

SCIENTIFIC REPORTS



OPEN

Species Level Description of the Human Ileal Bacterial Microbiota

Heidi Cecilie Villmones¹, Erik Skaaheim Haug², Elling Ulvestad^{3,4}, Nils Grude¹, Tore Stenstad⁵, Adrian Halland² & Øyvind Kommedal³

Received: 28 November 2017

Accepted: 6 March 2018

Published online: 16 March 2018

The small bowel is responsible for most of the body's nutritional uptake and for the development of intestinal and systemic tolerance towards microbes. Nevertheless, the human small bowel microbiota has remained poorly characterized, mainly owing to sampling difficulties. Sample collection directly from the distal ileum was performed during radical cystectomy with urinary diversion. Material from the ileal mucosa were analysed using massive parallel sequencing of the 16S rRNA gene. Samples from 27 Caucasian patients were included. In total 280 unique Operational Taxonomic Units were identified, whereof 229 could be assigned to a species or a species group. The most frequently detected bacteria belonged to the genera *Streptococcus*, *Granulicatella*, *Actinomyces*, *Solobacterium*, *Rothia*, *Gemella* and *TM7(G-1)*. Among these, the most abundant species were typically streptococci within the *mitis* and *sanguinis* groups, *Streptococcus salivarius*, *Rothia mucilaginosa* and *Actinomyces* from the *A. meyeri/odontolyticus* group. The amounts of Proteobacteria and strict anaerobes were low. The microbiota of the distal part of the human ileum is oral-like and strikingly different from the colonic microbiota. Although our patient population is elderly and hospitalized with a high prevalence of chronic conditions, our results provide new and valuable insights into a lesser explored part of the human intestinal ecosystem.

The human gut microbiota has been extensively investigated in recent years owing to its impacts on human health and disease^{1–3}. Most research on host-microbe interactions are based on studies of the faecal microbiota. The feasibility of this practice is questionable^{4–6}. The small bowel is responsible for most of the body's nutritional uptake and for the development of intestinal and systemic tolerance towards microbes. It has protruding villi, making its surface area approximately fifteen times greater than that of the colon⁷. The small intestines also have a thinner, more vulnerable mucin barrier than the colonic epithelium, offering closer contact between the luminal content and mucosal enteroendocrine and immunological cells^{8,9}. Peyer's patches are present mainly in the distal jejunum and ileum and increase to a maximum together with the Paneth cells in the distal ileum^{10–12}.

A recent study on rats have concluded that faecal sampling misrepresents the microbiota at different intestinal locations¹³. Similar studies have been difficult to perform in humans, due to major challenges in the procurement of representative and uncontaminated samples from intestinally healthy individuals. Sampling techniques have included naso-ileal catheters¹⁴, capsules^{15,16} and retrograde colonoscopy^{16,17}. Investigations have also been performed on ileostomy effluent samples^{12,14,18,19}, autopsies²⁰ and samples from patients suffering from inflammatory bowel disease or in need of emergency surgery^{21–23}.

Patients undergoing radical cystectomy, a treatment where the bladder is removed, have their urinary diversion created using a segment of ileum. Most of these patients have no known intestinal diseases, despite their high median age and bladder condition. In this study, by using uncontaminated samples collected directly from the distal ileal mucosal surface during this procedure, we aimed to characterize the microbiota of the ileum using massive parallel sequencing of the bacterial gene encoding 16S ribosomal RNA (16S rRNA gene). When the 16S rRNA gene provided insufficient species-level discrimination, we supplemented the characterization with targeted sequencing of alternative genes providing higher taxonomic resolution. Species-level identification is necessary to increase our understanding of the microbial ecosystem and host-microbe interactions, but also the microbiota's role in infections through leaking, translocation and haematogenous spread of potentially pathogenic organisms.

¹Department of Microbiology, Vestfold Hospital Trust, 3103, Tønsberg, Norway. ²Department of Urology, Vestfold Hospital Trust, 3103, Tønsberg, Norway. ³Department of Microbiology, Haukeland University Hospital, 5021, Bergen, Norway. ⁴Department of Clinical Science, University of Bergen, 5021, Bergen, Norway. ⁵Department of Infectious Medicine, Vestfold Hospital Trust, 3103, Tønsberg, Norway. Correspondence and requests for materials should be addressed to H.C.V. (email: heivil@siv.no)

Population characteristics	Years/ kg/m ²	(range)	Patient identification, sample number
Age, median years (min-max)	71	(53–85)	
BMI, median kg/m ² (min-max)	27	(21–40)	
BMI ≥ 30	5		4, 5, 10, 21, 29
BMI ≥ 40	1		29
Population characteristics	Number of patients	(%)	Patient identification, sample number
Sex, male	21	(78)	
Sex, female	6	(22)	9, 13, 14, 15, 16, 20
Neoadjuvant chemotherapy	8	(29)	
MVAC	5	(19)	3, 9, 14, 20, 28
GC	3	(11)	10,11,16
Antibiotic prophylaxis	27	(100)	
quinolone + metronidazole	25	(93)	
trimethoprim sulfa + metronidazole	1	(4)	13
furadantin + metronidazole	1	(4)	18
Indication, bladder cancer	24	(88)	
Indication, prostate cancer	2	(7)	2, 27
Indication, chronic cystitis	1	(4)	15
Chronic diseases none	5	(19)	8, 11, 14, 18, 30
Chronic diseases yes	22	(81)	
Diabetes mellitus 2	2	(7)	10, 22
Cardiovascular disease*	15	(56)	1, 2, 3, 4, 5, 9, 10, 13, 16, 19, 20, 22, 23, 24, 28
COPD/Asthma	5	(19)	5, 6, 12, 28, 29
Irritable bowel disease	1	(4)	20
Constipation	1	(4)	28
Cancer coli operata	1	(4)	6
Regular medications, none	5	(19)	8, 11, 14, 18, 30
Regular medications, yes	22	(81)	
Statins	15	(56)	1, 2, 4, 5, 10, 13, 16, 19, 21, 22, 24, 26, 27,28, 29
PPI	3	(11)	3, 26, 27
Antidiabetics	2	(7)	10, 22
Antibiotics prior to admission (last week)	5	(19)	11, 13, 15, 16, 19
Antibiotics prior to admission (last six months)	15	(56)	1, 2, 9, 11, 12, 13, 14, 16, 18, 19, 20, 26, 28, 29, 30

