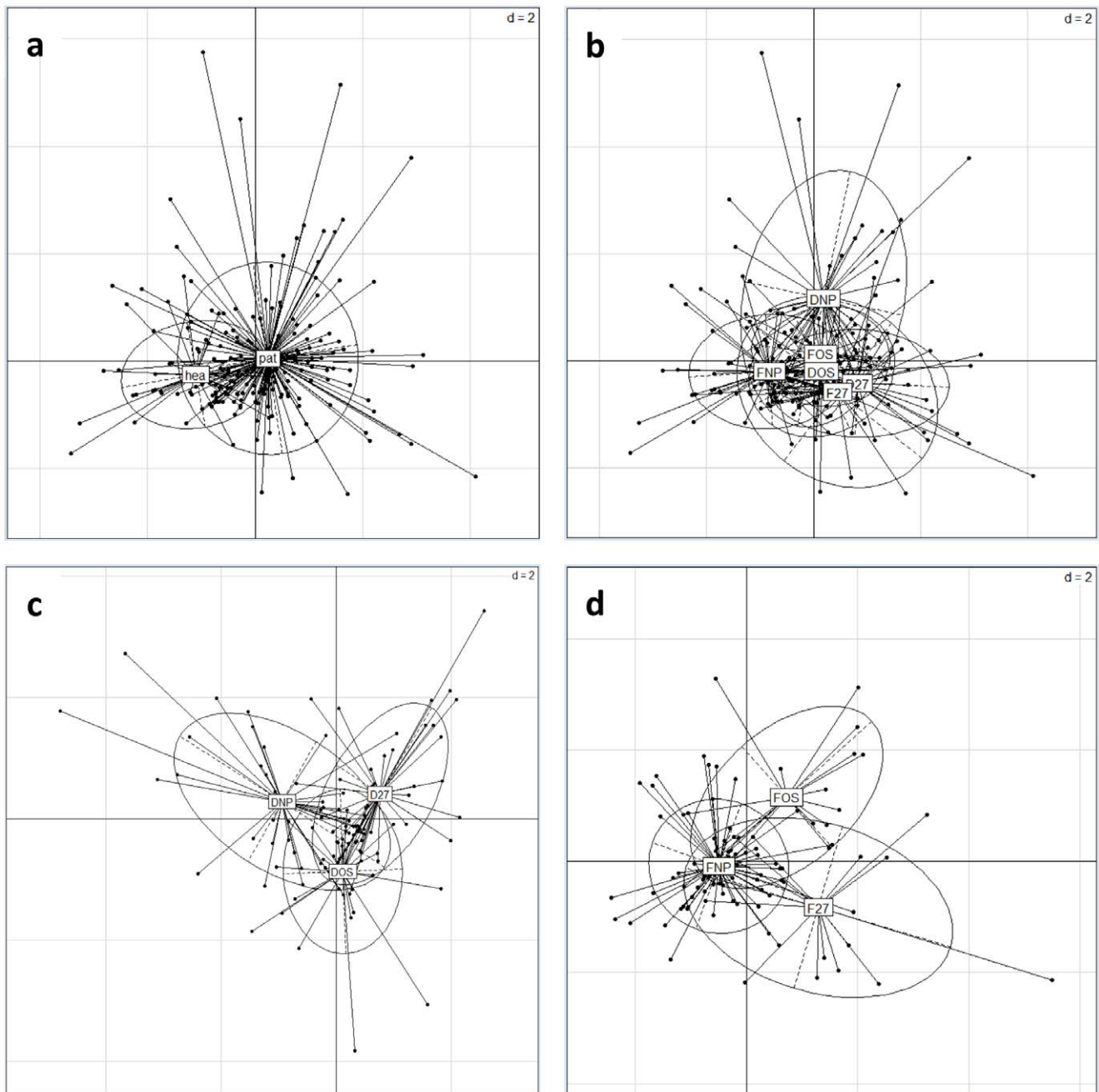


Figure 4. Principal component analysis of DHPLC profiles of dominant bacterial, fungal and archaeal species. Individual samples (represented by black dots) were clustered into sets according to *C. difficile* presence and type, origin and consistency (ellipses) and the centre of gravity computed for each set of samples. All sets of samples show large dispersion and overlap (ellipses). Figure 4a compares the samples sent for routine *C. difficile* testing (pat) and healthy donors (hea). Particularly the former set shows a great dispersion. Figures 4b, c and d compare different sets of samples from the patients. Figure 4b shows comparison of six different sets of routine samples; a greater shift in the centre of gravity between the formed and the diarrhoeal *C. difficile* negative samples (FNP/DNP), compared to the *C. difficile* positive samples (FOS/DOS, F027/D027). Fig 4c shows only the subset of different routine diarrheic samples. Fig 4d shows only the subset of different routine formed samples DNP - diarrhoeal/*C. difficile* negative samples, D27 - diarrhoeal/*C. difficile* ribotype 027 positive samples, DOS - diarrhoeal/*C. difficile* other ribotypes positive samples, FNP - formed stool/*C. difficile* negative samples, F27 - formed stool/*C. difficile* ribotype 027 positive samples, FOS - formed stool/*C. difficile* other ribotypes positive samples, pat - samples sent for routine *C. difficile* testing, hea - healthy donors.
doi:10.1371/journal.pone.0058005.g004

shows that the healthy donors had the highest number of bacterial groups (in average 16.54) while the highest average among the routine samples 12.69 was the one for formed/*C. difficile* negative stools. The lowest average number of bacterial groups was observed in the samples colonised with *C. difficile* ribotype 027