Table 1. Demographic and clinical data. BMI: body mass index, MVAC: Methotrexate, vinblastine, neomycin (Adriamycin), cisplatin. GC: Gemcitabine, cisplatin. COPD: Chronic Obstruction Pulmonary Disease, *including hypercholesterolemia and hypertension

Results

The mean and median age of the 27 patients was 71 years (range 54–86). Indications for cystectomy were bladder cancer for 24 patients, complications following prostate cancer for two patients and chronic cystitis for one. Nine patients received neoadjuvant chemotherapy prior to surgery but none had received radiotherapy or immunotherapy. Most suffered from additional chronic diseases and 22 out of 27 used medications on a regular basis. Fifteen out of 27 had at least one antibiotic treatment during the last six months prior to admission, mainly for urinary tract infections. Further demographic and clinical data are displayed in Table 1.

The mean number of reads was 318,742 per sample (Supplementary Table S1). After removal of short reads (<250 base pairs), small clusters (<10 reads) and chimeras, a mean of 100,276 valid reads remained (range 15,079 to 376,370, median 97,047) for each sample. The main reason for this substantial loss was short reads, probably due to premature truncation in the ion semiconductor sequencing technology²⁴. About 50% of the reads in all samples had a length below 250 base pairs.

By 16S rRNA sequencing, a total of 280 unique Operational Taxonomic Units (OTU's) were accepted, whereof 191 could be identified to the species level, 38 to a species group level, and 51 to the genus level (Supplementary Table S2). The detailed results from the targeted Sanger-sequencing of *gdh* and *rpoB*-genes, providing species level identification for some of the bacteria that could only be assigned to the group level by 16S, are provided in Supplementary Table S3. Whenever species level identifications obtained by these supplementary genes are used in the text, the gene is indicated in parenthesis after the species name. The mean species richness was 51 (range 15 to 124, median 50) with an average Shannon index of 2.84 (range 0.83 to 3.72, median 2.56). Most samples had species richness between 30 and 79 and a Shannon-index between 2.36 and 3.72 (Fig. 1a,b and Supplementary Table S4).

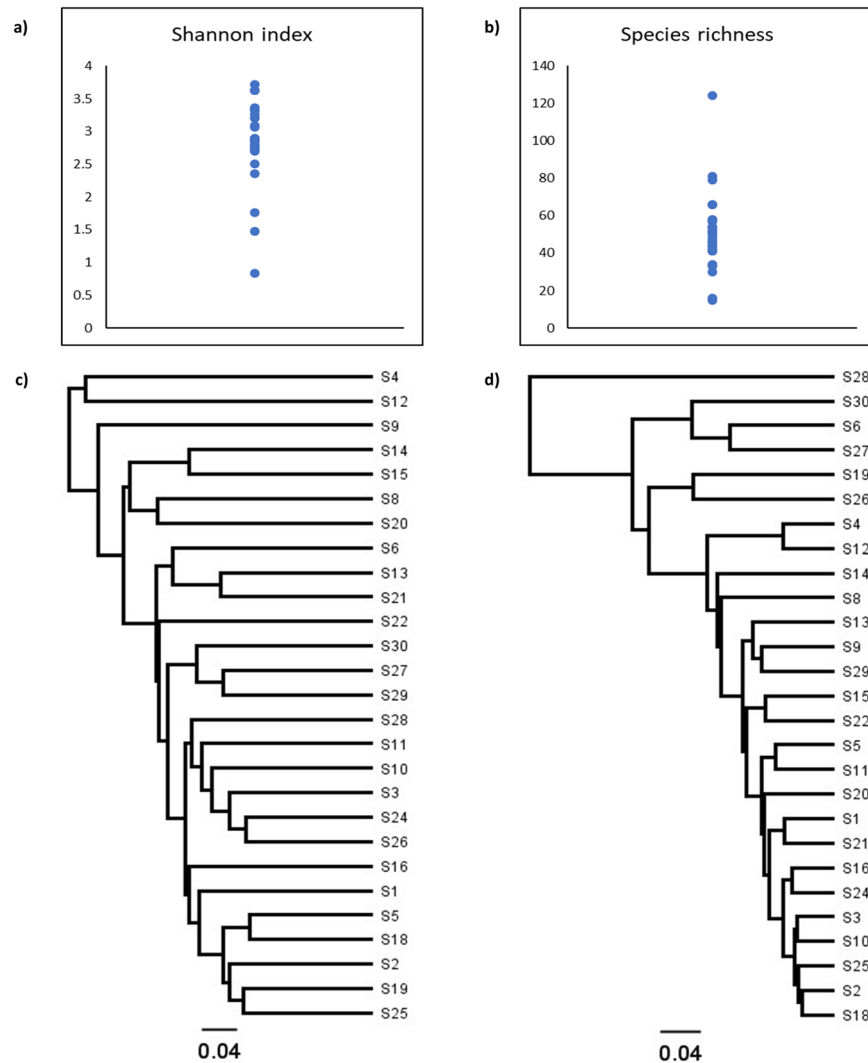


Figure 1. Alpha- and beta diversities. (a) Shannon index distribution. (b) Species richness distribution. (c) Unweighted (qualitative) UniFrac analyses. (d) Weighted (quantitative) UniFrac analyses.

Bacterial structure of the ileal samples. The most abundant phylum in the distal ileum was Firmicutes followed by Actinobacteria. Most samples also contained Candidate division TM7 (24/27), Proteobacteria (24/27) and Fusobacteria (23/27) (Fig. 2a). In four samples species from the latter three phyla flourished at the expense of both Firmicutes and Actinobacteria. These were *Fusobacterium periodonticum* in sample 30, *TM7(G-1) sp. oral taxon 352* in samples 6 and 27 and *Escherichia coli* in sample 28. Although outliers in the weighted UniFrac analysis, they remained within the main cluster in the unweighted UniFrac analysis (Fig. 1c,d), reflecting that the qualitative species compositions were not atypical. The most extreme outlier, sample 28 with 77% reads from *E. coli*, was from a patient with long term constipation that had been using a combination of high-osmotic and contact-laxative medications on a regular basis. We observed no congruence between clusters in the UniFrac analyses and BMI category, gender, use of statins, antibiotic treatment or neoadjuvant chemotherapy. Bacteroidetes were found in 11 out of 27 samples whereas species within the phyla Synergistetes, Tenericutes and Spirochetes were detected at low levels in only one sample each.

The distal ileum core microbiota at genus and species level, defined as genera and species that occurred in more than 50% of the individuals, are displayed in Fig. 3. The typical microbiota is dominated by the facultative anaerobic genera *Streptococcus*, *Actinomyces*, *Gemella*, *Granulicatella* and *Rothia* and in most cases also contain the genera *Atopobium*, *Lachnoanaerobaculum*, *Oribacterium*, *Solobacterium* and *TM7(G-1)* (Fig. 2b). On average, strict anaerobic bacteria constituted only 10% of the reads (range 0.5% to 29%, median 7%), whereof the most significant were *Atopobium parvulum*, *Oribacterium asaccharolyticum*, *Oribacterium sinus*, *Solobacterium moorei*, *Fusobacterium nucleatum*, *Fusobacterium periodonticum*, *Parvimonas micra* and *Bifidobacterium longum*. *Clostridium* species were detected only sporadically and at low levels, except from *Clostridium celatum*, which was found in seven samples. *Clostridium celatum* was also the dominant anaerobe in one individual. The newly described species *Romboutsia timonensis* (Clostridiales order), was recovered from six samples and among the dominant anaerobes in three. *Prevotella* and *Bacteroides* species were found in only six and four samples