(9.69 and 11.10 for formed and diarrhoeal stools respectively). The fungal microbiota showed no difference in the average number of fungi between the healthy donors and routine samples (Fig. 5). Simpsons reciprocal index of diversity showed similar differences between the sets of samples with healthy donors having

significantly higher microbiota diversity (10.03) than routine samples (from 4.47 to 8.38). In addition, the samples colonised with the ribotype 027 had significantly lower diversity (4.47 and 5.37, formed and diarrhoeal samples respectively) than samples colonised with other *C. difficile* ribotypes (6.25 and 8.38, formed and diarrhoeal samples respectively). The diversity indexes and P-values are given in Supplementary table S2.

After assessing differences between the average numbers of microbial groups, we have further analysed the differences on the level of specific microbial taxons with PCT (predictive clustering trees) analysis and with logistic regression. Several PCTs were built, relating microbiota composition to different combinations of independent variables. Here we present the trees built using the sample category only (Table 4) and the ribotype of *C. difficile* only (Table 5). The PCT models take into account only microbial groups present at high frequencies (above 50% in at least one sample set). We highlight the differences between healthy donors and routine samples (Table 4) and between routine samples with different *C. difficile* colonisation/ribotype status (Table 5), and only list the microbial groups that are present differentially (i.e., present in the majority of one set of samples and absent from the majority of the other set of samples).. When healthy donors were compared to all routine samples (not distributed into sets regarding stool consistency and *C. difficile* presence) by PCT, five microbial groups were identified to be more frequent in healthy donors, including different *Enterobacteriaceae* and *Bacteroides sp.* (Table 4).

Subsequently, all microbial groups were analysed by logistic regression to calculate the odds ratios for individual microbial groups when comparing sets of samples one to another (Tables 1, 2 and 3). In logistic regression, the routine samples were also further distributed into sets (*C. difficile* negative, 027 *C. difficile* positive, other than 027 *C. difficile* positive). Healthy donor microbiota differed from *C. difficile* negative routine samples mainly in bacterial composition (Table 1). On the other hand, differences

Table 4. Differences in microbial groups between healthy donors and routine samples as detected by the PCT model.

| Organism group | Healthy donor | Routine samples |
|---------------------------------|---------------|-----------------|
| <i>Enterobacteriaceae</i> 2 | 1 | 0 |
| <i>Enterobacteriaceae</i> 3 | 1 | 0 |
| <i>Escherichia/Shigella sp.</i> | 1 | 0 |
| <i>Bacteroides sp.1</i> | 1 | 0 |
| <i>Bacteroides uniformis</i> | 1 | 0 |

An entry of 1 means that the organism group (row) was present and an entry of 0 means that the organism group was absent in the majority of samples of the considered type (column).

doi:10.1371/journal.pone.0058005.t004

between the healthy donors and *C. difficile* positive routine samples were detected in bacterial, fungal and archaeal components (Table 2).

Changes of microbiota associated with *Clostridium difficile* presence in the faecal sample

As described before the average number of bacterial and fungal groups was not significantly different between the routine samples with a different *C. difficile* colonisation status and ranged from 9.69 bacterial groups in diarrhoeal/*C. difficile* ribotype 027 positive and to 12.69 bacterial groups in formed/*C. difficile* negative samples, respectively, and (on average) from 1.5 to 2.0 fungal groups per sample (Fig. 5).

When different subsets within the routine samples were compared by the PCT model, which takes into account only microbial groups present at high frequencies (above 50% in at least one sample set), we identified five microbial groups characteris-