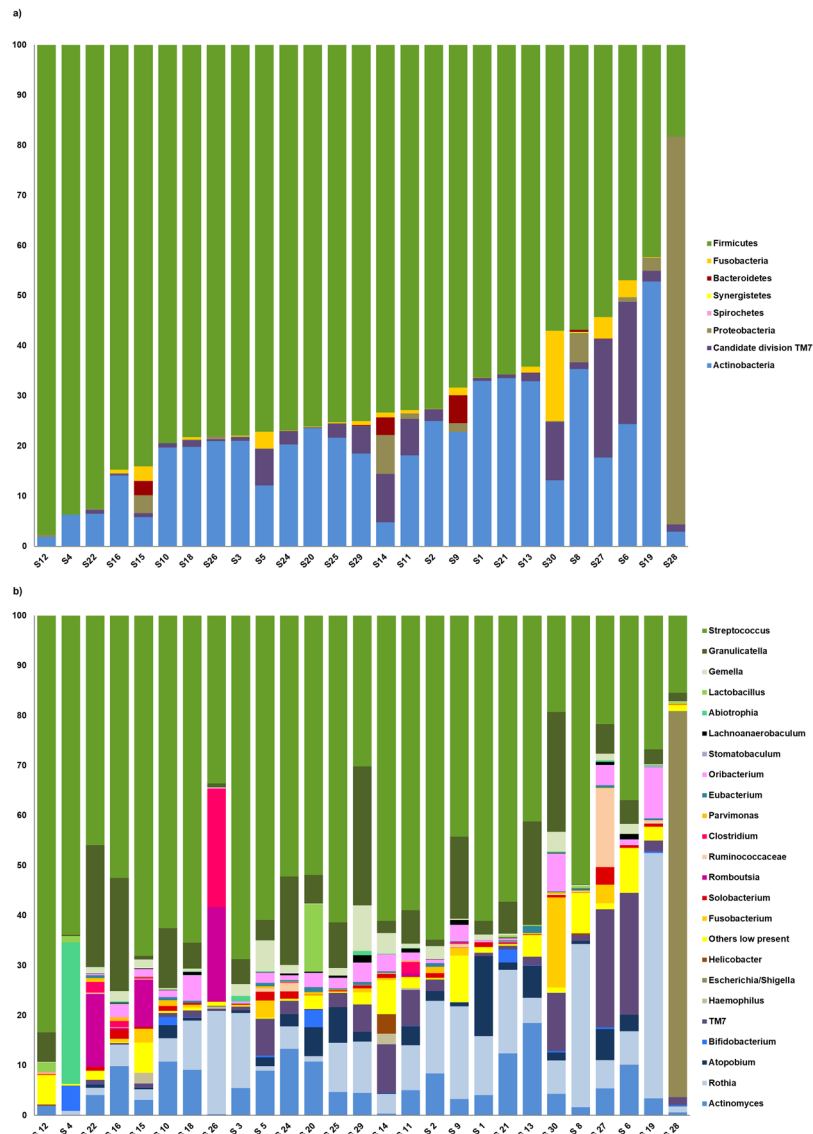


Figure 2. Phylum and genus distribution. **(a)** Phylum distribution. Relative abundances based on number of reads (%). **(b)** Distribution of most abundant genera. Relative abundances based on number of reads (%).

respectively. *Bacteroides fragilis* was only detected in one patient with 0.01% of the total reads. *Faecalibacterium prausnitzii*, considered a dominant and ubiquitous member of the human intestinal flora²⁵, was detected in only two patients with <0.1% of the total reads. Microaerophilic bacteria were represented by *Helicobacter pylori* and *Campylobacter* species, mainly *Campylobacter concisus*. Strict aerobic bacteria were present in small fractions in seven samples except for sample 19 with 2.0% of reads representing *Stenotrophomonas maltophilia*.

Discussion

In this study we have identified a group of patients that provides surgical access to the ileal lumen and that potentially represents an attractive population for future research on host-microbiota interactions. To obtain species level identification we chose a bioinformatic approach specifically developed for use in diagnostic microbiology.

The study reveals fundamental differences between the microbiota of the distal ileum and the colon. This contradicts previous studies based on samples collected by retrograde colonoscopy which indicate that the small bowel microbiota at the level of the distal ileum is similar to the colonic microbiota^{16,17}. Our findings also diverge significantly from the other small and inconsistent studies on samples from ileostomy patients or collected from healthy individuals using naso-ileal catheters^{12,14,15,18,19}. The colonic content is dominated by the Firmicutes and Bacteroidetes phyla^{4,26–28}. Although Firmicutes is the major phylum also in ileum, it is here represented mainly by facultative anaerobic species within the Bacilli class (Streptococcaceae, Lactobacillaceae, Aerococcaceae, Carnobacteriaceae) and not the strict anaerobic Firmicutes from the Clostridia class (Peptostreptococcaceae, Eubacteriaceae, Clostridiaceae, Lachnospiraceae, Ruminococcaceae) found in colon (Fig. 2b). In our samples, Clostridia account for less than 5% of the sequences whenever present. The second most dominant phylum in

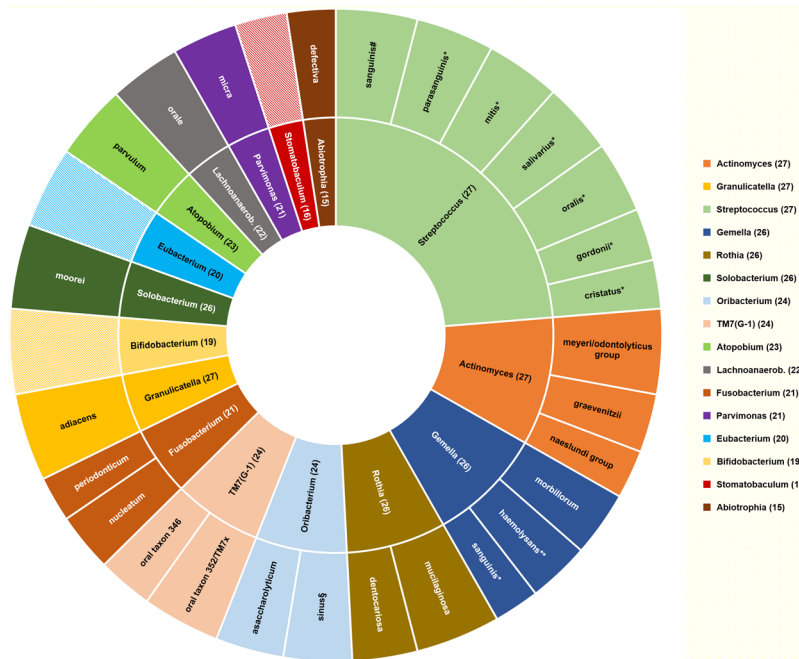


Figure 3. Core microbiota of the ileum. The inner circle represents the genus-level core microbiome defined as genera present in >50% of samples. The outer circle represents the species-level core microbiome defined as species present in >50% of samples. For the outer circle, the width of a segment is proportional to the observed incidence for that species. *Species level identification obtained with targeted *gdh* or *rpoB* Sanger sequencing. **Species level identification obtained with *rpoB* sequencing. Includes *Gemella haemolysans* sensu strictu (n = 7) and the newly proposed species *Gemella para-haemolysans* (n = 9) and *Gemella taiwanensis* (n = 6). †Only 0.7% distance to *Streptococcus sinensis*. Formally *S. sanguinis* (*S. sinensis*). ‡Only 0.7% distance to *Oribacterium parvum*. Formally *O. sinus* (*O. parvum*).