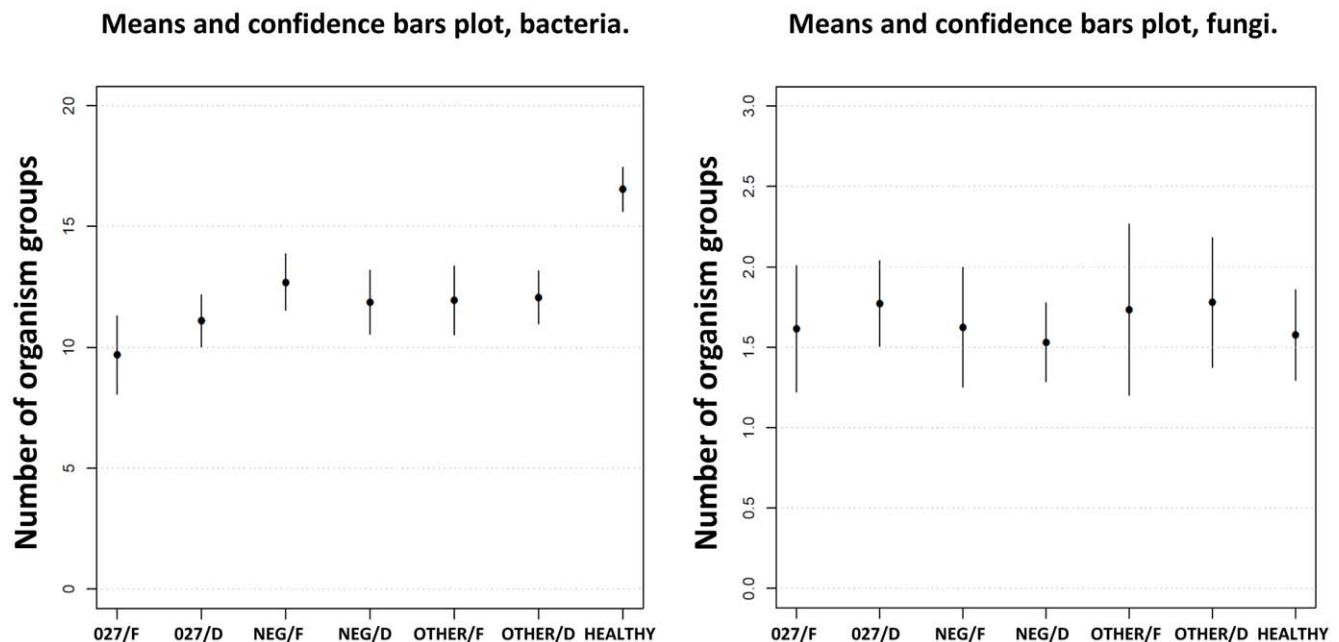


Figure 5. Average numbers of bacterial and fungal groups in different sets of faecal samples. The average numbers of present bacterial and fungal groups within each sample set are presented as means with confidence bars plot. 027/F: *C. difficile* 027 ribotype/formed stool; 027/D: *C. difficile* 027 ribotype/diarrhoea; NEG/F: *C. difficile* negative/formed stool; NEG/D: *C. difficile* negative/diarrhoea; OTHER/F: *C. difficile* non 027 ribotype/formed stool; OTHER/D: *C. difficile* non 027 ribotype; HEALTHY: healthy donors. doi:10.1371/journal.pone.0058005.g005

Table 5. Differences in microbial groups present between different types of *C. difficile* colonisation as predicted by the PCT model.

| Organism group | <i>C. difficile</i> negative | <i>C. difficile</i> non 027 ribotype | <i>C. difficile</i> 027 ribotype |
|-----------------------------------|------------------------------|--------------------------------------|----------------------------------|
| <i>Enterobacteriaceae</i> 2 | 1 | 1 | 0 |
| <i>Bacteroides</i> sp. 1 | 1 | 0 | 0 |
| <i>Escherichia/Shigella</i> sp. | 1 | 0 | 0 |
| <i>Saccharomyces cerevisiae</i> | 1 | 1 | 0 |
| <i>Methanobrevibacter smithii</i> | 1 | 0 | 0 |

An entry of 1 means that the organism group (row) was present and an entry of 0 means that the organism group was absent in the majority of samples of the considered type (column)

doi:10.1371/journal.pone.0058005.t005

tically present or absent (the majority of) in different groups of routine samples (Table 5). All five were predicted to be present in *C. difficile* negative routine samples and only two (*Enterobacteriaceae* group and *S. cerevisiae*) in routine samples colonized with *C. difficile* ribotype other than 027.

When comparing *C. difficile* negative or positive samples to healthy volunteers, bacteria and archaea were less frequent in routine samples than in healthy donors, while some fungi were more frequent in *C. difficile* positive samples (Tables 1 and 2).

When comparing diarrhoeal *C. difficile* negative samples to diarrhoeal *C. difficile* positive samples (regardless of the ribotype) significant differences were found in four bacterial groups (*Peptostreptococcaceae*, *Ruminococcus bromii*, *Streptococcus/Enterococcus* sp. 2 and *Bifidobacterium longum*) and one archaeal group (*Methanobrevibacter smithii*) (Table 3, last column). Among them, *Bifidobacterium longum* and *Methanobrevibacter smithii* showed lower frequencies in the *C. difficile* positive samples. However, these differences were mainly due to the diarrhoeal 027 positive *C. difficile* samples as it will be described below.