ileum was Actinobacteria whereas species from the Bacteroidetes phylum were detected only sporadically and at low levels. The microbiota of the distal ileum has a higher resemblance to the oral microbiota with dominance of viridans streptococci and high contributions of *Rothia*, *Gemella*, *Granulicatella* metagenomic sequencing was performed at the 1a and *Actinomyces* species. The major differences are higher abundances of *Neisseria*, *Haemophilus*, *Prevotella* and *Veillonella* in the oral cavity and higher levels of the TM7 phylum in the ileum^{28–32}. A study using 16S metagenomics to analyse capsule biopsies from jejunum¹⁵ also found similarity with the oral microbiota. In jejunum even the contributions of *Haemophilus*, *Prevotella* and *Veillonella* were preserved. The same study demonstrated an inverse relationship between *Streptococcus* and *Prevotella* abundances, supporting the almost complete absence of *Prevotella* in our *Streptococcus*-dominated ileal specimens.

Twenty-four individuals harboured TM7 genera (3–11 species each), representing almost a quarter of the reads in some samples. Although previously detected in the oral cavity and colon, the TM7-phylum has not previously been reported from the small intestine. The most likely explanation for this discrepancy is methodological. Microarrays designed for the gut-microbiome like the HITChip³³ or HuGChip³⁴ do not target the TM7 phylum. In sequence based studies, the lack of TM7-references in databases might have rendered TM7-derived reads unassigned or even erroneously assigned to other phyla. Only two reports exist on the successful cultivation of TM7-species from human samples. In the first study³⁵, several TM7-species were identified in mixed bacterial cultures together with *Actinomyces naeslundii* group, *Fusobacterium nucleatum*, *Parvimonas micra*, *Shuttleworthia satelles*, *Streptococcus gordonii* and *Veillonella parvula*, all frequently detected in this study as well. Dual-species biofilm experiments demonstrated synergistic biofilm formation with *F. nucleatum*, *P. micra* and *S. gordonii*. (gdh) Interestingly, the cellular shape of the TM7-species shifted between coccoid and filamentous depending on which bacterium it was co-cultivated with. In the second study³⁶, a TM7-species was recovered as an epibiont of a specific clone of *Actinomyces odontolyticus*. Together, this creates the impression of a phylum with potential for close interactions with its neighbours. Establishing the human ileum as a major TM7-reservoir represents an intriguing novel discovery that should stimulate research on this enigmatic phylum.

An interesting finding was the presence of *Helicobacter pylori* in three ileal samples, representing as much as 3.9% of the 16S sequences in sample 14. It has until now been thought to colonize only the gastric mucosa. Additional studies, including fluorescence *in situ* hybridization of biopsies, need to be undertaken to further illuminate *H. pylori* colonization of the ileum.

Colonic and ileal microbiota are reported to change with diet^{14,37,38}. A strength of our study is therefore that the collection of mucosa-near samples under standardized preoperative diet restrictions assured that microbiotas were compared under similar nutritional conditions. However, we only obtain a snapshot of the ileal microbiota under these specific conditions. Theoretically, other species could dominate after the intake of more protein-rich or fat-rich food.

Despite the almost complete absence of known intestinal illnesses in our patient population, the samples obtained are not necessarily representative of a normal ileal flora. The median age was high and most patients suffered from chronic conditions including COPD and cardiovascular diseases. These conditions themselves as well as some of the medications prescribed to treat them could affect the intestinal microbiota. Another concern is the impact of the preoperative antibiotic prophylaxis. Antibiotics are documented to impact gut microbiota^{39,40}. It can be argued that the luminal surface of the ileum is covered with a mucus layer²¹ likely to protect bacteria against both direct and systemic effects of antibiotics at least in the short-term. It can also be argued that results from DNA-based analyses are less vulnerable to short-term effects of antibiotics due to detection of DNA from non-viable bacteria and even persisting free DNA from lysed bacterial cells. However, the exact impact of the antimicrobial exposure in our population remains unknown. In the UniFrac analysis, patients that had received antibiotics during the last six months or during the last week prior to surgery did not cluster separately from the remaining population.

Patients undergoing cystectomy are currently the closest we might get to a “normal” population. Obtaining clean surgical samples for unbiased metagenomic characterization from this group, presents no major ethical dilemmas. However, high median age, underlying cancer, chronic illnesses and antibiotic prophylaxis could all impact microbiota composition. Some of the differences observed between this and previous studies could arise from these factors.

In conclusion, the distal part of ileum harbours a distinctive niche of the human gut ecosystem that differs more from the colonic than the salivary flora. These findings oppose the relevance of the bacterial flora in colon as proxy for the overall intestinal microbiota. In future studies of host-microbe interactions, it will be necessary to pay greater attention to the ileal microbiota.

Materials and Methods

Population and sample collection. Patients undergoing cystectomy with urinary diversion in Vestfold Hospital Trust (Tønsberg, Norway) are enrolled in a local quality registry based on broad informed consent. Informed consent was obtained from all subjects. Thirty patients were consecutively included in the current survey that was approved by the Regional Ethical Committee of Southern and Eastern Norway Regional Health Authority (2016/926 REK Sør-Øst D) and performed in accordance with European Association of Urology (EAU) and local guidelines and regulations.

All included patients were Caucasian. None of the patients had current gastrointestinal diseases except from patient number 20 with irritable bowel syndrome and patient number 28 with chronic obstipation (Table 1). Due to low output of reads after sequencing, three of the patients (number 7, 17 and 23) were later excluded (Supplementary Table S1). Patients were routinely fasting for solid food for 20 hours prior to surgery. They were given standardized carbohydrate drinks (PreOp Nutricia, Danone, The Netherlands) the evening before and the morning prior to surgery, and otherwise followed the Enhanced Recovery After Surgery (ERAS)-regimen⁴¹. Two patients with diabetes (patient 10 and 22) did not receive PreOp drinks. All patients received peroral ofloxacin 400 mg (25 patients) prophylaxis two hours before surgery or according to findings in preoperative urine-culture (Table 1). Metronidazole 1000 mg along with tranexamic acid 1000 mg was administered intravenously from the start of surgery as standard.

The surgeon collected the sample about 25 cm proximal to the ileocecal valve by rubbing the swab against the luminal wall, and then inserted the swab in a standardized Transwab medium (MWE, Medical Wire, England). Samples were frozen immediately at minus 70 °C for later DNA extraction.

Pre-PCR treatment of samples. The collected sample-suspension was diluted 1:2 with Qiagen ATL buffer (Qiagen, Hilden, Germany) and extracted using the “Pathogen complex kit” (Qiagen) according to the manufacturer’s instruction on a QIA Symphony platform (Qiagen). Negative controls from the relevant batches of Transwab media were extracted in the same manner.

16S metagenomic analysis. 16S metagenomic sequencing was performed at the Public Health Agency of Sweden that offers this as a commercial service for the Nordic countries. Briefly, a fragment containing the variable areas V3 and V4 was amplified using forward-primer 5'-CGG-CCC-AGA-CTC-CTA-CGG-GAG-GCA-GCA-3' and reverse-primer 5'-GCG-TGG-ACT-ACC-AGG-GTA-TCT-AAT-CC-3'⁴². An Ion Chef instrument (ThermoFisher, Foster City, California) was used for automated library preparation and bidirectional sequencing was done using the IonS5 sequencer (ThermoFisher). Barcode separated FASTQ-files were processed individually using the RipSeq NGS software⁴³ (Pathogenomix, Santa Cruz, California). Reads shorter than 250 base pairs were removed before de-novo clustering into OTUs using a similarity threshold of 99%. OTUs containing less than 10 sequences were rejected. For each of the remaining OTUs the most representative sequence variant was used for a BLAST-search against the curated “RipSeq. 16S human pathogen”-database (Pathogenomix) that currently contains more than 2500 non-redundant references for clinically relevant bacteria and commensals. OTUs that did not obtain a species-level match were re-analysed against GenBank (nr/nt-database) and the “Human oral microbiome” database (www.homd.org) assuring the highest possible level of identification for all clusters. RipSeq NGS flags the quality of a BLAST-result based on adjustable interpretation criteria. For unambiguous species-level identification, the Clinical and Laboratory Standard Institute (CLSI) guidelines for 16S sequence interpretation recommends $\geq 99\%$ homology with a high-quality reference combined with a minimum distance of $>0.8\%$ to the next alternative species⁴⁴. These recommendations are based on the V1-V3 segment. Due to the lower overall variability in the V3-V4 segment we adopted a more stringent cut off of $\geq 99.3\%$ while retaining the minimum distance to the next species at $>0.8\%$. OTUs with a similarity $>99.3\%$ but with a distance to next species of $\leq 0.8\%$ was assigned to a group of species (group-level identification) whereas OTUs with a similarity of 97–99.2% was assigned to the genus level. When appropriate, the provisional taxonomy developed by the

Human oral microbiome project (www.homd.org) was used for hitherto unnamed species⁴⁵. Chimeras and any species/OTU found in the negative controls were rejected. A comparison between results obtained by QIIME, the software package typically used in 16S rRNA metagenomic studies⁴⁶ and RipSeq NGS on a commercial Mock-up community is provided in supplementary table S5. This comparison demonstrates the better performance of the RipSeq NGS software in obtaining species level identification.

Comparative analyses were performed using the QIIME. Alfa-diversities were measured using the Shannon-index and weighted and unweighted UniFrac analyses were used to calculate beta-diversities^{47,48}.

Sanger sequencing of alternative genes. For discrimination among bacteria with too similar 16S-amplicons, we designed species-specific PCRs for genes with higher mutations rates (Supplementary Table S6). Due to limited volumes of DNA-eluate we prioritized species that consistently came out among the top-scoring organisms in frequently encountered ambiguous 16S clusters. For streptococci, the gene for glutamate dehydrogenase (*gdh*) was selected and we applied the taxonomical modifications proposed by Jensen *et al.*⁴⁹. Within the mitis group, we targeted *Streptococcus oralis* (including ssp. *dentisani*, *oralis* and *tigurinus*) and *Streptococcus mitis*. For *Streptococcus infantis* a single *gdh*-PCR could not be designed due to varying sequence patterns. Recent whole genome-based comparisons of strains within the *S. infantis* clade expose significant unresolved taxonomical matters⁴⁹. Within the parasanguinis group, PCRs were designed for *Streptococcus cristatus* (synonym: *S. oligofermentans*), *S. parasanguinis* and *S. gordonii*. *Streptococcus salivarius* and *S. vestibularis* were targeted by a shared PCR and discriminated by the nucleotide sequence. *S. sanguinis* was unambiguously identified based on the 16S rRNA-gene.