Archaea (*Methanobrevibacter smithii/Methanosphaera stadtmanae*) were associated with healthy donors (detected in 65% (*M. smithii*)/22% (*M. stadtmanae*) of samples) and *C. difficile* negative samples (52%/21% for formed stool and 51%/19% for diarrhoea). *C. difficile* positive samples (especially those colonized with non - 027 ribotype) were characterised by lower frequencies of archaea. Frequencies of *M. smithii/M. stadtmanae* in 027 positive *C. difficile* samples were (38%/19% for formed stool and 32%/17% for diarrhoea) and in non - 027 positive samples (29%/12% for formed stool and 26%/10% for diarrhoea). The diarrhoeal non - 027 positive *C. difficile* samples and diarrhoeal *C. difficile* negative samples significantly differed to each other only in *Methanobrevibacter smithii* (Table 3, first column).

A decision tree was constructed by the J48 machine learning approach, based on the presence or absence of microbial groups, to predict the outcome regarding *C. difficile* colonisation. The resulting decision tree splits the samples by category first. We only discuss the part that refers to the routine samples, shown in Figure 6. This decision tree was built automatically from the data on the collected samples. Each internal node is labeled with a test on the presence/absence of a microbial group, while each terminal node (leaf) is labeled with the most frequent class/outcome (*C. difficile* 027/*C. difficile* other ribotypes/*C. difficile* negative). The numbers in parentheses provide an indication of the accuracy of the prediction made by the leaf (total number of samples in the leaf/incorrectly classified samples). The tree therefore summarizes the data on the samples and the results of their analysis, relating microbial group composition to the outcome.

The decision tree can also be used as a predictive model, i.e., to make predictions regarding the *C. difficile* ribotype of new samples

based on their microbial groups. In this case, its accuracy would have to be estimated on unseen data (and not on the training data as done here) by using procedures such as cross-validation.

The tree identifies several different combinations of microbial groups associated with the same outcome. The presence as well as absence of microbial groups influenced the outcome importantly. The microbial groups identified by the model to be most related to the outcome were bacteria, but also fungi and archaea (Fig. 6). The most strongly related microbial group was *Bifidobacterium longum* the presence of which was linked to the *C. difficile* negative outcome in 18 out of 27 cases. The presence of *Prevotella* sp. was associated with both, *C. difficile* positive and negative status, depending on the presence or absence of other microorganisms. It was associated with *C. difficile* absence when found together with *Enterobacteriaceae* 1 and 2. This was also the pattern with the best classification accuracy, misclassifying only 2 out of 18 samples. On the other hand, *Prevotella* sp. was associated with *C. difficile* presence when found together with *Enterobacteriaceae* 3 and *Streptococcus* sp./*Enterococcus* sp. 1.

Specific changes in microbiota in the patients colonised with the *C. difficile* ribotype 027

Because a large number of *C. difficile* positive routine samples had ribotype 027 (57/105) we could study those samples as a separate set. We found most differences (8 bacterial groups) between the diarrhoeal *C. difficile* negative and diarrhoeal ribotype 027 positive samples (Table 3). Three bacterial groups (*Peptostreptococcaceae*, *Ruminococcus bromii* and *Streptococcus/Enterococcus* sp. 2) were more frequent and five bacterial groups (*Enterobacteriaceae* 1, 2, 3; *Bacteroides* sp. 1 and *Bifidobacterium longum*) were less frequent in *C. difficile* ribotype 027 positive samples. Among them *Bifidobacterium longum* had the highest statistical significance ($P = 0.0031$) (Table 3).

The analysis of the fungal groups showed that the diarrhoeal samples positive for the *C. difficile* ribotype 027 were more frequently colonised by *Candida albicans* and *Candida glabrata*, compared to stools from healthy donors (Table 2). However, no differences in fungi were detected among diarrhoeal samples positive for the *C. difficile* ribotype 027 and other diarrhoeal routine samples.

The machine learning approach identified four colonization patterns associated with the 027 ribotype (Fig. 6). They are different one to another and include bacteria, fungi and archaea. The pattern with the best predictive score, albeit associated with small number of samples (1 misclassified sample out of 5), involves the presence of *Enterobacteriaceae* 1 and 2 as well as *Candida glabrata*. In addition to the logistic regression analysis (Tables 2 and 3) and the J48 model (Fig. 6), specific changes in the samples colonised by the *C. difficile* ribotype 027 were detected also by the PCT model (Table 5). Table 5 shows only those five microbial groups for

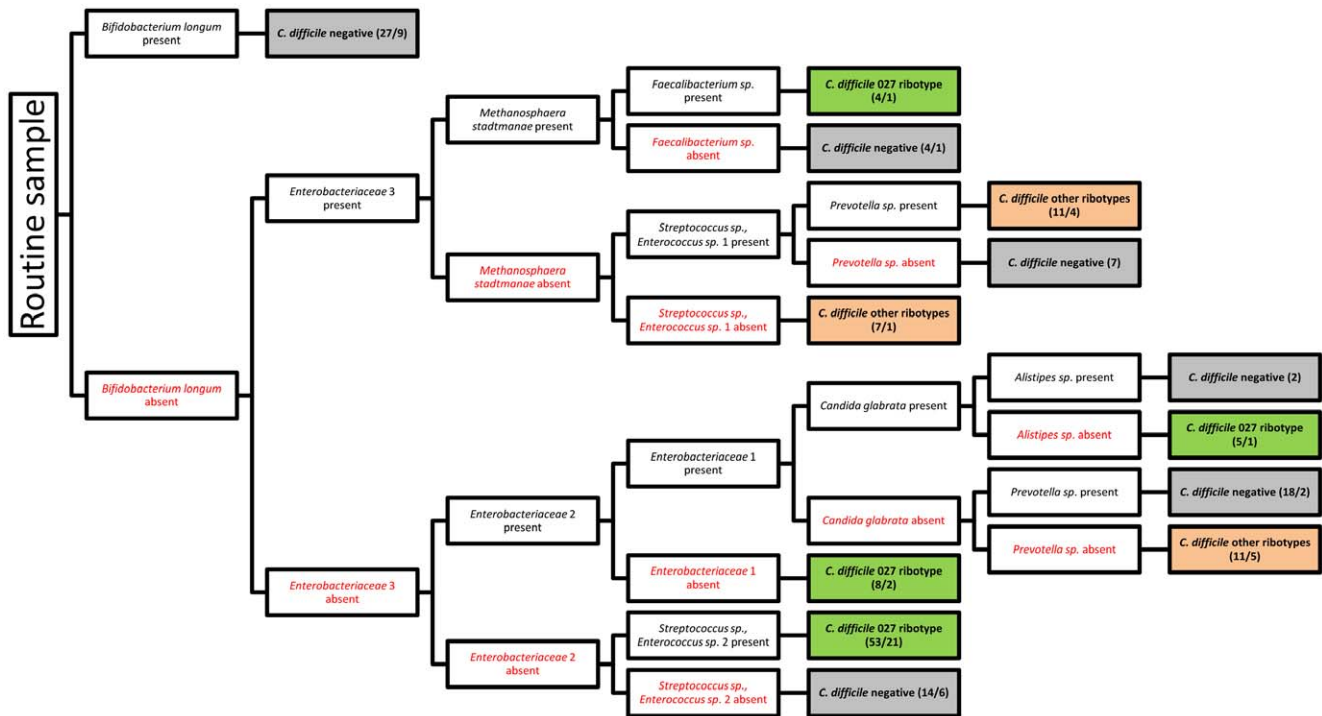


Figure 6. J48 decision tree describing the patterns of colonisation associated with presence or absence of *C. difficile*. Decision tree is showing only routine sample sets. The decision tree shows combinations of microbial groups that predict the outcome regarding *C. difficile* colonisation. The numbers in parentheses show the total number of samples and the number of incorrectly classified instances, respectively. doi:10.1371/journal.pone.0058005.g006

which the predicted outcomes between the sample sets were different. All of them are predicted to be present in *C. difficile* negative routine samples, but their absence is predicted in *C. difficile* 027 positive samples.

Discussion

In this study we have used the DHPLC method to detect changes in faecal microbial communities during *C. difficile* colonisation. Although the method is simpler than more laborious sequencing approaches and can detect only a small proportion of present microbial groups, we have shown here that it does detect differences between selected groups of faecal samples and that detected differences are in agreement with studies published to date.

The tested faecal samples were first distributed into those collected from healthy volunteers and routine samples with requested *C. difficile* testing. The latter were further divided into six sets according to stool consistency, *C. difficile* presence or absence and ribotype (027 or non-027) (Fig. 2 and 3).

In current study lower diversity was a common trait of routine samples (formed or unformed) and not necessarily connected only to *C. difficile* colonisation as average number of bacterial groups in all routine samples was significantly lower compared to healthy donors. Although 66 samples out of 171 routine samples were negative for *C. difficile*, patients did have changed microbiota (Fig. 5). The fact that samples were sent to the routine testing could be either due to diarrhoea (with *C. difficile* risk factors), while for formed routine samples the reason for test requirement could be untypical abdominal discomfort or retesting after CDI treatment. No clinical data were collected to confirm this possible explanation. Lower bacterial diversity was previously specifically attributed to different gastrointestinal diseases. In addition to

patients with a recurrent 8] or active CDI 11], lower bacterial diversity was also observed in patients suffering from Crohns disease (CD) 30], diarrhoea-predominant irritable bowel syndrome 31] and chemotherapy 32]. In addition, antibiotic treatment has been shown to cause short-term 33] and long-term 34] shifts and decreased diversity in gut microbiota.