For discrimination within the *Gemella haemolysans/sanguinis* cluster the gene for RNA-dependent polymerase beta (*rpoB*) was used. In addition to specific PCRs for *G. haemolysans* and *G. sanguinis* we designed a shared PCR for *G. haemolysans* and the newly proposed species *Gemella parahaemolysans* and *Gemella taiwanensis*⁵⁰.

All primers were aligned against available references in GenBank (nr/nt and wgs databases) to assure coverage of all known sequence variants. The PCR-products from positive reactions were sequenced using Sanger-sequencing and aligned against the GenBank nr/nt and wgs databases for confirmation. Due to lack of established cut-offs for a valid species level identification all species level assignments were supported by a pairwise comparison of the alignment table in GenBank using the “distance tree of results” function.

Primer sequences, primer concentrations and PCR conditions are listed in Supplementary Table S6. SYBR-green real-time PCRs were performed in 25 µl reaction tubes on a SmartCycler real-time apparatus (Cepheid, Sunnyvale, California). PCR mixtures consisted of 12.5 µl ExTaq SYBR master mix (TaKaRa, Otsu, Japan), 0.4 or 0.6 µM of each primer (corresponding to 1.0 or 1.5 µl from a 10 µl stock solution), 8.0 or 8.5 µl PCR-grade water (depending on primer-volume) and 2 µl extracted DNA. Thermal profiles included an initial polymerase activation step at 95 °C for 30 seconds, followed by 45 cycles of 95 °C for 10 seconds, PCR-specific annealing temperature for 10 seconds and 72 °C for 20 seconds.

The products from positive PCR reactions were spun out of the SmartCycler reaction tubes into a 1.5-ml Eppendorf tube and cleaned up using ExoSAP-IT enzymatic degradation (Affymetrix, Santa Clara, California). Cycle sequencing was run for 28 cycles with annealing temperature 50 °C for all amplicons. Sanger sequencing was performed using an ABI prism 1.1 Big Dye sequencing kit and an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, California 16S).

Data availability. Source data of this study are available from the corresponding author upon reasonable request. Not all patient data are publicly available due to restrictions from the Regional Ethical Committee (REK Sør-Øst D).

References

- Sonnenburg, J. L. & Backhed, F. Diet-microbiota interactions as moderators of human metabolism. *Nature* **535**, 56–64 (2016).
- Collins, S. M., Surette, M. & Bercik, P. The interplay between the intestinal microbiota and the brain. *Nature reviews. Microbiology* **10**, 735–742 (2012).
- Brestoff, J. R. & Artis, D. Commensal bacteria at the interface of host metabolism and the immune system. *Nature immunology* **14**, 676–684 (2013).
- Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nature reviews. Microbiology* **14**, 20–32 (2016).
- Raoult, D. & Henrissat, B. Are stool samples suitable for studying the link between gut microbiota and obesity? *European journal of epidemiology* **29**, 307–309 (2014).
- Raoult, D. The study of microbiota needs both microbiologists and medical doctors. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **23**, 500–501 (2017).
- Helander, H. F. & Fandriks, L. Surface area of the digestive tract - revisited. *Scandinavian journal of gastroenterology* **49**, 681–689 (2014).
- Johansson, M. E., Sjovall, H. & Hansson, G. C. The gastrointestinal mucus system in health and disease. *Nature reviews. Gastroenterology & hepatology* **10**, 352–361 (2013).
- Shan, M. *et al.* Mucus enhances gut homeostasis and oral tolerance by delivering immunoregulatory signals. *Science (New York, N.Y.)* **342**, 447–453 (2013).
- Moran, C., Sheehan, D. & Shanahan, F. The small bowel microbiota. *Current opinion in gastroenterology* **31**, 130–136 (2015).
- Santaolalla, R. & Abreu, M. T. Innate immunity in the small intestine. *Current opinion in gastroenterology* **28**, 124–129 (2012).
- van den Bogert, B., Meijerink, M., Zoetendal, E. G., Wells, J. M. & Kleerebezem, M. Immunomodulatory properties of *Streptococcus* and *Veillonella* isolates from the human small intestine microbiota. *PLoS one* **9**, e114277 (2014).
- Li, D. *et al.* Microbial Biogeography and Core Microbiota of the Rat Digestive Tract. *Scientific reports* **8**, 45840 (2017).
- Zoetendal, E. G. *et al.* The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME j* **6**, 1415–1426 (2012).
- Dlugosz, A. *et al.* No difference in small bowel microbiota between patients with irritable bowel syndrome and healthy controls. *Scientific reports* **5**, 8508 (2015).