Our results are in concordance with the previous findings at the level of individual bacterial taxa which showed lower levels of *Bifidobacterium longum*, *Prevotella sp.* and *Bacteroides sp.* in *C. difficile* positive samples [11,35,36]. In addition, in routine samples *Bifidobacterium longum* proved to be the most important predictor for the *C. difficile* negative status. Bacterial groups, which in our study were more frequent in *C. difficile* positive samples, include *Ruminococcus bromii*, the family *Peptostreptococcaceae* and *Streptococcus sp./Enterococcus sp. 2*. The first two groups confirm the previous studies which showed the increased numbers of firmicutes 36], particularly *Ruminococcus sp.* in symptomatic adults 11] and asymptomatic infants 37]; and increased levels of *Clostridium sp.* in *C. difficile* colonised subjects 8,35]. In our study, significant increase in colonisation frequencies of *Streptococcus sp./Enterococcus sp. 2* group were found in diarrhoeal *C. difficile* colonised samples. Previous studies found a positive correlation between a persistent asymptomatic *C. difficile* colonization and enterococci 38] and connected vancomycin resistant enterococcus (VRE) infection to more than 50% of *C. difficile* colonisations which prompted the authors to suggest an active surveillance for VRE in *C. difficile* colonised patients 39].

Very few studies have so far reported on the fungal microbiota in connection to *C. difficile* colonization or the gut microbiota in general. In *C. difficile* colonized hamsters treatment with monoclonal antibodies against *C. difficile* resulted in proliferation of fungus *Wickerhamomyces sp.* 16]. In humans we have here described

several fungi to be found in healthy or disturbed faecal microbiota. However, average number of fungal groups in faeces was low (1.5–2.0), which is similar to a previous study of fungal gut microbiota in faeces of healthy individuals reporting 1–3 fungal species per faecal sample [40]. Another culture independent study which investigated fungal diversity in patients with inflammatory bowel disease (IBD), showed a slightly higher number of mucosa-associated fungi (6 per patient) [41]. Comparison of diversity in patients with different types of intestinal inflammation (Crohn's disease (CD), ulcerative colitis, non-IBD colitis, unclear diarrhoea) showed, that only CD patients had a higher fungal diversity [41]. We did not find differences in fungal diversity between all *C. difficile* colonized and all non-colonized subjects. However, a specific subgroup of diarrhoeal patients, colonised with the 027 ribotype, was more frequently colonised with *Candida albicans* and *Candida glabrata* compared to the healthy donors. Our results suggest that while fungal diversity does not increase during CDI, there is a qualitative shift in the composition towards opportunistic pathogens.

This is the first study to look at the possible changes in archaeal microbiota in connection to *C. difficile* colonization. We found higher frequencies of a single methanogenic species in the healthy donors compared to *C. difficile* positive samples, but also when comparing the *C. difficile* negative samples, sent for the routine *C. difficile* testing, to the *C. difficile* positive samples. Our results are therefore linking *Methanobrevibacter smithii* to the microbiota of healthy individuals. Until recently, archaea have been a neglected part of gut microbiota studies. The importance of archaea for the homeostasis of the gut ecosystem and the role they play in the bacterial degradation of sugars has led to the increased studies of these microorganisms [42,43]. They have been implicated in obesity after gastric bypass surgery [44] and in periodontal diseases [45]. Because of the large number of samples colonised with the ribotype 027, we could analyse them separately. Similar as observed before on a much smaller set of samples [11], our results also show that bacterial diversity in faecal samples colonized with *C. difficile* 027 was lower than in other *C. difficile* positive samples. Additionally, we showed that also certain fungi and archaea were absent in faecal samples positive for *C. difficile* 027. Although so called hypervirulence of ribotype 027 has to be considered with caution [46], it seems that this particular strain has increased potential to spread in the environment and to colonize the gut. One of the possible explanations for this would be superior colonizing properties of 027 strains and hence reduced need for an

imbalanced microbiota. However, the association of ribotype 027 strains with the lowest microbiota diversity among all samples in this study potentially suggest that ribotype 027 is capable of actively affecting the composition of faecal microbiota and is therefore increasing its own colonization potential.

Current studies have characterised *C. difficile* specific changes in microbiota by evaluating relative frequencies and abundances for each gut microorganism separately. Here, in contrast, we have applied machine learning methods, and could show that *C. difficile* colonisation is associated with patterns in the gut microbiota composition. Some species are very strong predictors, as for example, *Bifidobacterium longum*, which was connected to the *C. difficile* negative outcome. Some other bacterial groups (e.g. *Prevotella*) could be associated with a *C. difficile* positive or negative status, which depends on the other microorganisms present in the microbiota. This indicates that not only the presence or absence of a single species/taxon but rather that certain combinations of gut microbes are important in *C. difficile* associated changes of microbiota.

Supporting Information

Table S1 The DHPLC separation conditions and peak analysis parameters. Conditions for the separation of mixed species 16S rDNA or ITS2 amplicons using the WAVE microbial analysis system and peak analysis parameters for the Navigator Software 2.2.0 (Build 25) are listed. (DOCX)

Table S2 Average reciprocal Simpsons indexes of diversity (1/D), standard deviations and P values of the compared groups of samples. (DOCX)

Acknowledgments

Authors would like to thank N. Oresic and I. Strumbelj for providing part of the routine samples and to M. Medved for technical assistance.

Author Contributions

Conceived and designed the experiments: JS MR. Performed the experiments: JS. Analyzed the data: JS SD BZ DM SG MR. Contributed reagents/materials/analysis tools: MR SD. Wrote the paper: JS SD BZ DM SG MR.

References

- Rupnik M, Wilcox MH, Gerding DN (2009) Clostridium difficile infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 7: 526–536.
- Johnson S, Samore MH, Farrow KA, Killgore GE, Tenover FC, et al. (1999) Epidemics of diarrhea caused by a clindamycin-resistant strain of Clostridium difficile in four hospitals. N Engl J Med 341: 1645–1651.
- Graul T, Cain AM, Karpa KD (2009) Lactobacillus and bifidobacteria combinations: a strategy to reduce hospital-acquired Clostridium difficile diarrhea incidence and mortality. Med Hypotheses 73: 194–198.
- Mattila E, Uusitalo-Seppala R, Wuorela M, Lehtola L, Nummi H, et al. (2012) Fecal Transplantation, Through Colonoscopy, Is Effective Therapy for Recurrent Clostridium difficile Infection. Gastroenterology 142: 490–496.
- Mellow MH, Kanatzar A (2011) Colonoscopic fecal bacteriotherapy in the treatment of recurrent Clostridium difficile infection--results and follow-up. J Okla State Med Assoc 104: 89–91.
- Ho N, Prasad V (2011) Clostridium difficile diarrhea and fecal transplantation. J Clin Gastroenterol 45: 742–743.
- Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, et al. (2012) Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to Clostridium difficile-induced colitis. Infect Immun 80: 62–73.
- Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, et al. (2008) Decreased diversity of the fecal Microbiome in recurrent Clostridium difficile-associated diarrhea. J Infect Dis 197: 435–438.
- De La Cochetiere MF, Durand T, Lalande V, Petit JC, Potel G, et al. (2008) Effect of antibiotic therapy on human fecal microbiota and the relation to the development of Clostridium difficile. Microb Ecol 56: 395–402.
- Rousseau C, Lemec L, Le Monnier A, Poilane I, Pons JL, et al. (2011) Prevalence and diversity of Clostridium difficile strains in infants. J Med Microbiol 60: 1112–1118.
- Rea MC, O'Sullivan O, Shanahan F, O'Toole PW, Stanton C, et al. (2012) Clostridium difficile Carriage in Elderly Subjects and Associated Changes in the Intestinal Microbiota. J Clin Microbiol 50: 867–875.
- Goldenberg O, Herrmann S, Marjoram G, Noyer-Weidner M, Hong G, et al. (2007) Molecular monitoring of the intestinal flora by denaturing high performance liquid chromatography. Journal of Microbiological Methods 68: 94–105.
- Wagner AO, Malin C, Illmer P (2009) Application of denaturing high-performance liquid chromatography in microbial ecology: fermentor sludge, compost, and soil community profiling. Appl Environ Microbiol 75: 956–964.
- Domann E, Hong G, Imirzalioglu C, Turschner S, Kuhle J, et al. (2003) Culture-independent identification of pathogenic bacteria and polymicrobial infections in the genitourinary tract of renal transplant recipients. Journal of Clinical Microbiology 41: 5500–5510.
- Rouge C, Goldenberg O, Ferraris L, Berger B, Rochat F, et al. (2010) Investigation of the intestinal microbiota in preterm infants using different methods. Anaerobe 16: 362–370.