16. Wang, M., Ahrne, S., Jeppsson, B. & Molin, G. Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS microbiology ecology* **54**, 219–231 (2005).
17. Dave, M. *et al.* A randomised trial of sheathed versus standard forceps for obtaining uncontaminated biopsy specimens of microbiota from the terminal ileum. *Gut* **60**, 1043–1049 (2011).
18. Booijink, C. C. *et al.* High temporal and inter-individual variation detected in the human ileal microbiota. *Environmental microbiology* **12**, 3213–3227 (2010).
19. van den Bogert, B. *et al.* Diversity of human small intestinal Streptococcus and Veillonella populations. *FEMS microbiology ecology* **85**, 376–388 (2013).
20. Hayashi, H., Takahashi, R., Nishi, T., Sakamoto, M. & Benno, Y. Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *Journal of medical microbiology* **54**, 1093–1101 (2005).
21. Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L. P. & Lochs, H. Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *Journal of clinical microbiology* **43**, 3380–3389 (2005).
22. Ahmed, S. *et al.* Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Applied and environmental microbiology* **73**, 7435–7442 (2007).
23. Gevers, D. *et al.* The treatment-naïve microbiome in new-onset Crohn's disease. *Cell host & microbe* **15**, 382–392 (2014).
24. Salipante, S. J. *et al.* Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Applied and environmental microbiology* **80**, 7583–7591 (2014).
25. Miquel, S. *et al.* Faecalibacterium prausnitzii and human intestinal health. *Current opinion in microbiology* **16**, 255–261 (2013).
26. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science (New York, N.Y.)* **308**, 1635–1638 (2005).
27. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
28. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
29. Wang, K. *et al.* Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus. *Scientific reports* **6**, 22943 (2016).
30. Yang, F. *et al.* Saliva microbiomes distinguish caries-active from healthy human populations. *ISME j* **6**, 1–10 (2012).
31. Zaura, E. *et al.* On the ecosystemic network of saliva in healthy young adults. *ISME j* **11**, 1218–1231 (2017).
32. Zhou, J. *et al.* Exploration of Human Salivary Microbiomes—Insights into the Novel Characteristics of Microbial Community Structure in Caries and Caries-Free Subjects. *PLoS one* **11**, e0147039 (2016).
33. Rajilic-Stojanovic, M. *et al.* Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environmental microbiology* **11**, 1736–1751 (2009).
34. Tottey, W. *et al.* The human gut chip “HuGChip”, an explorative phylogenetic microarray for determining gut microbiome diversity at family level. *PLoS one* **8**, e62544 (2013).
35. Soro, V. *et al.* Axenic culture of a candidate division TM7 bacterium from the human oral cavity and biofilm interactions with other oral bacteria. *Applied and environmental microbiology* **80**, 6480–6489 (2014).
36. Bor, B. *et al.* Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont Actinomyces. *Microbial ecology* **71**, 243–255 (2016).
37. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
38. Walker, A. W. *et al.* Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME j* **5**, 220–230 (2011).
39. Sullivan, A., Edlund, C. & Nord, C. E. Effect of antimicrobial agents on the ecological balance of human microflora. *The Lancet. Infectious diseases* **1**, 101–114 (2001).
40. Modi, S. R., Collins, J. J. & Relman, D. A. Antibiotics and the gut microbiota. *The Journal of clinical investigation* **124**, 4212–4218 (2014).
41. Cerantola, Y. *et al.* Guidelines for perioperative care after radical cystectomy for bladder cancer: Enhanced Recovery After Surgery (ERAS((R))) society recommendations. *Clinical nutrition (Edinburgh, Scotland)* **32**, 879–887 (2013).
42. Kommedal, O., Kvello, K., Skjastad, R., Langeland, N. & Wiker, H. G. Direct 16S rRNA gene sequencing from clinical specimens, with special focus on polybacterial samples and interpretation of mixed DNA chromatograms. *Journal of clinical microbiology* **47**, 3562–3568 (2009).
43. Kommedal, O. *et al.* Massive parallel sequencing provides new perspectives on bacterial brain abscesses. *Journal of clinical microbiology* **52**, 1990–1997 (2014).
44. Petti, C. A. *et al.* Interpretive Criteria for Identification of Bacteria and Fungi by DNA Target Sequencing; Approved Guideline. CLSI document MM18-A. *Clinical and Laboratory Standards Institute* **28**(12), 88 (2008).
45. Dewhirst, F. E. *et al.* The human oral microbiome. *Journal of bacteriology* **192**, 5002–5017 (2010).
46. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335–336 (2010).
47. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology* **71**, 8228–8235 (2005).
48. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS one* **5**, e9490 (2010).
49. Jensen, A., Scholz, C. F. & Kilian, M. Re-evaluation of the taxonomy of the Mitis group of the genus Streptococcus based on whole genome phylogenetic analyses, and proposed reclassification of Streptococcus dentisani as Streptococcus oralis subsp. dentisani comb. nov., Streptococcus tigurinus as Streptococcus oralis subsp. tigurinus comb. nov., and Streptococcus oligofermentans as a later synonym of Streptococcus cristatus. *International journal of systematic and evolutionary microbiology* **66**, 4803–4820 (2016).
50. Hung, W. C. *et al.* Gemella parahaemolyans sp. nov. and Gemella taiwanensis sp. nov., isolated from human clinical specimens. *International journal of systematic and evolutionary microbiology* **64**, 2060–2065 (2014).

Acknowledgements

This work was supported by Department of Microbiology, Vestfold Hospital Trust, Norway and Department of Microbiology, Haukeland University Hospital, Norway. We especially thank Inger Lill Anthonisen for the initial technical support and Dag Gullan for linguistic assistance.

Author Contributions

H.C.V., E.S.H., T.S. and N.G. initiated the study. H.C.V. and Ø.K. were responsible for the study design and E.S.H. and A.H. enrolled patients and collected samples. H.C.V. and Ø.K. performed the 16S metagenomic data analysis. Ø.K. designed and performed the Sanger-sequencing experiments. H.C.V., Ø.K. and E.U. wrote the manuscript with inputs from all the authors.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-23198-5>.

Competing Interests: ØK has contributed to the development of the RipSeq NGS software and is a minor shareholder in Pathogenomix. The remaining authors declare no competing financial interests.

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018