16. Peterfreund GL, Vandivier LE, Sinha R, Marozsan AJ, Olson WC, et al. (2012) Succession in the Gut Microbiome following Antibiotic and Antibody Therapies for *Clostridium difficile*. *PLoS One* 7: e46966.
17. Zidaric V, Zemljic M, Janezic S, Kocuvan A, Rupnik M (2008) High diversity of *Clostridium difficile* genotypes isolated from a single poultry farm producing replacement laying hens. *Anaerobe* 14: 325–327.
18. Bidet P, Barbut F, Lalande V, Burghoffer B, Petit JC (1999) Development of a new PCR-ribotyping method for *Clostridium difficile* based on ribosomal RNA gene sequencing. *FEMS Microbiol Lett* 175: 261–266.
19. Rupnik M, Avesani V, Janc M, von Eichel-Streiber C, Delmee M (1998) A novel toxinotyping scheme and correlation of toxinotypes with serogroups of *Clostridium difficile* isolates. *J Clin Microbiol* 36: 2240–2247.
20. Thompson JR, Marcelino LA, Polz MF (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Res* 30: 2083–2088.
21. Ovreas L, Forney L, Daae FL, Torsvik V (1997) Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl Environ Microbiol* 63: 3367–3373.
22. Borneman J, Hartin RJ (2000) PCR primers that amplify fungal rRNA genes from environmental samples. *Appl Environ Microbiol* 66: 4356–4360.
23. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: a guide to methods and applications* Academic Press, New York, USA: 315–322.
24. Turenne CY, Sanche SE, Hoban DJ, Karlowsky JA, Kabani AM (1999) Rapid identification of fungi by using the ITS2 genetic region and an automated fluorescent capillary electrophoresis system. *J Clin Microbiol* 37: 1846–1851.
25. Goldenberg O, Herrmann S, Adam T, Marjoram G, Hong G, et al. (2005) Use of denaturing high-performance liquid chromatography for rapid detection and identification of seven *Candida* species. *J Clin Microbiol* 43: 5912–5915.
26. Quinlan JR (1993) *C4.5: Programs for Machine Learning*. Morgan Kaufman, San Francisco, CA, USA.
27. Breiman L, Friedman J, Olshen R, Stone C (1984) *Classification and Regression Trees*. Wadsworth, Belmont, CA.
28. Witten IH, Frank E, Hall MA (2011) *Data mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufman, San Francisco, CA, USA, third edition.
29. Blockeel H, De Raedt L, Ramon J (1998) Top-down induction of clustering trees. In: *Proceedings of the Fifteenth International Conference on Machine Learning (ICML 98)* Morgan Kaufmann: 55–63.
30. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, et al. (2006) Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55: 205–211.
31. Carroll IM, Ringel-Kulka T, Siddle JP, Ringel Y (2012) Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Neurogastroenterol Motil* 24: 521–530, e248.
32. Zwieler J, Lassl C, Hippe B, Pointner A, Switzeny OJ, et al. (2011) Changes in human fecal microbiota due to chemotherapy analyzed by TaqMan-PCR, 454 sequencing and PCR-DGGE fingerprinting. *PLoS One* 6: e28654.
33. De La Cochetiere MF, Durand T, Lepage P, Bourreille A, Galmiche JP, et al. (2005) Resilience of the dominant human fecal microbiota upon short-course antibiotic challenge. *J Clin Microbiol* 43: 5588–5592.
34. Lofmark S, Jernberg C, Jansson JK, Edlund C (2006) Clindamycin-induced enrichment and long-term persistence of resistant *Bacteroides* spp. and resistance genes. *J Antimicrob Chemother* 58: 1160–1167.
35. Hopkins MJ, Macfarlane GT (2002) Changes in predominant bacterial populations in human faeces with age and with *Clostridium difficile* infection. *J Med Microbiol* 51: 448–454.
36. Manges AR, Labbe A, Loo VG, Atherton JK, Behr MA, et al. (2010) Comparative metagenomic study of alterations to the intestinal microbiota and risk of nosocomial *Clostridium difficile*-associated disease. *J Infect Dis* 202: 1877–1884.
37. Rousseau C, Levenez F, Fouqueray C, Dore J, Collignon A, et al. (2011) *Clostridium difficile* colonization in early infancy is accompanied by changes in intestinal microbiota composition. *J Clin Microbiol* 49: 858–865.
38. Ozaki E, Kato H, Kita H, Karasawa T, Maegawa T, et al. (2004) *Clostridium difficile* colonization in healthy adults: transient colonization and correlation with enterococcal colonization. *J Med Microbiol* 53: 167–172.
39. Fujitani S, George WL, Morgan MA, Nichols S, Murthy AR (2011) Implications for vancomycin-resistant *Enterococcus* colonization associated with *Clostridium difficile* infections. *Am J Infect Control* 39: 188–193.
40. Scanlan PD, Marchesi JR (2008) Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme Journal* 2: 1183–1193.
41. Ott SJ, Kuhbacher T, Musfeldt M, Rosenstiel P, Hellmig S, et al. (2008) Fungi and inflammatory bowel diseases: Alterations of composition and diversity. *Scand J Gastroenterol* 43: 831–841.
42. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444: 1027–1031.
43. Samuel BS, Gordon JI (2006) A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. *Proc Natl Acad Sci U S A* 103: 10011–10016.
44. Zhang H, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, et al. (2009) Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci U S A* 106: 2365–2370.
45. Eckburg PB, Lepp PW, Relman DA (2003) Archaea and their potential role in human disease. *Infect Immun* 71: 591–596.
46. Barbut F, Rupnik M (2012) Editorial commentary: 027, 078, and others: going beyond the numbers (and away from the hypervirulence). *Clin Infect Dis* 55: 1669–1672.